AN EVALUATION OF LANDSAT MSS DATA FOR ECOLOGICAL LAND CLASSIFICATION AND MAPPING IN THE NORTHERN CAPE

Andrew Alan Gubb

VOLUME 3

East London 1989

AN EVALUATION OF LANDSAT MSS DATA FOR ECOLOGICAL LAND CLASSIFICATION AND MAPPING IN THE NORTHERN CAPE

VOLUME 1

VOLUME 2

VOLUME 3

PAPER 3

CONTENTS

PAPER 3

VEGETATION CLASSIFICATION BY PRINCIPAL COMPONENTS ANALYSIS OF MULTITEMPORAL LANDSAT DATA

1. INTRODUCTION

During the period 1983 to 1986 the author was involved in the visual interpretation of Landsat Multispectral Scanner System (MSS) data as an aid to the classification and mapping of vegetation of the Northern Cape. In 1984 supervised, unsupervised and hybrid methods of computer classification of a single date Landsat scene were undertaken. It was found during visual interpretation technique investigations that each unit had to be investigated on the ground. This was found to be time-consuming and worthwhile for terrain/spectral units which were unique, but costly for units which were found to be similar in terms of classification and mapping units. The aim of computer classification investigation was to answer two questions: firstly, would computer classification result in quick and accurate extrapolations to surrounding, unsurveyed areas? Secondly, would computer classified terrain/spectral units be geographically and spatially more accurate than units demarcated using visual interpretation of 1:250 000 false colour composite images? These questions were not answered in the affirmative by the methods of computer classification used.

Computer classification occasionally resulted in confusion between physiognomically and structurally dissimlar field units which, on reflectance values, were classified as the same unit. Other minor (spatial) problems related to the confusion in classification between physiognomically similar but floristically dissimilar terrain units (more surprising than the former as it has been shown in Paper 1 that spectral unit differences correlate more accurately with physiognomic units than floristic units). A further aspect which came to light in computer classifications of single date Landsat MSS data was the high degree of heterogeneity of classified spectral units, i.e. no terrain unit was ever classified with a high degree of homogeneity of the pixels representing that terrain unit. Cross-checking these units showed that the terrain units were indeed relatively homogeneous (acceptably heterogeneous). The problem was ascribed to a "within" and "between" mixed pixel condition, where the satellite's resolution (57 X 79 pixel size) cell is larger than the intrinsic pattern of the vegetation, and, the fact that the computer is unable to register to spatial pattern (classifies one pixel at a time, see section 7 in Paper 2).

It was concluded after computer classifications of single date Landsat MSS data that if these problems could be solved, then the aims and objectives of the study would be effectively met. Several additional methods could be employed. This paper represents one of these methods, that of the use of multitemporal overlays. The objective of the multitemporal investigation was to improve upon the meaningful separability of spectral classes.

2. MULTITEMPORAL ASPECTS

Multitemporal overlays involve the precise superimposition of two or more sets of MSS data of the same area but of different dates. In the majority of applications, multitemporal overlays are used to detect change with time ("difference" images (Sabins 1978)), or less frequently, to gain additional information for the specific purpose of increasing classification accuracy, i.e. for enhancement reasons (Short 1982; Nelson & Hoffer 1979). Multitemporal overlays were investigated in this study for this latter purpose. Schreier *et al.* (1982) found contrast between vegetation units and surface material improved upon superimposition of temporal imagery.

In multitemporal analyses, spectral changes help to identify the land cover, providing discriminability of cover types not otherwise distinguishable (Swain 1983). It was hoped that the extra information content, in the form of the data from the second image as well as the differences between the two images, would improve classification accuracy. It is therefore desirable that the sets of MSS data to be superimposed should contrast as much as possible with each other and so maximise the potential of additional information. One method of attaining this goal is to use MSS data from different seasons. In the Northern Cape, there is maximum contrast in terrain features between winter and summer MSS data. A detailed discussion of this may be found in section 5.5 of Paper 1. Those features (terrain and atmosphere) which have changed from one date to the next are responsible for the improved contrast in multidate images. It was presumed that these changes would be additive. These facts are clearly reflected in the difference between winter and summer "photographic" false colour composite (FCC) images. It is always advisable to refer to FCC imagery in selecting MSS data for multitemporal analyses.

3. THE STUDY AREA

As stated above, several problems were encountered during the single date, single image computer classification investigation. The purpose of using multitemporal overlays was to negate the problems encountered during the single date computer classification investigation and thus comparison with the previous computer classified image was necessary for assessing accuracies and improvements. Therefore the position of one identical sub-scene, sub-scene no. 1 "Kalahari Thornveld", was used in this investigation (Fig. 4 and Plate 1 in Paper 1). Owing to the amount of data that the computer must handle in multitemporal analyses, the sub-scene was divided into four quarters and only the top left (North-west (NW) quadrat, Kalahari Thornveld) and bottom right (South-east (SE) quadrat, Ghaap Plateau) quarters were analysed.

4. SELECTION OF IMAGERY

In the use of multitemporal overlays, it was decided that only two sets of MSS data would be superimposed, one set representing data acquired during winter and the other during summer. Field reference data (over and above that data collected for the single date computer classification investigation) was collected during September 1984 and it was decided to select from MSS data acquired as close as possible to this period. The following MSS data sets, collected at about 09h30 over South Africa, were selected:

29/07/1984 (winter), ID 50150 - 07502, non-standard
173/78/79 - Landsat 5. Sun Azimuth = 45,55°
Sun Angle = 29,45°

16/04/1984 (summer), ID 40640 - 07474, non-standard
173/78/79 - Landsat 4. Sun Azimuth = 51,08°
Sun Angle = 36,18°

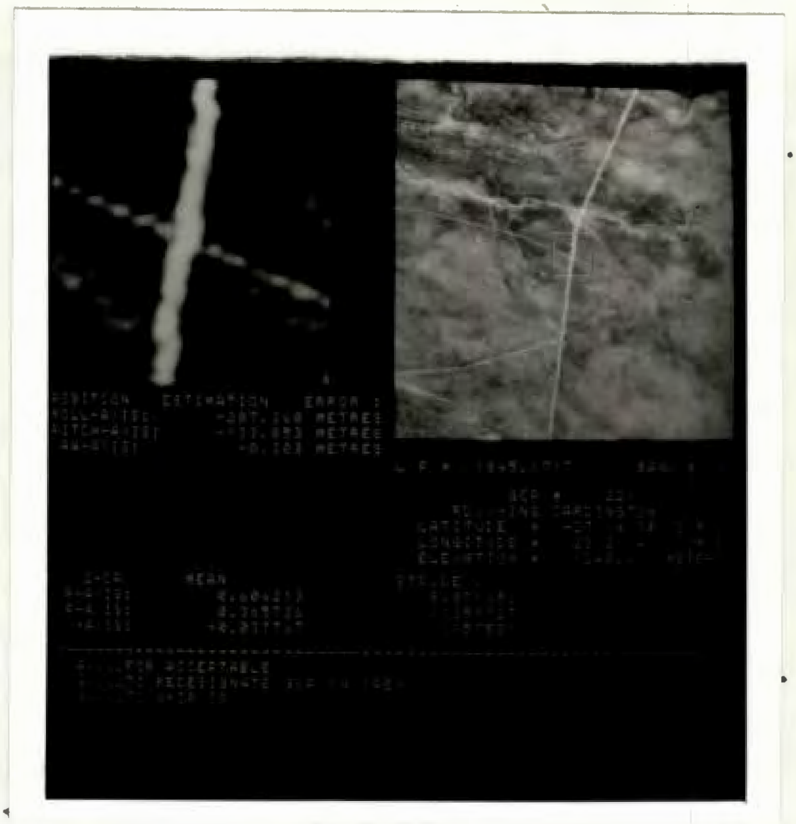
These two data sets were chosen for their good contrast after several computer compatible tapes were used to produce the image on screen for viewing and assessment.

5. PRECISION GEOMETRIC CORRECTION AND REGISTRATION OF THE IMAGES

Since two sets of MSS data were to be overlayed, the data sets require precision geometric correction (PGC) and then registration to each other. PGC was done on bands 4, 5, 6 and 7.

Precision geometric correction places the individual "reflectance values", i.e. pixels or picture elements, of each ground resolution unit in their correct latitudinal and longitudinal positions. PGC is done by means of selecting a series of ground control points (GCP's) within the area represented by the MSS data set. Graetz *et al.* (1986) referred to ground control points as "tie-points"; points which "tie" the two images together. GCP's are unique, sharp, clear and accurate surface features such as crossroads, aerodromes, etc. (Plate 1). The GCP's must also be "observable" in the MSS data, or observable on the image if hard copy prints are produced. GCP's should not be features which change with time, such as ploughed field boundaries which may change from year to year, or even with the ploughing seasons. In this study the winter data set was used for initial PGC The summer data set was therefore registered to it. The image to be registered is known as the "slave image" and the one it is registered to as the "master image" (Adeniyi 1985).

Initially GCP's are selected by scrutinising 1:250 000 topographical maps for likely points. Once a likely point is observed, its presence is checked in the data set or on the image. If the GCP is acceptable, then the exact latitude and longitude is read off the 1:50 000 topographical map and stored in the computer in the library file. In multitemporal analyses, not only is the GCP stored, but also a 32 X 32 pixel square with the actual GCP forming its central point. This 32 X 32 pixel square is known as a "chip". In this study, 24 GCP's were eventually chosen for the PGC process. It is important that GCP's be positioned and chosen so that all areas of the image are



<u>Plate 1</u> An example of a good ground control point (GCP) for use in geometric correction: calcrete-surfaced cross roads.

represented, otherwise areas well represented by GCP's will be more precisely geometrically corrected than other areas.

Once the master image has been precision geometrically corrected, then the slave image (MSS data set) is registered to it. In registration, one relates a point to a point while in PGC one relates a point to a latitude/longitude grid reference. The aim of registration is the perfect fit or superimposition of one data set on the other, i.e. each pixel or data "point" must fall on the same pixel or data "point" on which it is superimposed. The GCP "chips" are used again in the PGC. The "chips", and the same area of the data set to be registered, are compared with each other by the computer. If the total reflectance values of the 32 X 32 matrix are sufficiently similar, then the computer automatically correlates the two and registers them to each other. If the reflectance values are too dissimilar, then the two identical GCP's have to be registered manually by means of the cursor or track ball. In this study only 5 of the 24 GCP's used for PGC were automatically registered. The remaining 19 GCP's were manually registered using the following method. The data set on which the PGC was carried out (winter data set) was produced in image form on the comtal screen. A 2X zoom was applied to the area containing the G.C.P, the cursor was placed exactly on the GCP and a 7X zoom was applied. The identical area of the slave (summer) image was brought up and the two GCP's were superimposed.

After registration of the two images, the slave image is "flicked" on and off several times. If the image or part of the image gives the impression of shifting slightly with respect to the master image, then the two images are not exactly superimposed. More GCP's may be called for, scattered throughout the training area or in the area where registration is poor. A ground inaccuracy of 100 m between the two images is the acceptable upper limit, i.e. about two pixels in size. The two registered images are now 3240 X 3240 pixels in size instead of the usual 3240 X 2286 pixel matrix; the pixel size (resolution) is now 57 m X 57 m. It should be noted that the whole data set must be used for PGC and registration even though the study area may consist only of a specific section of that data set.

The two types of geometric corrections, PGC and registration may be summarized as follows:

a) Image to grid registration, i.e. image to the relevant 1:50 000 and 1:250 000 topographical maps using GCP's.

b) Image to image registration, i.e. the slave image to geometrically corrected master image.

The accuracy of the image to image registration is thus dependent on the accuracy of the image to grid registration (Adeniyi 1985).

6. PRINCIPAL COMPONENTS ANALYSIS

In multitemporal overlay analyses, e.g. winter and summer MSS data sets, four Landsat bands in each data set result in eight-dimensional data. Thus the amount of information to be synthesized by the computer is considerably increased. Furthermore, since the correlation between adjacent Landsat bands of single date MSS data, and correlation between the same bands on multidate MSS data is high, much of this information is redundant. A mathematical transformation of the data, such as Principal Components Analysis (P.C.A.) is required, so that most of the information (seen in terms of the spread of the data) may be extracted and the dimensionality of the data reduced without serious loss of information (Lasserre *et al.* 1983; Browne 1983; Pendock & Sears 1983; Lo *et al.* 1986; Turner 1988). Swain (1983) questioned the use of P.C.A. in cases where classification is the goal and class separability is more important than feature variability *per se*. The fundamental assumption underlying the P.C.A. dimensionality reduction approach is that variance (or covariance) is the most important information-bearing characteristic of the data.

