
RESEARCH AND METHODOLOGY

A Simple Method for the Analysis of Stratified Aerial Sample Counts

John Watkin¹ and Kes Hillman Smith²

¹African Conservation Centre, PO Box 62844, Nairobi, Kenya, ²Project Parc National de la Garamba, PO Box 15024, Nairobi, Kenya

ABSTRACT

Aerial sample count methods are widely used to accurately determine numbers and distribution of animal populations. In this paper we present a relatively simple procedure to depict aerial survey results on a spread sheet whose inner contour lines resemble the sampled area and give estimates of numbers or densities in individual cells. The procedure is based on the standard method of Jolly for the provision of the population estimate and a measure of the error (Jolly, 1969). Our method aims at improving the accuracy of the population estimate and at deriving simplified raster maps from spreadsheet programs such as Quattro Pro, Excel or Lotus.

RESUMEE

Les méthodes aériennes de comptage d'échantillon sont largement utilisées pour déterminer avec exactitude la taille et la répartition de populations animales. Nous présentons dans cet article une procédure relativement simple pour décrire les résultats de recensements aériens dans un tableau dont les lignes de contour intérieures représentent la surface échantillonnée et donnent une estimation des nombres ou densités à l'intérieur de chaque cellule. La procédure est basée sur la méthode standard de Jolly pour l'évaluation de la population et la mesure de l'erreur (Jolly, 1969). L'objectif de cette méthode est d'améliorer l'exactitude de l'estimation des populations et de dériver des cartes de quadrillage simplifiées à partir de tableurs comme Quattro Pro, Excel ou Lotus.

INTRODUCTION

Aerial sample count methods - perfected in the 1960s - have been comprehensively described by Norton-Griffiths (1978) and more recently by Mbugua (1996). These methods are widely applied for the census of wild and domestic mammals throughout Africa. Jolly's II method for unequal sized sampling units (Jolly, 1969) has been consistently used to provide both the population estimate and a measure of the error of the estimate. Norton-Griffiths (1978) also explains how stratification of

the census zone serves to reduce the margin of error in the population estimate. However, although both the practical aspects of the methods undertaking the sample count and the mathematics required to calculate the population estimate have been adequately described, there is little information on the methods that can be used to undertake a series of calculations to produce the population estimate and confidence limits, and figures to represent the distribution of wildlife in the count zone.

Several researchers have developed computer programs to calculate the Jolly's II population estimate

and map distributions (Burrill and Douglas Hamilton, unpubl.; Campbell, unpubl.; Western, 1976). However, these programs have often been tailored for individual areas or are not widely available and cannot be easily adjusted to other count zones.

This paper describes how readily available spreadsheet programs can be used to provide:

- a simple method of data entry
- an instant total of each species in each stratification block
- a simple calculation of both the population estimate and the 95% confidence limit
- figures depicting the distribution of species within the count zone - of both the numbers of each species per subunit or the density of each species per subunit.

This method of data analysis is of value to researchers who are undertaking the analysis of sample count data and who would like to improve on the accuracy of the population estimate and produce basic figures representative of the distribution of wildlife surveyed without the need for Geographic Information Systems. Other programs do exist for both of these aspects, but this method makes use of commonly available software (Fox, 1998) and does not require the most up-to-date computer hardware to undertake more elaborate analysis.

The paper is based on the use of QuattroPro 4.1 (Borland, 1992), but the approach can be used in all other similar spreadsheet programs eg Microsoft Excel, Lotus 1-2-3. In other spreadsheet programs most of the commands will be identical or require only limited changes.

Many aspects of this method of analysis can best be illustrated through tables. The majority of the tables attempt to depict spreadsheet windows with letters and numbers giving the co-ordinates for columns and rows respectively.

This method was developed in 1993 as part of the Garamba National Park Monitoring Programme (Smith et al., 1993) and has been thoroughly tested since (Hillman Smith et al., unpubl. 1995a,b).

