the cost of some 30 days, ten darts and limited quantities of drug. Under the conditions in the Sengwa Wildlife Research Area, drug immobilization did not provide a viable alternative to catching warthog at holes.

Acknowledgments

I thank the Director, Department of National Parks and Wild Life Management, Rhodesia, for permission to publish this work. The development of this technique was a team effort and it is a pleasure to acknowledge the enthusiastic co-operation of the Game Scouts at Sengwa, without whose help this work would not have been possible. Other members of staff assisted in the field at various times and to them I extend my thanks. I am also grateful to Dr J. Hanks for his constructive criticism of this paper in manuscript.

References

CUMMING, D.H.M. (1970 *A contribution to the biology of warthog* (Phacochoerus africanus *Gmelin*) *in the Sengwa region of Rhodesia.* Unpublished Ph.D. thesis, Rhodes University, Grahamstown.

CUMMING, D.H.M. (1975) A field study of the ecology and behaviour of warthog. *Mem. natn. Mus. Rhodesia.* (In press.)

HARTHOORN, A.M. (1965) Application of pharmacological and physiological principles in restraint of wild animals. *Wildl. Monogr.* No. 14, 78 pp.

PIENAAR, U. de V. (1968a) Capture and immobilizing techniques currently employed in South African National Parks and Reserves. In: *Productivity of Large herbivores* (Ed. by F. B. Golley and H. K. Buechner), I.B.P. Handbook No. 7, pp. 132–144. Blackwell Scientific Publications, Oxford.

PIENAAR, U. de V. (1968b) Recent advances in the field immobilization and restraint of wild ungulates in South African National Parks. *Acta. Zool. et Path. Antv.* 46, 17–38.

(*Manuscript received* 17 *December* 1973)

E. Afr. Wildl. J., 1975, Volume 13, pages 121–140

The numbers and distribution of large mammals in Ruaha National Park, Tanzania*

M. NORTON-GRIFFITHS† *Serengeti Ecological Monitoring Programme, Serengeti Research Institute*

Summary

Two aerial sample counts were carried out in Ruaha National Park, Tanzania, in the dry season of 1972 and the wet season of 1973. Population estimates of elephant, buffalo and other large mammals were made, and trend surface analysis was used to identify the basic gradients and contours of density, diversity and biomass within the Park. A cluster analysis, based on similarities between principal component scores, and validated by a multiple discriminant analysis, produced a classification of the Park into three Regions which differed significantly in densities, diversity and biomass.

Ruaha Park carries year round a high density of elephants (1·7 km^{-2}). Seasonal movements into and out of the Park were not found, although movements within and between Regions could be detected.

Introduction

Ruaha National Park lies in central Tanzania just west of the Iringa highlands. The Park is bordered by the Great Ruaha river to the south and by the Njombe river to the north, and it lies immediately adjacent to, and slightly south-east of, the Rungwa Game Reserve. Ruaha National Park is approximately 10 000 km^2. The main geographical feature (Fig. 1) is a low escarpment (*c.* 200 m) that runs parallel to the Great Ruaha River. The rift valley portion of the Park, i.e. that area below the escarpment, lies at about 3000′ (900 m) a.m.s.l. while the plateau above the escarpment rises in places to approximately 4500′ (1370 m) a.m.s.l. In the western part of the Park there are numerous hills and a small, but spectacular, montane area rising to some 6500′ (1980 m) a.m.s.l.

The Park is almost entirely wooded with the relatively few areas of open grassland restricted to the mbugas on the central part of the plateau. The western third of the Park consists of miombo woodlands (the vegetation boundary is shown in Fig. 1) so the Park contains part of the major ecotone found in this part of Tanzania between the woodlands dominated by *Brachystegia* and those dominated by *Acacia*. The

---

* S.R.I. Publication Number: 161.

† Present Address: Department of Zoology, Oxford University, England.

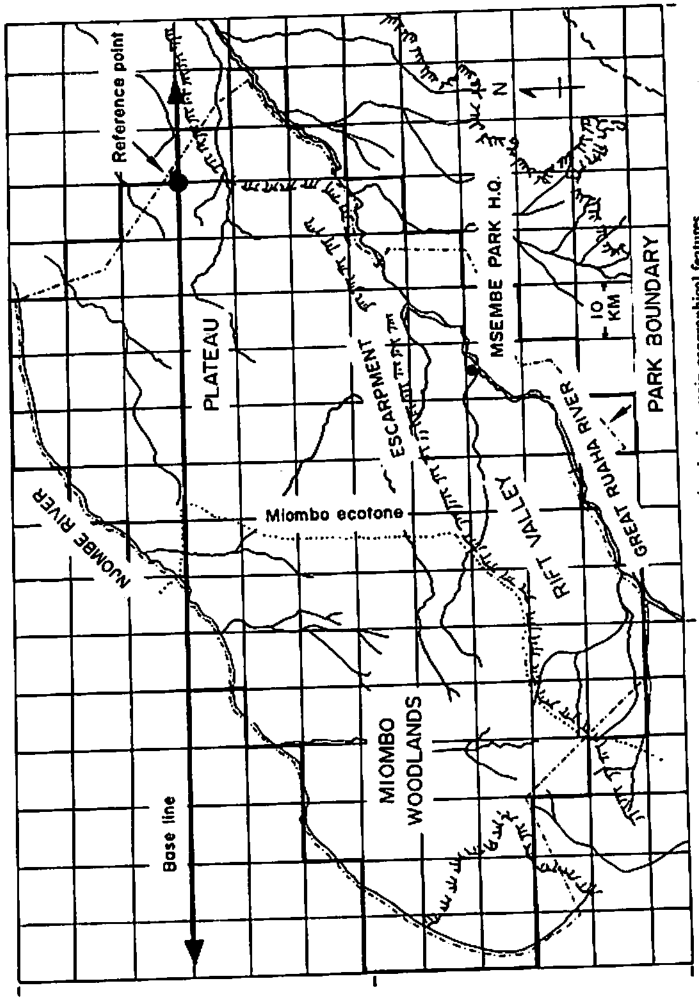

Fig. 1. Ruaha National Park, Tanzania, showing main geographical features.

plateau contains mixed *Acacia* and *Combretum* woodlands with areas of Baobab along the top of the escarpment, while the rift valley part contains extensive *Commiphora* woodlands, with areas of *A. tortilis* and Baobab in the area around the Park Headquarters at Msembe. A striking feature of the vegetation in this area is the mature stands of *A. albida* along the banks of the Great Ruaha River.

Like many other National Parks in East Africa, Ruaha is apparently suffering from an elephant 'problem'. Reports from the incumbent Chief Park Warden of severe damage to the vegetation by elephant in the Msembe area culminated in requests for a reduction cropping scheme to conserve the vegetation, especially the mature *A. albida* stands along the river. The Park Headquarters, the tourist camp site and the main tourist circuits were all apparently in danger of total deforestation.

