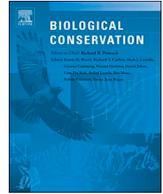




ELSEVIER

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Comparing an automated high-definition oblique camera system to rear-seat-observers in a wildlife survey in Tsavo, Kenya: Taking multi-species aerial counts to the next level

Richard Lamprey^{a,*}, Frank Pope^a, Shadrack Ngene^b, Michael Norton-Griffiths^c, Howard Frederick^d, Benson Okita-Ouma^a, Iain Douglas-Hamilton^a

^a Save the Elephants, P.O. Box 54667-00200, Nairobi, Kenya

^b Kenya Wildlife Service, P.O. Box 40241-00100, Nairobi, Kenya

^c P.O. Box 15227-00509, Nairobi, Kenya

^d Tanzania Wildlife Research Institute, PO Box 661, Arusha, Tanzania

ARTICLE INFO

Keywords:

Africa
Elephants
Aerial count
Wildlife
Multi-species
Census
Observers
Cameras
Imaging
Bias

ABSTRACT

In aerial wildlife counts, human observers often fail to detect animals. We conducted a multi-species sample-count in Tsavo National Park, Kenya, with traditional rear-seat-observers (RSOs) and an automated 'oblique-camera-count' (OCC) imaging system to compare estimates of 23 wildlife species derived from these two survey methods. An aerial Total Count of elephant, buffalo and giraffe, conducted a month previously, provided a further comparison. In the Tsavo Core (9560 km²), which harbours 80% of Tsavo's elephants, the OCC system acquired 81 000 images for interpretation, of which 67 000 were obtained in parallel with RSO-counting along 3004 km of flight line. The Tsavo outer blocks (24 171 km²) were surveyed using the OCC system without RSOs to acquire a further 84 000 images. A random sample of 11 553 images were re-interpreted to derive species-specific probabilities of detection and correction factors. Using 'Jolly II', non-parametric and Bayesian analyses, and applying correction factors, we demonstrate that the RSOs did not detect 14% of elephants, 60% of giraffe, 48% of zebra and 66% of the large antelopes. For comparison, the Total Count observers did not detect 27% of elephant, 33% of buffalo, 57% of giraffe and 85% of carcasses. The OCC method raises the elephant population estimate to 16 681 ± 4047 (95% ci) from the 12 722 counted in the Total Count ($Z = 1.917$, $p = .0276$). These results suggest that RSO-based methods have significantly undercounted wildlife populations. To align with improved counting methods, previous results need to be re-calibrated.

1. Introduction

In eastern and southern Africa the standard method for counting wildlife and livestock over large areas is the systematic reconnaissance flight (SRF), whereby light aircraft are flown along systematically spaced transects over the terrain at low level whilst 'rear-seat-observers' (RSOs) count animals in sample strips defined on each side of the aircraft (Andere, 1981; Caughley, 1977; Craig, 2012; Grimsdell and Westley, 1981; Norton-Griffiths, 1978; Ogotu et al., 2016; Ottchilo et al., 2000; PAEAS, 2014). The 'strip-transects' are the sample units, and the species population estimate is derived from the ratio of the total area to the sampled survey area (Jolly, 1969). An alternative approach, often applied in South Africa, Australia and the USA, is the 'line-transect' method, where the population is estimated as a function of the

distance of animals from a survey line (Buckland et al., 2004; Burnham et al., 1985; Eberhardt, 1978; Kruger et al., 2008).

Observers in both strip-transects and line-transect sampling often fail to detect animals with the result that populations are underestimated (Caughley, 1974; Cook and Jacobson, 1979; Gasaway et al., 1986; Jachmann, 2002; Lee and Bond, 2016; Norton-Griffiths, 1976; Pollock and Kendall, 1987; Schlossberg et al., 2016; Whitehouse et al., 2001). Animals might not be detected either because they are overlooked, or because they are 'unavailable for detection', being in dense vegetation cover for terrestrial animals (Bayliss and Yeomans, 1989; Jachmann, 2001; Jacques et al., 2014) or underwater for marine mammals (Marsh and Sinclair, 1989; Pollock et al., 2006). Determining the number of unavailable animals requires that surveys are implemented in areas where the 'true' population is known (Jachmann,

* Corresponding author.

E-mail address: lamprey.richard@gmail.com (R. Lamprey).

<https://doi.org/10.1016/j.biocon.2019.108243>

Received 11 December 2018; Received in revised form 4 September 2019; Accepted 9 September 2019

Available online 16 December 2019

0006-3207/ © 2019 Elsevier Ltd. All rights reserved.

2002; Lubow and Ransom, 2016; Tracey et al., 2008; Whitehouse et al., 2001), or where availability is modelled through the detection of radio-tagged or marked animals during sample counts (Jacques et al., 2014; Zabransky et al., 2016).

In consideration of 'available animals', the SRF method is subject to a wide range of biases relating to survey factors such as aircraft type, ground speed, altitude, sample strip width, and observer's experience, interest and fatigue (Beasom et al., 1981; Caughley, 1974; Fleming and Tracey, 2008; Jachmann, 2001; Norton-Griffiths, 1976; Pennycuik and Western, 1969; Steinhilber and Samuel, 1989). 'Environmental factors' also cause biases, including animal size, group size, colour against background, multi-species groups, reaction of animals to the aircraft, species habitat preference, vegetation type, cover and seasonal phenology, sun-angle, illumination and shadowing, topography and weather conditions (Anderson and Lindzey, 1996; Griffin et al., 2013; Jachmann, 2002; Jacques et al., 2014; Lubow and Ransom, 2016; McConville et al., 2009; Ransom, 2012; Rice et al., 2009; Strobel and Butler, 2014; Wal et al., 2011).

All of these factors are influenced by one major limitation of the SRF method, this being the short window of time that an observer has to scan the terrain and count. This effect can be somewhat mitigated if the counting platform can hover, but helicopters are often unaffordable or unavailable, especially in Africa where large-area counts are needed. Therefore, fixed-wing aircraft with high-wings are used, with cheaper operating costs. In SRF surveys conducted to normal standards with fixed-wing aircraft (Caughley, 1977; Craig, 2012; Norton-Griffiths, 1978; PAEAS, 2014) the aircraft is flown at low level, for example 350 ft (107 m), above ground level and at 160–190 km/hr. The aircraft must be operated at an airspeed no slower than about 40% to 60% above the stall speed to remain safe, and there can be no deviation from the flightpath without distorting the sampling frame. RSOs have the continuously changing elements of the landscape mosaic – woodland patches, waterholes, open glades – in sight for just 5–7 seconds, and within this 'scan-time' the RSO must detect and count animals within the strip. RSO training is essential, involving animal identification, rapid enumeration, recording the estimate and referencing the location in some prescribed way (Craig, 2012; PAEAS, 2014). If large herds are encountered, they must be estimated and photographed for later counting, and the photograph frame code and number also recorded. In a wildlife-rich environment, the workload of an RSO is therefore high, whilst in areas of low wildlife density fatigue and boredom set in; both scenarios lead to bias in estimation (Caughley, 1974; Fleming and Tracey, 2008; Jachmann, 2001; Norton-Griffiths, 1976).

For these reasons it has long been proposed that counts can be improved by replacing observers with camera-only systems (Caughley, 1974; Leedy, 1948; Siniff and Skoog, 1964). Probably the first rigorous study to apply vertical aerial photography was a stratified random count of wildebeest in the Serengeti in Tanzania (Norton-Griffiths, 1973). Here, a seasonal stratum was defined that encompassed most of the population, which was then sampled by vertical 'aerial-point-sampling' (APS) photography along transects. Since then, APS has been used in a number of large-area or large-population surveys, from regular wildebeest counts in the Serengeti (Hopcraft et al., 2015; TAWIRI, 2010), to caribou in Canada (Couturier et al., 1994) and desert antelopes in Mongolia (Norton-Griffiths et al., 2015). Such methods now have great potential applications for UAV-based surveys, when UAV endurance and lightweight camera systems can be improved (Vermeulen et al., 2013). However, whilst vertical imaging such as APS may be used in open areas with low tree canopy cover and/or high contrast backgrounds, for example open plains, deserts or snowfields, it is not suitable for areas of higher cover where herds may be clustered under tree canopies or in thick cover. Here, the oblique viewing approach, such as the strip-transect count, remains the most suitable method. The potential of oblique continuous imaging to supplement RSOs has only recently been explored, the first reported example being in a marine environment off Greenland for a single species, narwhal

(*Monodon monoceros*), where it was concluded that in this marine environment and for a single species, there was no significant differences in RSO-based and image-based estimates (Bröker et al., 2019).

To date there have been no reported experiments in complex terrestrial environments in Africa or elsewhere to acquire continuous oblique imagery of the sample strips normally scanned by RSOs and interpreting this imagery for multiple species. This largely removes the scan-time constraint of the SRF method, since interpreters can spend as long as they need on any one scene to ascertain if animals are present. We ask the question: if interpreters have an unlimited time to detect animals within the sample strip, do they find more animals than aerial observers who have just seconds to detect animals within the same strip? This paper presents the results of an experimental multispecies wildlife count of a large protected area in Kenya, where we operate an 'oblique-camera-count' (OCC) system simultaneously with RSO-counting. We visually interpreted the resulting 165 000 images to determine how effectively RSOs have detected animals. We also compare the OCC estimates with a total count conducted a month previously. We determine whether bias is significant in RSO enumeration of available animals and explore image-based methods for improved counting.

2. Methods

The OCC survey method was adapted from techniques developed in Uganda (Lamprey, 2016). For Tsavo, a standard RSO-based SRF was implemented in parallel with camera systems imaging the same strips, and the number of animals counted by both methods was compared. The survey test area in Kenya was the 'Tsavo Conservation Area', a vast expanse of semi-arid grassland and bushland where wildlife densities are relatively low. We selected Tsavo since the OCC method could be tested in conjunction with two other wildlife counts that were already planned in early 2017, thereby reducing aircraft operation costs. Tsavo is a challenging area for aerial counting; large areas of this arid landscape are monotonous, and at low flight levels the ambient air temperature often exceeds 35 °C, with accompanying strong winds and turbulence. Recent total count and SRF sample count estimates for elephants in Tsavo have shown significant variation (Chase et al., 2014; Kyale et al., 2014; Ngene et al., 2011, 2017), and more accurate and precise population estimates are needed for elephants and a wider range of species to determine whether current conservation measures are effective.

The surveys were conducted in collaboration with the government's wildlife agency, the Kenya Wildlife Service (KWS). Counting methods were compared across three surveys conducted in February–March 2017. The first, the 2017 Tsavo Total Count was conducted by KWS: as its name implies, it was conducted to count all individuals of certain key species. The second survey ('TS1'), was a simultaneous OCC and RSO-based sample count of the Tsavo Core area. The third ('TS2'), was an OCC-only sample count of the outer Tsavo strata. The OCC counts generated 165 000 images for interpretation, of which 67 000 images in TS1 coincided with simultaneous RSO strip-counting from the same aircraft.

For the RSO count, methods follow current guidelines of RSO-based SRF counts for elephants (Craig, 2012; PAEAS, 2014). Calibration exercises were carried out to determine the width of the imaged strips as a function of aircraft flying height. For our implementation of the OCC method, image interpretation was carried out by a team of 8 interpreters, who visually scanned the imagery to identify and count animals, and entered the data into a database. We test for bias between interpretation teams, and then compare transect counts and population estimates derived by the RSO sample count, the OCC sample count and the Total Count.