The digital number (DN) data for any spectral feature are usually plotted on multidimensional axes that represent the potential range of intensities (brightness) recorded for each band of a multispectral scanner. These axes are quantified by the full range of values for each spectral band. The data normally scatter about their means in a broad elliptical distribution in the two-dimensional case. Fig. 1 portrays such a case when band 4 is plotted against band 6 in a typical scatter plot diagram. Here the spread, or variance, of the data may be loosely regarded as the quantity of information along a line (a) passing through the elliptical scatter plot. The initial

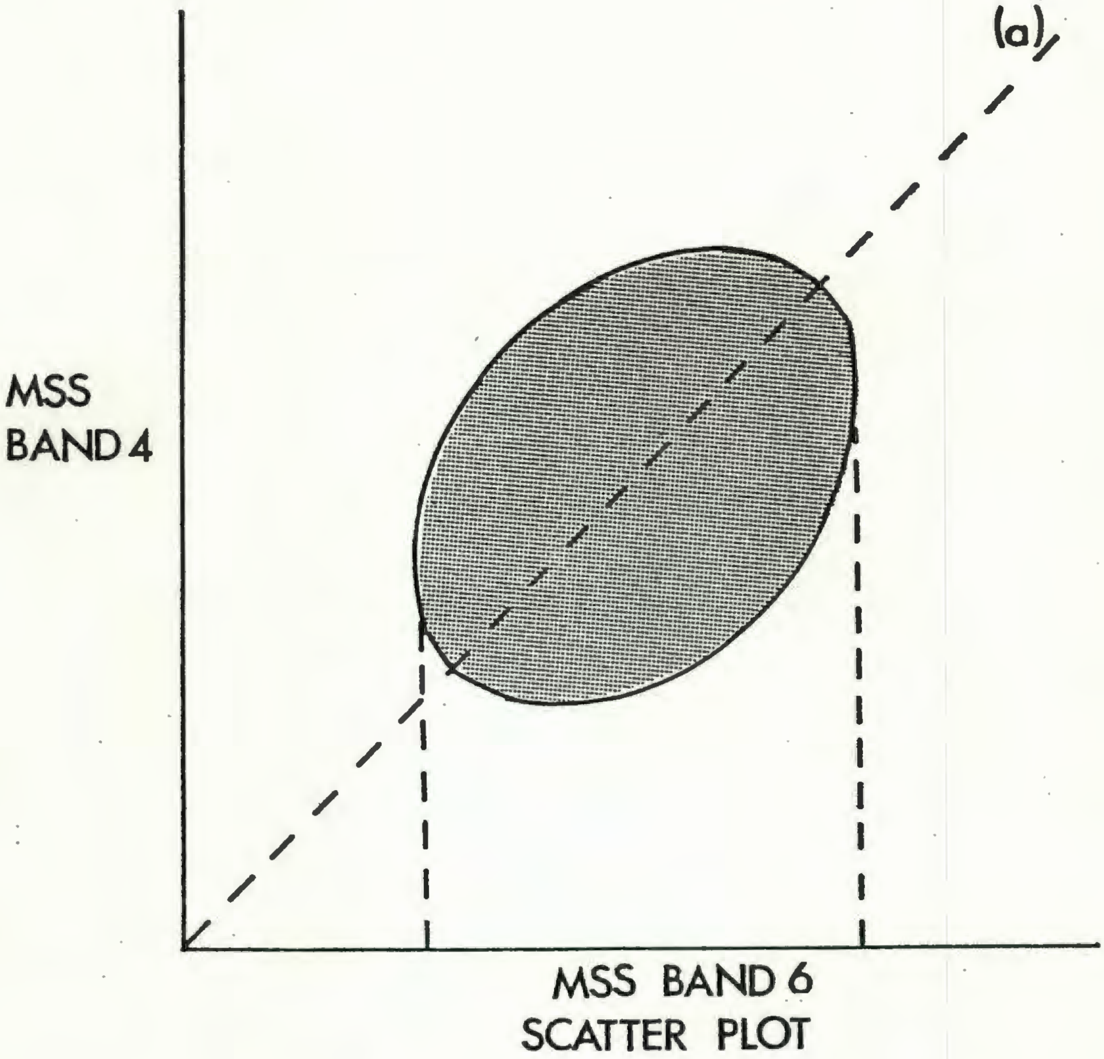


Figure 1 Scatter plot diagram of band 4 against band 6. The digital numbers (DN) data scatter about their means in a broad elliptical distribution in this two-dimensional case. The quantity of information is shown as the spread (variance) of the data along the line (a).

measurement axes are not necessarily the best arrangement in multivariate space for expressing data variations and hence optimising information content This can be understood by the distance Z_1 to Z_2, whereas the distance along the axis (a) of the spread of the information is greater. The strategy is to find a mathematical transformation that redistributes all DN values with respect to a new set of axes.

The principal component method makes use of the fact that for a single Landsat image or data set there is a great deal of redundancy (e.g. between MSS bands 6 and 7); in addition, in multitemporal overlays, there may be redundancy between the same bands on different dates. One objective of P.C.A. is to eliminate redundancy. P.C.A. is thus concerned with the structure of the variability and interdependence between random variables (Steffens 1983). Fig. 2 is a simple geometric illustration of P.C.A. as applied to the two-dimensional case. Again a scatter plot diagram is produced from MSS band 4 data plotted against band 6 data. The first set of axes is X_1 and X_2 and the points, producing the elliptical pattern, are positioned according to their spectral reflectance values. If a second set of axes, Y_1 and Y_2, is placed in the same direction as the first set of axes, but with the origin moved to the centroid (the point located at the mean of each band), then the quantity of information (the spread or variance of the data along the two axes) in the data set is better expressed (Gauch 1982). If a third set of axes, P_1 and P_2, is created by means of a rigid rotation around the centroid so that P goes through the major extension of the data points, i.e. accounts for the greater variability, then the information content is fully optimized. P_1 is perpendicular to P_2 and accounts for the remaining maximal variance, though in this two-dimensional case the placement of P_2 is trivial. In the multidimensional case this is not so.

P.C.A. may be algebraically illustrated as well. The scatter diagram in Fig. 2 shows how the two sets of reflectance measurements are correlated. The numerical description of the correlation represented in this scatter plot comes from the covariance matrix. The variance of a single variable measures the spread of the set of data about its mean while the covariance of two variables (two sets of data) measures their joint variation about the respective means. In this two-dimensional illustration, the covariance

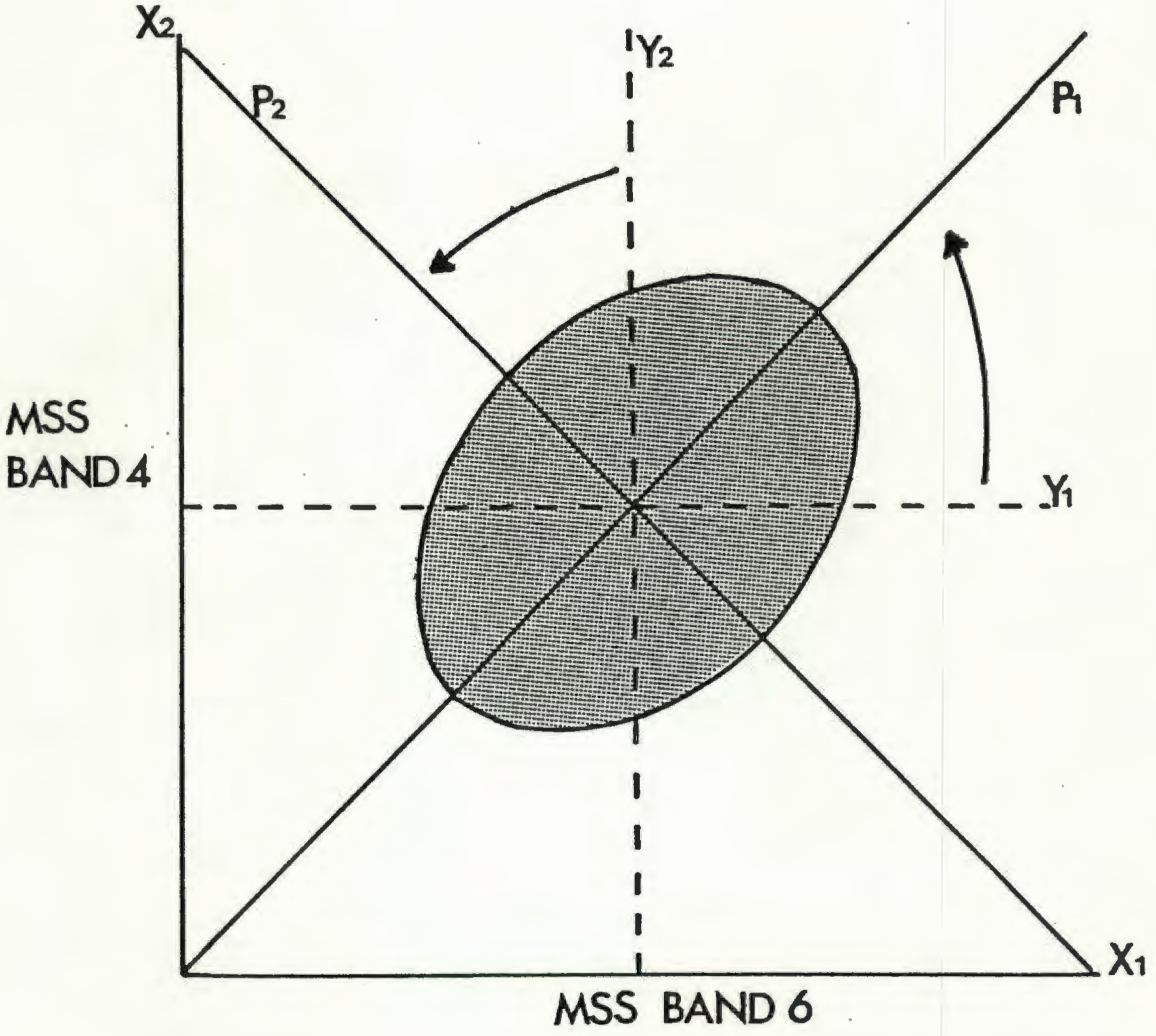


Figure 2 A simple geometric illustration of principal components analysis as applied to the two-dimensional case.

matrix of two variables is written and calculated as follows:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

where S_{11} = variance of variable X_1
S_{22} = variance of variable X_2
and $S_{12} = S_{21}$ is their covariance

This would be the case for a 2 X 2 matrix. In the multidimensional case of P variables, S would be a P X P matrix.

Once the covariance matrix has been calculated then a procedure called eigenanalysis is performed on the matrix. This analysis yields eigenvectors and eigenvalues. The lines onto which the original measurements are projected (P_1 and P_2 in the two-dimensional example given above) are called the eigenvectors or principal axes. The eigenvalues, which represent the variance along the principal axes, are always maximized (Sokal & Rohlf 1978). The original measurements (X_1, X_2) are projected onto the eigenvectors to obtain Y_1, Y_2, which comprise the principal components, e.g.
$Y_1 = a_1X_1 + a_2X_1$ and $Y_2 = a_1X_2 + a_2X_2$
where the a's are constants termed elements of the eigenvectors.

The line along which the projections of the original measurement have a maximum spread is known as the first principal component or eigenvector. The axis at right angles to this is known as the second principal component or eigenvector (Clifford & Stephenson 1978). The principal components are linear combinations of the original measurements. They are ranked according to their eigenvalues. The result of P.C.A. ordination is a sequence of axes of diminishing importance (Gauch 1982).

Each principal component may contain information from all classes. In the case where there is a high correlation between the original observations, most of the variation is accounted for in the first few components and little information is lost by ignoring the remaining components. If, say, 90% of the variation is found in the first two components, then only 10% of the variation is lost by ignoring the remaining components; thus a useful parameter is the variance of each component expressed as a percentage of the total variance.

Principal components, then, are new variables obtained from the original observed variables. All principal components are uncorrelated with each other (Browne 1983). The method of P.C.A. is useful only when all variables are measured on the same scale.

In the case of Landsat MSS data, the approximate correlations between bands 4, 5, 6 and 7 would be as follows: -

CORRELATION MATRIX (from Piper 1981)

	4	5	6	7
4	1,0			
5	0,7977	1,0		
6	0,6891	0,8913	1,0	
7	0,6503	0,8382	0,9681	1,0

Thus adjacent bands are highly correlated, e.g. band 6 with band 7 = 0,9681 (highly significant with a sample size of 7500 pixels), while non-adjacent bands have a lower correlation, e.g. band 6 with band 4 = 0,6891. In the above example the percentage variance of each component was found to be:

PC1	88,92%
PC2	7,69%
PC3	2,63%
PC4	0,78%

Thus most of the useful data from the four bands is contributing to the first component, and a lot less, in descending order, to components 2, 3 and 4. A black and white image can be produced for each principal component and forms a useful image for visual interpretation with its excellent impression of ruggedness (Newton & Boyle 1988). In general the first component image most nearly resembles a standard image. For heavily vegetated terrain, the first component image provides a sensitive vegetation indicator. Newton (1983) reported that the first component illustrated the topography of the area well. Byrne *et al.* (1980) found that first and second order component images represent unchanged landcover, an important result in change-detection studies.

7. RESULTS OF PRINCIPAL COMPONENT TRANSFORMATIONS OF THE NORTH-WEST QUADRAT (KALAHARI THORNVELD) AND SOUTH-EAST QUADRAT (GHAAP PLATEAU), WINTER AND SUMMER

A principal component transformation, the Karhunen-Loeve transformation, was executed on the winter and summer data of the NW and SE quadrats. The results are given in Table 2.

TABLE 2

RESULTS OF THE PRINCIPAL COMPONENT ANALYSIS

NORTH-WEST QUADRAT (KALAHARI THORNVELD)

	PRINCIPAL COMPONENTS							
	WINTER				SUMMER			
EIGENVECTOR	1	2	3	4	1	2	3	4
MSS 4	0,35	0,58	0,53	0,51	0,35	0,64	0,49	0,47
5	-0,78	0,88	0,67	-0,73	-0,76	0,51	0,20	-0,34
6	0,73	0,33	-0,43	-0,43	0,29	0,57	-0,73	-0,24
7	0,58	-0,74	0,30	0,12	0,46	-0,79	0,43	-0,78
EIGENVALUE	289,8	41,24	20,32	12,27	353,3	26,70	24,30	16,20
% of total variance	79,79	11,34	5,59	3,37	84,02	6,34	5,78	3.86
cumulative % of total variance	79,79	91,04	96,63	100,0	84,02	90,35	96,14	100,0

SOUTH-EAST QUADRAT (GHAAP PLATEAU)

	PRINCIPAL COMPONENTS							
	WINTER				SUMMER			
EIGENVECTOR	1	2	3	4	1	2	3	4
MSS 4	0,35	0,55	0,56	0,52	0,31	0,61	0,53	0,48
5	-0,86	0,62	0,68	-0,73	-0,12	-0,58	0,80	-0,95
6	0,45	0,62	-0,47	-0,44	-0,18	0,50	0,25	-0,95
7	0,82	-0,56	0,95	-0,56	-0,93	-0,18	-0,31	-0,33
EIGENVALUE	249,6	42,35	19,67	7,60	296,2	31,70	24,46	18,49
% of total variance	78,19	13,27	6,16	2,38	79,86	8,54	6,59	5,00
cumulative % of total variance	78,19	91,50	97,62	100,0	79,87	88,41	95,00	100,0

For both principal component transformations, the percentage variance of components one and two accounts for about 90% of the total variance and from 78% to 84% of the total variance is accounted for by the first component.