Since there were no data on elephant numbers, distribution or movements within the Park as a whole it seemed premature to consider a cropping programme until some base-line information had been collected. An aerial sample count of elephant was therefore carried out in the dry season of 1972 (November), and because a surprisingly large number of elephant were found the count was repeated in the wet season of 1973 (April). The wet season count was to establish whether the Park always contained so many elephant or whether it acted as a dry season concentration area for the surrounding country.

The counts were designed to give data on distribution as well as numbers. Both elephant and buffalo were counted each time, and a number of other large mammals was included in the wet season count. A study of the population structure of elephant was made during the dry season count, the results of which will appear independently from this report. The species counted during the two censuses are listed below: Elephant *Loxodonta africana* (Blumenbach); buffalo *Syncerus caffer* (Sparrman); impala *Aepyceros melampus* (Lichtenstein); zebra *Equus burchelli* (Gray); giraffe *Giraffa camelopardalis* (L.); eland *Tauretragus oryx* (Pallas); rhinoceros *Diceros bicornis* (L.); kudu spp. *Tragelaphus* spp.

(I) The aerial census

(i) *Maps*

The design and execution of the two sample counts were confounded by a lack of good maps. The eastern two thirds of the Park were mapped at a scale of approximately 1 : 50 000 by the German Geological Mission in 1965 and this map was essentially correct. The western third was unmapped although a simple sketch map was available. The Park boundaries were not known accurately, neither were they marked physically on the ground, so there was often uncertainty as to where the Park actually was. The maps were improved upon whilst flying by marking in hills, etc., so by the second count a fair approximation to the ground was achieved. Nevertheless, the accurate measurement of area was not possible.

(ii) *Sample design*

Since distribution data were as important as numbers, and since the maps were of low quality, a complete systematic sample design was chosen. Aerial transects were spaced systematically across the Park and each transect was sub-divided along its length into sub-units of 5 km. An east–west base line of 159 km was constructed on the map (Fig. 1) so that it embraced the whole Park area, and the transects were located systematically along this base line from an arbitrarily selected 'reference point'. The transects were oriented north-south so that they cut across the major drainage systems of the Park and thus evened out any effect of elephant clumping along these rivers. This orientation had the additional advantage that shadow meters could be used for height control (see (iv) below).

The Park was divided into three strata for the dry season count. The eastern stratum consisted of the area for which there was an adequate map. The transects were spaced 1·5 km apart to give a sample fraction of 36% ($n=67$). The western stratum consisted of the area for which only the sketch map was available. The transects were spaced every 3 km (it was considered impossible to navigate more accurately than this) to give an 18% sample fraction ($n=18$) The third stratum consisted of the small area south of the Great Ruaha River. The transects were also 3 km part (sample fraction =38%, $n=13$).

The Park was sampled as one stratum in the wet season count. The transects were located every 5 km along the same base line and from the same reference point, and they were oriented in the same direction. The sample fraction was 5% ($n=32$).

(iii) *Aircraft and crews*

Two Cessna 182s were used in the dry season count, one in the wet season. Each aircraft had a crew of four consisting of pilot (navigation and height control), two observers in the back, and a recorder in the front right-hand seat who wrote down the observations called out by the observers. A Piper Cruiser (PA-12) carried out the study of the population structure of elephant in the dry season, and the block counts in the wet season (see II (ii) below).

(iv) *Flying height, strip widths and counting*

The flying height aimed for was 300' (91·4 m) above the ground. In the dry season count height control was by reference to a shadow meter (Pennycuick, 1973) by which height above ground is judged by comparing the size of the aircraft's shadow against markers placed on the wing strut. Both aircraft were equipped with shadow meters and these were calibrated against pressure altimeters to give a mean flying height for both aircraft of 310' (94·5 m) above the ground. A radar altimeter was used during the wet season count and the mean height was calculated from 1893 observations written down by the recorder at intervals unknown to the pilot. The mean height was found to be 303' (standard deviation of 32') (92·4 m with standard deviation of 9·8 m). An over-reading by the radar altimeter of 15' (4·6 m) was found from calibrations against the pressure altimeter so therefore the mean flying height was 288' (87·8 m).

The transect strips were defined by streamers attached to the wing struts. These were set up on the ground by the method described by Pennycuick & Western (1972), and they were calibrated by flying at known heights across large markers laid out on the ground. The mean strip width for the count was then calculated from the mean flying height. In the dry season count the mean strip width for both aircraft was 272 m either side (544 m total width). A much narrower strip was used in the wet season count because the vegetation was by then considerably thicker. Mean strip width was found to be 128 m either side of the aircraft (256 m total width).

The total number of transects that could have been located along the base line was found by dividing the length of the base line by the total strip width.

The observers counted all animals seen within the streamers and photographed any group of elephant larger than five individuals.

The pilot flew along a transect until a known boundary was reached, or until he considered, on the basis of time, that he had reached the approximate position of the Park boundary. The 5 km sub-units were demarcated by ground features where possible, otherwise by time flown. The pilot called out the beginning of each sub-unit and records were kept separately for each one.

(II) **Biases and errors**

(i) *Sampling design*

Although the transects were located systematically along the base line they were treated in the analysis as if they had been randomly located. The effect of this is to give an unbiased estimate while overestimating the variance.

The number of animals counted on each side of the aircraft was summed for each transect to give a transect total. These totals were then treated using Jolly's (1969a) Method 1 for equal-sized units, even though the units (i.e. the transects) differed widely in size. This had to be done because neither the transect across nor the total area under survey could be measured with any degree of confidence from the maps. Both of these have to be known to use Jolly's alternative Method for unequal sized units. The effect of treating the transects as being of equal size is to overestimate the variance while leaving the estimate unbiased (Norton-Griffiths, 1973).

These two biases in the sampling design therefore both overinflate the variance. The standard errors of the population estimates given in Table 2 are therefore overestimated.

(ii) *Counting*

In the dry season count groups of elephant numbering more than five individuals were photographed as well as being estimated and these estimates were later checked against the counts off the photographs. Due to poor quality of photography, and to some films being lost in the developing process, only 93 out of the 285 sets of photographs of elephant groups could be used. These 93 sets were therefore treated as being a sample from the 611 groups seen during the whole count. An overall counting bias was calculated using Jolly's (1969b) method of comparing the estimates against the photographic counts. A highly significant negative bias of – 7·25% was found, and this counting bias was corrected for by dividing the transect totals by 0·928 to give a corrected transect total. These corrected totals were then used in the ensuing calculations. Correcting for counting bias in this way can not of course correct for groups missed altogether (Norton-Griffiths, 1974).