Since aerial wildlife counts are often implemented by pilot-biologists, we use the standard aviation units of feet for altitude and knots for airspeed since the measurement instruments in the aircraft are calibrated in these units. Acronyms used in this paper are indicated in

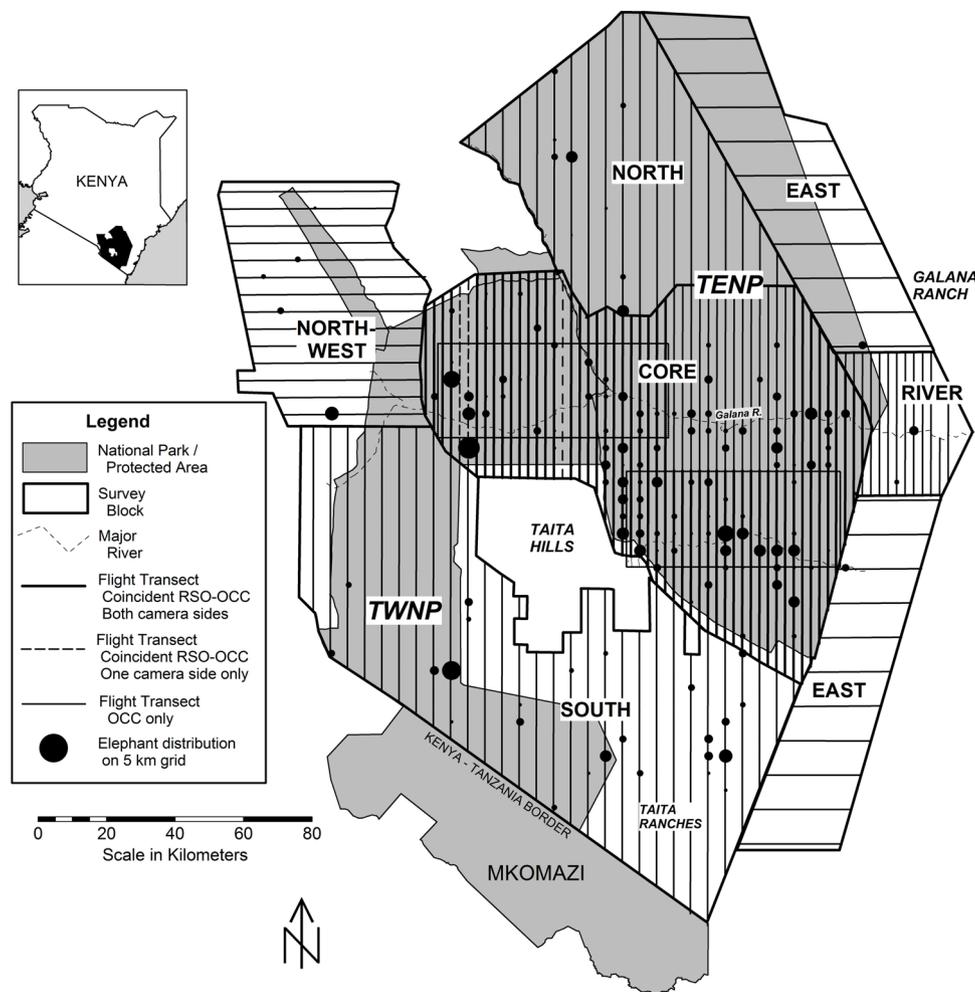


Fig. 1. Map of the Tsavo survey area, with protected areas, strata and survey transects. The coincident RSO-OCC transects are shown in the Core. TENP and TWNP are Tsavo East and Tsavo West National Parks respectively. The inset map shows the location in Kenya. Mkomazi in Tanzania was not surveyed, but few elephants were recorded here in the total count. For clarity, the elephant distribution across the ecosystem is plotted on the ecosystem-wide 5 km grid, rather than the Core only 2.5 km grid. The rectangles in Core stratum delimit the equal length transects used for RSO-OCC count comparisons.

Appendix A. To open wildlife agency minds long-steeped in the ‘sanctity of p-values’ to the concepts of inference, likelihood and weights-of-evidence (Burnham and Anderson, 2014), we undertake simple Bayesian testing to compliment the 50-year old Jolly II analysis and standard parametric methods for comparing results from aerial counts.

2.1. Survey area

The survey covered 33 731 km² of Tsavo East and West National Parks, the Chyulu Hills National Park and edges of Galana Ranch to the east, see Fig. 1. This vast region of semi-arid *Acacia-Commiphora* bushland and grassland was thought to harbour an estimated 12 500 elephants (Ngene et al., 2011; Thouless et al., 2016), which constitute about 60% of Kenya’s elephants. This population was afflicted by drought and poaching in the 1970s and 80s (Cobb, 1976; Olindo et al., 1988; Ottichilo, 1987), reaching a low point of about 7000 in 1988 and recovering slowly since then. Most of Tsavo’s elephants are now located within a central zone, the Tsavo Core, which for this survey covered 9560 km² and included elephant areas north of the Galana River. The Tsavo Core was surveyed in TS1, which combines the OCC and RSO methods. Transects were spaced at 2.5 km intervals to provide a high intensity ~12% sample according to PAEAS guidelines (PAEAS, 2014), and were orientated north-south across the Galana River to minimize variance. Surrounding the Core are the North, North-West, South and East strata, framed as in the previous sample count of 2014 (Chase

et al., 2014). These strata were surveyed using the OCC method only, with east-west transects spaced at 5 km apart for a medium intensity (~6%) sample, and 10 km apart in the elephant-sparse East stratum. The TS1, TS2 and Total Count surveys were conducted at the end of a long dry spell when previous ‘short rains’ of October–November 2016 had been below average.

2.2. Aircraft and camera specifications

For TS1 a 6-seat Cessna 206 aircraft (registration 5Y-AKP) was used, equipped for an RSO count with intercom systems and counting rods. Through standard ground measurements procedures with observers seated in their comfortable positions the counting rods were mounted to define a strip of 150 m each side of the aircraft (Norton-Griffiths, 1978; PAEAS, 2014). In clearing the aircraft undercarriage, the counting rod alignment resulted in the RSO viewing angle of 57° from vertical through the centre line of the markers. The cameras were mounted at this angle to align to the RSO counting strip. The cameras were unmodified Nikon D3200 digital-single-lens-reflex cameras of 24-megapixel CMOS sensors, with 18–70 mm zoom lenses. Images were captured in standard JPG format, since higher definition RAW formats are too large for data storage in multi-hour survey missions.

To capture the counting strip, the required lens focal length was calculated as 42 mm, this being a function of field-of-view (Edin, 2014), camera inclination angle (57°) and height-above-ground-level (HAGL)

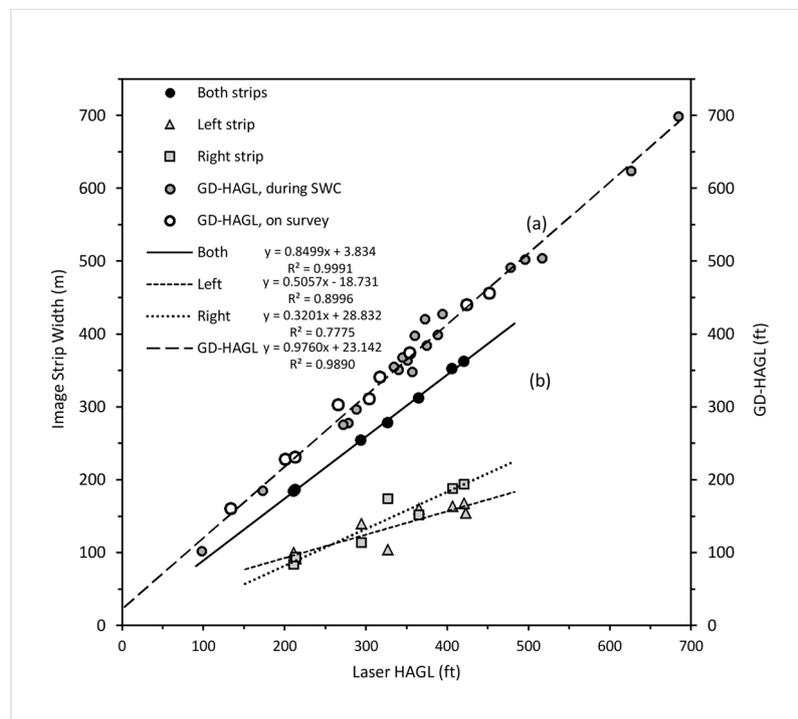


Fig. 2. (a) Correlation of GD-HAGL with laser HAGL during the strip-width-calibration (closed circles) and during the TS1 survey (open circles), and (b) correlation of OCC image strip width with laser-HAGL for left, right, and both cameras combined.

(350 ft). This combination resulted in a theoretical ground-sampling-distance (Neumann, 2008) of 1.5 cm at the frame inner edge, increasing to 2.7 cm at the outer edge. In practice this ‘resolution’ is reduced in JPG image formats by about 50% to 3 cm and 5.4 cm respectively by the camera sensor Bayer array (Aerial-Survey-Base, 2014; Bull, 2014), and possibly further to ~8 cm by the forward velocity of the aircraft (O’Connor et al., 2017). The cameras were mounted on a frame at the same eye level as the RSOs, with lenses protruding through open ventilation ports in the third seat-row windows. In this position they did not interfere with the counting operations of the RSOs seated in the second row of seats. No inertial measurement unit for roll or pitch measurement was installed, since the cramped cabin arrangement with four crew and much equipment precluded this option and emphasis was in simplicity of operation. With camera ISO adjusted to \leq ISO 400, images were acquired at shutter speed of \leq 1/1000th of a second (\leq 0.001 s). Images were acquired at 2 s intervals to provide overlapping coverage without gaps at the relatively slow groundspeed of 89 knots \pm 5.3 knots (SD, $n = 78\ 830$) (165 km/h) to conform to requirements for RSO counting (Craig, 2012; PAEAS, 2014).

For TS2 we used a Cessna 182 aircraft (registration 5Y-ATS) with a large camera port for the camera installation. This was a single-crew operation without observers, where camera functions, GPS recording and height management were under the control of the pilot (the lead author). Here, the flying height, camera inclination and focal length specifications followed those of the Uganda surveys (Lamprey, 2016), with HAGL set at 500 ft (152 m) and cameras inclined at 45°. The image acquisition was also set at 2 s interval, and groundspeed at 105 knots (194 km/h), to acquire continuous strips in this specific configuration. For TS2 the lens focal length of 35 mm was set to define a strip-width of 130–140 m, with the more exact strip being defined at strip-width-calibration. The theoretical ground-sampling-distance for frame inner and outer edges was determined as 2.1 cm and 3.3 cm respectively (Aerial-Survey-Base, 2014), with Bayer degradation to approximately 4.2 and 6.6 cm respectively.

The cameras were set up to operate automatically for at least 4 h with external power supplies, intervalometers and data storage cards

sufficient for 8000 images. At the end of the survey, all images were renamed using a single code structure defining date, time, aircraft, camera side (left/right), camera folder and frame number. Before each mission, the camera clocks were synchronized to the aviation Global Positioning System (GPS) receiver clock to within 1 s for later georeferencing in Universal Transverse Mercator (UTM) coordinates against GPS tracklogs, in order to create a geographical information system (GIS) shapefile of all image locations (‘photopoints’). In TS1, due to camera problems linked to the high ambient air temperature, two days of transects had to be repeated, but this time as OCC-only without the RSOs. Out of the total of 81 000 images acquired for TS1, 67 000 coincided with RSO recording. These simultaneous RSO-OCC transects are shown in Fig. 1.

2.3. Height control and strip-width calibration

For TS1, the HAGL of 350 ft (107 m) was adopted according to CITES-MIKE and PAEAS guidelines for RSO surveys (Craig, 2012; PAEAS, 2014), and the pilot maintained HAGL with reference to a logging laser altimeter recording at 1 s intervals. The logger failed after 5 days of flying, but standard 30 s visual recording of the laser read-out by the front-seat-observer was maintained throughout the simultaneous RSO survey according to PAEAS guidelines. This laser HAGL record was used in the RSO-survey Jolly II analysis and calculation of densities for paired transect analysis.