METHODS

The analysis involves seven basic steps prior to and after flying the count.

Prior to flying the count:

1. Defining the count zone and individual strata on the spreadsheet (here after called MAPfile)
2. Preparing the spreadsheet for calculating the population estimate (referred to as JOLLY'SII file)

After flying the count:

3. Flying the calibration flights and producing the regression equation.
4. Using the altitude above ground level and the regression equation to calculate the strip width and N.
5. Calculating the area sampled per subunit from the altitude above ground level
6. Calculating the area sampled in each stratum
7. Transposing the totals for each stratum into the Jolly's II calculation.

Figure 1 gives an example of how the MAPfile relates to the area censused during the aerial count of Garamba National Park and the surrounding reserves.

So that this method of analysis can be applied to other aerial survey sites, this paper uses an example of a generic "count zone" comprising three strata A, B and C (Figure 2). From the description provided it is hoped that the approach can be used to create a mapfile and JOLLY'SII file that pertain to specific areas. These primary files can then be used in any subsequent counts of the same count zone.

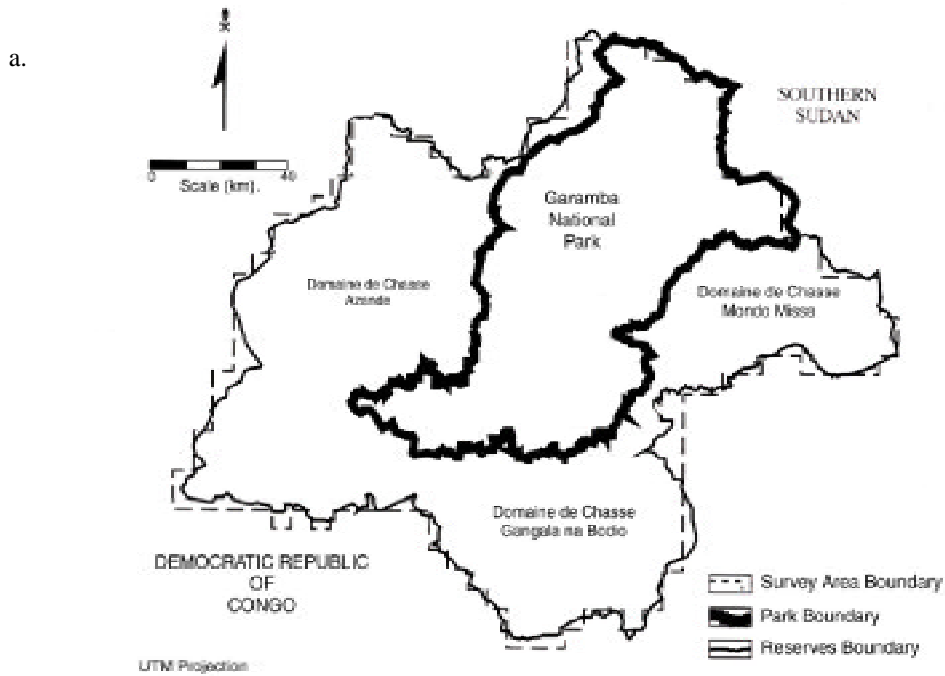
Prior to the flying the count

Definition of the survey area in the spreadsheet (mapfile).