Photography was not used in the wet season count and instead independent block counts were made by a second aircraft with pilot and observer. The blocks were located in different Park areas and different vegetation types (one block in the rift valley, one on the plateau, and two in the miombo woodlands) and they were made large enough so that they embraced a number of sub-units in neighbouring transects. The block was intensively total counted, and the estimate of density was compared with that obtained from the sub-units falling in the block. The sub-unit density was found by averaging the observations within each block after compensating for actual strip width (from radar altimeter readings). The two estimates of density from each block were then compared using a t-test as shown in Bailey (1959) where

$$t = \frac{\bar{d} - D}{s/n}$$

where $\bar{d}$ is the estimate of density in each sub-unit with standard deviation s and sample size n, and D is the estimate of density from the block total count. The observed differences between all blocks were then tested using a paired comparisons t-test (Bailey, 1959). The results of this analysis are shown in Table 1 where it can be seen that no significant differences exist between the two estimates of density.

Table 1. Tests for counting bias from block counts of elephant

Block	1	2	3	4
Area km^2	328	179	81	56
Number of sub-units from transects	15	10	3	4
D = density from block count	4·22	3·64	1·38	1·42
$\bar{d}$ = density from sub-units	6·15	3·24	1·50	1·48
Difference	1·93	−0·40	0·12	0·06
t-value	0·69	0·40	0·08	0·16
	ns	ns	ns	ns

From paired comparisons t test, t=0·84, ns; block 1 was located in the rift valley, block 2 on the plateau, and blocks 3 and 4 in the miombo woodlands.

Although this does not demonstrate a lack of counting bias it does at least indicate that the transect estimates are as reliable as those using another method.

The counting biases for other species were more difficult to correct. No corrections were made for those occurring at low densities and in small groups (giraffe, zebra, rhinoceros, eland, kudu spp.), while for impala and buffalo, corrections were based on figures obtained in the Serengeti. For impala, ninety-eight comparisons between estimates and photographic counts made during a multi-species sample count (Sinclair, 1972) showed a counting bias of −39·9%. The buffalo counting bias was calculated from data given in Sinclair (1973) only for groups not larger than 25 individuals (slightly larger than the largest group encountered in the Ruaha transects). The bias was found to be −30·3%. The transect totals for impala and buffalo were therefore divided by 0·601 and 0·697 respectively to give corrected transect totals. These corrections should strictly be based on data obtained in Ruaha itself. The estimates given in Table 2 can always be corrected later should such data become available.

(III) Numbers

The transect totals (corrected where possible) were treated using Jolly's (1969a) Method 1 for equal-sized sampling units. The population estimates, along with their standard errors and 95% confidence limits, are given in Table 2.

The two population estimates for elephant were not significantly different from each other, and they have therefore been combined (using the methods shown by Norton-Griffiths, 1973) to give a mean estimate of 16 355 elephant with 95% confidence limits of ±12%. Ruaha therefore carries year round a very large number of elephant, and there is no evidence for seasonal immigration/emigration on a Park wide basis. The difference between the wet and dry season estimates for elephant (9%) is the same as that for buffalo, which suggests that the difference was caused by the narrower counting strip used in the wet season rather than by immigration, for buffalo do not migrate.

Table 2. Population estimates from wet and dry season counts

Species	Count	Population estimate	Standard error	95% confidence limits
Elephant	Wet	17623	2110	4136 (23%)
	Dry	15966	1172	2297 (14%)
	Combined	16355	1021	2001 (12%)
Buffalo	Wet	14985	6054	11866 (79%)
	Dry	13886	1765	3459 (25%)
	Combined	13972	1695	3322 (24%)
Impala	Wet	13228	2699	5290 (40%)
Zebra	Wet	6025	1741	3412 (57%)
Giraffe	Wet	2430	514	1007 (41%)
Eland	Wet	2080	362	710 (34%)
Kudu spp.	Wet	797	339	664 (83%)
Rhinoceros	Wet	447	123	241 (54%)

The two estimates for buffalo were also combined to give a mean estimate of 13 972 with 95% confidence limits of ±24%. The population estimates for all other species should be treated as order of magnitude estimates only, for their standard errors are very high. This is not surprising considering the narrow counting strip and the small number of transects counted (32).

(IV) Distribution

(i) *Methods*

With two sets of data for elephant and buffalo and only one set for each of the other species it is only possible to make generalized statements about distribution within the Park. A grid of 10×10 km squares was superimposed upon the Park map so that distribution could be studied with respect to a spatial framework of sample units. This divided the Park into 102 squares, and the numbers of animals counted in the sub-units falling within each of them was averaged, corrected for counting bias where possible, and converted to give estimates of density (expressed as n km^{-2}). The wet and dry season observations of elephant and buffalo were averaged, whilst for impala, giraffe and zebra only wet season observations were available. Species diversity, total large mammal density and large mammal biomass were also calculated. Diversity was measured by summing the number of species seen in each grid square (the data being too sparse to permit a more elaborate analysis); total density was found by summing the estimates of density within each grid square; and biomass was calculated by summing the products between density and the average weights given by Lamprey (1964) for elephant, by Sinclair (1970) for buffalo and by Sachs (1967) for impala, giraffe and zebra. Biomass was expressed in kg ha^{-1}.

Trend surface analysis was used to identify the basic contours and gradients of each of these variables across the Park. Trend surface analysis is essentially a three-dimensional multiple regression that uses polynomials of increasing power to fit surfaces of increasing complexity to values measured at points located on a grid co-ordinate system. The fundamentals of this approach may be found in Davis (1973), King (1969) and Doornkamp and King (1971), while Gittins (1968) and Harris (1972) discuss applications of the method to ecological data. The computer programme used for this analysis uses a least squares method to fit polynomial equations of increasing order from linear up to quintic. At each stage the surface is calculated and an analysis of variance carried out.

The analysis of variance could be used to determine whether any significant reduction in the residual sum of squares had been achieved by the addition of an extra order to the equation. The highest order equation which gave a significant reduction in the sum of squares (i.e. a significant *F* ratio) was taken as that which fitted the data best. Table 3 shows the order of the surface of best fit for each of the eight variables, with the *F* ratio and significance level for the addition of that order, and also the coefficient of determination which expresses the percentage of the total sum of squares which has been accounted for.

Table 3. Results of trend surface analysis

Variable	Surface of best fit	Coefficient of determination	Component's *F* ratio	*P* level
Elephant	Quartic	21%	3·12	<0·05
Buffalo	Quartic	22%	3·41	<0·05
Impala	Quartic	34%	3·49	<0·05
Giraffe	Quartic	21%	2·96	=0·05
Zebra	Quartic	21%	3·86	<0·05
Diversity	Quartic	55%	6·00	<0·01
Total density	Cubic	35%	3·01	<0·05
Biomass	Quartic	29%	3·70	<0·05

(ii) *Results*

It is immediately apparent from the coefficients of determination in Table 3 that local variations are extremely important here, for these coefficients are low for such high order polynomials. Table 3 shows that the trends, although real, account for a relatively small % of the variations in density across the Park.