In TS2, height was controlled using the GPS-DEM method (Lamprey, 2016) which does not depend on the installation of laser or radar altimeters; only an aviation-quality GPS is needed. GPS-DEM measures HAGL as the difference in elevation above mean-sea-level between the aircraft navigation GPS and the terrain below, the latter determined from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) (SRTM version 4.1) (Jarvis et al., 2008). We refer to this HAGL estimator as GPS/DEM-HAGL, or ‘GD-HAGL’. This method is made possible by recent developments in GPS technology, the GPS Standard Position Service (Kaplan and Hegarty, 2006; NTSB, 2018), the estimated accuracy of the base SRTM model (Mukul et al., 2015;

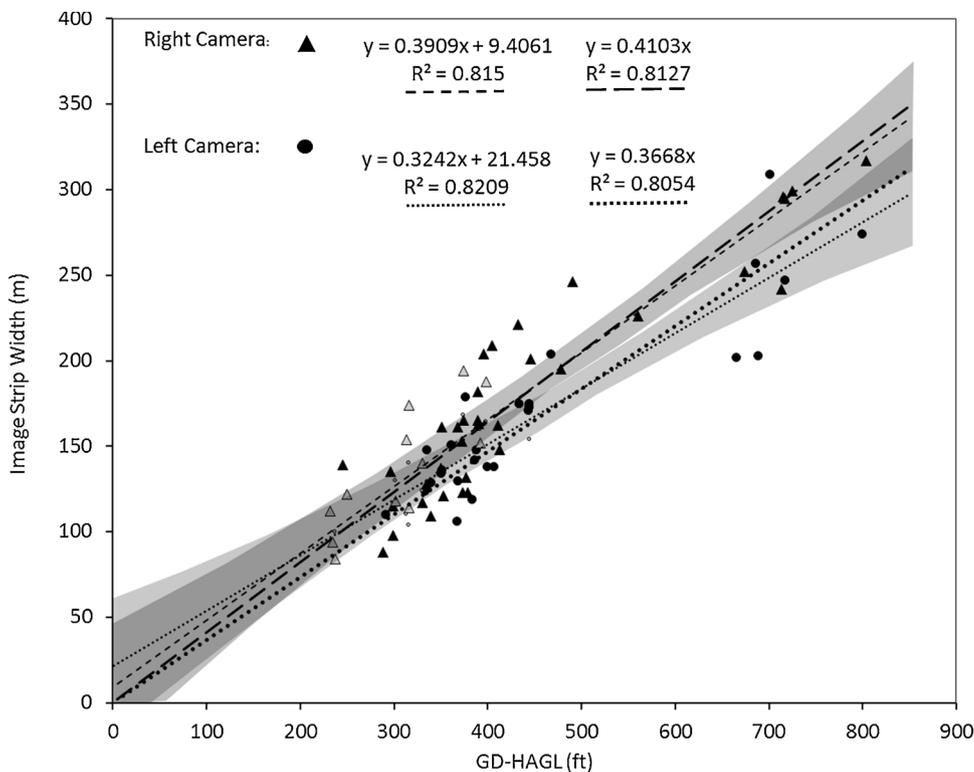


Fig. 3. The image strip-width-calibration (SWC) for left and right cameras using GD-HAGL. Open symbols indicate data received in the SWC, closed black symbols indicate data derived from image footprints plotted on Google Earth. The plot shows the 95% CI of the left and right camera regressions, based on the slope and intercept. The regressions through the origin were adopted for the transect sample area analysis.

Rodríguez et al., 2006), and the incorporation of the EGM96 and EGM08 geoid into the model (Lemoine et al., 1998; Pavlis et al., 2012). GPS accuracy, expressed as low ‘dilution of precision’, is particularly high in central and eastern Africa (NTSB, 2018).

Strip-width-calibration (SWC) was conducted according to standard methods (PAEAS, 2014) with flight overpasses at increasing HAGL across ground-marks laid down at 20 m intervals on the base runway at Voi airfield. HAGL was recorded from the ‘zeroed’ pressure altimeter, laser altimeter, and, in post-hoc analysis, by GD-HAGL. Fig. 2 shows the relationship between laser-HAGL, GD-HAGL and image strip width from the SWC exercises of 12 March, where both cameras were precisely synchronized. As shown, the GD-HAGL to strip width relationship proved consistent and precise. The continuous unbroken GD-HAGL record was used as the post-hoc height estimator for all OCC imagery of both TS1 and TS2. Fig. 3 shows the Tsavo Core (TS1) SWC using GD-HAGL, each camera side shown separately. Although there is scatter due to turbulence-induced ‘wing-tilt’, the correlations are highly significant, and the 95% confidence limits of the regression line intersect the origin; we therefore use the simple coefficients of right strip (m) = 0.4013 GD-HAGL (ft) and left strip (m) = 0.3668 GD-HAGL (ft) for all computations of sample size in the Jolly II analysis. In TS1 the achieved GD-HAGL was 367 ft \pm 39.4 (SD, n = 81 007).

From the RSO and the OCC strip-width regressions, we determined that the RSO strip (left and right combined) was 5.7% wider than the OCC image strip, but for paired-transect comparisons based on animal density this was compensated for by the separate strip-width calibrations for the RSO and the OCC methods. Fig. 4 shows the geometry of the counting strip in TS1 at 350 ft (107 m) HAGL, as defined through the SWC and GE measurements, with the standard image footprint covering 1.9 ha. Further details of HAGL, strip width and image footprint calibrations are given in Appendix B.

2.4. Image interpretation

Eight airphoto interpreters were employed, four university graduates from Uganda who had previous experience in interpretation for

wildlife census and land use surveys (Lamprey, 2005, 2016; Marshall et al., 2017), and four recruited in Kenya who received training. All eight interpreters undertook a guided fieldtrip to Tsavo for familiarization with species and habitats.

The ‘key species’ for enumeration in TS1 and TS2 were elephant (*Loxodonta africana*), buffalo (*Sincerus caffer*), Masai giraffe (*Giraffa camelopardalis* ssp. *tippelskirchii*) and elephant carcasses, the latter to determine carcass ratios as an indicator of mortality and poaching pressure (Douglas-Hamilton and Burrell, 1991; PAEAS, 2014). Carcasses were recorded in two stages of decomposition, recent and old, rather than the more common four stages (Douglas-Hamilton and Burrell, 1991), and were confirmed as elephant carcasses only if the skull or lower jaw was present. The ‘supplementary species’ are those commonly observed in Tsavo (Cobb, 1976; Leuthold and Leuthold, 1976), these being common zebra (*Equus quagga* ssp. *burchelli*), wildebeest (*Connochaetes taurinus*), Coke’s hartebeest (*Alcelaphus buselaphus* ssp. *cokii*), eland (*Taurotragus oryx*), lesser kudu (*Tragelaphus imberbis*), greater kudu (*Tragelaphus strepsiceros*), oryx (*Oryx beisa* ssp. *beisa*), impala (*Aepyceros melampus*), waterbuck (*Kobus ellipsiprymnus* ssp. *ellipsiprymnus*), gerenuk (*Litocranius walleri*), Grant’s gazelle (*Gazella granti*), Thomson’s gazelle (*Gazella thomsoni*), warthog (*Phacochoerus africanus*) and ostrich (*Struthio camelus*). Livestock species enumerated included cattle (*Bos indicus*), camels (*Camelus dromedarius*), sheep (*Ovis aries*) and goats (*Capra hircus*).

For image analysis, teams were divided into sub-teams of two interpreters designated Uganda A/B and Kenya A/B. To present a systematic spatial allocation of transects to each interpreter, the assignment of images followed a repeating roster, see Table 1. This assignment is firstly by transect, sequentially allocated between Kenya and Uganda teams; secondly by camera side, alternating between in-country sub-teams A and B; and thirdly by even-odd number frames alternating between the two members of the sub-team. Interpretation was conducted using standard software for viewing JPG images and EXIF files and for image annotation. For each assigned image the interpreters in each sub-team entered the metadata to the hardcopy datasheet; file and folder, date, time, transect, interpretation date, team.

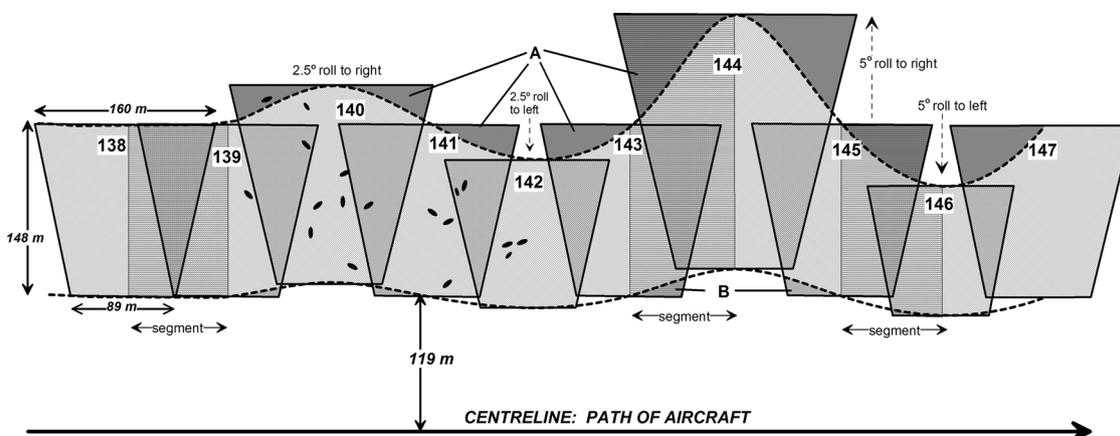


Fig. 4. Diagrammatic image footprint string, left side camera, drawn to scale, for camera angle of 57°, TS1. Curved lines show the limits of the strip, without correction for tilt-induced footprint spread. Randomized ‘segments’ as shown are selected and measured on Google Earth to determine correction factors for tilt.

They then visually scanned images for animals in a systematic sweep-and-enlarge pattern, conferred with each other on species identification and number, and annotated large herds for ease of counting. To avoid double-counting in frame overlaps, all images with even-number coding were total-counted, whilst odd-number images filled the gaps between. The principle is shown in Fig. 4, where the animals shown as black dots are assigned as 1 animal to image 1–139, 9–140, 3–141 and 5–142. The image meta-data and species numbers were entered into hardcopy data sheets, for transfer into Excel spreadsheets.

Fig. 5 show how this alternate-frame partition works in practice with consecutive images of an elephant herd. Fig. 6 shows a complex image with three species, indicating the challenges facing interpreters: in this typical image, additional animals (eg giraffe calf) were detected only after 10 min of re-interpretation.

2.5. Determining bias in image interpretation

In aerial wildlife surveys, bias causes a directional shift in enumeration or estimation, whilst errors are random and without direction (Caughley, 1974; Gasaway et al., 1986; Norton-Griffiths, 1978). For Tsavo, bias in estimation of ‘available animals’ can occur when an interpreter (or RSO) (a) fails to see an animal(s) when an animal is present (false negative or FALSE –ve); (b) reports an animal when there is no animal (FALSE +ve); (c) consistently reports animals of one species as another species (misidentification, with both FALSE +ve’s and FALSE –ve’s); or (d) double-counts animals in a frame overlap (double-counting).

Double-counting of animals in image overlaps is a potentially significant source of bias, and we assessed this bias first, and in two ways. Firstly, for key species, we checked each and every record against its image to determine if the interpreters had correctly identified the

animals, had not double counted any animals in forward or backward overlaps, and had correctly assigned the animals to even- or odd-number images. This second count was implemented by a wildlife specialist formally trained in airphoto-interpretation (the lead author). Secondly, for the supplementary species, we determine whether double-counting is significant by comparing the Jolly II estimates for the full and alternate frame image datasets, the latter being point samples with full-counts and no overlap.

The third step in detection of biases followed PAEAS (2014) guidelines, where chi-squared tests are prescribed to test for differences in species encounters between interpreters or interpreter teams. Mis-counting may also be identified by comparing the total number of animals of a species counted by interpreter teams; here, in order to escape assumptions of distribution, we use the Mann-Whitney U-test, and the Bayesian independent sample t-test. However, caution is needed in drawing conclusions for species that occur in clumped distributions, for example buffalo which are found in a relatively small number of large herds.

Determining FALSE –ve’s in a dataset of 165 000 images requires that a random sample of the images are re-interpreted for comparison with the primary interpretation. Interpreters were re-assigned 11 553 random even-number images of TS1 and TS2 (a sample of 14% of even-number images). There was an equal chance for any interpreter out of the eight to be assigned a particular image. Let S_1 = the animals seen in the image by the first interpreter, but missed by the second; S_2 = the animals seen in the image by the second interpreter, but missed by the first; B = the number of animals seen by both observers. For S_1 or S_2 for both key and supplementary species, where one interpreter saw the animal(s) but the other did not, each contested image was re-examined by a skilled interpreter/ ecologist with a third ‘vote’ to assign the animal to FALSE –ve or FALSE ve. For B-counts, if both interpreters were

Table 1

Example of assignment roster of images for interpretation teams Uganda A, Kenya A, Uganda B, Kenya B. UA1 and UA2 are the two interpreters of sub-team Uganda A, and similar coding for the three other sub-teams. Strings of images are selected between transect start and end times. L/R = left or right camera; O/E = even/odd image.