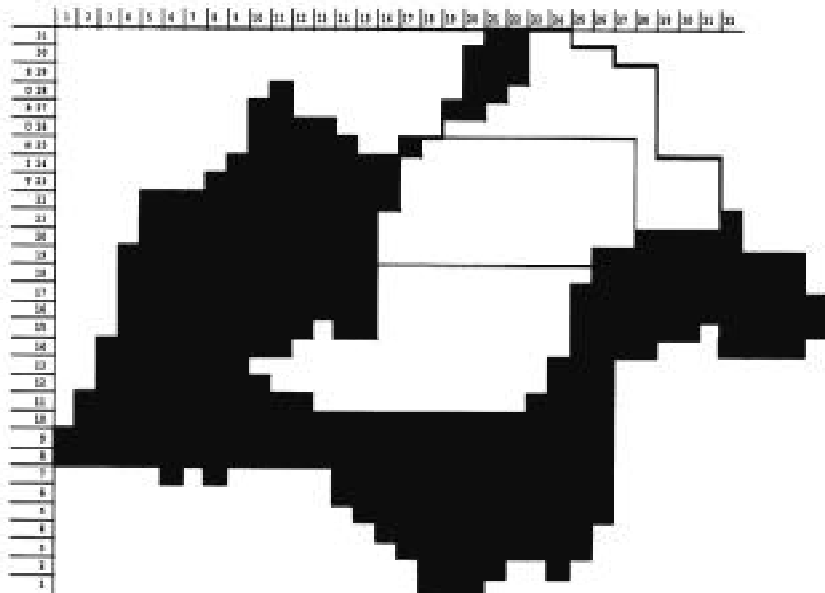
The analysis hinges upon a primary MAPfile which needs to be prepared in advance of the count. This file uses the spreadsheet columns as the transects and the rows as subunits or visa versa depending upon the direction that the transects are flown. The transects and subunits are numbered and the survey area defined with lines and shading for clarity. The individual strata should also be sketched in using lines (Figure 1b).

Once the MAPfile has been defined it is useful, when entering data, to lock the titles of both the columns and rows so that these are always displayed (freeze panes). The survey area defined in the MAPfile only occupies part of the grid established by the transect/subunit matrix. Many of the procedures in this analysis will involve writing an equation in the top left hand corner of the survey area (cell B2 in Figure 2) and then copying this formula across the whole sur-

Figure 1. Similarity between a digitised map (a.) and the Mapfile (b.) produced using the cells of the spreadsheet as subunits.



b.



vey area. Using Figure 2 as an example the source block would be cell B2, and the destination, B2..J10. The **Edit/cOpy Special/Contents** function is very powerful as it does not overwrite the formatting (lines and shading) used to define the survey area.

Stratification bar

Below the MAP defined on the spreadsheet, a series of rows which total the cells in each transect above need to be defined using the @SUM command. Cells falling into individual strata should be totalled in separate rows. Figure 2 illustrates how the stratification line relates to the strata defined in the MAP above.

This MAPfile forms the basis of the data entry, analysis and figures depicting the wildlife distributions. Several back-ups of this original file should be made in case the original file is overwritten.

Jolly'SII Population estimate spreadsheet (JOLLY'SIIfile)

The JOLLY'SIIfile is composed of two parts (Tables 1a-c): an upper section where data are received from the Mapfiles (Tables 1a and b) and a lower section in which the formulae are predefined to calculate the population estimates and confidence limits (Table 1c). The left hand block receives the area data calculated from the height AGL for each strata. Below these are the formulae used to calculate the areas used in the derivation of the population estimates.

To the right of the area data in the block G2 to J10, figures for the number of each species are copied from the species MAPfile. Below these columns are the calculations to produce the population estimate and confidence limits for each species by strata and the stratified total.

Tables 1a-c detail the steps involved in trans-

posing data from the MAPfile but also show the location of the formulae required to calculate the population estimate for each stratum, the 95% confidence limit and, hence, the stratified total. These formulae are detailed in Table 2. When adapting these formulae to another survey area the formula must include all the values in the column above. With the @SUMPRODUCT command, a blank cell in the column above will return an error (ERR) command. Blank values need to be replaced with a zero value to make the above column a continuous list of numbers.

The flight plan for an aerial sample count should provide the values of "N" (the total number of units in the population from which the transects were drawn) and "n" (the number of sample units in the sample). Many of the formulae in Tables 2 and 3 are anchored to specific cells, containing these values rather than having numbers contained in the formulae. QuattroPro uses the "\$" sign before the column and row values to fix on to one cell. This enables new values for n, N, and area sampled to be inserted into the MAPfile for subsequent counts. If these values need to be changed all the calculations will adjust automatically.