The images of the four principal components were contrast stretched to portray as much detail as possible. Images of the principal components of NW quadrat (Kalahari Thornveld), winter and summer, and SE quadrat (Ghaap Plateau), winter and summer, were produced, i.e. 16 images. Plates 2 to 5 show the four images produced from the principal components data of the NW quadrat.

PC1 has outstanding clarity and ruggedness. A simple "eye" delineation of vegetation units (using the vegetation map as a reference) indicates that PC2 and then PC1 of winter, and PC3 and then PC4 of summer correspond better with vegetation units. In using P.C.A. as an image-enhancement technique for studies on vegetation cover in desert conditions, Hielkema (1979) found that virtually only two classes were represented by PC3; the vegetation and all classes of the abiotic environment in one single class. The topography is clearest on PC1. In all probability PC1 and PC2 correspond largely with the soil/geology complex of the training areas, as vegetation projected canopy cover (PCC) accounts for approximately 35% cover of the terrain, the rest mostly being exposed soil and to a minor extent, geological formations. PC4 of winter contains mostly atmospheric "noise".

8. COMPUTER CLASSIFICATION OF THE PRINCIPAL COMPONENT DATA

The colour display system can deal with only four spectral channels at one time. In both the NW and SE quadrats, the data sets from the first and second components of summer and winter were used for the final computer classification. This solitary FCC image produced from this data set bears strong resemblance to a standard 3-band false colour composite, but shows many more subtle differences that discriminate real variations among features more effectively, as well as defining new feature patterns on the surface (Plates 6 and 7).

Using the Satellite Applications Centre's (SAC) Image Analysis System (IAS), an unsupervised computer classification was performed on each of these two principal component data sets (NW quadrat - Kalahari Thornveld and SE quadrat - Ghaap Plateau).

<u>Plate 2</u> PC1, NW quadrat.

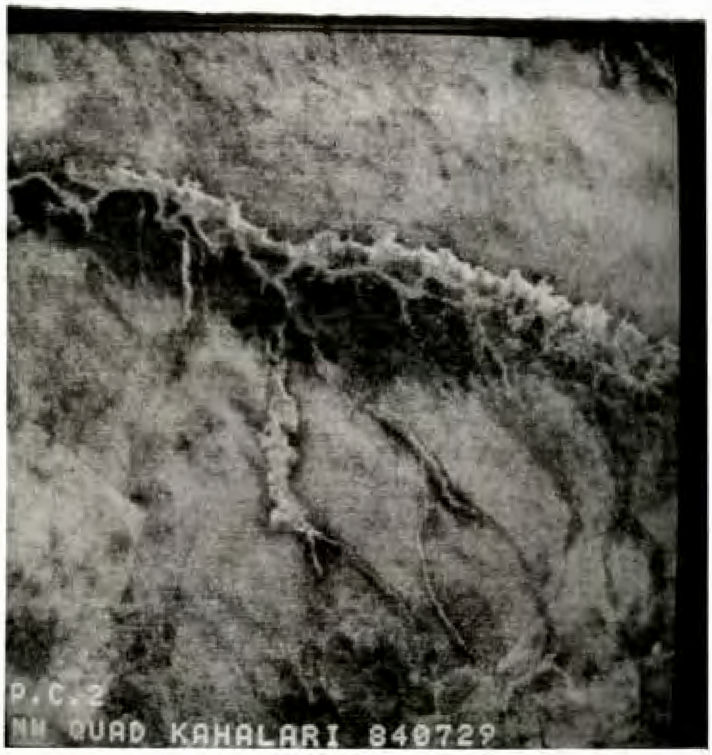


Plate 3 PC2, NW quadrat.

Plate 4 PC3, NW quadrat.

Plate 5 PC4, NW quadrat.

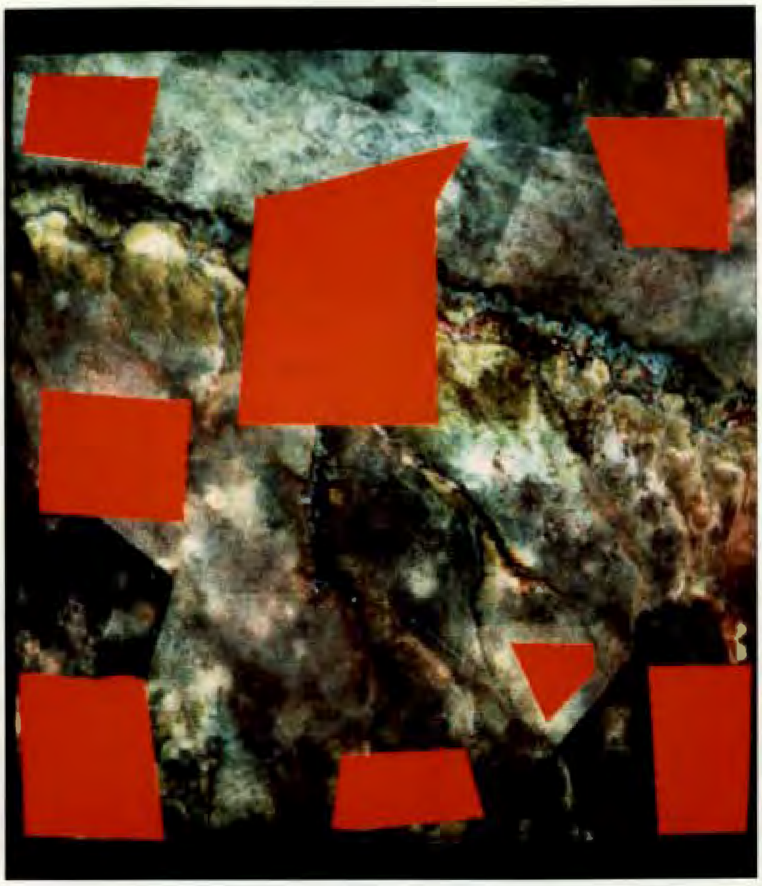


Plate 6 False colour composite image derived from principal components of NW quadrat. The positive masks indicate areas used for unsupervised classification.

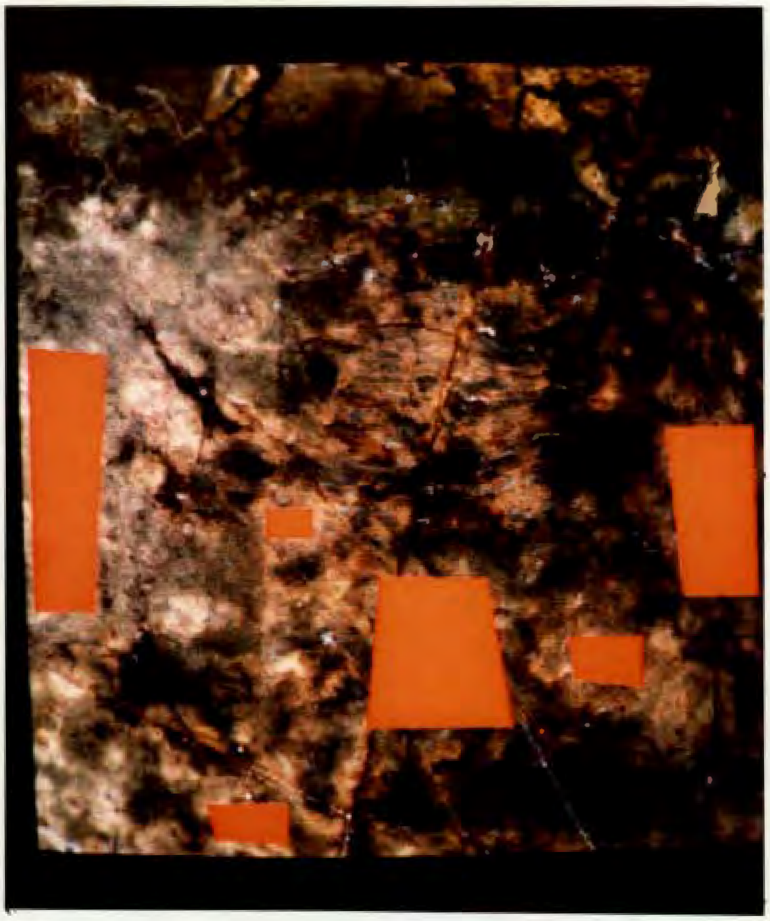


Plate 7 False colour composite image derived from principal components of SE quadrat. The positive masks indicate areas used for supervised classification.

8.1 NORTH-WEST QUADRAT (KALAHARI THORNVELD)

The greater part of this training area consists of eight major vegetation units. Eight areas representative of the vegetation units in the quadrat were positively masked and an unsupervised classification was performed on the eight masks. Initially sixteen classes were generated from this procedure but two classes were combined to make fifteen classes. The entire quadrat was then classified.

8.2 SOUTH-EAST QUADRAT (GHAAP PLATEAU)

Once the northern 20% of the SE quadrat has been masked, this training area consists of six major vegetation units. Six areas representative of the vegetation units in the quadrat were masked and unsupervised classification was performed on the six masks. Originally eight classes were produced from the unsupervised classification. Class eight was split to produce a total of nine classes.

The negative mask was applied to exclude the vegetation units belonging to another ecological region; a group of vegetation types typically on elevated and rugged sites and forming a natural "divide" between the Kalahari Thornveld vegetation units in the north and the Ghaap Plateau shrublands lying to the south. The results of the unsupervised classifications of both quadrats are given in Table 3.

9. RESULTS AND DISCUSSION OF THE COMPUTER CLASSIFICATION

9.1 NORTH-WEST QUADRAT (KALAHARI THORNVELD)

Plate 8 shows the final colour-coded, computer-generated classification. (Unfortunately, this plate does not show the blue of class 15. This would fill the "black" patch in the extreme south-west corner and a portion of the "black" patches in the north-west section of the quadrat.) The colour-coding lacks general cohesion but there is good pattern in colour-code combinations. This will be discussed in detail in section 11 below. The area has a measure of topographical complexity. It is extremely difficult to assign meaningful vegetal descriptions to the fifteen classes generated during the unsupervised classification. The vegetation contains weak gradients within and between units, but units are clearly distinguishable, more in terms of species composition than structure. Pink to red Kalahari sands occur almost throughout. The field vegetation units represented by the various colour-codes are described below. A description of the vegetation-landscape units (V-L) is given in Table 1.

TABLE 3

RESULTS OF THE UNSUPERVISED CLUSTER ANALYSIS

NORTH-WEST QUADRAT (KALAHARI)			SOUTH-EAST QUADRAT (GHAAP PLATEAU)		
CLASS	NO. OF PIXELS	%	CLASS	NO. OF PIXELS	%
unclas.	914	0,35	unclas.	683	0,34
1	742	0,28	1	13609	6,74
2	2609	1,00	2	23258	11,52
3	5425	2,07	3	31282	15,49
4	12087	4,61	4	37699	18,67
5	16664	6,36	5	39224	19,42
6	25962	9,90	6	29181	14,45
7	41813	6,68	7	15237	7,55
8	48602	18,54	8	7970	3,95
9	31207	11,90	9	3797	1,88
10	25838	9,86			
11	17502	6,68			
12	8645	3,30			
13	3420	1,30			
14	14897	5,68			
15	5817	2,22			

1 (dark grey):	a reasonably clearly defined area of V-L 5; some of V-L 18; small patches of V-L 8. Both V-L's are on high ground and both have a thin layer of red, base sands.
2 (light grey):	parts of V-L's 5 and 18; most of V-L 8 (see 1 above).
3 (yellow):	parts of V-L's 1 and 18; a cohesive part of V-L 5. There is some similarity in the species composition and structure of V-L's 5 and 1.
4 (puce):	parts of V-L's 1, 2 and 8. V-L's 2 and 8 are adjacent to each other; both are low woodland formations on pink to pale-pink sands.

5 (pink): mostly in parts of V-L's 1 and 18, but scattered in V-L's 5, 2 and 8.

6 (light brown): lack of cohesion and pattern; widely scattered throughout several V-L's.

7 (light blue): scattered throughout, but mostly in parts of V-L 1 and a cohesive patch in V-L 3. Both V-L's are closed woodlands with the same co-dominant species.

8 (green): scattered in V-L 1; a little on outer edges of V-L 3.

9 (red): mostly in V-L's 9 and 2, but lacking in cohesion; small patches in V-L 1 and outer sections of V-L 3. V-L's 9, 2 and 3 lie adjacent to each other; all are low-lying areas and are low woodland formations.

10 (dark blue): major, cohesive portion of V-L 4; small patches in V-L's 1 and 3.

11 (cyan): mostly in V-L's 9, 2, 3 and 1; "artificially structured" V-L 4 (see discussion in section 10)(see 9 above).

12 (dark orange): small, cohesive patches scattered throughout; a fairly extensive section of V-L 4.

13 (white): small, cohesive patches in V-L's 2, 3 and "artificially structured" 4; a fairly large, cohesive patch in SW corner of V-L 1 adjoining V-L 5.

14 (purple): cohesive patches of V-L 4

15 (blue): small cohesive patches in V-L's 4 and 5. Both are short woodland formation types on red, base sands.

Separation of classes is indicated by the divergence matrix in Table 4. A number of adjacent classes had a divergence of less than 3, viz. 2 and 3, 3 and 4, 4 and 5, 5 and 6, 6 and 7, 7 and 8, and 8 and 9, indicating possible overlap. There are a number of non-adjacent classes with low divergence measures. The class centres (means) of PC1 winter and PC2 winter, and PC1 summer and PC2 summer were plotted on two dimensional diagrams known as bispectral plots (Figs. 3 and 4). The spheroids have been drawn using one standard deviation of the mean and thus encompass 68% of the digital values for each class. This cluster diagram method of analysis indicates which classes cluster together and which are separate and is a method that was used by May (1986). A number of adjacent classes with a divergence measure of less than 3 are shown to be reasonably separate in both the winter and summer diagrams. There is less overlap in the case of summer principal components.