The contour maps (Figs. 2a–h) were drawn from the surfaces printed out by the trend surface programme. The outline of the Park, the position of the rift wall and the miombo ecotone are shown as well. The most striking feature of all the maps is the marked gradient of increasing density away from the miombo woodlands towards the eastern part of the Park. A second striking feature is the marked north-east/south-west ellipsoid nature of the contours, running parallel to the rift wall and often centred over it.

Each of the maps have individual characteristics as well. Elephants (Fig. 2a) have a ridge of very high density running along the rift wall, whilst buffalo (Fig. 2b) show two such ridges, one along the eastern rift wall and one along the Njombe river in the north of the Park. Impala (Fig. 2c) are very much restricted to the rift valley area of the Park and they show a steep gradient of decreasing density towards the north. Giraffe (Fig. 2d) and zebra (Fig. 2e) have similar distributions with ridges of high density along the plateau. The species diversity map (Fig. 2f) shows a ridge of high diversity along the rift wall and along the northern end of the miombo ecotone, whilst the maps for total density and biomass (Fig. 2g, h) show high ridges along the rift wall.

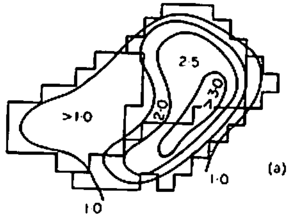

Fig. 2. Contour maps produced from trend surface analysis. Contours represent density km^{-2} in

All sightings (without density estimates) of rhinoceros and kudu spp. are shown in Fig. 2i. Rhino occur widely throughout the rift valley but have a peculiar distribution around the edges of the plateau. Kudu spp. seem to occur most in the eastern part of the Park above the rift wall, though many were also seen along the Njombe river in the north.

(V) A regional classification of Ruaha National Park

(i) *Methods*

The results of the trend surface analysis indicated that local effects were very important in explaining the variations in density, diversity and biomass within the Park. This was shown by the low coefficients of determination: by the similar basic gradients and contours in the maps; and by some marked similarities between different areas of the Park (for example, the rift valley wall, in stark contrast to the miombo woodlands, is characterized by high densities, diversity and biomass). These indications of high regional diversity within Ruaha made it worthwhile to attempt a Regional Classification that might identify areas that differed fundamentally in their general ecological characteristics.

The methods used are essentially the same as those described in detail by Doornkamp & King (1971). A clustering analysis cast the grid squares into groups in which the individuals were more similar to each other than they were to individuals in other groups. The statistical validity of the clustering was then tested with multiple discriminant analysis.

The raw data were the densities of elephant, buffalo, impala, giraffe and zebra; species diversity; total large mammal density; and total large mammal biomass in each of the 102 grid squares. One immediate problem was that since some of these variables were of a very different 'kind' (e.g. densities as *n* km^{-2}, biomass as kg ha^{-1}) they could not therefore be used validly within the same classification procedure (see Grigg, 1967). There were also marked differences in scale between many of the variables. These problems were overcome by first using principal components analysis and then carrying out the clustering on the component scores. The main advantage of principal components here is that differences in 'kind' and scale between the variables no longer matter (because the components are extracted from a correlation

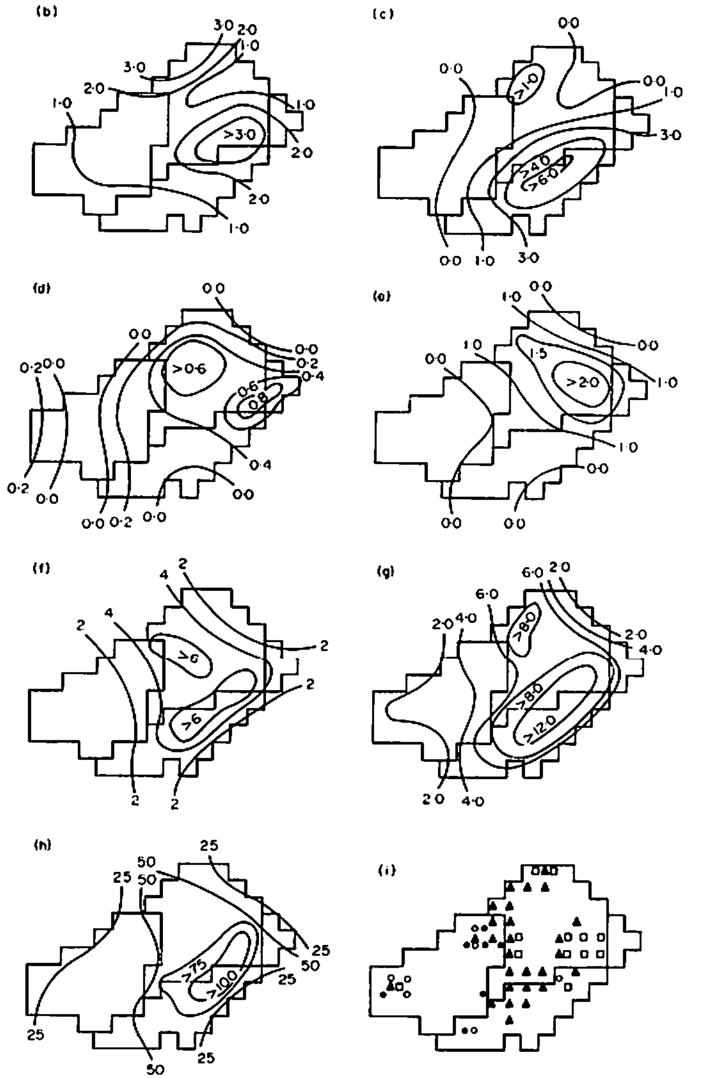

Fig. 2a–e, and g; number of species seen in Fig. 2f; and biomass in kg ha^{-1} in Fig. 2h. Fig. 2i shows sightings only. (a) Elephant; (b) buffalo; (c) impala; (d) giraffe; (e) zebra; (f) diversity; (g) total density; (h) biomass; (i) rare species: □, kudu spp.; ▲, rhinoceros; ○, roan antelope; ●, sable antelope.

matrix), but there are additional advantages in that the original data set is collapsed into fewer variables (components), and that interrelationships between the variables can be identified and described.

The result of the principal components analysis is shown in Table 4. Of the eight components, the first five each accounted for more than 10% of the total variation, and together they accounted for 93% of the total variation. These first five components were therefore used in the subsequent analysis. Although it would not matter in the subsequent analysis whether or not the components could be identified, the variable loadings in fact made this possible. The first component, accounting for some 40% of the total variation, is a generalized component that expresses the species diversity, total density and biomass within each grid square. The second component expresses elephant density and biomass. The third has heavy loadings for giraffe and zebra, whilst the fourth and fifth have heavy loadings for buffalo and zebra respectively.