Once the calculations for the first animal population estimate have been completed it is simply a matter of copying the whole block across horizontally directly underneath the first column of the next stratum. In the case of Table 2 this involves copying the block G12..K25 across to L12. Block copying these formulae across reduces the chance of introducing errors.

Table 2 is a simplified version of the formulae which relate to the Mapfile and the data transposed

Figure 2. The grammar of the formula for each stratum in the stratification bar below the map. Note the use of commas as well as specified ranges of cells to include all cells of an individual stratum in the column above.

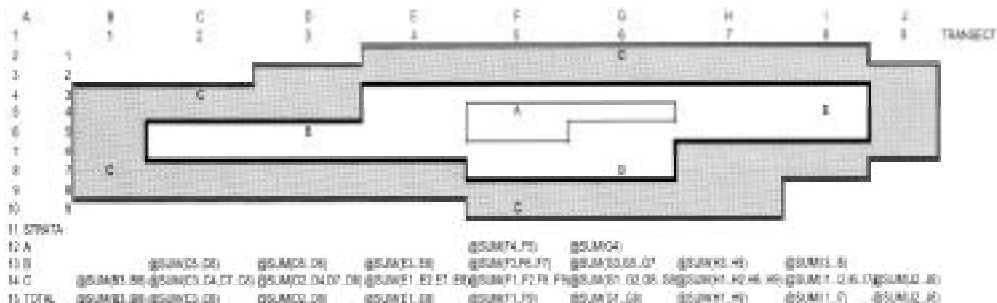


Table 1a. Original AREAfile.

	A	B	C	D	E	F	G	H	I	J
1	Transect	1	2	3	4	5	6	7	8	9
11	Strata									
12	A					3.4	1.7			
13	B		3.7	3.6	7.3	5.1	7.4	5.2	5.5	
14	C	9.7	7.8	9.5	7.2	7.5	7.1	11.7	7.2	9.6
15	Total	9.7	11.5	13.1	14.5	16	16.2	16.9	12.71	9.6

Table 1b. Original SPECIESfile.

	A	B	C	D	E	F	G	H	I	J
1	Transect	1	2	3	4	5	6	7	8	9
11	Strata									
12	A						37	9		
13	B			1	41	29	47	37	49	18
14	C		5	41	49	44	25	19	22	37
15	TOTAL		5	42	81	64	109	55	71	36

Table 1c. Transposed AREA and SPECIESfile into JOLLY'Sfile.

A	B	C	D	E	F	G	H	I	J	K	
STRATA					STRATA						
1	TRANSECT	A	B	C	TOTAL	A	B	C	TOTAL		
2	1				9.7				5	5	
3	2		3.7		7.8			1	41	42	
4	3		3.6		9.5			41	40	81	
5	4		7.3		7.2			20	44	64	
6	5	3.4		5.1	7.5			37	47	25	
7	6	1.7		7.4	7.1	16.2		9	37	9	
8	7			5.2	11.7	16.9			49	22	
9	8			5.5	7.2	12.7			18	37	
10	9				9.6	9.6				36	
STRATA					STRATA						
11	A	B	C	TOTAL	Total	A	B	C	TOTAL		
12	Total	SEE TABLE 1 FOR DETAILS OF THE FORMULAE IN THESE CELLS			Total	SEETABLE 1 FOR DETAILS OF THE FORMULAE IN THESE CELLS					
13	Sum of Squares				Sum of Squares						
14	n				Sum(Z*y)						
15	N				R=Sy/Sz						
16	Area				Vary						
17	Var Z										
18											
19					Covar zy						
20					Pop est						
21											
22					SE(Y)						
23					95%CL						
24					95% as %						
25											

Table 2. The spreadsheet for calculating Jolly's II population estimate. Columns wise calculations shown in Norton-Griffiths (1978) for an individual species the stratified sample estimate for the species.