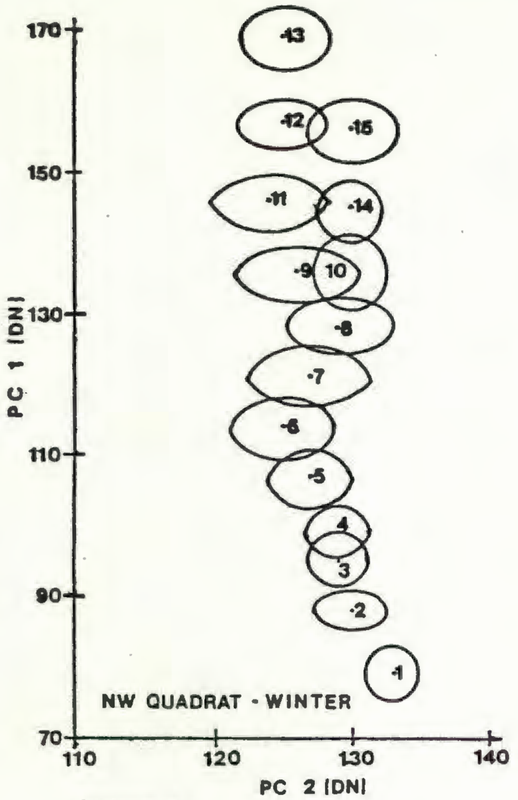


FIGURE 3

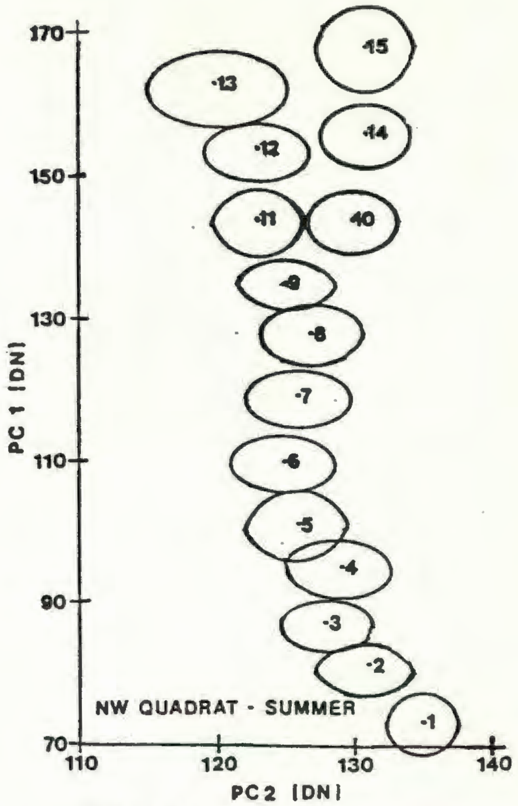


FIGURE 4

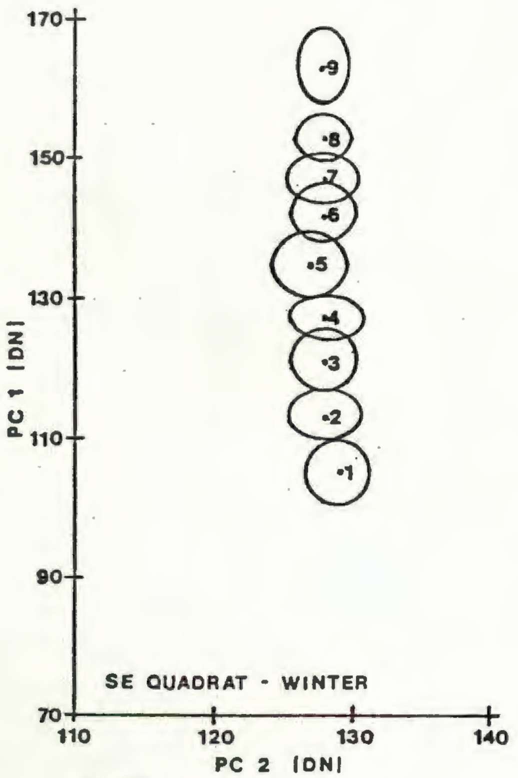


FIGURE 5

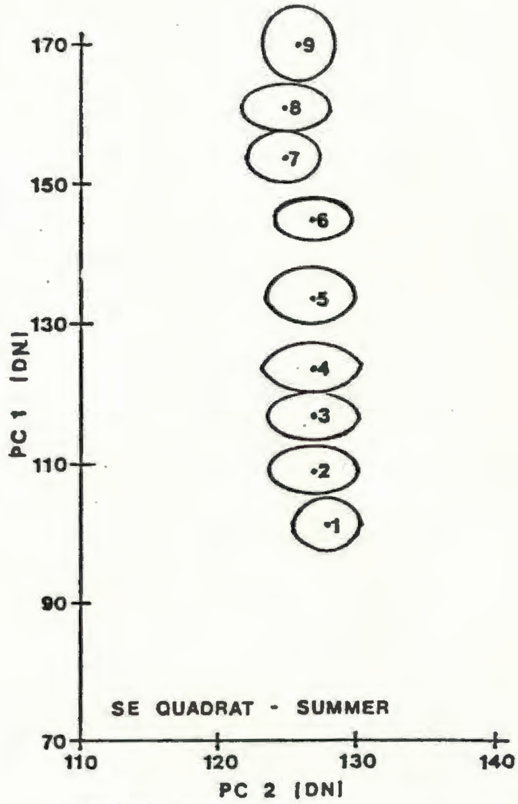


FIGURE 6

Figures 3 to 6 Mean digital number and standard deviation for each class derived from principal components 1 and 2, summer and winter, for both quadrats.

Fig. 3 shows considerable overlap of classes 3 and 4, and class 10 overlaps with 14, 9 and 8; non-adjacent classes 12 and 15 also overlap, but are well-separated in Fig. 4. Had the spheroids been drawn using two standard deviations, so including 95% of the digital values for each class, then the overlap between classes would have been considerably more.

TABLE 4

DIVERGENCE MEASURE BETWEEN CLASSES

NORTH-WEST QUADRAT (KALAHARI)

CLASS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	CLASS
1	0	3,62	6,27	7,74	10,56	13,39	15,76	18,68	22,08	22,19	24,79	28,34	28,24	26,17	27,24	1
2		0	2,66	4,66	7,53	10,47	13,17	16,24	19,66	20,08	22,60	26,38	26,66	24,24	25,68	2
3			0	2,57	5,21	8,39	11,32	14,41	17,81	18,80	20,82	24,66	25,06	23,03	24,52	3
4				0	2,69	5,33	7,88	10,70	13,72	14,53	16,58	20,09	20,83	18,34	20,00	4
5					0	2,78	5,54	8,36	11,28	12,44	14,25	17,84	18,89	16,21	18,09	5
6						0	2,97	5,83	8,63	10,28	11,70	15,32	16,63	14,06	16,12	6
7							0	2,84	5,58	7,29	8,75	12,31	13,97	11,00	13,30	7
8								0	2,80	4,50	6,15	9,77	11,88	8,15	10,80	8
9									0	3,13	3,54	7,31	9,78	6,24	8,79	9
10										0	3,64	6,16	8,60	3,64	8,98	10
11											0	3,69	6,47	4,02	6,15	11
12												0	3,26	3,95	3,85	12
13													0	6,20	4,10	13
14														0	3,48	14
15															0	15

9.2 SOUTH-EAST QUADRAT (GHAAP PLATEAU)

Plate 9 shows the final colour-coded, computer-generated classification. (Unfortunately the grey of signature 1 and the brown of signature 2 do not appear on the classification. The large cohesive "black" patches would be grey, and the scattered small patches, brown. This obviously excludes the masked area in the north.) Once again there is a general lack of cohesion though to a lesser extent than in the NW quadrat. The vegetation in this quadrat displays strong mosaicism and the units are closely related but are relatively easy to distinguish from one another because of physiognomic dissimilarities. Soil/geological units are also easily distinguishable. Once again, it is difficult to assign meaningful vegetation-landscape units to the classes generated in the computer classification. In this case, though, the classes do contain fewer vegetation-landscape units as is evident from the descriptions given below:

1 (grey): cohesive patches in parts of V-L's 33 and 16b; small patches in the southern extensions of V-L 36a; V-L's 33 and 16b are both high shrublands on brown, base soils and have similar dominant and co-dominant species.

2 (salmon): a large, cohesive patch in V-L 33; almost all of V-L 16b. The pink possibly represents areas of dense to very dense formation classes.

3 (brown): very small scattered patches in V-L's 33 and 36a.

4 (white): large, fairly cohesive patches in V-L's 33, 36a and 17a; some in 17d. Possibly represents intershrub grass patches.

5 (pink): cohesive in V-L's 36a and 17a; a fairly large amount scattered in V-L 17d (see 4 above).

6 (yellow): large, cohesive amounts in V-L's 4 and 17d (western regions); some small, scattered patches in V-L 17a. Possibly the effect of Kalahari element, importance of grass layer and exposed soils.

7 (red): moderately cohesive in V-L 17d and to a lesser extent in V-L 4; some scattered in V-L 17a (see 6 above).

8 (blue): cohesive patches in V-L's 4 and 17d (see 6 above).

9 (green): cohesive patches in V-L's 4 and 17d (see 6 above).

TABLE 5

DIVERGENCE MEASURES BETWEEN CLASSES

SOUTH-EAST QUADRAT (GHAAP PLATEAU)

CLASS	1	2	3	4	5	6	7	8	9
1	0	3,07	5,79	8,12	10,93	14,45	18,03	19,83	18,37
2		0	2,86	5,52	8,62	12,36	16,05	18,11	17,20
3			0	2,85	6,11	9,88	13,53	15,75	15,42
4				0	3,21	6,77	10,16	12,40	12,64
5					0	3,40	6,51	8,81	9,74
6						0	3,07	5,60	7,31
7							0	2,82	5,36
8								0	3,09
9									0

The descriptions of the V-L's may be found in Table 1. To a large extent, V-L's are represented by combinations of colours rather than single colours. Separation of classes is indicated by the divergence matrix in Table 5, and the bispectral plots of PC1 winter against PC2 winter and PC1 summer against PC2 summer (Figs. 5 and 6).

The following pairs of adjacent classes had divergence measures of less than 3: 2 and 3, 3 and 4, and 7 and 8. This indicates better class separability than in the NW quadrat, but possibly still some confusion between classes with low divergence measures. The bispectral plots in Figs. 5 and 6 definitely indicate less overlap between classes than was the case with the NW quadrat. There is no class overlap at all in the case of PC1 summer against PC2 summer (Fig. 6). In the winter diagram it is evident that there is some overlap between adjacent classes, but none between non-adjacent classes. Once again, if two standard deviations had been used to draw the spheroids, the overlap would have been greater. Class separability is better in this quadrat than in the NW quadrat.

10. DISCUSSION

10.1 CLASSIFICATION OF THE VEGETATION OF THE TRAINING AREA

Information on the vegetation units in the training area for this investigation was extracted from independent studies on the classification and mapping of the vegetation of the Northern Cape. For the purposes of ascertaining the usefulness of Landsat data in the attainment of this objective, a landscape/pseudo-Braun-Blanquet method was used on selectively extracted vegetation and abiotic environmental data (Paper 2). Owing to the need for much less detail for the interpretation of Landsat surrogate spectral classes and their correlation with recorded ground reference data, the complete range of survey data was not required, e.g. total floristic composition. The vegetation map depicted in Fig. 5 in Paper 2 is the result of such a landscape/pseudo-Braun-Blanquet interperetation and visual interpretation of Landsat FCC imagery. The original maps are at a scale of 1:250 000. The numbering system follows that used in the original classification interpretation and Table 1 contains relevant, descriptive information extracted for each of the vegetation units.

The international syntaxonomic nomenclature customarily used for hierarchical classifications of vegetation is not used here. The classification terminology is kept as simple, though meaningful, as possible. The units are interpreted and discussed in relation to computer classification comparisons of various methods used on Landsat MSS digital data (which relates primarily to landscape units and not solely to floristically derived Braun-Blanquet vegetation units). As mentioned in both Papers 1 and 2, Landsat MSS digital data relates primarily to the soil-geological forms and the physiognomy/structure of the vegetation. The latter is most often controlled by the former in the semi-arid Northern Cape. For this reason the classes derived from the computer classifications must be compared with a vegetation map which is based on landscape units. Vegetation-landscape units are based on vegetation structure and physiognomy, dominant and co-dominant species (floristics), soil types, geological formations and terrain morphological types. The relevance of landscape or "natural" units to phytosociological vegetation units is discussed in detail in Paper 5.

Reference to the vegetation-landscape map of the training area (Fig. 5 in Paper 2) and the simplified classification and tabulation of relevant environmental data (Table 1) emphasizes the diversity in the training area. The same ground reference data set was used for both the production of the vegetation-landscape map and multitemporal computer classification interpretations. The multitemporal data, besides representing two different seasons, was recorded by two different satellites, viz. Landsat 4 and Landsat 5. In the discussion that follows the multitemporal overlay computer classification will be compared to the vegetation-landscape map.

10.2 ECOLOGICAL REGIONS OF THE TEST AREA

From Table 1 it can be seen that four ecological regions are involved, viz. the Kalahari Thornveld, Kalahari-Ghaap Plateau Watershed, Kuruman Sourveld and Ghaap Plateau Ecological Regions. The Kalahari Thornveld Ecological Region is uniformly covered by deep, red to pink, aeolian Kalahari sands and is generally dominated by *Acacia* woodlands, with a strong graminoid layer. This ecological region lies mainly in the north but some impure subtypes have penetrated southwards. The Kalahari-Ghaap Plateau Watershed Ecological Region, though not extensive, is of importance because annual rainfall north of it (Kalahari Thornveld Ecological Region) is about 80% of that to the south (Kuruman Sourveld and Ghaap Plateau Ecological Regions).