Date	Transect	Sub-Team		Uganda A		Kenya A		Uganda B		Kenya B	
		Start Time	End Time	UA1	UA2	KA1	KA2	UB1	UB2	KB1	KB2
12/3	76	06:55:32	07:01:28	L-O	L-E	R-E	R-O				
12/3	77	07:02:36	07:11:32					R-E	R-O	L-O	L-E
12/3	78	07:15:58	07:25:16	R-E	R-O	L-O	L-E				
12/3	79	07:27:16	07:34:56					L-O	L-E	R-E	R-O
12/3	80	07:37:04	07:46:04	L-E	L-O	R-O	R-E				
12/3	81	07:48:32	07:59:24					R-O	R-E	L-E	L-O
12/3	82	08:01:50	08:13:58	R-O	R-E	L-E	L-O				
13/3	83	05:04:32	05:17:08					L-E	L-O	R-O	R-E



Fig. 5. Principle of total count in alternate (even-number) images: the images have been cropped and enlarged for clarity. The elephants in image 0779 (frame 2017-03-23-ATS-1-105-0779), are divided by the border line (black line) with elephants in even-number image 0780 (frame 2017-03-23-ATS-1-105-0780). Only 7 elephants in frame 0779 are counted since the remainder of the herd (8 elephants) are counted in frame 0780.

in agreement on species and numbers, the record was confirmed as TRUE +ve; the third vote was used for the key species B-counts as re-confirmation. To test for bias, we calculate the original unchecked numbers of animals claimed by both primary and secondary interpreters (the original TRUE +ve's) as $TP_O = (2 * B) + S_1 + S_2$. We

compare with the same calculation after all third-party checking to generate 'corrected TRUE + ve's', where $TP_C = TP_O + \text{all FALSE -ve's} - \text{all FALSE +ve's}$. A ratio of TP_C / TP_O approximating to unity indicates that the double count yields similar estimates, and therefore any misidentifications are errors without direction. A major departure from



Fig. 6. Image 2017-03-15-ATS-105-0724, an aggregation of three animal species under trees. These animals would not be detected in vertical aerial images.

unity indicates that an interpreter might be biased in continuously identifying one species as another.

For correction factors we followed double-observer ‘capture-recapture’ methods for RSOs developed in Australia (Caughley and Grice, 1982; Graham and Bell, 1989; Magnusson et al., 1978; Marsh and Sinclair, 1989) to determine the ‘probability of detection’. The probability of detection of a species by the primary interpreters is defined as $\hat{P}_1 = B/(B + S_2)$. We then extend this process to include into the analysis those photos where the secondary interpreters detected animals on images where the primary interpreters did not, where $\hat{P}_2 = B/(B + S_1)$. With each interpreter, primary and secondary, having an equal chance of finding animals on the same image, the detection factor, \hat{P}_d , then becomes the average of \hat{P}_1 and \hat{P}_2 . The correction C_d to be applied to the Jolly II estimate is then $1/\hat{P}_d$.

In our sample of 11 553 re-interpreted images we encountered elephants in just 44 of them, between B , S_1 and S_2 above. For the calculation of \hat{P}_d for supplementary species, we generated a sample of 80 images of encounters by aggregating hartebeest, eland, oryx and waterbuck into a group class termed ‘large brown antelopes’ (LBAs), and a larger sample of 97 images of ‘all brown antelopes’ that include LBAs, plus impala and the gazelles.

2.6. Implementation of comparative RSO surveys

The simultaneous RSO sample-count was conducted according to current accepted standards for aerial strip-transect sample counts (Craig, 2012; PAEAS, 2014), with professional RSOs working for government wildlife agencies and livestock programmes. Recording was by the ‘sub-unit method’, where animal observations were assigned by RSOs to 2.5 km subunits along the transect, called out by the front-seat-observer. Large groups of key species, eg herds ≤ 10 animals, were also photographed by the RSOs with window-mounted cameras for later counting. The subunits became the smallest unit of comparison of RSO and OCC counting, which presents a constraint to close matching of RSO to OCC records. The subunit in our case is a 2.5 km segment traversed in just 50 s, and an overloaded RSO might delay a record until the next subunit. In general, in considering the subunit as the closest match, for example in determining animal visibility by vegetation type, there will be enough coincident observations for comparison. Where more precision is needed in comparing numbers counted, albeit with a smaller number of samples, we use the transect as the paired sample unit.

For TS1 and TS2, a further comparison is the 2017 Tsavo Total Count, implemented to methods that have evolved over three decades. In 2017 the target species were elephants, buffalos and giraffes, and elephant carcasses in four decomposition categories (Douglas-Hamilton and Burrill, 1991; Douglas-Hamilton et al., 1994; Ngene et al., 2017). The Tsavo ecosystem was divided into 91 blocks which were systematically searched at 1 km transect intervals, at 400–600 ft (122–183 m) altitude by 10 aircraft and crews. Animal sightings were recorded using GPS receivers and large herds were photographed for later counting. Being a count of all animals (as attempted), total count population estimates have no estimates of precision – we cannot calculate confidence intervals for them.

Table 2

Results of re-interpretation and revision of all records for key species.

	Number of original images with species	Original count	Records revised	New Count	% change
Elephant	466	1 597	110	1 494	–6.45%
Buffalo	162	1 383	47	1 263	–8.68%
Giraffe	182	310	18	299	–3.55%

2.7. Comparison of RSO and OCC methods

Initial analysis was conducted using the standard Jolly II ratio-method, where the strip-transects are the sample units (Caughley, 1977; Jolly, 1969; Norton-Griffiths, 1978), with area calculated as length x average width. For OCC we assume there are no significant gaps in the imaged transect that would reduce this sampled area. In our second approach, allied to aerial point sampling, only even-number photos are included, and the sample area for transect $Z = \sum (A_i...)$ where A is the area (hectares) of the i th even-numbered image. The disadvantage of the point sample approach is that the area of each frame footprint must be derived from a careful footprint / GD-HAGL calibration, and that some 40% of the sample, the overlap area, is not used in the estimate.

For survey TS1, we initially compare RSO and OCC population estimates using Jolly II. However, this approach is coarse since Jolly II measures the variance between transects of an independent survey, rather than between paired transects. We therefore test for differences between the paired transects or subunits; the former provide a sufficient sample size. To avoid complications of sample-area / animal-count covariance, we initially used 49 equal-length transects measuring 27.5 km x 0.286 km (11 subunits x the mean strip width) within the TS1 dataset (see Fig. 1). This yields 2×1347.5 km of coincident RSO-OCC recording, based on 32 000 images. We test for differences using the basic paired t -test and the non-parametric Wilcoxon signed-ranks test.

We then compare estimates using the entire simultaneous RSO-OCC dataset of 67 000 images captured in 2×3004 km of sample strips, making no assumptions about the distribution of the data. We repeat the Wilcoxon signed-ranks test, and conduct the Bayesian inference t -test, computed in JASP as an interface to R, which computes a ‘weight’ of probabilistic evidence (the Bayes Factor - ‘BF’) derived from continued inclusion of paired data (Ellison, 1996; Kruschke, 2013; Marsman and Wagenmakers, 2017). We conduct the analysis on the basis of a ‘prior’ that OCC counts are higher than RSO counts.

For TS2 the only comparison to be made with RSO counting is with 2017 Tsavo Total Count. Here, we determine the probability of the Total Count results lying within the 95% confidence limits of the OCC count.

3. Results

Initially we tested the OCC results for interpretation biases that might compromise accuracy and precision. We then compared the RSO counts and OCC counts by transect using standard parametric, non-parametric and Bayesian paired-sample methods. Counts from both methods were processed using the standard Jolly II procedure (Jolly, 1969) to generate population estimates and standard errors for the survey area for further comparison as independent samples. For key species the results of both the OCC and RSO counts were then compared with the 2017 Total Count that had been carried out a month previously.

3.1. OCC counting bias and probability of detection

Ahead of any analysis, for the key species we determined correct identification and evidence of double-counting by re-checking every

Table 3

Testing for differences in species encounter rates and numbers between the Uganda and Kenya country teams, using χ^2 on encounters and the Mann-Whitney *U* test and Bayesian independent sample *t*-test for tests on numbers. The Bayes *t*-test determines evidence for the null hypothesis of H_0 (no difference), in this case with no prior assumption of a directional difference between RSO and OCC.

Row Labels	Kenya Encounters	Uganda Encounters	Chi-squared		Kenya, Number	Uganda, Number	Mann-Whitney		Bayes Inference t-test	
			χ^2	p			U-statistic	p	BF ₁₀	Evidence for H ₀
Elephant	221	211	0.28	NS	763	705	955	NS	0.250	Moderate
Carcass-all	38	104	30.68	< 0.001	39	107	273	0.028	3.515	Moderate reject
Buffalo	67	69	0.03	NS	670	593	288	NS	0.285	Moderate
Giraffe	73	101	4.51	NS	118	181	414	NS	0.634	Anecdotal
Zebra	116	125	0.34	NS	644	661	680	NS	0.249	Moderate
Hartebeest	38	48	1.16	NS	123	144	138	NS	0.465	Moderate
Eland	24	30	0.67	NS	72	100	112	NS	0.336	Moderate
Oryx	59	40	3.65	NS	228	229	238	NS	0.297	Moderate
Waterbuck	17	65	28.1	< 0.001	35	130	163	NS	0.382	Anecdotal
Impala	21	43	7.56	< 0.01	125	225	155	NS	0.336	Anecdotal
Grant's G.	40	32	0.89	NS	112	58	109	NS	1.760	Anecdotal reject
Thom. G.	11	12	0.04	NS	36	37	39	NS	0.494	Anecdotal
Gerenuk	25	50	8.33	< 0.01	64	70	193	NS	0.691	Anecdotal
Warthog	8	105	83.27	< 0.001	15	172	55	0.020	1.470	Anecdotal reject

record against its image and neighbouring images. This indicated almost 100% correct identification, but removal due to double-counting reduced the elephant counts by 6.45%, buffalos by 8.68% and giraffes by 3.55%, see Table 2. Checking was also conducted for supplementary species with herd size ≤ 10 animals. Double counting for both key and supplementary species is again tested in the even-frame point sample approach for Jolly II, indicated in the next section.

We then compared encounter rates of different species between the Uganda and Kenya country teams, see Table 3. Applying the PAEAS observer-comparison guidelines (PAEAS, 2014), differences are not significant for elephant (encounters 211 and 222 respectively, $\chi^2 = 0.28$, $p = 0.597$) and buffalo (69 vs 67 encounters, $\chi^2 = 0.03$, $p = 0.864$). The Uganda team had more encounters than the Kenya team for giraffe (101 vs 73 encounters, $\chi^2 = 4.51$ $p = 0.034$), but after Holm-Bonferroni correction (Holm, 1979) this may not be considered significant (adjusted $p^* = 0.304$). Differences were significant for carcasses, waterbuck, gerenuk, impala and warthog. For actual numbers counted, where we tested our null hypothesis of no-difference using Mann-Whitney *U*-tests and Bayes independent sample *t*-tests, differences between teams were not significant for most species except for carcasses, ($U = 273$, $p = 0.028$, $BF_{10} = 3.515$), warthogs ($U = 55$, $p < 0.020$, $BF_{10} = 1.470$), and Grants' gazelle ($BF_{10} = 1.760$), where again the Uganda team detected more of these targets. In general, we

are comfortable with the consistency of interpretation between the two teams for the key variables of Tsavo, with the exception of carcasses, for which correction factors may be applied.

For the 11 553 random image re-interpretation of even-number images, all cases of S_1 , S_2 and B (both) observations for the key species were independently re-scrutinized. No FALSE -ve's, of key species were identified. The probability of detection of available animals, P_d , is 96.3% for elephants, 92.6% for buffalo, 90.9% for giraffe, 78.2% for zebra, 67.7% for LBAs and 67.1% for all antelopes, see Table 4. Carcass detection was lower at 60.0%. Overall, these results indicate that key species detection by interpreters is generally satisfactory, but that some 33% of antelopes are missed and 40% of carcasses.