Table 4. Variable loadings on the first five principal components

	Components						
	1	2	3	4	5	Σ1–5	Total
Eigenvalue	3·17	1·46	1·24	0·83	0·78	7·48	8·00
% of variation	40%	18%	15%	10%	10%	93%	100%
Elephant	0·29	0·47	−0·42	0·06	0·08		
Buffalo	0·24	−0·49	−0·02	0·70	0·27		
Impala	0·33	−0·42	−0·20	0·53	−0·37		
Giraffe	0·22	0·28	0·59	0·23	−0·45		
Zebra	0·23	0·07	0·51	−0·40	0·70		
Diversity	0·43	0·00	0·30	0·06	−0·29		
Total density	0·50	−0·32	−0·09	−0·06	0·08		
Biomass	0·42	0·42	−0·28	0·09	0·10		

The scores on the five components in each grid square were found by summing the products between the standardized variables and their component loadings. The original $n \times m$ data matrix (for $n=1, 2 \ldots 102$ grid squares, and $m=1, 2 \ldots 8$ variables) was thus transformed into an $n \times k$ matrix of component scores (for $k=1, 2 \ldots 5$ scores). The numeric taxonomic distance (Sokal & Sneath, 1963) was used to calculate the 'similarity' between each grid square and all the others on the basis of their five component scores. This statistic measures the Euclidean distance, d_{ij}, between two points in k-dimensional space, and it is computed by

$$d_{ij}=\left(\sum^{k}(X_{ik}-X_{jk})^2/k\right)^{\frac{1}{2}}$$

where d_{ij} is the distance between the ith and the jth grid square, X_{ik} is the kth component score on individual i and X_{jk} is the kth component score on individual j. A small value of d_{ij} shows that two squares are quite similar to each other whilst a large value of d_{ij} shows that they are very dissimilar.

A single linkage cluster analysis of the Euclidean distances cast the grid squares into groups. Single linkage clustering is the simplest one available and it seems to suffer from as many disadvantages as do any of the myriad alternative methods The problem with any clustering analysis is to decide when to stop, i.e. to decide how many groups to use, for this can be very subjective. In Fig. 3, the number of clusters, and the number of individuals that have been clustered, are plotted against decreasing similarity (i.e. increasing Euclidean distance). There is initially a sharp increase in the number of clusters followed by a decrease until, by similarity level 70, the clustering stabilizes at three clusters. It is not until similarity level 290 that all the individual squares become clustered into one group (which they inevitably must do). The number of individuals clustered shows a quite different pattern, rising evenly to similarity level 70 at which point 90% of the squares belong to one of three clusters. This similarity level of 70 was therefore chosen to differentiate between the clusters.

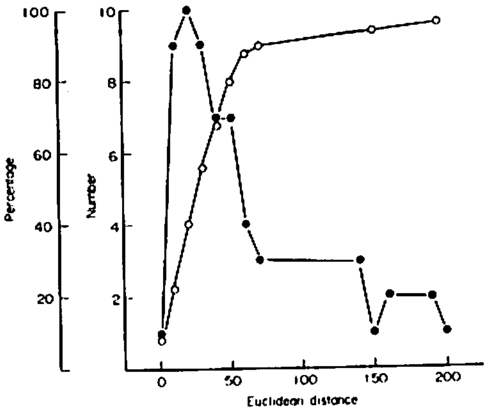

Fig. 3. Relationship between the number of clusters (●) and the percentage of individuals clustered (○) with increasing Euclidean distance (i.e. decreasing similarity).

These three clusters were considered to be 'proto-groups' that represented fundamental differences between the grid squares. The mean scores for the five components were calculated for each proto-group and the distance statistic was once more used to compare each square, including the unclustered ones, against these three sets of mean component scores. Each square was assigned to the proto-group which it most resembled. This resulted in the formation of three groups that accounted for all of the grid squares.

The statistical validity of this clustering was examined with a multiple discriminant analysis of the component scores which tested the null hypothesis that the vectors of group means come from the same multivariate population. The results (Table 5) show clearly that the three groups are highly significantly different from one another on the criteria of both Wilks' Lambda and the Mahalanobis' distance. From this it may be concluded that three statistically valid groups of grid squares can be distinguished within the Park.

(ii) *Results*

The location of the three groups of grid squares (Fig. 4) show enough spatial coherence to allow the term Region to be used, although they do not form such distinct Regions as perhaps one might have hoped for. Region I is located mainly in the west of the Park, especially in the miombo woodlands, while Region II, lying more to the east, occupies most of the plateau. Region III lies along the rift wall and accounts for most of the rift valley, but it also includes the southern part of the plateau and much of the miombo ecotone. Part of Region I also lies slap in the middle of Region III.

Table 5. *F* ratios from Mahalanobis' *D*-squared statistic and Wilks' Lambda in multiple discriminant analysis of principal component scores

	Group 2	Group 3
Group 1 versus:	$F=22{\cdot}7$ (d.f. 5, 69)	$F=76{\cdot}9$ (d.f. 5, 55)
Group 2 versus:		$F=23{\cdot}5$ (d.f. 5, 50)

F ratio from Wilks' Lambda = 32·7, d.f. 1 = 10, d.f.2 = 178.

The Regions differ markedly in their mean component scores (Table 6) with Region I having low scores on all components and Region III having the highest scores. Region II is intermediate. It is interesting that the Regional boundaries approximate to ellipses with north-east/south-west axes, and this probably accounts for the similar patterns seen in many of the contour maps. In addition, the low coefficients of determination of the trend surfaces probably arise from the marked mosaic nature of the spatial distribution of these three Regions.