TABLE 1 : DESCRIPTIONS OF THE VEGETATION - LANDSCAPE UNITS PRESENT IN THE S E AND N W QUADRATS OF THE KALAHARI THORNVELD TRAINING AREA

UNIT NO.	ECOLOGICAL REGION	VEGETATION TYPE	FORMATION CLASS	SOIL/GEOLOGY	TOPOGRAPHY	DOMINANT SPECIES
1	Kalahari	Kalahari Thornveld	short, closed woodlands	pink to yellow, deep, base sands	high-lying, gentle, extensive slopes	*Terminalea sericea* *Acacia erioloba* *Boscia albitrunca* *Dichrostachys cinerea*
2	Kalahari	Kalahari Thornveld	low, open woodlands	pale pink, deep, base sands	lower slopes and low undulating dunes	*Acacia haematoxylon* *Acacia erioloba*
3	Kalahari	Kalahari Thornveld	low, closed woodlands	pale pink to yellow, deep, base sands	low-lying, gentle, extensive slopes	*Acacia haematoxylon* *Terminalea sericea* *Acacia erioloba*
4	Kalahari	Kalahari Thornveld	short, open to sparse woodlands	deep, red, base sands; quasi-alluvium (young)	plains and gentle, extensive slopes - partial alluvium	*Acacia erioloba* *Acacia mellifera* *Tarchonanthus camphoratus* *Grewia flava* *Lycium hirsutum*
5	Kalahari	Kalahari Thornveld	short, open to closed woodlands	red, base sands overlying quartzite	high, rounded hills	*Acacia erioloba* *Acacia mellifera* *Ziziphus mucronata* *Grewia flava* *Dichrostachys cinerea* *Terminalea sericea*
6	Kalahari	Kalahari Thornveld	short, closed woodlands/ thickets	red to pink, base sands	plains	*Acacia erioloba* *Tarchonanthus camphoratus* *Grewia flava* *Lycium hirsutum*

UNIT NO.	ECOLOGICAL REGION	VEGETATION TYPE	FORMATION CLASS	SOIL/GEOLOGY	TOPOGRAPHY	DOMINANT SPECIES
8	Kalahari	Kalahari Thornveld	low, sparse, grassy woodlands	pink, calcareous sands and soils, overlying calcrete (occ. low calc. banks)	wide, dry riverbeds	Kalahari grass species *Acacia hebeclada* *Acacia haematoxylon* *Acacia erioloba* *Diospyros lycioides* *Lycium hirsutum* *Acacia karroo*
9	Kalahari	Kalahari Thornveld	low, closed shrubby woodlands	shallow, red, base sands, overlying calcrete (occ. conglomerate, high-level gravels and calc. spurs)	low-lying, uneven slopes	*Acacia mellifera* *Tarchonanthus camphoratus* *Acacia erioloba* *Acacia haematoxylon*
14	Kalahari	Kalahari Thornveld Ecotonal Woodlands	short, open woodlands	red, base sands overlying calccrete and dolomite bedrock	gentle slopes and plains	*Acacia erioloba* *Acacia mellifera* *Tarchonanthus camphoratus*
13	Kalahari Ghaap Plateau Water shed	Kalahari-Ghaap Plateau Water shed Ecotonal Shrublands	high, closed woody shrublands	eutrophic, red soils overlying andesite lava; occ. quartzitic hills (occ. exposed)	plains and scattered, low hills	*Tarchonanthus camphoratus* *Acacia erioloba* *Acacia karroo* *Grewia flava* *Lycium hirsutum*

UNIT NO.	ECOLOGICAL REGION	VEGETATION TYPE	FORMATION CLASS	SOIL/GEOLOGY	TOPOGRAPHY	DOMINANT SPECIES
18	Kalahari - Ghaap Plateau Watershed	Woody Hill Shrublands	tall, closed, woody shrublands interspersed between high-lying grasslands	shallow, red sands and loamy sands overlying porphyritic granite (rarely exposed)	steep hills	*Dichrostachys cinerea* *Tarchonanthus camphoratus* *Euclea undulata* *Rhus tenuinervis* *Rhus ciliata* *Grewia flava* *Ziziphus mucronata* Kalahari grass species
20	Kalahari - Ghaap Plateau Watershed	Woody Hill Shrublands	short to high, closed, woody shrublands	shallow, red soils with exposed andesite lava and porphyritic granite rocks and boulders	rugged hills	*Tarchonanthus camphoratus* *Acacia tortilis* *Acacia robusta* *Vangueria infausta* *Acacia karroo* *Grewia flava* *Ehretia rigida* *Rhus ciliata*
17a	Kalahari	Kalahari-Kuruman Sourveld Transitional Grasslands	short, closed grasslands (mixed)	shallow, red, base sands overlying calcrete bedrock	plains	*Themeda triandra* *Schmidtia pappophoroides* *Stipagrostis uniplumis* *Anthephora pubescens* *Eragrostis pallens*
17d	Kuruman Sourveld	Kuruman Sourveld Grasslands	short, closed grasslands	shallow, brown, base soils with chert pebbles overlying dolomite bedrock slabs and very occ. calc. outcrops	undulating plains	*Themeda triandra* *Enneapogon* spp. *Stipagrostis uniplumis* *Cymbopogon plurinodis* *Elionurus argenteus* *Eragrostis lehmanniana*

UNIT NO	ECOLOGICAL REGION	VEGETATION TYPE	FORMATION CLASS	SOIL-GEOLOGY	TOPOGRAPHY	DOMINANT SPECIES
16a	Kalahari	Kalahari-Ghaap Plateau Mixed, Transitional Woody Shrublands	high, open to closed shrublands	shallow, red sands overlying calcrete bedrock; occ. raised calcrete faults	gentle slopes and plains	*Tarchonanthus camphoratus* *Acacia erioloba* *A. mellifera* *Grewia flava* *Rhus ciliata* *R. lancea*
16b	Kuruman Sourveld	Kuruman Sourveld Shrublands	high, open to closed shrublands	brown, base soils mixed with banded ironstone chips and chert pebbles and occ. exposed dolomite bedrock	gentle slopes and plains	*Tarchonanthus camphoratus* *T. minor* *Rhus ciliata* *Lebeckia macrantha* *Rhus tridactyla* *R. undulata* *Grewia flava*
33	Kuruman Sourveld	Kuruman Sourveld Shrublands	high, closed, woody shrublands	shallow, brown, base soils overlying dolomite bedrock slabs	plains and occasional calcrete lineaments (raised)	*Tarchonanthus camphoratus* *Rhus lancea* *R. ciliata* *Acacia karroo* *Grewia flava* *Ziziphus mucronata* *Diospyros austro-africana*
36a	Kuruman Sourveld	Kuruman Sourveld Grassy Shrublands plus Kalahari element	tall, sparse to open, grassy shrublands to shrubby grasslands	shallow, red, base sands overlying dolomite bedrock slabs, interbedded with calcrete; calc. faults	plains	*Tarchonanthus camphoratus* *Diospyros lycioides* *Acacia erioloba* *A. karroo* Mixed Kalahari and Kuruman grasses

UNIT NO	ECOLOGICAL REGION	VEGETATION TYPE	FORMATION CLASS	SOIL-GEOLOGY	TOPOGRAPHY	DOMINANT SPECIES
38	Ghaap Plateau	Ghaap Plateau Woodlands (Southern form)	short, open to closed, shrubby woodlands - thickets	shallow, brown soils (frequently absent) overlying calcrete bedrock	gentle slopes (associated with marshy "rivers")	*Olea europaea* ssp.*africana* *Rhus lancea* *R. pyroides* *Tarchonanthus camphoratus* *Ziziphus mucronata* *Diospyros lycioides* *Protasparagus laricinus* *Rhus ciliata* *Grewia flava* *Maytenus heterophylla* *Euclea crispa* *Euclea undulata* *Tarchonanthus minor*

This "divide" forms a series of ridges and hills of several geological formations, often covered by stretches of wind-deposited Kalahari sands, dipping deep below the sands to the west of the training area and exposed to the east. The "divide" has prevented the Kalahari sands from shifting further southwards during geologically drier periods, but in places where the "divide" is lower, the sands have been carried over it and deposited in the south by the prevailing north-west winds. The ecological region south of the watershed is more uneven, with geological formations close to the surface, or even exposed in places. The Kuruman Sourveld Ecological Region is generally covered by grasslands, grassy shrublands and shrubby grasslands, on brown, base soils, mixed with chert and banded ironstone pebbles or overlying solid dolomite bedrock slabs. This ecological region is very much more extensive to the south of the training area. The Ghaap Plateau Ecological Region is poorly represented in the training area, though some of the vegetation types in the Kuruman Sourveld Ecological Region could be regarded as transitional to this ecological region, e.g. vegetation types 33 and 36a. The Ghaap Plateau Ecological Region is dominant in the south-east of the training area, and consists of *Olea* and *Rhus* woodlands, shrubby woodlands and tall shrublands, mainly on calcrete, but infrequently on mixed calcrete-dolomite substrates. Surface soils are very shallow, calcareous or even almost absent.

10.3 COMPARISONS BETWEEN MULTITEMPORAL COMPUTER CLASSIFICATIONS AND VEGETATION-LANDSCAPE MAPS

As stated previously, one of the training areas used to test single date computer classification, the Kalahari Thornveld Training area, was used to test the multitemporal overlay computer classifications. Owing to the time and cost involved, only the NW and SE quadrats were investigated in detail.

10.3.1 NORTH-WEST QUADRAT (KALAHARI THORNVELD)

Plate 8 shows the results of the unsupervised classification of the PC1 and PC2, winter and summer. Fig. 7 is the vegetation-landscape map of the equivalent ground area. In the map the solid line encloses the NW quadrat as it appears in the training area of single date computer classification and the dotted line encloses that area of the vegetation map equivalent to NW quadrat multitemporal overlay computer classification. Where the squares overlap, the single date amd multitemporal overlay computer classifications may be compared (Paper 4).

Plate 8 Final colour-coded, computer-generated classification performed on principal components of NW quadrat.

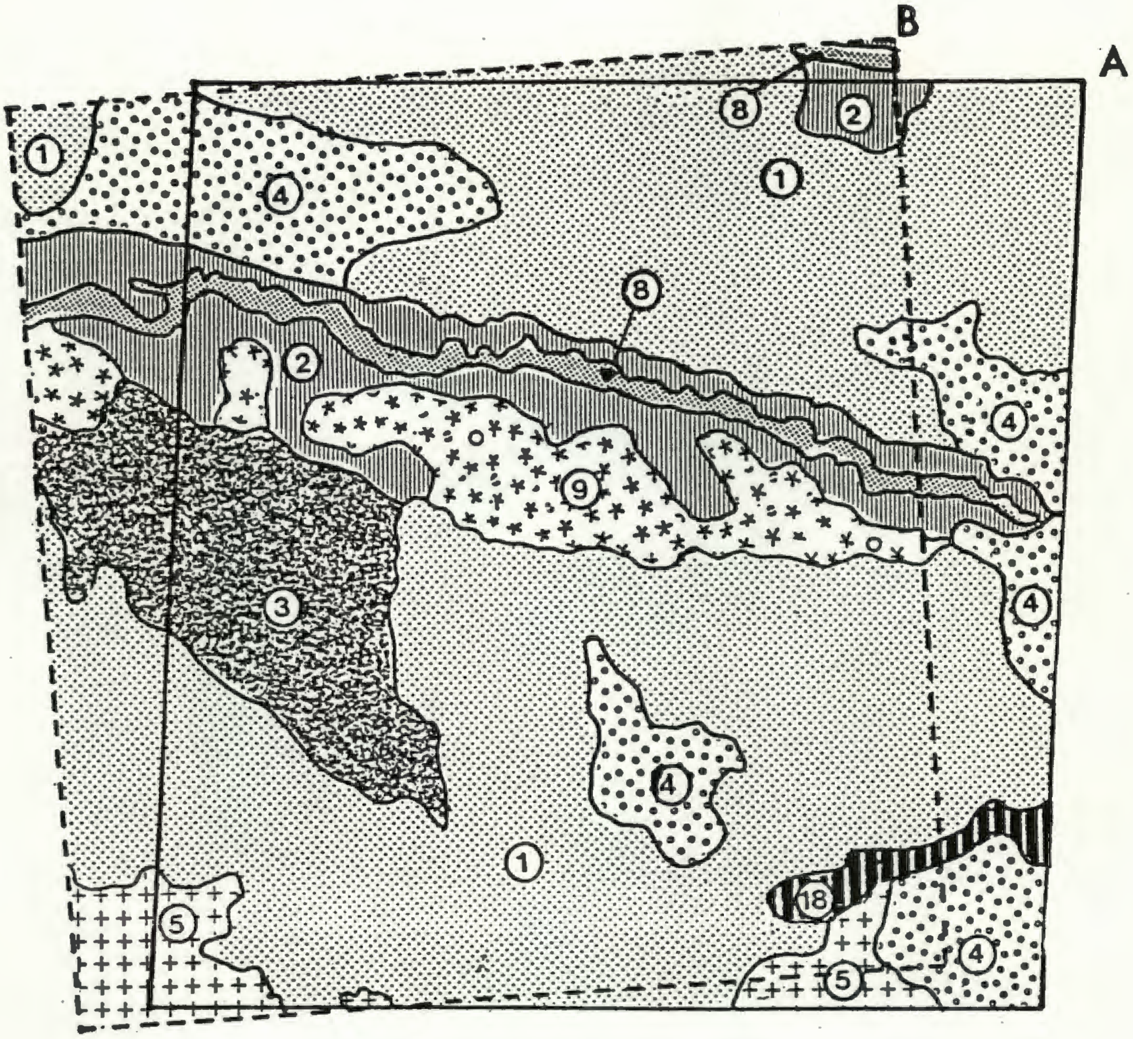


Figure 7 Map showing the vegetation-landscape units (V-L's) in the NW quadrat of the Kalahari Thornveld Training Area. The alpha-numeric codes refer to the vegetation-landscape units described in Table 1. The solid line (A) indicates the NW quadrat of the single date computer classification. The dotted line (B) indicates the NW quadrat of the multitemporal classification.