	A	B	C	D	E	F	G
11		A	B	C	TOTAL		A
12	Total	@SUM(B6..B7)	@SUM(C3..C9)	@SUM(D2..D10)	@SUM(E2..E10)		@SUM(G6..G7)
13	Sumsqu	@SUMPROD (B6..B7,B6..B7)	@SUMPROD (C3..C9,C3..C9)	@SUMPROD (D2..D10,D2..D10)	@SUMPROD (E2..E10,E2..E10)		@SUMPROD (G6..G7,G6..G7)
14						Sum(Z*y)	@SUMPROD (\$B\$6..\$B\$7,G6..G7)
15	n	23	17	15	37		
16	N	343	255	223	500	R=Sy/Sz	@SUM(G12/\$B\$12)
17	AREA	5500	3550	1950	9675	Vary	@VARS(G6..G7)
18	Varz	@VARS(B6..B7)	@VARS(C3..C9)	@VARS(D2..D10)	@VARS(E2..E10)		
19						Covarzy	@SUM(1/\$B\$15-1) *(G14-(\$B\$12*G12) /\$B\$15))
20							SPECIES
21						Popest(Y)	(G16*\$B\$17)
22							
23						SE(Y)	@SQRT(@SUM(\$B\$16 *(\$B\$16-\$B\$15)/\$B\$15) (G17-2*G16*J19 +(G16^2)*\$B\$18))
24						95%C.L.	2.04*G23
25						Percentage	@SUM(G24/G21*100)

the constant values. Define the survey area block and enter the full value for the constant as the search string but do not enter anything as the replace string. This should remove all the values outside the MAP area.

Calculating n and N

Variations from the flight plan and differences in terrain mean that values for both N and n need to be calculated for each stratum from the actual strip width recorded for each sub-unit during each transect (Norton-Griffiths, 1978).

In the STRIPWIDTHfile the stratification bar summarises the strip width in each stratum which determines the value of N for each stratum. To

calculate N the @SUM command needs to be replaced with @AVG (Figure 3) This is easily done with the Edit/Search/Replace routine as above.

Calculating the area of each stratum

The regression equation giving the strip width will also provide the area surveyed in each subunit. The equation is:

$$\text{AREA SURVEYED} = ((mx) + c) * \text{subunit length}/1000$$

This calculation again should be entered into the block B2 and then using the Edit/cOpy Special/Contents function copied to the range e.g. B2..J11. The stratification bar below will total up the area surveyed in each stratum in each transect.

B to E calculate the area surveyed for each stratum. Columns G to J follow the step-in each stratum. Column K combines the estimates from individual stratum to reach

H	I	J	K
B	C	TOTAL	
@SUM(H3..H9)	@SUM(12..110)	@SUM(J2..J10)	
@SUMPROD	@SUMPROD	@SUMPROD	
(H3..H9,H3..H9)	(12..110,I2..I10)	(J2..J10,J2..J10)	
@SUMPROD	@SUMPROD	@SUMPROD	
(\$C\$3..\$C\$9,H3..H9)	(\$D\$2..\$D\$10,I2..I10)	(\$E\$2..\$E\$10,J2..J10)	
@SUM(H12/\$C\$12)	@SUM(I12/\$D\$12)	@SUM(J12/\$E\$12)	
@VARS(H3..H9)	@VARS(12..110)	@VARS(J2..J10)	
@SUM((1/\$C\$15-1)	@SUM((1/\$D\$15-1)	@SUM((1/\$E\$15)	START
(H14-(\$C\$12*H12)	*(I14-(\$D\$12*I12)	*(J14-(\$E\$12*J12)	TOTAL
/\$C\$15))	/\$D\$15))	/\$E\$15))	
(H16*\$C\$17)	(I16*\$D\$17)	(J16*\$E\$17)	@SUM(G521,H21,I521)
@SQRT(@SUM(\$C\$16	@SQRT(@SUM(\$D\$16	@SQRT(@SUM(\$E\$16)	@SQRT(@SUM
*(\$C\$16-\$C\$15)/\$C\$15)	*(\$D\$16-\$D\$15)/\$D\$15)	*(\$E\$16-\$E\$15)/\$E\$15)	(G23^2,H23^2,I23^2))
*(H17-2*H16*J19)	*(I17-2*I16*J19)	*(J17-2*J16*J19)+	
+(H16^2)*\$C\$18))	+(I16^2)*\$D\$18))	(J16^2)*\$E\$18)	
2.04*H23	2.04*I23	2.04*J23	@SUM(K23*1.96)
@SUM(H24/H21*100)	@SUM(I24/I21*100)	@SUM(J24/J21*100)	@SUM(K24/K21*100)