Table 6. Mean component scores in each group of grid squares

		Group 1 ($n=40$)	Group 2 ($n=35$)	Group 3 ($n=27$)
Component 1	Mean	13	44	64
	se	1·5	1·2	7·3
Component 2	Mean	9	34	56
	se	1·2	1·4	6·5
Component 3	Mean	7	25	39
	se	4·9	4·4	4·5
Component 4	Mean	3	8	12
	se	0·4	0·6	1·6
Component 5	Mean	2	9	14
	se	0.3	0·5	1·6

Of more interest are the mean values of the original eight variables in each Region. Differences between Regional means are difficult to interpret at this stage for there is not enough information yet on vegetation types, productivity, etc. However, comparisons between Regional means and the overall mean (for the whole Park) are of importance for this shows whether the variables are distributed randomly with respect to the area of the three Regions. If, for example, elephants were distributed randomly within the Park then their densities within each Region would be the same and would be equal to the overall density, for the numbers in each Region would have been proportional to the area of each Region. Any departure from a random distribution would be shown by significant differences between a Regional

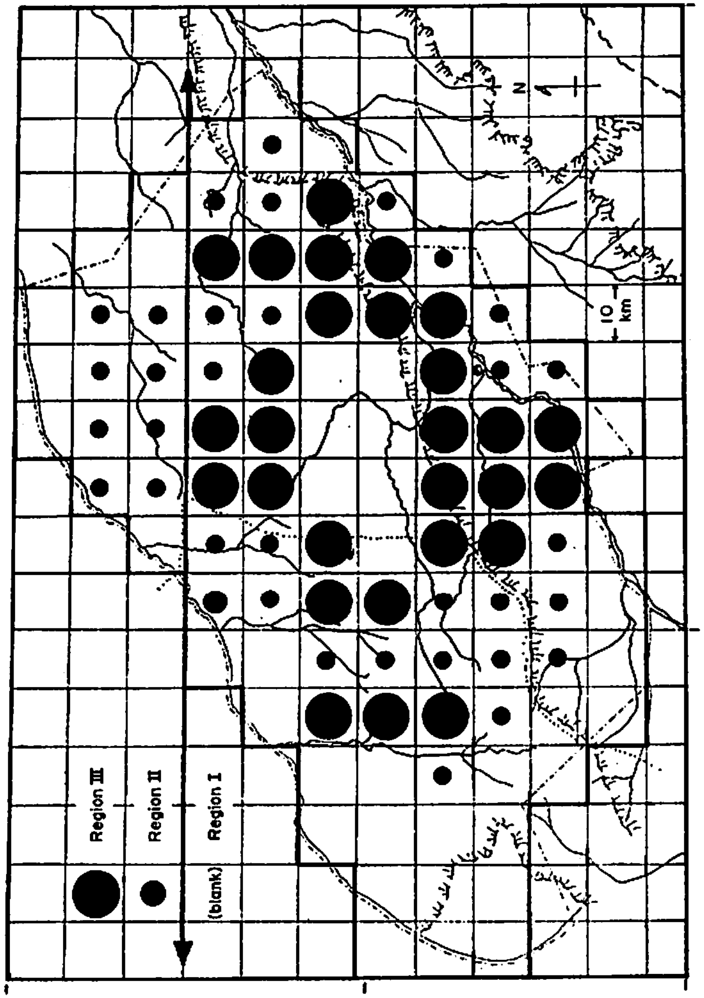

Fig. 4. Location of the three Regions within Ruaha Park.

mean and this overall mean. A simple one-way analysis of variance (Sokal and Rohlf 1969) is the most straightforward method for showing this for any significant departure from randomness will result in a significant *F* ratio.

Table 7 gives the results from a simple one-way anovar of each variable. Of the eight variables, six have significant *F* ratios. In general, the means of Region III tend to be the highest and those of Region I the lowest, Region II having intermediate values. The one exception here is the mean density of buffalo which is higher in Region II than in Region III or I (the resourceful reader can carry out his own tests of significance). These Regional means thus reflect the pattern of mean component scores.

Table 7. One-way anovar of each variable

			Overall		Region I		Region II		Region III	
	F	*P*	$\bar{x}$	se	$\bar{x}$	se	$\bar{x}$	se	$\bar{x}$	se
Elephant	48·0	<0·001	1·7	0·1	0·4	0·2	1·6	0·2	3·7	0·3
Buffalo	3·4	<0·05	1·5	0·3	0·8	0·5	2·5	0·5	1·1	0·5
Impala	4·4	<0·01	1·3	0·4	0·3	0·6	1·1	0·6	3·0	0·7
Giraffe	0·6	ns	0·4	0·1	0·3	0·1	0·4	0·1	0·5	0·2
Zebra	2·0	ns	0·6	0·2	0·2	0·3	0·8	0·3	1·1	0·3
Diversity	7·0	<0·01	2·9	0·2	2·2	0·3	3·2	0·2	3·6	0·3
Total density	12·0	<0·01	5·6	0·5	2·5	0·9	6·4	0·9	9·2	1·0
Biomass	74·0	<0·001	45·7	3·0	10·9	5·1	40·4	5·0	104·1	5·8
			(n=102)		(n=40)		(n=35)		(n=27)	

Means (x) and standard errors (se) were computed from the analysis of variance tables; F=regional MS/residual MS; d.f. 1=1; d.f. 2=99; P=probability level.

The nature of the differences between the Regions are illustrated in Table 8 where the means have been transformed into percentage for easier interpretation. The percentage of the total area of the Park falling into Regions I, II and III is 39%, 34% and 26% respectively. If the variables had been randomly distributed with respect to the Regions then the percentage in each Region should be the same as these overall percentages. The observed percentages are shown, the test for statistical difference being a *t*-test that compared each Regional mean against the appropriate

Table 8. Percentage distribution of each variable within the three Regions

	Region I	Region II	Region III
% of total area	39%	34%	26%
Elephant	9%(–)	33%(O)	58%(+)
Buffalo	21%(O)	59%(+)	20%(O)
Impala	9%(–)	29%(O)	62%(+)
Giraffe	30%(O)	35%(O)	35%(O)
Zebra	12%(O)	43%(O)	45%(O)
Diversity	30%(–)	38%(O)	32%(+)
Total density	17%(–)	39%(O)	44%(+)
Biomass	10%(–)	30%(O)	60%(+)

(–) and (+) indicate that the regional means are significantly lower or higher than the overall mean at $P = <0{\cdot}05$ (t-test); (O) shows no significant difference.

grand mean.* Region III has significantly high proportions of elephants, impala, diversity, total density and biomass, while Region I has significantly low proportions of these same variables. Region II has intermediate proportions (i.e. the variables are distributed randomly with respect to this Region) with the exception of buffalo which have a significantly high proportion in this Region.

The distribution of zebra and giraffe are random with respect to the three Regions. Although this may at first sight appear surprising, it is unlikely that any classification procedure will produce groups that differ significantly on all the variables used. The densities of these two species are low and are therefore relatively unimportant in comparison to the higher densities and biomass of the other species.

(VI) Movements of elephant

With the caveat that the following analysis is based on two counts only, movements both within and between Regions can be detected. Local movements within Regions are shown from correlating the observed wet and dry season densities in each grid square. Although Region I shows a random correlation ($r=-0{\cdot}17$, d.f.$=38$, ns), both Regions II and III show strong negative correlations of $r=-0{\cdot}71$ (d.f.$=33$, $P=<0{\cdot}001$) and $r=-0{\cdot}49$ (d.f.$=25$, $P=<0{\cdot}01$) respectively. The parts of these two Regions that are occupied during the wet season are therefore not occupied to the same extent during the dry season. This is supported by a series of counts made by J. Savidge (personal communication) during 1965 and 1966 in a block of 320 km^2 in the rift valley portion of Region III. These showed a highly seasonal movement of elephant (Fig. 5).