V-L 1 is acceptable as a major unit consisting of a mosaic of colour-codes: light blue and pink (major portion), and green with small amounts of puce, red, dark blue, white and yellow. V-L 1 is a short, closed woodland on pink to yellow, deep, base sands on high-lying, gentle, extensive slopes with the dominant species being *Terminalia sericea* and *Acacia erioloba*, often mixed, but commonly with one or the other more dominant in areas (Plate 5 in Paper 2). Here most of the light blue colour-code relates to areas dominated by *Terminalia sericea*. Note that the burn present in 1984 MSS data and in the field, has been obscured by regrowth and the development of new foliage on the large shrubs and trees; it would have been present in the extreme south-east corner. Farms with a high PCC are dominated by pink and puce.

V-L's 2 and 9, though mosaics of colour-codes, are separable. V-L 2 is exceptionally mixed, consisting of small groups of pixels of every colour used, except the purple and dark blue of V-L 4 in the north-west corner and the yellow of V-L's 1, 5 and 18. This presumably relates to the high degree of topographical diversity of the undulating pre-alluvium terrain, (drainage lines running into the Moshaweng River and sand dunes), as the vegetation is structurally and floristically simple. V-L 2 also has patches of white colour-code which relate to pure stands of Kalahari grassland where even the scattered, low, shrubby *Acacia haematoxylon* are absent. The small area of V-L 2 in the extreme north-east corner, lying next to a small section of the dry Lolwaneng riverbed is similar to that just discussed.

V-L 9, the low, closed shrubby woodland on red, base sands shallowly overlying calcrete spurs, and occasional areas of conglomerate and high-level gravels, is dominated by red, cyan and light blue, and is confused with V-L 3 and "artificial" V-L 4. V-L 9 is a mixture of areas of *Acacia mellifera* with some *Tarchonanthus camphoratus*, and vice versa.

V-L 8, the dry Moshaweng River banks and floodplain, with its pink, calcareous, alluvial sands and its high steep calcrete tufas on the north bank, is clearly dominated by light blue colour-code for the sandy, flat riverbed, and pink for the calcrete tufas. The narrow, true riverbed is largely dark blue and red, with a pink and yellow central portion. The major portion of the floodplain "riverbed" consists of grassland with very occasional single or small clumped *Acacia*, *Hebeclada* and *Rhigozum trichotomum* patches. The calcrete banks support an arid, low shrubby woodland of *Acacia* *mellifera*, *Rhigozum trichotomum*, *R. obovatum*, *Boscia albitrunca* and *Tarchonanthus camphoratus*. The central portion of the actual riverbed supports a high, closed woodland formation with *Acacia karroo*, *A. erioloba*, *Ziziphus mucronata*, *Diospyros lycioides*, *Lycium hirsutum*, etc. The small strip of the Lolwaneng River, consisting simply of a grassy (sparse to open) linear depression in the north-east corner, is colour-coded light blue. The flood plain area colour-codes as for the Moshaweng River.

V-L 3, although relatively uniquely classified using single date MSS data, is almost completely confused with V-L's 1, 9 and "artificial" 4 in the multitemporal analysis. V-L 3 consists of colour-codes dark blue, cyan, red, light brown, green and some white, with a fairly large amount of light blue, especially in the northern section. Although there would seem to be some pattern to the white, light brown, cyan, light blue, red and green, there is no explanation, other than that the white, light brown, cyan, red and green occur on higher ground than the light blue colour-coded areas, as the topography drops slowly in altitude, towards the Moshaweng River in the north. V-L 3 is a low, closed woodland (*Acacia haematoxylon*, *Terminalia sericea* and *Acacia erioloba*, mixed as co-dominants, but in areas *Terminalia sericea* forms low "thickets"), on pale pink to yellow, deep, base sands (plate 6 in Paper 2). The light blue would seem to occur either in more open areas with a good grass layer, or areas dominated by low, shrubby *Terminalia sericea*. The other colour-codes occur in areas dominated by *Acacia haematoxylon* and *Acacia erioloba*.

V-L 4 is a short, sparse to open woodland consisting of *Acacia erioloba* and *A. haematoxylon*, with species such as *Tarchonanthus camphoratus*, *Grewia flava* and *Lycium hirsutum* usually associated with the canopies of the *Acacia erioloba* trees (Plate 3 in Paper 2). In certain areas the density of *Acacia erioloba* is intermediate rather than open, especially in the transitional regions between V-L 4 and V-L 1 north of the river, where V-L 4 is colour-coded dark blue and purple. The black within the purple represents areas of blue colour-code (not shown) and a small quantity of unclassified pixels. The transitional zone between V-L 4 and V-L 1 has been classified with V-L 4 as dark blue. The small patch of V-L 4 on the eastern border, north of the Moshaweng River has been correctly classified. The two small areas of V-L 4 south of the Lolwaneng River in the NW quadrat have been classified differently to those in the north; the extreme south-east corner has colour-codes pink with some light brown and light blue in the centre. This is an *Acacia erioloba* woodland with *Tarchonanthus camphoratus.* The other small patch of V-L 4 (referred to as "artificial" V-L 4) has been structurally created by anthropogenic clearance of wood by nearby villagers. This is a sparse woodland of *Acacia erioloba* without *Tarchonanthus camphoratus.*

V-L 1 in the extreme south, while structurally falling into the same group, and also high-lying, occurs on red base sands overlying part of the quartzites of V-L 5 (similar to Plate 5 in Paper 2). This may explain why it is colour-coded pink here with some light brown and yellow instead of the colour-codes typical of the rest of V-L 1 to the north. There is a second possible explanation: both these areas, in the south-west and south-east lie between white-owned farms (southwards of the training area) and Bophuthatswana (south central and northern half of the training area). This can be clearly seen by the abrupt and "fence-line" changes in colour from pink, light brown and yellow to the light blue, green, cyan, red and dark blue. Generally the white-owned farmlands are well covered by vegetation which does not necessarily relate to vegetation in good condition as much of the vegetation cover is due to bush encroachment. There are, however, scattered farms which have been managed for good cover of highly palatable, nutritionally rich, subclimax to climax stage vegetation, especially graminoids where domestic stock is involved. The black-owned area generally lacks fencing and cultural, selective grazing is practised. This leads to bare pocket areas around villages, kraals and watering points (white and some of the light blue colour-coding) and a vegetation structure which is generally open, e.g. sparse to open woodlands with reasonably "good" grass cover.

V-L 5, is a short open to closed woodland on red, base sands overlying quartzite. V-L 5 occurs on high rounded hills, which have been largely covered by shifting aeolian Kalahari sands. The vegetation is a mosaic of communities, depending on topography and depth of the sands. Typical dominant species are *Acacia erioloba, A. mellifera, Ziziphus mucronata, Grewia flava, Dicrostachys cinerea* and *Terminalia sericea.* These species occur in varying combinations with either dominant *Terminalia sericea* or *Acacia erioloba*, and with dominant *Acacia mellifera* in the case of bush encroachment. V-L 5 has been colour-coded pink, yellow and light and dark grey with small areas of light blue. There is confusion between V-L 5 and southern parts of V-L 1 (where the Kalahari sands shallowly overlie the geological formations typical of these hilly ridges).

V-L 18 is a tall, closed, mixed, woody shrubland interspersed with high-lying grasslands (Plate 7 in Paper 1). These steep hills are variously covered by shallow red sands overlying porphyritic granite. On high areas the granite is exposed as relatively flat, rock-strewn plains and slopes supporting a floristically distinct closed grassland or grassland with very widely scattered large shrubs or solitary trees. To the north (along the north-facing edge of V-L 18) the ground falls away precipitously as an escarpment to the flat Kalahari stretching off to the northern horizon. Only the western section of V-L 18 is in the training area. The computer classification has failed to distinguish between V-L 18 and V-L 5. V-L 18 appears colour-coded pink, light brown and yellow, and is not separated from the high ground of the sand covered quartzite hills of V-L 5. Typical species of V-L 18 are *Dicrostachys cinerea, Tarchonanthus camphoratus, Euclea undulata, Rhus tenuinervis, R. ciliata, Grewia flava, Ziziphus mucronata* and several graminoid species. Both structurally and floristically V-L 's 18 and 5 are quite different from each other and it is thus surprising that the two have been classified so similarly. Common parameters of the two veld types are: (a) the shallow covering of red Kalahari sands over most of their distributions and (b) both occur on high ground which forms a physical "divide" between the Kalahari Thornveld to the north and Ghaap Plateau grasslands and tall shrublands to the south.

10.3.2 SOUTH-EAST QUADRAT (GHAAP PLATEAU)

Plate 9 shows the results of the unsupervised classification of the PC1 and PC2, winter and summer. Fig. 8 is the vegetation map prepared from field data. The solid boundary represents the SE quadrat of the single date computer classification and the dotted boundary encompasses the area of the SE quadrat of the multidate computer classification. The dotted line appearing in the north of the multitemporal vegetation map specifies the southern boundary of the area marked for the classification, i.e. V-L's 5, 20, 13 and northern distributions of V-L's 17a and 4. The mask was used to exclude the quartzitic and andesitic, elevated Kalahari-Ghaap Plateau Watershed (the "divide") from the computer classification. Only the V-L's lying south of this ridge were of interest, owing to their confusion in the single date

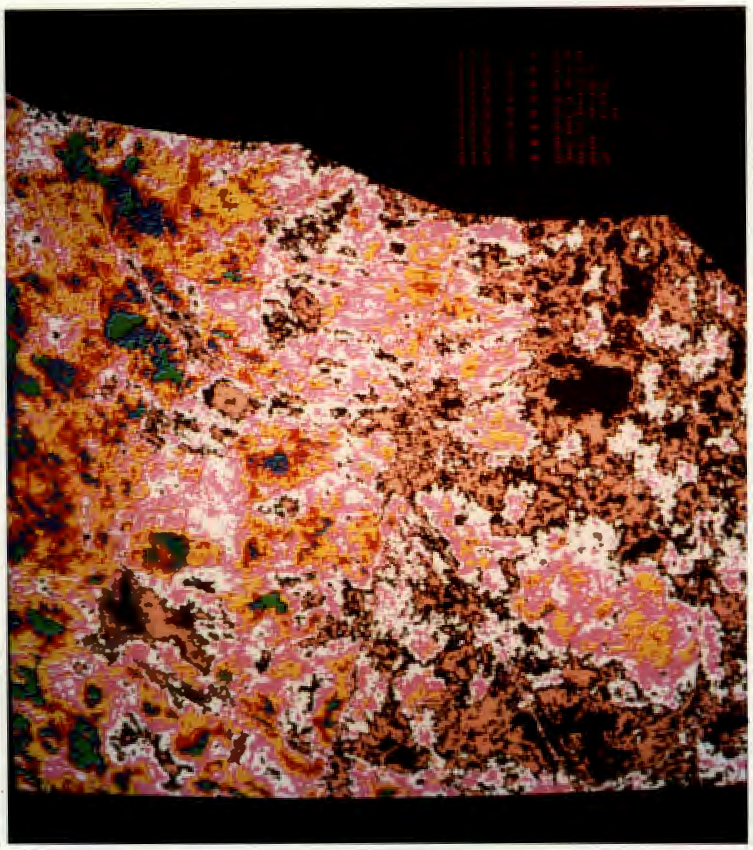


Plate 9 Final colour-coded, computer-generated classification performed on principal components of SE quadrat.

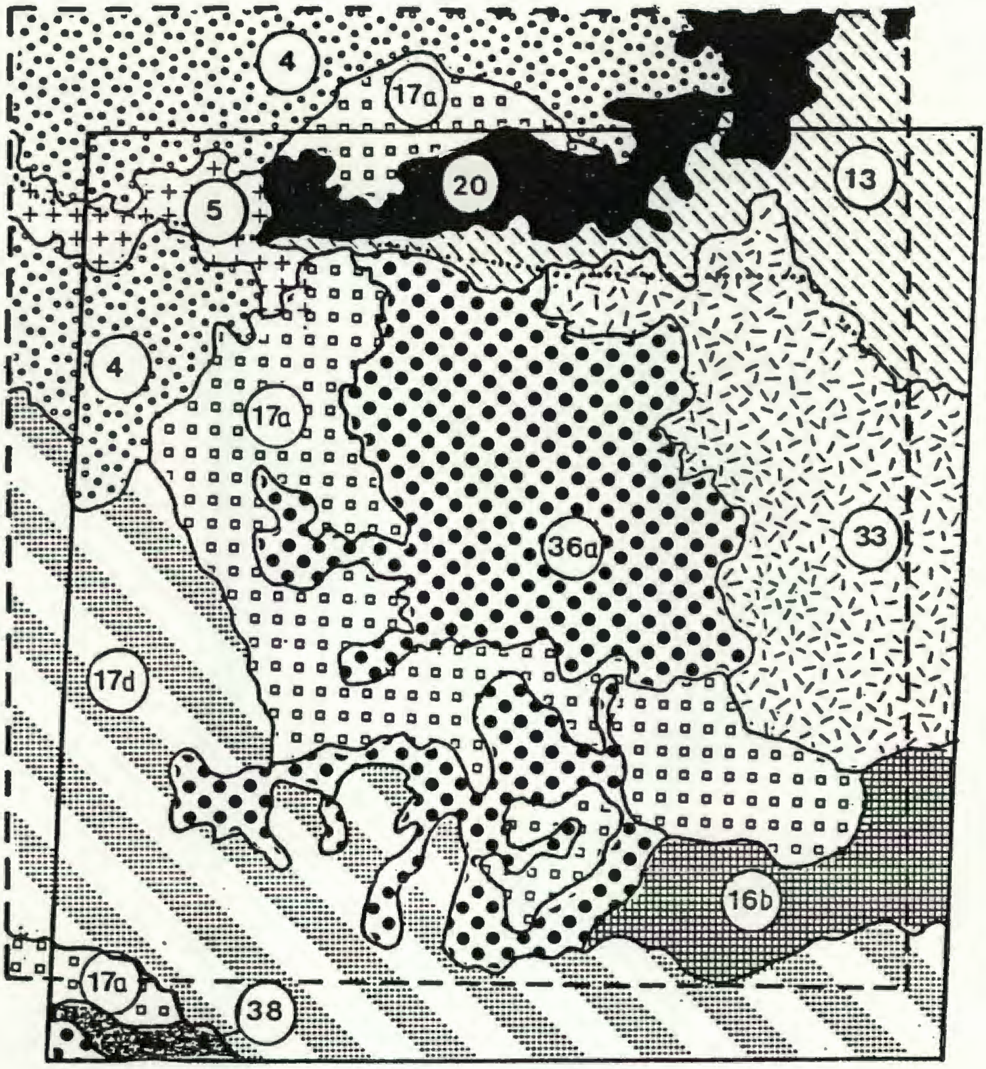


Figure 8 Map showing the vegetation-landscape units (V-L's) in the SE quadrat of the Kalahari Thornveld Training Area. The alpha-numeric codes refer to the vegetation-landscape units which are described in Table 1. The solid line (A) indicates the NW quadrat of the single date computer classification. The dotted line (B) indicates the NW quadrat of the multitemporal classification. The dotted line in the north of (B) specifies the northern boundary of the area marked for classification, i.e. the northern portion was not classified.

computer classification. This was probably a result of the vegetation "gradient" in mosaic units of grassland to shrubland.