Data entry

Once the MAPfile has been finalised, with the stratification bar it is ready for entering species data. A new file is needed for each species.

Open the MAPfile and rename this file immediately with a name that refers to the count and the species to be entered. Data from the observers should be entered in the appropriate cell corresponding to the transect and subunit indicated on the count data sheets. It is useful if the left and right observed figures are entered in the same cell but as a part of a formula.

$$\begin{array}{rclcl} \text{Left Obs.} & & \text{Right Obs.} & & \\ 30 & + & 28 & = & 58 \end{array}$$

If only one observer recorded a sighting, enter a zero value for the other observer. This practice makes checking data entries far easier. If tests have been undertaken to correct for observer bias these values can be included in the formula at this point. The totals for the numbers of each species observed in each stratum should automatically be recorded in the stratification bar below.

Transposing data

Transposing rearranges data from rows into columns. This can be done either on the same spreadsheet or data can be transposed from one open file onto another open file. The latter procedure is exemplified in Table 1c.

The values for the area surveyed need to be transposed for N/S transects or copied for E/W from the area file into the JOLLY'SII file as do the values from the stratification bar for each animal species. It is less complicated if there are only two windows open at the same time. There are two steps involved in transposing data.

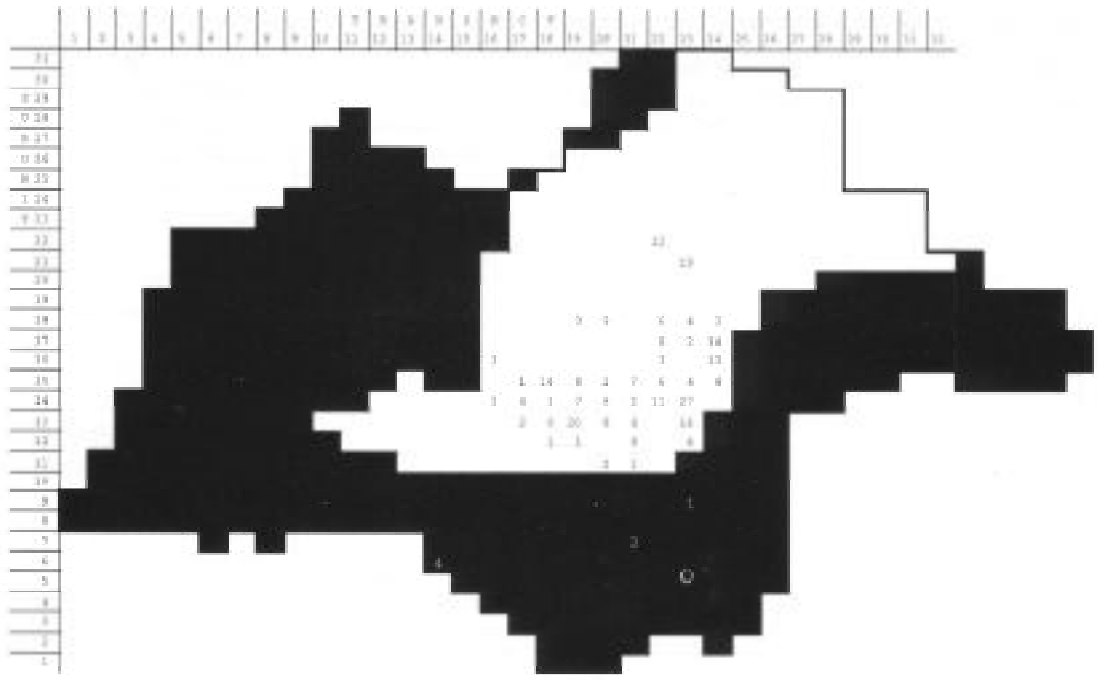
- The formula solutions in the stratification bar have to be converted into their true values. Using the **Edit/Values** command convert all the stratified totals for the range into their mathematical values, overlaying these values on to the original, or place them slightly below the stratification bar.
- These values have to be transposed into the JOLLY'SII file spreadsheet. Define the range of cells to be transposed i.e. the whole stratification bar. On Figure 2 this corresponds to B12..J15. When prompted for the destination, swap between open files to the JOLLY'SII spreadsheet. You need only define the top left hand corner of the block for where the columns need to be placed, and not the whole area for the columns.