The derived correction factors C_d are also indicated in Table 4, and the Jolly II estimates \hat{Y}_O (see below) may be corrected as $\hat{Y}_{OC} = C_d \times \hat{Y}_O$. The boxes indicate the misidentification of a single herd of 21 eland as oryx by one interpreter (S_1), with the result that on the 3rd party vote these were ascribed as eland FALSE -ve's, and hence added as eland TRUE +ve's. If we adjust for this misidentification, the correction factor C_d for oryx then drops from 2.2 to 1.9, whilst for eland TP_R/TP_O is recalculated as 109%. The correction factors C_d increase all OCC estimates, with oryx being a significant case in point where even interpreters miss half of them. However, we use only the uncorrected \hat{Y}_O for comparisons of counts, unless otherwise stated.

Table 4

Results of re-interpretation of random 11 553 even-number images to determine TRUE +ve, FALSE +ve, FALSE -ve. S_1 = the animals seen by first interpreter, but missed by second; S_2 = the animals seen by the second interpreter, but missed by first; B = animals seen by both interpreters. P_d = the average probability of detection between the two interpreters, C_d = correction factor for the Jolly II estimates. TP_O = animals originally recorded in first interpretation, TP_C = corrected TP_O after addition of FALSE -ve's and subtraction of FALSE +ve's. The italicized numbers include one herd of 21 eland mis-identified as oryx by the S_1 interpreter (see text).

	S_1	S_2	B	TRUE +ve TP_O	FALSE +ve	FALSE -ve	TRUE +ve TP_C	TP_C/TP_O	P_d	C_d
Elephant	6	6	154	320	0	5	325	102%	96.3%	1.04
Carcass	8	14	16	54	4	19	69	128%	60.0%	1.67
Buffalo	2	4	37	80	2	3	81	101%	92.6%	1.08
Giraffe	6	5	55	121	1	10	130	107%	90.9%	1.10
Zebra	56	56	201	514	40	29	503	98%	78.2%	1.28
Hartebeest	14	40	67	188	14	12	186	99%	72.7%	1.38
Eland	7	26	67	167	8	45	204	122%	81.3%	1.23
Oryx	42	17	22	103	23	22	102	99%	45.4%	2.20
Waterbuck	7	6	10	33	4	8	37	112%	60.7%	1.65
Impala	14	20	77	188	17	34	205	109%	82.0%	1.22
Grant's G.	22	27	15	79	18	51	112	142%	38.1%	2.62
Thom. G.	11	2	1	15	11	5	9	60%	20.8%	4.80
LBAs	70	89	166	491	49	87	529	108%	67.7%	1.48
All antelopes	117	138	259	773	95	177	855	111%	67.1%	1.49
All Animals	187	209	706	1808	138	224	1 894	105%	78.1%	1.28

Table 5

Jolly II estimates (\hat{Y}) and standard errors (SE) for OCC (\hat{Y}_O) and RSO (\hat{Y}_R) in survey TS1 (Tsavo Core), with comparison using the t-statistic (Norton-Griffiths, 1978). Estimates and standard errors in parentheses are derived from the even-frame (point sample) estimation. RSO 'non-detection' of animals available on the OCC imagery is calculated as $(\hat{Y}_O - \hat{Y}_R) / \hat{Y}_O$ (%) with non-detection also indicated where OCC estimates are corrected as $\hat{Y}_{Oc} = \hat{Y}_O \times C_d$ (see Table 4). The last column indicates the difference of corrected OCC estimates over the RSO estimates, where $\Delta\% = (\hat{Y}_{Oc} / \hat{Y}_R) - 1$ ('how much greater are OCC estimates?').

Variable	OCC estimates		RSO estimates		t-statistic	p	RSO non-detection with \hat{Y}_O	RSO non-detection, with \hat{Y}_{Oc}	OCC vs RSO, $\Delta\%$
	\hat{Y}_O	SE	\hat{Y}_R	SE					
Elephant	12 357 (12 258)	1 814 (2 204)	11 074	1 347	0.57	0.573	10%	14%	16%
Carcass, recent	91	29	17	12	2.36	0.022	81%		
Carcass, old	1 117 (1 076)	155 (184)	1 529	227	1.50	0.140	-37%		
Carcass, all	1 208	162	1 546	225	1.22	0.228	-28%	23%	30%
Buffalo	10 453 (10 618)	2 529 (2 858)	13 110	3 848	0.58	0.566	-25%	-16%	-14%
Giraffe	2 475 (2 510)	386 (413)	1 090	184	3.24	0.002	56%	60%	150%
Zebra	10 801 (11 086)	1 550 (1 665)	7 173	1 355	1.76	0.084	34%	48%	93%
Hartebeest	2 210 (2 261)	438 (454)	1 278	294	1.77	0.083	42%	58%	138%
Eland	1 424	431	611	361	1.45	0.154	57%	65%	187%
Kudu	687	159	94	27	3.68	0.001	-		
Oryx	3 782 (3 875)	775 (988)	2 148	675	1.59	0.118	43%	74%	288%
Waterbuck	1 366	264	434	133	3.15	0.003	68%	81%	419%
L.B.Antelopes	8 782	1035	4 471	852	3.22	0.002	49%	66%	190%
Impala	2 897 (2 634)	527 (677)	315	132	4.75	< 0.001	89%	91%	1022%
Grant's gazelle	1 407	313	1 065	281	0.81	0.420	24%	71%	247%
Thomson's gaz.	604	166	229	82	2.03	0.048	64%	92%	1166%
Gerenuk	1 109	175	34	20	6.10	< 0.001	97%	-	-
Warthog	1 548	194	239	61	6.44	< 0.001	85%	-	-
Ostrich	662	223	603	131	0.23	0.820	9%	-	-
Cattle	42 284	13 273	37 168	17 414	0.23	0.816	12%	-	-
Shoats	18 183	3 360	12 313	3 253	1.26	0.215	32%	-	-

3.2. TS1, Jolly II analysis and OCC-RSO comparison

Table 5 presents the Jolly II analysis for TS1 for both the RSO and OCC methods. RSO estimates (\hat{Y}_R) are based on the laser HAGL calibration, OCC estimates (\hat{Y}_O) on GD-HAGL calibration. We conclude that OCC estimates \hat{Y}_O for giraffe, zebra, kudu, waterbuck, impala, Thomson's gazelle, gerenuk and warthog are significantly higher ($\alpha = 0.05$) than the corresponding RSO estimates \hat{Y}_R . For giraffe, the difference in estimates is highly significant, where the RSOs missed 56% of the giraffe ($(\hat{Y}_O - \hat{Y}_R) / \hat{Y}_O$), this increasing to 60% if we consider the corrected estimates \hat{Y}_{Oc} . Re-analysis of the TS1 OCC data based on laser-HAGL and strip-width calibration (see Fig. 2) rather than GD-HAGL, indicates no significant difference with the Jolly II estimates, with for example the Core elephant population now indicated as $12\,714 \pm 1775$ (SE) and giraffe as $2\,544 \pm 383$ (SE).

Table 5 also shows, for the main species, the estimates derived from even-frame estimation to test for double-counting. Even-image point sample estimates are very close to all-image estimation. This suggests that the post-hoc checking process for key species has essentially eliminated double-counts, whilst for supplementary species the process of allocation to even-number frames at the interpretation stage, without subsequent correction, has also been effective.

Selecting the equal-length transects in TS1 as the paired samples for density (see Fig. 1), the Shapiro-Wilks test indicates significant departures from normality ($p < 0.001$) for all species, although for elephants this is the least significant ($W = 0.953$, $p = 0.047$). Table 6 shows, for pair-wise comparison of species densities (animals per unit area), the paired sample t-tests for equal-length coincident RSO-OCC transects. In combination the t-tests (2-tailed) and Wilcoxon-signed rank tests indicate that OCC density estimates for elephant, giraffe, zebra, eland, kudu, oryx, waterbuck, impala, Grant's gazelle, warthog and the LBAs are significantly higher (at $\alpha = 0.05$) than RSO estimates.

Differences in carcasses, buffalo and hartebeest were not significant. Fig. 6, an OCC TS2 image, shows graphically how even large animals may be virtually invisible in the landscape, whilst Fig. 7 shows an example from TS1 where RSOs have missed a herd of oryx recorded in the OCC imagery.

Table 6 also shows the Wilcoxon paired-sample tests, and the Bayes factors for prior $OCC > RSO$ (BF_{+0}) for the full OCC-RSO coincident dataset. Both tests demonstrate that the OCC method counts significantly more elephants than the RSO method, with Wilcoxon W statistic = 1242, $p = 0.003$, and the Bayes Factor $BF_{+0} = 16.95$, indicated as 'strong'; in other words, 'for elephants it is nearly 17 times more likely that OCC counts $>$ RSO counts, compared with the null hypothesis of no difference'. In support of $OCC > RSO$, differences for giraffe, zebra, hartebeest, eland, kudu, waterbuck and impala are now highly significant. The difference for the LBA combination, with Wilcoxon W statistic = 1479, $p < 0.001$, and Bayes Factor $BF_{+0} = 4472.85$, is now indicated under Bayes terminology as 'extreme'. Carcasses differences are estimated as 'strong' in the opposite way ($RSO > OCC$), with RSOs estimating more than OCC, but this is accounted for by the lower carcass detection of the Kenya interpreters, who detected 37% of the Uganda team total (see Table 3). For buffalo, the weight is also in favour of higher estimation by RSOs; this may be down to double-counting on RSO image strings, but this needs to be examined further. Bayes Factor prior and posterior plots for elephants, giraffe, zebra and LBAs are shown in Appendix C.

As a first stage in analysing the effect of group size on detection, we compare elephant counts in subunits with only OCC-detection with counts with joint OCC-RSO detection. For the sample of OCC-only 'elephant positive' subunits, the average group is 4.26 elephants ± 1.141 (95% cl, $n = 78$), compared with joint RSO-OCC detection with the higher group size of 5.95 ± 1.069 (95% cl, $n = 116$) ($t = 2.065$, $p = 0.040$, $df = 192$). This implies that elephant group size

Table 6

Comparison of RSO and OCC densities for the equal-length and all coincident transects using *t*-tests (df = 48 and df = 67 respectively), Wilcoxon signed-ranks test and the Bayes inference *t*-test. The raw count numbers are indicated, these are then converted into densities for the tests. Cattle and shoats were not encountered in the equal-length transects in TS1.

Species	Equal-length transects						All coincident transects					
	Numbers counted		t-test		Wilcoxon		Numbers counted		Wilcoxon		Bayes inference t-test	
	OCC	RSO	t-test	p	W	p	OCC	RSO	W	p	BF ₀	Evidence
Elephant	610	494	2.674	0.010	738	0.013	1022	891	1242	0.003	16.95	Strong
Carcass, old	27	49	-2.324	0.024	175	0.061	85	144	427	0.003	0.03	V. strong, null
Buffalo	489	589	-1.338	0.187	66	0.251	772	920	233	0.776	0.06	Strong, null
Giraffe	119	62	2.689	0.010	322	0.001	208	105	643	< .001	43.05	V. strong
Zebra	434	254	2.316	0.025	309	0.004	1076	736	977	< .001	34.98	V. strong
Hartebeest	98	62	1.787	0.080	61	0.092	200	118	241	0.036	4.03	Moderate
Eland	60	16	2.723	0.009	74	0.007	130	59	248	< .001	19.70	Strong
Kudu	23	3	2.183	0.034	69	0.021	44	6	231	< .001	36.08	V. Strong
Oryx	91	58	1.47	0.148	131	0.05	331	205	316	0.034	1.17	Anecdotal
Waterbuck	97	37	2.289	0.027	249	0.005	138	45	618	< .001	29.96	Strong
Impala	71	17	2.137	0.038	76	0.036	170	20	258	< .001	85.73	V. strong
Grant's	120	53	1.936	0.059	155	0.017	157	101	398	0.087	0.92	Anecdotal
Warthog	94	12	4.073	< .001	416	< .001	170	25	997	< .001	59 814	Extreme
LBAs	354	180	3.655	< .001	652	< .001	843	433	1479	< .001	4 472	Extreme
Cattle			N/A				4 965	4 222	353	0.014	0.43	Anecdotal
Shoats			N/A				2 097	1 399	348	0.005	1.97	Anecdotal

influences detectability by RSOs (Schlossberg et al., 2016). Further tests on disaggregated subunit data are required to explore group size as a factor of visibility for a range of species.