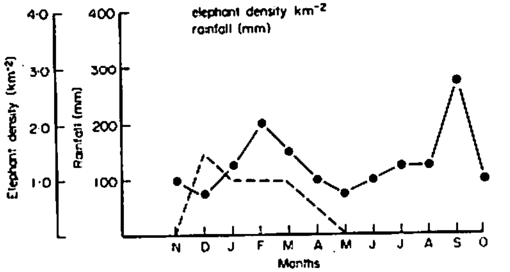

Fig. 5. Seasonal variations in elephant densities (——) within a 320 km^{-1} block in Region III (the block is situated between the two rivers west of Msembe). Rainfall (- - - - -)

Movements between Regions can be revealed through a two-way anovar (Sokal & Rohlf, 1969; Kirk, 1968). Table 9a gives the mean densities in each Region for the wet and dry seasons, the average wet and dry season densities, and the average Regional densities. These densities have been transformed into percentage of the total elephant seen on both counts in Table 9b, whilst Table 9c gives the results of the analysis of variance. The null hypothesis in this anovar is that the elephant are distributed randomly with respect to the area of each Region in both seasons of the year, and that therefore the Regional (and Seasonal) means are equal, and are also equal to the grand mean. Significant *F* ratios thus indicate departures from this hypothesized random distribution. Two-tailed tests of significance can be used since no predictions are made as to the direction of any observed differences.

Table 9. The analysis of the seasonal movements of elephant

(a) Mean densities

	Region I	Region II	Region III	Seasons
Wet season	0·135	1·391	4·556	1·736
Dry season	0·625	1·809	2·889	1·630
Regions	0·380	1·600	3·722	1·683

(b) Densities expressed as % of total elephant seen on both counts (% in ( ) are those expected on the hypothesis of random distribution)

	Region I	Region II	Region III	Seasons (50%)
Wet Season	2% (19%)	14% (16%)	36% (13%)	52%
Dry Season	7%	18%	23%	48%
Regions	9% (39%)	32% (34%)	59% (26%)	100%

(c) Results from two-way anovar

Source	*SS*	d.f.	*MS*	*F*	*P*
1 Seasons	0·600	1	0·6	0·165	ns
2 Seasons in R I	4·802	1	4·802	1·319	ns
3 Seasons in R II	3·045	1	3·045	0·837	ns
4 Seasons in R III	37·5	1	37·5	10·302	<0·01
5 Regions	361·0	2	180·5	49·588	<0·001
6 Regions in wet	321·33	2	60·665	44·139	<0·001
7 Regions in dry	84·31	2	42·155	11·581	<0·01
8 Interaction	44·4	2	22·2	6·099	<0·01
Error	684·0	188	3·64		
Total	1090·0				

Of the two main effects (Table 9c, rows 1 and 5), only the Regional one is significant. This is exactly as expected, for the two counts indicated that there was no major seasonal movement into or out of the Park, while the one-way anovar of Section V has already demonstrated the Regional effect. However, the Intersection effect (Table 9c, row 8) was significant, which shows that the strength of the Regional effect must vary depending upon the season.

Given this significant Interaction effect, the simple main effects can be investigated in detail. Firstly, there is an overall significant difference between Regions in both wet and dry seasons (Table 9c, rows 6 and 7). Secondly, the overall seasonal effect

---

* It is interesting to compare the anovar method and the more normally used χ^2 method for showing significant departures from random distribution with respect to area. The χ^2 method compares the number of animals observed in each area against the number expected on the hypothesis of random distribution. In this study the χ^2 method would show that every Regional mean was significantly different from its grand mean. The χ^2 method, however, ignores any variation within an area, which the anovar method takes into account. The anovar method thus seems to be the more appropriate numerical method to use.

is significant only in Region III (Table 9c, row 4) but not in Regions I or II (Table 9c, rows 2 and 3).

These results show that elephants are not distributed randomly with respect to the three Regions in either the wet or the dry seasons. Although they always avoid Region I and select Region III,* the strength of this Regional selection is dependent upon the season. There is a significant decrease in elephant density in Region III from wet to dry season, although the increases in density in Regions I and II over the same period cannot be shown to be significant. It therefore appears that Region III acts as a wet season concentration area for elephant, and that there is a dispersal out of Region III during the dry season.

(VII) Discussion

(i) *Elephants*

These two counts show that Ruaha National Park carries year round a density of elephants of some $1{\cdot}7\ \text{km}^{-2}$. Marked Regional differences in density are apparent, in that elephants avoid Region I (mainly miombo woodlands) but select Region III (mainly the rift wall and rift valley area of the Park). This regional selection is present in both wet and dry seasons, and Region III seems to act as a wet season concentration area. There is a dispersion out of this Region III during the dry season with the elephants apparently moving equally to the other two Regions.

There is, as yet, insufficient information to estimate the potential for immigration from outside the Park should the elephant densities within the Park be reduced naturally or artificially. Movement across the northern boundary into and out of the Rungwa Game Reserve must undoubtedly occur, and survey flights should be extended in future well outside the boundary to study this. There is also insufficient information to decide whether or not there are unit populations of elephant within the Park. This will have to await further survey flights, and an analysis based on a finer grid square system than 10 × 10 km.

(ii) *General*

Some general aspects of the ecology of Ruaha have been revealed. The marked gradients of increasing densities, diversity and biomass away from the miombo woodlands and towards the eastern side of the Park, with high values often centred along the rift wall, are clarified in the Regional analysis. Region III, occupying most of the rift valley, the rift wall and parts of the plateau and the miombo ecotone, is characterized by high densities, diversity and biomass, while Region I, lying mainly in the miombo woodlands, has low values. The ecological factors underlying these Regional characteristics (e.g. rainfall, productivity, type and structure of the vegetation) are as yet unknown, but the analysis has undoubtedly differentiated three fundamentally distinct Regions within the Park.

(iii) *Statistical methods*

The statistical methods used here seem quite suitable for these types of data. Trend surface analysis is a generally useful technique for describing all kinds of distribution patterns, although it might appear that in this particular example some alternative model such as double Fourier analysis might have given better results. Principal components analysis is a straightforward method for handling multivariate data and for describing and identifying significant interrelationships within sets of variables, and the anovar methods seem particularly well suited for simple problems concerning distribution with respect to area.

The Regional analysis also has general applications. In this example it was possible to define and describe three Regions, that were obviously fundamentally different from one another, on the basis of this aerial survey data. Aerial surveys are now widely used in East Africa because ground access in many places is difficult or expensive. Ruaha National Park is a case in point, for the road system covers only a fraction of the Park, and cross country driving is impossible. Numerical methods such as those used here not only result in more information being extracted from the surveys (no bad thing in itself considering how expensive they are), but also enable a meaningful stratification of an area on which ground work, if required, may be efficiently planned.