In Table 1c the values for the area surveyed in each-transect were transposed from the range B12..J15 to cell B2. Likewise the count data for the first species was transposed from range B12..J15 into cell G2. One column needs to be left for calculating the stratified total and the next set of species data would be transposed into cell L2.

Table 3. A worked example of a stratified sample (Smith et al. 1993). Values for the area respective files into the relevant columns. The area and the variance for each transect are then calculated

Trans	Park	North	South	Domain
1				3.3
2				7.1
3				11.6
4				20.8
5				22
6				27.8
7				25.2
8				28
9				28.9
10	1.7		1.7	32.4
11	3.4		3.4	31.1
12	4.9		4.9	26.9
13	8.7		8.7	24
14	6.7		6.7	26.6
15	6.6		6.6	27.7
16	18.7	6.7	12	17
17	24.2	12.4	11.8	15.2
18	24.3	13.1	11.3	15.5
19	24.7	13.5	11.1	18.7
20	26.5	14.6	11.8	18.4
21	28.7	16.6	12.1	23.2
22	31.5	19.4	12.1	18.8
23	33.1	22.8	10.3	14.3
24	29.3	22.9	6.4	20.8
25	21.4	21.4		26.6
26	19	19		25.1
27	15.8	15.8		10.5
28	15.2	15.2		11.6
29	6.5	6.5		9.8
30	6.1	6.1		10.2
31	6.8	6.8		7.2
32	2	2		7.1
33				
12.3				
34				7.5
35				8.1
36				7.9
37				3.3
Total	365.7	234.9	130.8	652.3
Sumsqu	8,234.60	3,891.30	1,314.80	14,128.20
n	23	17	15	37
N	343	255	223	500
AREA	5500	3550	1950	9675
Varz	110	40.4	12.4	73

Figure 4. Garamba National Park aerial sample count May 1993. Point densities of elephants per 25km² sub unit.



```
A    B
1
2    @SUM([SPECIES-file]B2/[AREAfile]B2)
```

- Using the **Edit/cOpy Special/Contents** command copy this equation over the whole range of the survey area e.g. B2..J11. Cells outside the survey area, which have no values for the area or species, return an ERR message. Cells where the species were not present will return a zero value
- Using the **Edit/Values** function all the equations can be converted to their numeric values. The error (ERR) messages are also converted to text.
- The **Edit/Search** and replace function can be used to remove all the ERR symbols from the block. Define the whole MAP range. Enter ERR as the search text but leave the replace box blank. To remove zero values a supervised search and replace needs to be undertaken, confirming each replacement; otherwise, zero values contained in the density values will also be removed.

This should leave only the density values remaining in each cell for where the species was recorded. Figure 4 is an example of such a distribution map from the 1993 aerial survey of Garamba National Park and surrounding hunting areas.