3.3. Comparison of TS1 and TS2 Jolly II estimates with the 2017 Tsavo Total Count

For TS2 we have no comparative RSO sample count data, and therefore we use these Jolly II stratum estimates in an ecosystem-wide comparison of the OCC results (uncorrected) with the 2017 Total Count which enumerated elephant, buffalo, giraffe, and elephant carcasses, see Table 7. The Total Count estimate of elephant and buffalo fall near the lower 95% cl of the OCC \hat{Y}_O estimates, and giraffe fall well below. Z-tests indicate a low probability of the Total Count elephant estimate lying within the Jolly

II confidence intervals for \hat{Y}_O ($Z = 1.917$, $p = .0276$), whilst the probability for buffalo in Core is also low ($Z = 1.937$, $p = .0262$), but higher for the entire ecosystem ($Z = 1.279$, $p = 0.1000$). Using a 1000-iteration bootstrap with replacement (Diciccio and Romano, 1988) for the elephant Jolly II estimates for TS1, the lower percentile-based 95% cl estimate is indicated as 9410 elephants, in comparison with the Total Count estimate of 9363 elephants; the bootstrap approach however, requires further investigation. The conclusion is that OCC \hat{Y}_O estimates are significantly higher than those of the Total Count by 31% for elephant, 39% for buffalo, 112% for giraffe and 293% for carcasses. The corrected OCC \hat{Y}_{Oc} estimates are higher by 37%, 54%, 137% and 532% respectively for these species. Finally, Table 8 shows, for completeness, the results for all other species for the combined strata of the Tsavo Conservation Area.



Fig. 7. In cloud-shadow conditions, this group of nine oryx (yellow circles) were not detected by the RSOs (image 2017-03-15-AKP-1-117-0147). By subunit and camera side, RSOs missed 38% of oryx by number, 24% by group.

Table 7

OCC estimates (\hat{Y}_O) and 95% confidence limits by stratum for elephants, buffalo, giraffe and elephant carcasses, for comparison with the 2017 Tsavo Total Count estimates. Figures in parentheses are the corrected estimates \hat{Y}_{Oc} , using the correction factors of Table 4.

	Elephant			Buffalo			Giraffe			Elephant Carcass		
	\hat{Y}_O	95% cl	Total Count	\hat{Y}_O	95% cl	Total Count	\hat{Y}_O	95% cl	Total Count	\hat{Y}_O	95% cl	Total Count
Core	12 357	3 556	9 363	10 453	4 957	5 555	2 475	756	1 055	1 208	318	573
South	3 002	1 760	2 548	1 197	1 023	2 297	3 789	1 329	1 583	2 091	510	338
North	491	475	293	175	192	544	316	382	118	140	102	60
N-West	302	359	195	0	0	125	1 207	657	1 030	78	49	20
East	529	528	323	0	0	1	465	320	105	828	278	115
TOTAL	16 681	4 047	12 722	11 826	5 065	8 522	8 251	1 737	3 891	4 345	680	1 106
\hat{Y}_{Oc}	(17 331)	(4 205)		(12 777)	(5 262)		(9 075)	(1 805)		(7 242)	(706)	

Table 8

OCC Jolly II population estimates (\hat{Y}_O) and standard errors (SE) of supplementary species for the Tsavo ecosystem.

Species	\hat{Y}_O	SE
Zebra	28 871	3 788
Wildebeest	758	277
Hartebeest	8 259	1 966
Eland	3 563	729
Kudu, lesser	1 564	270
Kudu, greater	617	148
Oryx	8 894	1 422
Waterbuck	2 105	594
Impala	6 639	1 360
Gazelle, Grant's	4 982	820
Gazelle, Thomson's	1 055	248
Gerenuk	2 204	339
Warthog	4 528	734
Ostrich	2 428	533

4. Discussion

We describe an aerial strip-transect method where a camera system is used to simultaneously image the entire sample strip observed by RSOs. The imaging system comprises off-the-shelf 24-megapixel digital cameras that are set up with lens focal lengths, angles and shutter speeds to capture a standard SRF strip with a ground-sampling-distance of approximately 3–6 cm, sufficiently small to resolve animals down to gazelle size. This method permits a very careful scrutiny of the strip for animals; the image may be frozen for minutes or even hours whilst interpreters may literally 'look under every bush'.

We rigorously tested the interpretation procedure for biases before comparing the method with RSO counts. For the majority of key and supplementary species we find no significant differences between two interpreter teams, working separately in two countries, in encounters of herds or numbers counted. Using protocols for assigning wildlife to alternate ('even-number') frames, we show that interpreters effectively avoided double-counting in frame overlaps. However, in random repeat counts, we also show that whilst interpreters were detecting over 90% of the key species, this dropped to 67% for the large antelopes. Clearly there are many available antelopes 'out there in the imagery' that the interpreters did not detect. From the probability of detection, we derive correction factors (C_d) to offset these species-specific biases.

Applying the correction factors to derive corrected OCC Jolly II estimates (\hat{Y}_{Oc}), we then determine that the sample-count RSOs did not detect 14% of elephants, 60% of giraffe and 66% of the large antelopes, hartebeest, eland, oryx and waterbuck. Again using \hat{Y}_{Oc} , we conclude that the Total Count observers did not detect 27% of elephant, 35% of buffalo, 58% of giraffe and 84% of carcasses.

We do not believe these differences to be an anomaly of performance of our sample-count RSOs. Our observers generated elephant estimates that are similar to, or higher than, other recent sample-count surveys. For example, for Tsavo Core our RSO-count estimated 11 074

elephants ± 1347 (SE), compared with 10 716 ± 1315 (SE) estimated by the Great Elephant Census survey of 2014 (Chase et al., 2014). The 2011 and 2014 government-implemented SRFs of the entire Tsavo ecosystem generated RSO-based elephant population estimates of 11 358 ± 2350 (SE) and 10 958 ± 2565 (SE) respectively (Ogutu et al., 2016). Our RSO zebra and oryx estimates, extrapolated to the ecosystem by density, are at least 50% higher than these two counts, whilst giraffe estimates are similar.

It is well known that SRF counts using RSOs produce highly variable results. Whilst this research points to the possibility of applying universal or retrospective correction factors for RSO-counts, this should be done with great caution. Firstly, the visibility characteristics of species vary in different environments and in different seasons. In the green landscape of Murchison Falls National Park, Uganda, for example, both RSO- and OCC-counts enumerated Rothschild's giraffe closely to the known population, all members of which had been identified and counted through coat markings (Lamprey, 2016; Wanyama et al., 2014). In the arid bushlands and plains of Tsavo in 2017, Maasai giraffe blended into the beige landscape and RSOs missed over half of them. The low detection of giraffe by RSOs has also been recorded in semi-arid environments in Tanzania (Lee and Bond, 2016). The second reason for caution is that performance is highly variable between observers. In the Tsavo RSO-count, for example, most species were enumerated fairly consistently between left and right RSOs, but one RSO saw eight groups of impala (total 566 animals), the other just one (81 animals). With regard to historical SRFs, the counts of individual RSOs are usually lost to the record, and we have no possibilities for correcting for individual observer bias.

In addition to the advantages of improved accuracy and precision, the entire survey record can be revisited at later stages by third parties for further checking, re-interpretation, processing and certification. Another advantage, also indicated in similar studies (Erwin, 1982; Frederick et al., 2003; Terletzky and Ramsey, 2016; Xue et al., 2017), is that interpretation does not require highly skilled practitioners. During the nine months of interpretation, three of the eight interpreters departed and were replaced by incoming university graduates who, after short training, performed as well as their more experienced colleagues with no biases that could be detected. This shows that the paired-interpreter team approach works well for mentoring incoming interpreters.

Imaging methods also have applications for line-transect surveys, to meet the primary assumption of 100% detectability on the line. In aerial distance-sampling in Kruger National Park in South Africa by fixed-wing aircraft, difficulties in assessing probability of sighting on the line and in defining species-specific distance visibility profiles have cast some doubt on the usefulness of aerial line-transects for complex multi-species counts in Africa (Kruger et al., 2008). With imaging techniques, we have more certainty that all, or almost all, animals on the line are detected, improving the accuracy of models. In addition, follow-up analyses may interpret the images for new variables, to determine visibility with respect to group size, vegetation cover,

topography, ambient light and sun-angle conditions (Jacques et al., 2014; Ndaimani et al., 2017; Norton-Griffiths et al., 2015).

The main disadvantage of the OCC method in its current stage of evolution is that it is labour intensive. For the 165 000 images of TS1 and TS2, the 8 interpreters worked for 9 months, each interpreting 200–230 images per day. Labour expenses increased the cost above that of the traditional RSO-based SRF, but by no more than 27% if student-rates are applied. At the same time, much of this labour is expended on searching TRUE –ve's, the images with nothing in them. For example, in Tsavo TS1 only some 1814 out of 81 008 images, or just 2.2%, had any target wildlife species in them. In total only 435 images, or 0.54%, had elephants. Therefore, this experimental OCC, while producing greater estimates for most species, should be regarded as an initial step towards more automated counting by machine learning, where computers can flag the 'presence-of-animal' for further investigation (Laliberte and Ripple, 2003; McMahon et al., 2014; Terletzky and Ramsey, 2016; Yang et al., 2014): great progress has already been made in the machine processing of camera trap imagery for animal identification (Norouzzadeh et al., 2018; Tabak et al., 2019), which can be improved further when this is incorporated with citizen science methods (Willi et al., 2019). Further analysis of the Tsavo dataset, can explore subsampling the records to determine how the volume of imagery can be reduced for visual interpretation (Norton-Griffiths et al., 2015). The even-image point-sample approach used in this study is a first step in this investigation. A further practical step towards improving the OCC method is the use of gimbals to stabilize cameras, and therefore improve the stability of the strip width.

The OCC method has applications for the use of unmanned aerial vehicles (UAVs) in aerial wildlife counting. Soon lightweight cameras will be available that will have the high pixel-densities required to achieve a ground-sampling-distance of 5 cm or less when operated in drones flying at low level, for example at 400 ft above-ground-level. There has been some success in using UAVs for image-based strip-transect counts for elephants in West Africa (Vermeulen et al., 2013), but the overall constraint remains their endurance, speed and reliability. As calculated by sampled area for the West Africa trial area of 940 km², the cost of implementing the UAV surveys was 10 times higher than using conventional light aircraft, without human resource costs included. For the vast area of Tsavo (34 000 km²), a UAV might not complete a single transect, but increases in UAV endurance suggest that smaller areas of perhaps 200 km² might be surveyed at useful sample intensities of 5–10% at costs similar to a light aircraft.

The implications of image-based counting methods for inventorying wildlife populations are significant. Recent assessments of the state of Kenya's wildlife, based on 40 years of RSO-based surveys, indicate major declines across most rangeland areas (Ogutu et al., 2016). This situation is particularly serious for Taita-Taveta County, of which Tsavo National Park occupies 62% of the land area, where for example it is reported that eland and oryx have declined by 43.5% and 56.5% respectively since 1977. Our results provide some more positive news for certain species. For example, giraffe are now categorized as Vulnerable on the IUCN red list (Muller et al., 2016), and the OCC results for Tsavo increase Kenya's known population of the Maasai giraffe subspecies (*Giraffa camelopardalis* ssp. *tippelsckirchi*) by 41% from 12 717 (Ministry of Tourism and Wildlife, 2018) to 17 890. Similar results might be obtained if image-based sample counts were to be conducted in the northern Kenya range of the Endangered (IUCN-Redlist) reticulated giraffe (*Giraffa camelopardalis* ssp. *reticulata*), with an estimated population of 11 048. We believe that image-based aerial survey methods will enable us to reset baselines for the future monitoring of wildlife over Kenya's rangelands and protected areas.