---

* The industrious reader may calculate the appropriate standard errors, and may carry out the relevant t-tests for himself. However, reference to Table 9b should suffice.

Acknowledgments

The costs of both counts were met from funds made available for the Serengeti Ecological Monitoring Programme by the African Wildlife Leadership Foundation of Washington D.C. Aviation petrol was supplied by Tanzania National Parks who also loaned the use of an aircraft for the dry season count. Our thanks are due to the Director TNP for permission to carry out the counts in Ruaha, and to the SPW of Ruaha for his helpful co-operation on both occasions.

Many individuals helped with these counts. The pilots were the author, Dr H. F. Lamprey and Dr C. J. Pennycuick. The observers were Dr T. Mcharo, Dr A. R. E. Sinclair, Dr B. C. R. Bertram, H. Hoek and D. Kreulen, all of SRI, and Dr J. Ellis from the Pawnee Grassland Research Programme, who finally learned the meaning of large mammal biomass. Recorders were F. Kurji, J. Bunning, Dr L. Pennycuick and Mrs A. Kreulen, and Mrs R. Lamprey, Mrs A. Kreulen and Dr L. Pennycuick did the cooking. Dr L. Pennycuick developed the computer programme for the trend surface analysis.

My thanks to all these people are manifold.

References

Bailey, N.T.J. (1959) *Statistical Methods in Biology*. London: The English Universities Press.

Davis, J.C. (1973) *Statistics and Data Analysis in Geology*. New York: John Wiley & Sons.

Doornkamp, J.C. & King, A.M. (1971) *Numerical Analysis in Geomorphology*. London: Edward Arnold.

Gittins, R. (1968) Trend surface analysis of ecological data. *J. Ecol.* **56**, 845–869.

Grigg, D. (1967) Regions, models and classes. In: *Integrated Models in Geography*. (Ed. by Chorley and Haggett.) London: Methuen.

Harris, L.D. (1972) An ecological description of a semi-arid East African ecosystem. Colorado State University: *Range Science Department Science Series* No. 11.

Jolly, G.M. (1969a) Sampling methods for aerial censuses of wildlife populations. *E. Afr. agric. For. J.* **34**, 46–49.

Jolly, G.M. (1969b) The treatment of errors in aerial counts of wildlife populations. *E. Afr. agric. For. J.* **34**, 50-55.

King, L.J. (1969) *Statistical Analysis in Geography*. London: Prentice Hall.

Kirk, R.E. (1968) *Experimental Design: Procedures for the Behavioural Sciences*. Belmont, Calif.: Brooks/Cole.

LAMPREY, H.F. (1964) Estimation of the large mammal densities, biomass and energy exchange in the Tarangire Game Reserve and Masai Steppe in Tanganyika. *E. Afr. Wildl. J.* **2**, 1–46.

NORTON-GRIFFITHS, M. (1973) Counting the Serengeti migratory wildebeest using two-stage sampling. *E. Afr. Wildl. J.* **11**, 135–149.

NORTON-GRIFFITHS, M. (1974) Reducing counting bias in aerial censuses by using photography. *E. Afr. Wildl. J.* **12**, 245–248.

PENNYCUICK, C.J. (1973) The shadowmeter: a simple device for controlling an aircraft's height above the ground. *E. Afr. Wildl. J.* **11**, 109–112.

PENNYCUICK, C.J. & WESTERN, D. (1972) An investigation of some sources of bias in aerial transect sampling of large mammal populations. *E. Afr. Wildl. J.* **10**, 175–191.

SACHS, R. (1967) Live weights and body measurements of Serengeti game animals. *E. Afr. Wildl. J.* **5**, 24–36.

SINCLAIR, A.R.E. (1970) *Studies on the ecology of the East African buffalo.* D. Phil. Thesis: Oxford University.

SINCLAIR, A.R.E. (1972) Long term monitoring of mammal populations in the Serengeti: census of non-migratory ungulates, 1971. *E. Afr. Wildl. J.* **10**, 287–298.

SINCLAIR, A.R.E. (1973) Population increases of buffalo and wildebeest in the Serengeti. *E. Afr. Wildl. J.* **11**, 93–107.

SOKAL, R.R. & ROHLF, F.J. (1969) *Biometry.* San Francisco: W. H. Freeman.

SOKAL, R.R. & SNEATH, P.H.A. (1963) *Principles of Numerical Taxonomy.* San Francisco: W. H. Freeman.

(*Manuscript received* 13 *July* 1973)

E. Afr. Wildl. J., 1975, Volume 13, pages 141–143

Notes and Records

Weights and measures of lions*

Introduction

It is often important to know the weights of large mammals in the field. Studying food requirements, calculating energy budgets, or determining the dose rates of drugs administered all depend on knowing how heavy animals are. But it is not always considered possible or convenient to weigh them, even when they are immobilized and at the scientist's feet. Estimates of weight are made instead, and in the absence of calibration these may become extremely inaccurate. This note attempts to reduce such inaccuracies: first, by describing a convenient weighing system; second, by providing a rough calibration of weight against linear dimensions for lions; and third, by outlining a scale of visual estimates of stomach size for assessing the stomach weight of animals which eat rarely but hugely.

Weighing system

Weighing large animals does not necessarily require huge tripods, trees, spring balances and teams of assistants. I am grateful to Dr J. M. King for suggesting the use of bathroom scales for weighing immobilized lions. I carried six lengths of angle iron and four wooden planks 30 cm wide; all were 120 cm long, and so fitted conveniently into a small vehicle. These components could be bolted together in 4 min to produce a platform roughly 120 cm by 200 cm. This was placed close to the back of the immobilized lion, which was then rolled over onto it and pushed to the centre of the platform. A set of low flat bathroom scales was placed underneath each end. With the platform with lion then balanced on the two sets of scales, the reading of each scale was taken; their sum, minus the weight of the platform, gave the weight of the lion.

With this system, it was possible for me to weigh a lion of 200 kg alone and without assistance, and with a minimum of disturbance. A slightly larger platform with four sets of scales would enable one to weigh considerably heavier animals.

Weights and linear measurements

For lions immobilized or dead, I took measurements of length of head, body, tail, and limbs, and circumference of neck and chest. Some of these lions were also weighed, and I have tried to find linear measures, or derivatives of them, which would serve as reasonably reliable indicators of weight.

Limb measurements proved unsatisfactory since they were very unreliable except with completely relaxed animals; even then they were somewhat inaccurate and arbitrary. Body length or total length depended greatly on the posture of the animal; repeat measurements of the length of the same animal were not consistent. Head length

---

* S.R.I. Publication No. 151.