CONCLUSION

The most difficult aspect of this method is defining the count MAPfile with the stratification bar and ensuring that the formulae in the JOLLY'sHfile include all the values in the column above and are linked to the relevant cells. This is a short-term problem and can be overcome with some attention to detail. In the long term, once established, the basic files can be used in analysing future counts of the same area. Other aspects of the analysis rely on sure footedness in using spreadsheet packages.

Data entry and analysis is overt and easy to verify. In addition the data from the figures depicting the distribution of species can be transferred to a raster or vector format for further analysis and

mapping using geographic information systems as illustrated in Hillman Smith (1995c).

The benefits of calculating the stratified population estimate can be seen in Table 3 where the 95% confidence limit is reduced from 56.9% to 35.0% of the population estimate.

ACKNOWLEDGEMENTS

These methods were developed as part of the Garamba Ecosystem Monitoring Programme which is part of the Garamba National Park Project (GNPP) with the Institute Congolaise pour la Conservation de la Nature. The Programme has been funded jointly by the World Wide Fund For Nature (WWF), Wildlife Conservation Society (WCS), Wildlife Conservation Fund, The World Conservation Union (IUCN) and the United Nations Educational, Scientific and Cultural Organisation (UNESCO). The Cessna 206 aircraft used belongs to, and is supported by, Frankfurt Zoological Society (FZS). The authors are indebted to reviewers who made constructive comments on the draft manuscript.

The authors would like also to acknowledge the support of the pilots and observers who have been involved in the aerial sample counts throughout the duration of the project.

REFERENCES

- Borland (1992) QuattroPro 4.1. Borland International. USA.
- Burrill, A. and Douglas-Hamilton, I. (unpubl.) Aerocount, a computer programme for analysis of aerial census data. Typescript 1984.
- Campbell, K. (unpubl.) SRF: Systematic reconnaissance flight software for aerial survey analysis. Computer software 1993.
- Fox, B. (1998) Keep it simple, keep it safe. *New Scientist* 2119, 46.
- Hillman Smith, A.K.K., Myayma, A., Likango, M., Smith, F., Ndey, A. and Panziama, G. (unpubl.) Parc National de la Garamba et Domaines de Chasse. General aerial count 1995 and evaluation of the status and trends of the ecosystem. GNPP Report, 1995a.
- Hillman Smith, A.K.K., Watkin, J.R., de Merode, E. and Smith, F. (unpubl.) Parc National de la Garamba et Domaines de Chasse. General Aerial Counts, manual of methods and analysis. GNPP Report, 1995b.
- Hillman Smith, A.K.K., de Merode, E., Nicholas, A., Buls, B. and Ndey, A. (1995c) Factors affecting elephant distribution at Garamba National Park and surrounding reserves, Zaire, with focus on human elephant conflict. *Pachyderm* 19, 39-48.
- Jolly, G.M. (1969) Sampling methods for aerial censuses of wildlife populations. *East African Agriculture and Forestry Journal* 34, 46-49.
- Mbugua, S. (1996) Counting elephants from the air - sample counts. In: *Studying elephants*. Handbook No. 7 in a series of Handbooks on Techniques currently used in African Wildlife Ecology. (Ed. K. Kangwana K.) AWLF, Nairobi, Kenya, pp. 21-27.
- Norton-Griffiths, M. (1978) *Counting Animals*. Handbook No. I in a series of Handbooks on Techniques currently used in African Wildlife Ecology. (Ed. J.J. Grimsdell). AWLF, Nairobi.
- Smith, A.K.K., Smith, F., Mbayma, A., Monungu, L., Watkin, J.R., de Merode, E., Amube, N. and Eza, K. (1993) Garamba National Park, General Aerial Count 1993. WWF/FZS/IZCN/IUCN/UNESCO Report.
- Western, D. (1976) An aerial method of monitoring large mammals and their environment, with a description of a computer programme for survey analysis. Project Working Document 9, UNDP/FAO Kenya Wildlife Management Project (KEN/71/526), Nairobi.