5. Conclusions

We use high-resolution camera systems to acquire images of sample-strips in tandem with rear-seat-observers in an aerial wildlife count in

Tsavo, Kenya. The primary finding is that for large-area systematic reconnaissance flight (SRF) counts in Kenya, imaging methods generate wildlife population estimates that are significantly higher than observer-based methods, ranging from 17% higher for elephant to 150% higher for giraffe, zebra and large antelopes. Comparisons with total counts derive similar results. This implies that observer-based counting methods conducted in East Africa over the last 60 years have significantly underestimated some wildlife populations. Further work is needed to streamline the acquisition and interpretation process, with image subsampling, machine learning and the continued development of UAVs being important avenues for research. At the same time, efforts should be made to recalibrate historical observer-based SRF datasets to reset baseline estimates of wildlife populations.

Acknowledgements

This work was funded by Save the Elephants. We thank the Kenya Council for Science and Technology for granting us Research Clearance (Number NACOSTI/P/17/24999/16096), and the Kenya Wildlife Service for authorizing and participating in the experimental survey. We are particularly grateful to the two rear-seat-observers, the eight airphoto interpreters and four data entry specialists for their hard work and dedication. We are also grateful to Alexis Peltier who expertly piloted the aircraft in TS1, and Dr Ian Games who assisted in image archiving. We would like to thank three anonymous reviewers for their helpful comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.biocon.2019.108243>.

References

- Aerial-Survey-Base, 2014. GSD Calculator. Retrieved July 27, 2019, from <https://www.aerial-survey-base.com/gsd-calculator/>.
- Anderer, D.K., 1981. The Kenya Rangeland Ecological Monitoring Unit. In: Grimsdell, J.J.R., Westley, S.B. (Eds.), *Low-Level Aerial Survey Techniques* (59–68). Monographs 4, International Livestock Centre for Africa, ILCA Monograph 4, International Livestock Centre for Africa, Addis Ababa.
- Anderson, C.R., Lindzey, F.G., 1996. Moose sightability model developed from helicopter surveys. *Wildlife Society Bulletin* (1973–2006), vol. 24. pp. 247–259. Retrieved from: <http://www.jstor.org/stable/3783114>.
- Bayliss, P., Yeomans, K.M., 1989. Correcting bias in aerial survey population estimates of feral livestock in northern Australia using the double-count technique. *J. Appl. Ecol.* 26 (3), 925–933.
- Beasom, S.L., Hood, J.C., Cain, J.R., 1981. The effect of strip width on helicopter censusing of deer. *J. Range Manag.* 34 (1), 36–37.
- Bröker, K.C.A., Hansen, R.G., Leonard, K.E., Koski, W.R., Heide-Jørgensen, M.P., 2019. A comparison of image and observer based aerial surveys of narwhal. *Mar. Mamm. Sci.* 2019 <https://doi.org/10.1111/mms.12586>. open.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L., 2004. *Advanced Distance Sampling*. Oxford University Press, Oxford, United Kingdom.
- Bull, D.R., 2014. *Communicating Pictures: a Course in Image and Video Coding*. Academic Press, Elsevier, Netherlands.
- Burnham, K.P., Anderson, D.R., 2014. P-values are only an index to evidence: 20th- vs 21st-century statistical science. *Ecology* 95 (3), 627–630.
- Burnham, K.P., Anderson, D.R., Laake, J.L., 1985. Efficiency and bias in strip and line transect sampling. *J. Wildl. Manage.* 49 (4), 1012–1018. <https://doi.org/10.2307/3801387>.
- Caughley, G., 1974. Bias in aerial survey. *J. Wildl. Manage.* 38 (4), 921–933. <https://doi.org/10.2307/3800067>.
- Caughley, G., 1977. Sampling in aerial survey. *J. Wildl. Manage.* 41 (4), 605–615. <https://doi.org/10.2307/3799980>.
- Caughley, G., Grice, D., 1982. A correction factor for counting emus from the air, and its application to counts in Western Australia. *Wildl. Res.* 9 (2), 253–259. <https://doi.org/10.1071/WR9820253>.
- Chase, M., Frederick, H., Mukeka, J., Ngene, S., 2014. *Dry Season Fixed-wing Aerial Survey of Elephants and Other Large Mammals in the Tsavo and Amboseli Ecosystems, Kenya*. Great Elephant Census, a Paul G. Allen Project, Vulcan Inc., Seattle, USA.
- Cobb, S., 1976. *The Distribution and Abundance of the Large Herbivore Community in Tsavo National Park*. D.Phil Thesis. University of Oxford.
- Cook, R.D., Jacobson, J.O., 1979. A design for estimating visibility bias in aerial surveys.

- Biometrics 35 (4), 735–742.
- Couturier, S., Courtois, R., Crépeau, H., Rivest, L.-P., Luttich, S., 1994. Calving photo-census of the Rivière George caribou herd and comparison with an independent census. In: Prince George, 1–4 March. Proceedings of the 6th North American Caribou Workshop, vol. 1994. pp. 283–296.
- Craig, G.C., 2012. Aerial survey standards for the CITES-MIKE programme, version 2. CITES-MIKE Programme. UNEP/DEL/C, Nairobi, Kenya.
- Diciccio, T.J., Romano, J.P., 1988. A review of bootstrap confidence intervals. *J. R. Stat. Soc. Ser. B* 50 (3), 338–354. Retrieved from: <http://www.jstor.org/stable/2345699>.
- Douglas-Hamilton, I., Burrell, A., 1991. Using elephant carcass ratios to determine population trends. In *African Wildlife: Research and Management*. International Council of Scientific Unions, pp. 98–105.
- Douglas-Hamilton, I., Gachago, S., Litoroh, M., Mirangi, J., 1994. Tsavo Elephant Count 1994. Report to Kenya Wildlife Service, Nairobi.
- Eberhardt, L.L., 1978. Transect methods for population studies. *J. Wildl. Manage.* 42 (1), 1–31. <https://doi.org/10.2307/3800685>.
- Edin, H., 2014. Field-of-view Calculator for DSLR and SLR Cameras. Retrieved May 15, 2015, from <https://www.howardedin.com/articles/fov.html>.
- Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecol. Appl.* 6 (4), 1036–1046.
- Erwin, R.M., 1982. Observer variability in estimating numbers: an experiment. *J. Field Ornithol.* 53 (2), 159–167.
- Fleming, P.J.S., Tracey, J.P., 2008. Some human, aircraft and animal factors affecting aerial surveys: how to enumerate animals from the air. *Wildl. Res.* 35 (4), 258–267. <https://doi.org/10.1071/WR07081>.
- Frederick, P.C., Hylton, B., Heath, J.A., Ruane, M., 2003. Accuracy and variation in estimates of large numbers of birds by individual observers using an aerial survey simulator. *J. Field Ornithol.* 74 (3), 281–287. <https://doi.org/10.1648/0273-8570-74.3.281>.
- Gasaway, W.C., DuBois, S.D., Reed, D.J., Harbo, S.J., 1986. Estimating Moose Population Parameters From Aerial S. Report 22, Institute of Arctic Biology. University of Alaska, Fairbanks, Alaska.
- Graham, A., Bell, R., 1989. Investigating observer bias in aerial survey by simultaneous double-counts. *J. Wildl. Manage.* 53 (4), 1009–1016. <https://doi.org/10.2307/3809603>.
- Griffin, P.C., Lubow, B.C., Jenkins, K.J., Vales, D.J., Moeller, B.J., Reid, M., et al., 2013. A hybrid double-observer sightability model for aerial surveys. *J. Wildl. Manage.* 77 (8), 1532–1544. <https://doi.org/10.1002/jwmg.612>.
- Grimsdell, J.J.R., Westley, S.B., 1981. In: Grimsdell, J.J.R., Westley, S.B. (Eds.), *Low-Level Aerial Survey Techniques*. Monograph 4, International Livestock Centre for Africa, ILCA Monographs 4, Addis Ababa, Ethiopia.
- Holm, S., 1979. A simple sequential rejective method procedure. *Scand. J. Stat.* 6 (2), 65–70.
- Hopcraft, J.G.C., Holdo, R., Mwangomo, E., Mduma, S., Thirgood, S., Borner, M., et al., 2015. Why are wildebeest the most abundant herbivore in the serengeti ecosystem? In: Sinclair, A.R.E., Metzger, S.A.R., Mduma, S.A.R., Fryxell, J.M. (Eds.), *Serengeti IV: Sustaining Biodiversity in a Coupled Human Natural System*. University of Chicago Press, Chicago, IL, pp. 35–71.
- Jachmann, H., 2001. *Estimating the Abundance of African Wildlife*. Kluwer, Dordrecht, The Netherlands.
- Jachmann, H., 2002. Comparison of aerial counts with ground counts for large African herbivores. *J. Appl. Ecol.* 39 (5), 841–852. Retrieved from: <http://www.jstor.org/stable/827210>.
- Jacques, C.N., Jenks, J.A., Grovenburg, T.W., Klaver, R.W., Deperno, C.S., 2014. Incorporating detection probability into northern Great Plains pronghorn population estimates. *J. Wildl. Manage.* 78 (1), 164–174. <https://doi.org/10.1002/jwmg.634>.
- Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled SRTM for the Globe Version 4, Available From the CGIAR-CSI SRTM 90m Database. Retrieved February 28, 2017, from <http://srtm.csi.cgiar.org>.
- Jolly, G.M., 1969. Sampling methods for aerial censuses of wildlife populations. *East Afr. Agric. For. J.* 34 (1), 46–49. <https://doi.org/10.1080/00128325.1969.11662347>.
- Kaplan, E.D., Hegarty, C.J., 2006. *Understanding GPS. Principles and Applications*, 2nd ed. Artech House, Norwood, MA, USA.
- Kruger, J.M., Reilly, B.K., Whyte, I.J., 2008. Application of distance sampling to estimate population densities of large herbivores in Kruger National Park. *Wildl. Res.* 35 (4), 371–376. <https://doi.org/10.1071/WR07084>.
- Kruschke, J.K., 2013. Bayesian estimation supersedes the t test. *J. Exp. Psychol. Gen.* 142 (2), 573–603. <https://doi.org/10.1037/a0029146>.
- Kyale, D., Mukeka, J., Kimitei, K., Munyao, M., Lala, F., 2014. Tsavo-Mkomazi Total Aerial Census for Elephants and Large Animals. Technical Report to Kenya Wildlife Service, Nairobi.
- Laliberte, A.S., Ripple, W.J., 2003. Automated wildlife counts from remotely sensed imagery. *Wildlife Society Bulletin (1973-2006)* 31 (2), 362–371. Retrieved from: <http://www.jstor.org/stable/3784314>.
- Lamprey, R.H., 2005. Trial Integrated Land Use Survey of Four Districts of Uganda. Report to Uganda Bureau of Statistics and Belgian Technical Cooperation, Kampala, Uganda.
- Lamprey, R.H., 2016. Aerial Surveys of Murchison Falls National Park and Bugungu Wildlife Reserve, Uganda, 2015-16. Fauna & Flora International, Cambridge, UK. Report to Total E&P Uganda, Kampala, Uganda.
- Lee, D.E., Bond, M.L., 2016. Precision, accuracy, and costs of survey methods for giraffe (*Giraffa camelopardalis*). *J. Mammal.* 97 (3), 940–948. <https://doi.org/10.1093/jmammal/gyw025>.
- Leedy, D.L., 1948. Aerial photographs, their interpretation and suggested uses in wildlife management. *J. Wildl. Manage.* 12 (2), 191–210.
- Lemoine, F.G., Kenyon, S.C., Factor, J.K., Trimmer, R.G., Pavlis, N.K., Chinn, D.S., et al., 1998. The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96. Report NASA/TP-1998-206861. National Aeronautics and Space Administration, Greenbelt, Maryland.
- Leuthold, W., Leuthold, B.M., 1976. Density and biomass of ungulates in Tsavo East National Park, Kenya. *Afr. J. Ecol.* 14 (1), 49–58. <https://doi.org/10.1111/j.1365-2028.1976.tb00151.x>.
- Lubow, B.C., Ransom, J.L., 2016. Practical bias correction in aerial surveys of large mammals: validation of hybrid double-observer with sightability method against known abundance of feral horse (*Equus caballus*) populations. *PLoS One* 11 (5), 1–15. <https://doi.org/10.1371/journal.pone.0154902>.
- Magnusson, W.E., Caughley, G.J., Grigg, G.C., 1978. A double-survey estimate of population size from incomplete counts. *J. Wildl. Manage.* 42 (1), 174–176. <https://doi.org/10.2307/3800708>.
- Marsh, H., Sinclair, D.F., 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *J. Wildl. Manage.* 53 (4), 1017. <https://doi.org/10.2307/3809604>.
- Marshall, M., Norton-Griffiths, M., Herr, H., Lamprey, R., Sheffield, J., Vagen, T., Okotto-Okotto, J., 2017. Continuous and consistent land use/cover change estimates using socio-ecological data. *Earth Syst. Dyn.* <https://doi.org/10.5194/esd-8-55-2017>.
- Marsman, M., Wagenmakers, E.-J., 2017. Bayesian benefits with JASP. *Eur. J. Dev. Psychol.* 14 (5), 545–555. <https://doi.org/10.1080/17405629.2016.1259614>.
- McConville, A.J., Grachev, I.A., Keane, A., Coulson, T., Bekenov, A.B., Milner-Gulland, E.J., 2009. Reconstructing the observation process to correct for changing detection probability of a critically endangered species. *Endanger. Species Res.* 6 (3), 231–237. <https://doi.org/10.3354/esr00166>.
- McMahon, C.R., Howe, H., van den Hoff, J., Alderman, R., Brotsma, H., Hindell, M.A., 2014. Satellites, the all-seeing eyes in the sky: counting elephant seals from space. *PLoS One* 9 (3), e92613. <https://doi.org/10.1371/journal.pone.0092613>.
- Ministry of Tourism and Wildlife, 2018. The National Wildlife Conservation Status Report, 2015-2017. Kenya Wildlife Service and Ministry of Tourism and Wildlife, Nairobi, Kenya.
- Mukul, M., Srivastava, V., Mukul, M., 2015. Analysis of the accuracy of shuttle radar topography mission (SRTM) height models using international global navigation satellite system service (IGS) network. *Journal of Earth Systems Science* 124 (6), 1343–1357.
- Muller, Z., Bercovitch, F., Brand, R., Brown, D., Brown, M., Bolger, D., et al., 2016. Giraffa camelopardalis (amended version of 2016 assessment). The IUCN Red List of Threatened Species. Retrieved April 16, 2019, from <https://www.iucnredlist.org/species/9194/136266699>.
- Ndaimani, H., Murwira, A., Masocha, M., Gara, T.W., Zengeya, F.M., 2017. Evaluating performance of aerial survey data in elephant habitat modelling. *Afr. J. Ecol.* 55 (3), 270–281. <https://doi.org/10.1111/aje.12348>.
- Neumann, K.J., 2008. Trends for digital aerial mapping cameras. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII, pp. 551–554.
- Ngene, S., Ihwagi, F., Nzisa, M., Mukeka, J., Njumbi, S., Omondi, P., 2011. Total Aerial Census of Elephants and Other Large Mammals in the Tsavo-mkomazi Ecosystem. Technical Report to Kenya Wildlife Service, Nairobi.
- Ngene, S., Lala, F., Nzisa, M., Kimitei, K., Mukeka, J., Ihwagi, F., et al., 2017. Aerial Total Count of Elephants, Buffaloes and Giraffe in the Tsavo-mkomazi Ecosystem, February 2017. Kenya Wildlife Service, Nairobi, and Tanzania Wildlife Research Institute, Arusha.
- Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc. Natl. Acad. Sci.* 115 (25), E5716–E5725. <https://doi.org/10.1073/pnas.1719367115>.
- Norton-Griffiths, M., 1973. Counting the Serengeti migratory wildebeest using two-stage sampling. *Afr. J. Ecol.* 11 (2), 135–149. <https://doi.org/10.1111/j.1365-2028.1973.tb00079.x>.
- Norton-Griffiths, M., 1976. Further aspects of bias in aerial census of large mammals. *J. Wildl. Manage.* 40 (2), 368–371. <https://doi.org/10.2307/3800445>.
- Norton-Griffiths, M., 1978. *Counting Animals. African Wildlife Foundation Handbook Number 1*, Washington DC, USA.
- Norton-Griffiths, M., Frederick, H., Slaymaker, D.M., Payne, J., 2015. Aerial Census of Wildlife and Livestock in the Oyu Tolgoi Area of the Gobi Desert, Mongolia may - July 2013. Final Report to Oyu Tolgoi LLC, Ulaanbaatar, Mongolia.
- NTSB, 2018. Global Positioning System, Standard Positioning Service; Performance Analysis Report. Report to FAA, January 2017. Retrieved May 12, 2018, from http://www.ntsb.tc.faa.gov/reports/PAN96_0117.pdf.
- O'Connor, J., Smith, M.J., James, M.R., 2017. Cameras and settings for aerial surveys in the geosciences: optimising image data. *Prog. Phys. Geogr.* 41 (3), 325–344. <https://doi.org/10.1177/0309133317703092>.
- Ogutu, J.O., Piepho, H.-P., Said, M.Y., Ojwang, G.O., Njino, L.W., Kifugo, S.C., Wargute, P.W., 2016. Extreme wildlife declines and concurrent increase in livestock numbers in Kenya: what are the causes? *PLoS One* 11 (9), e0163249. <https://doi.org/10.1371/journal.pone.0163249>.
- Olindo, P., Douglas-Hamilton, I., Hamilton, P., 1988. The 1988 Tsavo Elephant Count. Report to Wildlife Conservation and Management Department, Nairobi.
- Ottichilo, W.K., 1987. Recent Elephant Population Estimates and Distributions in Kenya, 1982-1987. KREMU, for GEMS/United Nations Environment Programme, Nairobi.
- Ottichilo, W.K., Grunblatt, J., Said, M.Y., Wargute, P.W., 2000. Wildlife and livestock population trends in the Kenya rangeland. In: Prins, H.H.T., Grootenhuys, J.G., Dolan, T.T. (Eds.), *Wildlife Conservation by Sustainable Use*. Springer, Dordrecht, Netherlands, pp. 203–218. https://doi.org/10.1007/978-94-011-4012-6_10.
- PAEAS, 2014. Aerial Survey Standards and Guidelines for the Pan-african Elephant Aerial Survey. Vulcan Inc, Seattle, USA.
- Pavlis, N.K., Holmes, S.C., Kenyon, S.C., Factor, J.K., 2012. The development and