In general V-L's 4 and 17d and some areas of V-L 17a (north and parts of central distribution) are classified together, with colour-codes yellow, red, green and blue, with small intervening patches of pink and white. The major portion of this combination consists of V-L 17d. V-L 17d is a short, closed grassland on shallow, brownish base soils mixed with chert pebbles and overlying dolomite bedrock slabs (Plate 7 in Paper 2). V-L 4 is a short, open to sparse woodland on deep, red, base sands (Plate 3 in Paper 2). In both V-L's, very widely scattered solitary trees and shrubs occur, although the grass layer is dominant in both V-L's. This may account for some of the misclassification as the soils, and vegetation structure and dominant floristics are different. It is thus surprising that there is not more misclassification between V-L's 4 and 17a, where both at least belong to the same ecological region, the Kalahari Thornveld. It is pleasing to note that V-L's 17a and 17d, both grasslands, are largely distinguishable as two separate vegetation-landscape units. There is more confusion between V-L's 4 and 17d than V-L's 17d and 17a, especially in terms of white and pink colour-codes in the extreme north-west corner of the classified quadrat.

V-L's 17a and 36a (northern section), much of the south central area of 17d and the northern parts of V-L 4 have similar colour-coding, viz. pink and white, with small amounts of yellow, lesser amounts of red and very small amounts of blue and green. V-L 17a (Plate 4 in Paper 2), a short, closed grassland on shallow, red, base sands overlying calcrete bedrock is largely confused with 36a (northern section), a tall, sparse to open grassy shrubland to shrubby grassland. The two vegetation-landscape units form a structural gradient, where it is often difficult to position a boundary between the two units. Within V-L 17a there are many open shrubland patches and in V-L 36a there are many pure grassland patches, and at a mapping scale of 1:250 000 these patches are below the acceptable resolution level. As mentioned above, V-L's 17a and 17d, although structurally almost identical, differ markedly in terms of soil-geological types and dominant - co-dominant species. Both are plains though V-L 17d is more undulating.

V-L's 33, 36a (southern distribution), part of V-L 36a (northern section) and V-L 16b are confused in their classification and have been colour-coded salmon and grey, though a rather large portion of V-L 33 also appears white with some pink. V-L 33 is a high, closed, woody shrubland on shallow, brown soils overlying dolomite (Plate 8 in Paper 2). V-L 36a is a sparse to open, grassy shrubland on shallow, red sands overlying dolomite. V-L 16b is a high, open to closed shrubland on brown, base soils mixed with banded ironstone chips and chert pebbles (Plate 8 in Paper 1). Thus, in terms of vegetation structure and soil-geology, V-L's 33 and 16b are similar, but V-L 36a is quite different. In terms of topography, all three V-L's are similar; plains and/or very gentle slopes. In terms of floristics all three V-L's have *Tarchonanthus camphoratus* as the dominant species though in V-L 36a, sparse to open shrublands, this species is widely to occasionally scattered. Both V-L's 16b and 33 are rather woody, with species such as *Rhus ciliata, R. tridactyla, R. undulata, R. lancea, Grewia flava* and *Ziziphus mucronata,* while the grass layer is of less importance. In V-L 36a, the grass layer is important and the other woody species are widely scattered and solitary, e.g. *Acacia karroo, A. erioloba* and *Diospyros lycioides.* The more rugged areas, the raised terrain and the exposed geology (banded ironstone, chert and dolomite) of V-L's 36a and 16b are colour-coded salmon and grey. Viewed in this way, these areas correlate well with the well-covered, mixed woody areas of V-L 33. Calcrete "faults" or lineaments and calcrete-surfaced farm roads are quite obvious in the classification and are colour-coded as speckled pink and white "lines".

In summary, the two grasslands, V-L's 17d and 17a, though structurally similar, are classified largely separately (except in the north-west) and shrubby grasslands are confused with grassy shrublands. There is overlap between adjacent vegetation-landscape units in an east-west direction, owing to the gradual gradient in all environmental parameters in the same direction and red, base sands dominate vegetal differences. Many of the pink and white colour-coded areas are grasslands, areas where the soil is exposed and exposed calcrete lineaments and roads. Many of the salmon and grey colour-coded areas constitute elevated sites, sites where geology is exposed and well-vegetated areas with high, woody shrublands. Tall, closed grasslands in good condition are often colour-coded yellow.

11. CONCLUSION

The production of a thematic map from the results of the computer classification of the multitemporal overlay would be a futile exercise. In most instances, multiple explanations would be required to interpret a given colour-code, and the classes generated in the classification cannot be regarded as meaningful in terms of the objectives of this study. Despite this, there is general cohesion in the classification, and when groups of colours are considered in combination, a clearer pattern emerges in terms of vegetation-landscape units.

As a result of the semi-arid conditions, interseasonal change in the Northern Cape and Karoo is largely limited to variability in cover and colour of natural vegetation. In considering multitemporal overlays in which contrast is of importance, summer and winter data files must be integrated in order to obtain maximum information. In addition, the summer data should be chosen with care as it is likely that data acquired during very dry periods would not contrast greatly with the winter data. Adequate registration of the two images can make a fundamental difference to the success of the classification (May 1986). Poor or insufficient ground control points can lead to the failure of the method (Boyle *et al.* 1988). This must be borne in mind prior to the integration of MSS data files. Byrne *et al.* (1980) suggested that a slight mis-registration of images served to enhance the resulting image by emphasizing some features ("stereoscopic" effect). In this study, good and adequate ground control points were used to register the images to within a 50 m margin of error.

In this study, the multitemporal overlay of summer and winter images and subsequent data dimensionality reduction by means of a principal components transformation, certainly resulted in an enhanced image with more information content than single date images from either season. The colour-contrast and range of brightness were good, and topographical features were enhanced. It is felt that good results would be obtained with the visual interpretation of an FCC image produced from the multitemporal overlay. Multitemporal overlays lead to sophistication and detail way beyond the simple addition of the two sets of single date MSS data.

When dealing with data from two seasons and in four bands a multitude of comparisons and analyses is possible, which, owing to constraints in cost and computer time could not be undertaken in this study. There was no in-depth investigation of single PC data or analyses combining spectral bands of different dates. Possibly an analysis of a different combination of the data would have yielded better results. A supervised or hybrid method of computer classification would probably have resulted in the generation of more meaningful classification classes. The fact that the bispectral plots of PC1 and PC2, winter, indicated more class overlap than those of summer has implications for alternative data analyses. One of the disadvantages of the multitemporal techniques is that it is costly.

It may be concluded that the computer classification of the multitemporal overlay yielded a better result than that of single date MSS data. Fontanel *et al.* (1975) reported similar results. Once again, it was found that the level of detail resulting from the classification was far beyond that required for reconnaissance level mapping. This level of detail implies that far more detailed ground reference data must be collected in order to interpret the classification accurately. Since the computer classification is intrinsically at a larger scale than reconnaissance level, it, in a sense, dictates the scale at which the investigator must work and interpret results. The multidate classification led to better discrimination in terms of vegetation-landscape units. Its performance was superior to that of a single date classification but it is doubtful whether the result, in terms of the objectives of this study, justified the costs involved.

REFERENCES

ADENIYI, P.O. 1985. Digital analysis of multitemporal Landsat data for land-use/land-cover classification in a semi-arid area of Nigeria. *Photogr. E.R.* 51(11): 1761-1774.

BOYLE, T.P., MONNIG, N.H. & VAN DER ZEL, D.W. 1988. A forest map of southern Africa with the aid of Landsat imagery. *S. Afr. J. Photogr., Rem. Sens. & Cartog.* 15(1): 61-71.

BROWNE, M.W. 1983. Principal components analysis: a review. In: Proc. Sem. P.C.A. in the Atmos. and Earth Sci., CSIR, Pretoria.

BYRNE, G.F., CRAPPER, P.F. & MAYO, K.K. 1980. Monitoring land-cover change by principal component analysis of multitemporal Landsat data. *Remote Sens. Environ.* 10: 175-184.

CLIFFORD, H.T. & STEPHENSON, W. 1975. An introduction to numerical classification. Academic Press, New York.

FONTANEL, A., BLANCHET, C. & LALLEMAND, C. 1975. Enhancement of Landsat imagery by combination of multispectral classification and principal component analysis. In: Proc. NASA Earth Resources Survey Symp., Houston.

GAUCH, H.G. 1982. Multivariate analysis in community ecology. Cambridge Univ. Press, Cambridge.

GRAETZ, R.D., PECH, R.P., GENTLE, M.R. & O'CALLAGHAN, J.F. 1986. The application of Landsat image data to rangeland assessment and monitoring: the development of a land image-based resource information system (LIBRIS). *J. Arid Environ.* 10: 53-80.

HIELKEMA, J.U. 1979. Application of a principal components analysis on Landsat multispectral data for studies on vegetation cover under desert conditions. In: Machine processing of remotely sensed data, ed. Lafayette, W. Proc. 5th Ann. Congr. on Machine Processing of Remotely Sensed Data, FAO, Rome.

LASSERRE, M., MALAN, O.G. & TURNER, B. 1983. The application of principal components analysis to Landsat MSS data. In: Proc. Sem. P.C.A. in the Atmos. and Earth Sci., CSIR, Pretoria.

LO, T.H.C., SCARPACE, F.L. & LILLESAND, T.M. 1986. Use of multitemporal spectral profiles in agricultural land-cover classification. *Photogr. E. R.* 52(4): 535-544.

MAY, L.N. 1986. An evaluation of Landsat MSS digital data for updating habitat maps of the Louisiana coastal zone. *Photogr. E. R.* 52(8): 1147-1158.

NELSON, R.F. & HOFFER, R.M. 1979. Computer-aided processing of Landsat MSS data for classification of forestlands. LARS Technical Report 102678.

NEWTON, A.R. 1983. An experiment in rock differentiation using digital processing of Landsat imagery. Proc. Int. Symp. Earth Data Info. Systems, Pretoria, ed. Martin, C., S. Afr. Soc. Photogr., Rem. Sens. and Cartog., Pretoria.

NEWTON, A.R. & BOYLE, T.P. 1988. Digital comparisons of multitemporal and multispectral Landsat MSS images. *S. Afr. J. Photogr., Remote Sens. & Cartog.* 15(1): 43-60.

PENDOCK, N. & SEARS, M. 1983. Reducing spectral dimension of remotely sensed data using principal components analysis and a peano scanning. In: Proc. Sem. P.C.A. in the Atmos. and Earth Sci., CSIR, Pretoria.

PIPER, S.E. 1981. Remote sensing course. Dept. of Survey and Mapping, Univ. of Natal, Durban. Unpublished.

SABINS, F.F. 1978. Remote sensing, principles and interpretation. W.H. Freeman & Co., San Francisco.

SCHREIER, H., GOODFELLOW, L.C. & LAVKULICH, L.M. 1982. The use of digital multi-date Landsat imagery in terrain classification. *Photogr. E. R.* 48(1): 111-119.

SHORT, N.M. 1982. The Landsat tutorial workbook: Basics of satellite remote sensing. NASA Publ. 1078, Washington D.C.

SOKAL, R.R. & ROHLF, F.J. 1978. Biometry - the principles and practice of statistics in biological research. W.H. Freeman & Co., San Francisco.

STEFFENS, F.E. 1983. What is principal components analysis? In: Proc. Sem. P.C.A. in the Atmos. and Earth Sci., CSIR, Pretoria.

SWAIN, P.H. 1983. Remote sensing: surveying the frontiers in the information age. *S. Afr. J. Photogr., Rem. Sens. & Cartog.* 13(5): 305-313.

TURNER, B. 1988. Landsat MSS temporal-spectral profiles of crops and grass on the highveld. *S. Afr. J. Photogr., Rem. Sens. & Cartog.* 15(1): 23-34.

PAPER 4

CONTENTS

PAPER 4

COMPUTER CLASSIFICATION OR VISUAL INTERPRETATION, MULTITEMPORAL OR SINGLE DATE DATA SETS? - A DISCUSSION

The objective as outlined in the preceding papers was to test the usefulness of Landsat imagery as a tool in the classifying and mapping of the vegetation of semi-arid areas. To this end a number of Landsat products as well as techniques of analysis were investigated. The results varied. This paper examines the strengths and weaknesses of the products and techniques in so far as has not been done in the preceding papers. Several issues concerning the future use of Landsat imagery are also discussed.

1. A COMPARISON OF THE COMPUTER CLASSIFICATIONS SINGLE DATE AND MULTITEMPORAL DATA SETS

The computer classifications of the single date Landsat data set and the multitemporal overlay were not regarded as successful in terms of the objectives of the study; each had its relative strengths and weaknesses and neither should necessarily be rejected for other vegetation analyses using Landsat data. The two classifications are compared with reference to Plate 2 in Paper 2 and Plate 8 and Plate 9 in Paper 3. The scales of the photographs are not the same. Plate 8 in Paper 3 corresponds with the upper left quarter of Plate 2 in Paper 2. Likewise, Plate 9 in Paper 3 corresponds with the lower right quarter of Plate 2 in Paper 2. The vegetation maps (Figs. 1, 2 & 3) are used as the standard for comparison. Only the areas of the single date classification of the Kalahari Thornveld Training area which correspond to the NW and SE quadrats of the multitemporal overlay will be discussed.