- evaluation of the Earth Gravitational Model 2008 (EGM2008). *J. Geophys. Res. Solid Earth* 117 (B4), 1–28. <https://doi.org/10.1029/2011JB008916>.
- Pennyquick, C.J., Western, D., 1969. An investigation of some sources of bias in aerial transect sampling of large mammal populations. *Afr. J. Ecol.* 10 (3), 175–191. <https://doi.org/10.1111/j.1365-2028.1972.tb00726.x>.
- Pollock, K.H., Kendall, W.L., 1987. Visibility bias in aerial surveys: a review of estimation procedures. *J. Wildl. Manage.* 51 (2), 502–510. <https://doi.org/10.2307/3801040>.
- Pollock, K.H., Marsh, H.D., Lawler, I.R., Alldredge, M.W., 2006. Estimating animal abundance in heterogeneous environments: an application to aerial surveys for dung. *J. Wildl. Manage.* 70 (1), 255–262. Retrieved from. <http://www.jstor.org/stable/3803568>.
- Ransom, J.I., 2012. Detection probability in aerial surveys of feral horses. *J. Wildl. Manage.* 76 (2), 299–307. <https://doi.org/10.1002/jwmg.204>.
- Rice, C.G., Jenkins, K.J., Chang, W.-Y., 2009. A sightability model for mountain goats. *J. Wildl. Manage.* 73 (3), 468–478. Retrieved from. <http://www.jstor.org/stable/40208552>.
- Rodríguez, E., Morris, C.S., Belz, J.E., 2006. A global assessment of the SRTM performance. *Photogramm. Eng. Remote Sensing* 72 (3), 249–260.
- Schlossberg, S., Chase, M.J., Griffin, C.R., 2016. Testing the accuracy of aerial surveys for large mammals: an experiment with African savanna elephants (*Loxodonta africana*). *PLoS One* 11 (10), e0164904. <https://doi.org/10.1371/journal.pone.0164904>.
- Siniff, D.B., Skoog, R.O., 1964. Aerial censusing of caribou using stratified random sampling. *J. Wildl. Manage.* 28 (2), 391–401. <https://doi.org/10.2307/3798104>.
- Steinhorn, R.K., Samuel, M.D., 1989. Sightability adjustment methods for aerial surveys of wildlife populations. *Biometrics* 45 (2), 415–425. <https://doi.org/10.2307/2531486>.
- Strobel, B.N., Butler, M.J., 2014. Monitoring whooping crane abundance using aerial surveys: influences on detectability. *Wildl. Soc. Bull.* 38 (1), 188–195. <https://doi.org/10.1002/wsb.374>.
- Tabak, M.A., Norouzzadeh, M.S., Wolfson, D.W., Halseth, J.M., Sweeney, S.J., Vercauteren, K.C., et al., 2019. Machine learning to classify animal species in camera trap images : applications in ecology. *Methods Ecol. Evol.* 10, 585–590. <https://doi.org/10.1111/2041-210X.13120>.
- TAWIRI, 2010. *Aerial Census in the Serengeti Ecosystem-wet Season*. Tanzania Wildlife Research Institute, Arusha, Tanzania.
- Terletzky, P.A., Ramsey, R.D., 2016. Comparison of three techniques to identify and count individual animals in aerial imagery. *J. Signal Inf. Process.* 7 (August), 123–135. <https://doi.org/10.4236/jsip.2016.73013>.
- Thouless, C.R., Dublin, H.T., Blanc, J.J., Skinner, D.P., Daniel, T.E., Taylor, R.D., Maisels, F., Frederick, H.L., Bouché, P., 2016. African elephant status report 2016: an update from the African Elephant Database. *Occasional Paper Series of the IUCN Species Survival Commission*, No. 60 IUCN / SSC Africa Elephant Specialist Group. IUCN, Gland, Switzerland vi + 309pp.
- Tracey, J.P., Fleming, P.J.S., Melville, G.J., 2008. Accuracy of some aerial survey estimators: contrasts with known numbers. *Wildl. Res.* 35 (4), 377–384. <https://doi.org/10.1071/WR07105>.
- Vermeulen, C., Lejeune, P., Lisein, J., Sawadogo, P., Bouché, P., 2013. Unmanned aerial survey of elephants. *PLoS One* 8 (2), e54700. <https://doi.org/10.1371/journal.pone.0054700>. Retrieved from.
- Wal, E.V., McLoughlin, P.D., Brook, R.K., 2011. Spatial and temporal factors influencing sightability of elk. *J. Wildl. Manage.* 75 (6), 1521–1526. Retrieved from. <http://www.jstor.org/stable/41418189>.
- Wanyama, F., Elkan, P., Grossman, F., Mendiguetti, S., Mwedde, G., Kato, R., et al., 2014. *Aerial surveys of murchison Falls protected Area*. Pan-African Aerial Elephant Census. Vulcan Inc., Seattle, USA.
- Whitehouse, A.M., Hall-Martin, A.J., Knight, M.H., 2001. A comparison of methods used to count the elephant population of the Addo Elephant National Park, South Africa. *Afr. J. Ecol.* 39 (2), 140–145. <https://doi.org/10.1046/j.1365-2028.2000.00285.x>.
- Willi, M., Pitman, R.T., Cardoso, A.W., Locke, C., Swanson, A., Boyer, A., et al., 2019. Identifying animal species in camera trap images using deep learning and citizen science. *Methods Ecol. Evol.* 10 (1), 80–91. <https://doi.org/10.1111/2041-210X.13099>.
- Xue, Y., Wang, T., Skidmore, A., 2017. Automatic counting of large mammals from very high resolution panchromatic satellite imagery. *Remote Sens. (Basel)* 9, 878–894.
- Yang, Z., Wang, T., Skidmore, A.K., De Leeuw, J., Said, M.Y., Freer, J., 2014. Spotting East African mammals in open savannah from space. *PLoS One* 9 (12), 1–16. <https://doi.org/10.1371/journal.pone.0115989>.
- Zabransky, C.J., Hewitt, D.G., Deyoung, R.W., Gray, S.S., Richardson, C., Litt, A.R., Deyoung, C.A., 2016. A detection probability model for aerial surveys of mule deer. *J. Wildl. Manage.* 80 (8), 1379–1389. <https://doi.org/10.1002/jwmg.21143>.