1.1 NORTH-WEST QUADRAT (KALAHARI THORNVELD)

V-L'S 2, 9 AND 8:

SINGLE DATE: These three V-L's were largely classified as one unit. In addition they have been confused with other V-L's in the training area.

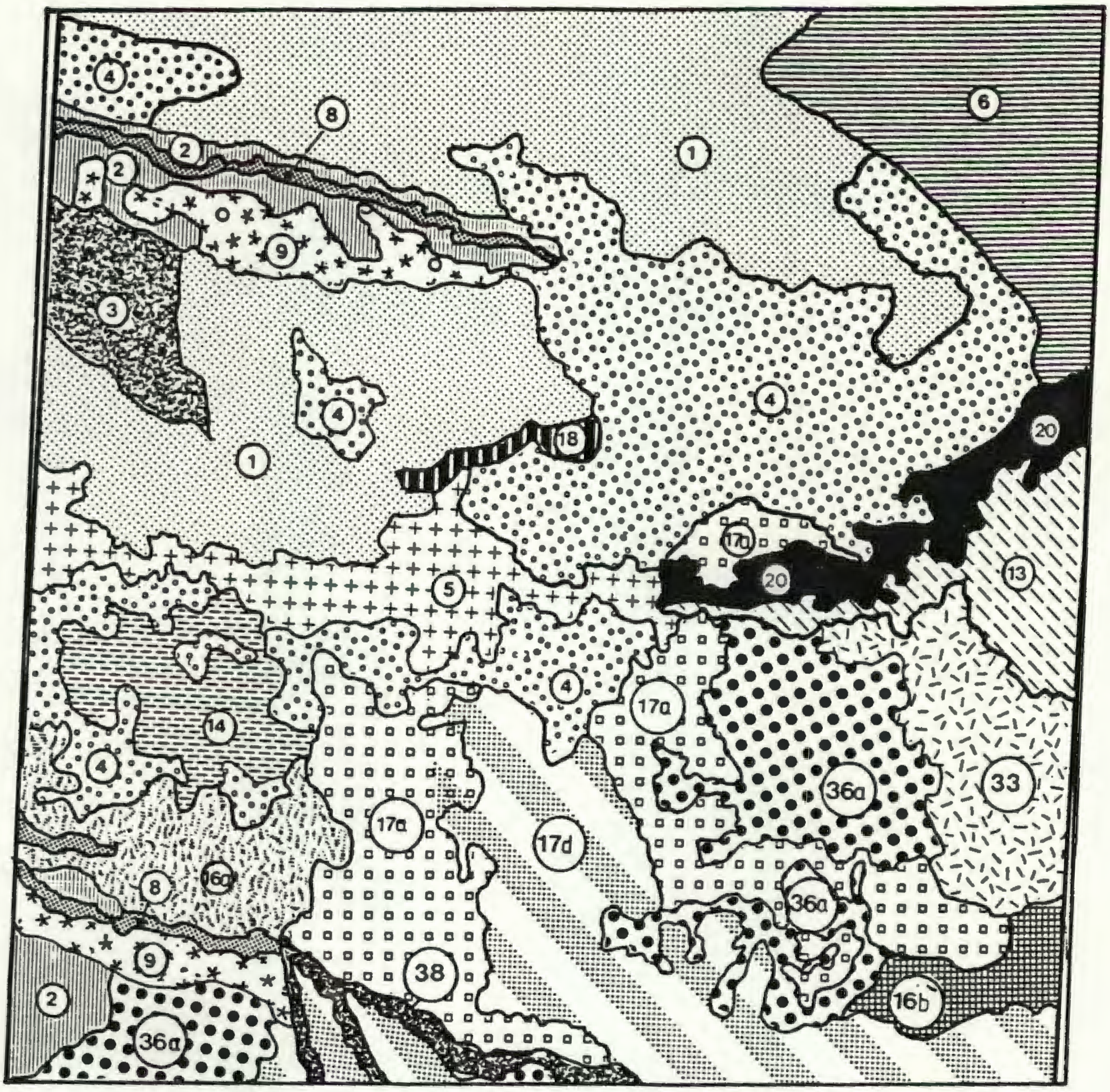


Figure 1 Map showing the vegetation-landscape units (V-L'S) in the Kalahari Thornveld Training Area. The alpha-numeric codes refer to the vegetation-landscape units which are described in Table 1.

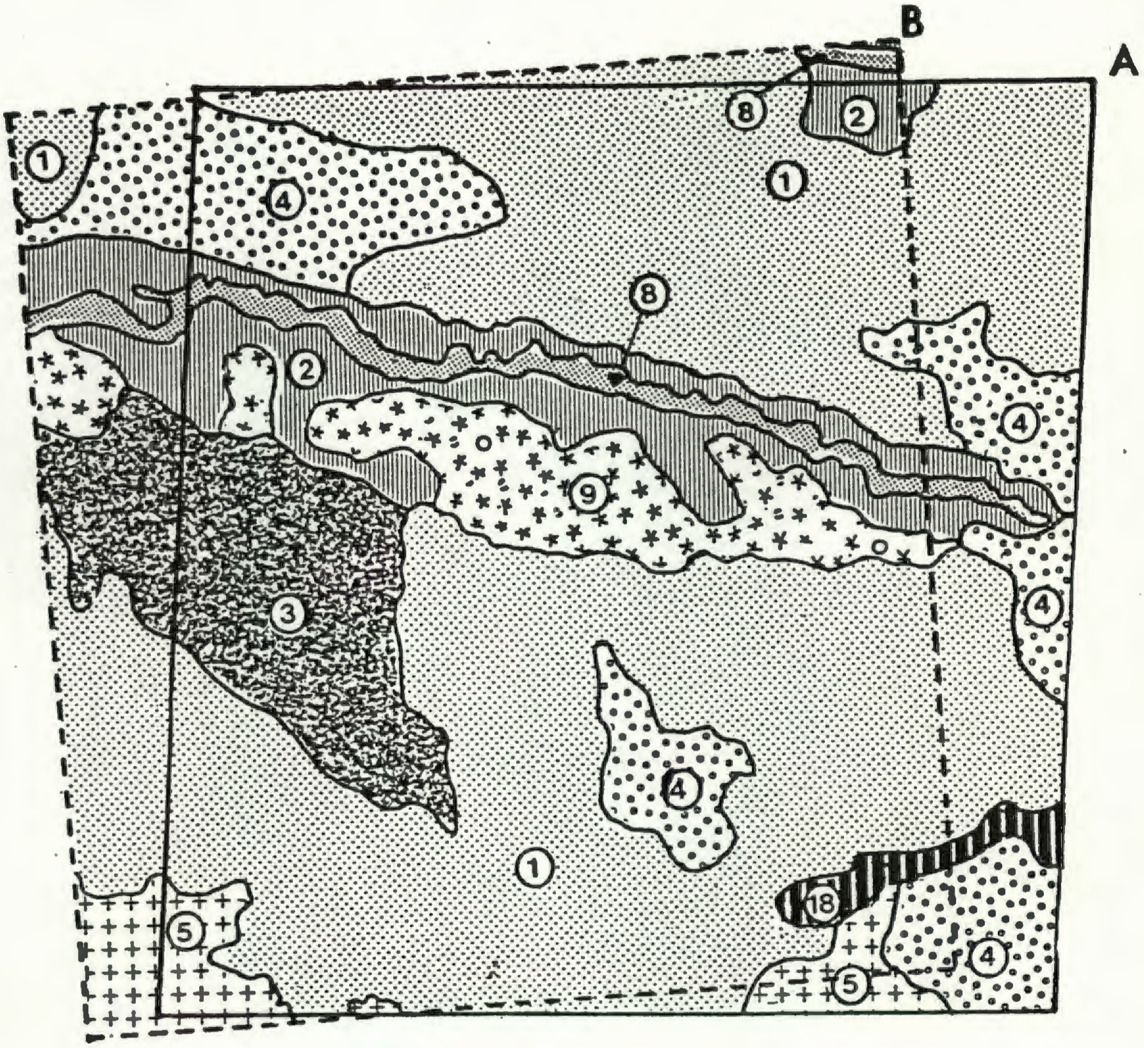


Figure 2   Map showing the vegetation-landscape units (V-L's) in the NW quadrat of the Kalahari Thornveld Training Area. The alphanumeric codes refer to the vegetation-landscape units described in Table 1. The solid line (A) indicates the NW quadrat of the single date computer classification. The dotted line (B) indicates the NW quadrat of the multitemporal classification.

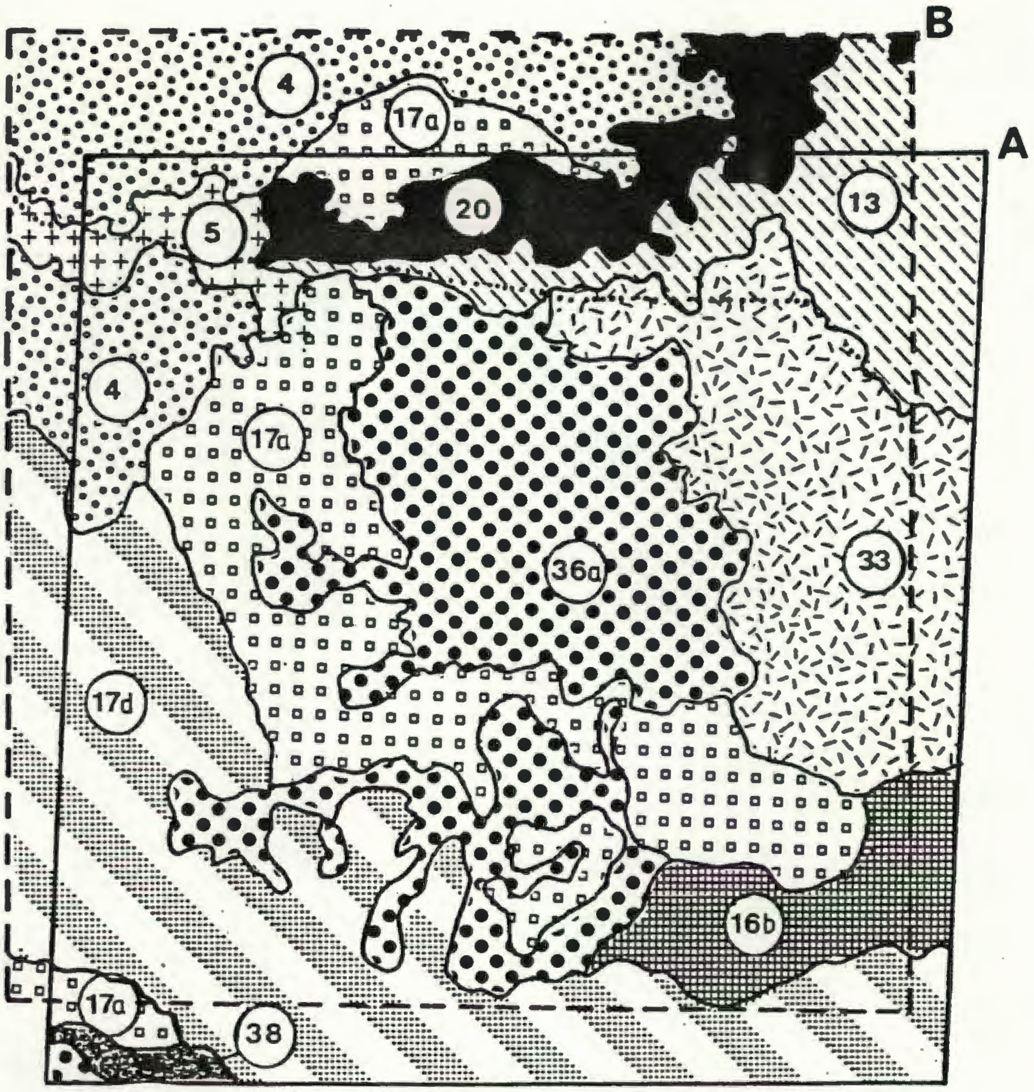


Figure 3 Map showing the vegetation-landscape units (V-L's) in the SE quadrat of the Kalahari Thornveld Training Area. The alpha-numeric codes refer to the vegetation-landscape units which are described in Table 1. The solid line (A) indicates the NW quadrat of the single date computer classification. The dotted line (B) indicates the NW quadrat of the multitemporal classification. The dotted line in the north of (B) specifies the northern boundary of the area marked for classification, i.e. the northern portion was not classified.

<u>MULTITEMPORAL</u>: There was some separation of the three V-L's. There was some confusion with other V-L's, but if V-L's 2, 9 and 8 are viewed in terms of distinctive combinations of colours, they form an excellent grouping.

The multitemporal classification is preferred.

<u>V-L 1</u>:
<u>SINGLE DATE</u>: V-L 1 has been unsatisfactorily isolated, consisting of many colour-codes that are typical of other V-L's.
<u>MULTITEMPORAL</u>: This classification is more successful especially if regarded as a specific combination of colours. There is some confusion with V-L 3 and to a lesser extent V-L 5.

The multitemporal classification is viewed as more successful in classifying this highly variable unit. This is as a result of the increased information obtained from the overlay method.

<u>V-L 4</u>:
<u>SINGLE DATE</u>: This is isolated to some extent as a combination of colour-codes. There is much confusion with V-L 1 and to a lesser extent with V-L 3.
<u>MULTITEMPORAL</u>: V-L 4 is a well-defined, cohesive unit of two colour-codes, except the "artificial" V-L 4 in the south. There is a small amount of confusion with V-L's 1 and 3.

While there are problems with both classifications, the multitemporal classification is preferred.

<u>V-L 3</u>:
<u>SINGLE DATE</u>: Seen as a combination of colour-codes, this V-L is reasonably well isolated.
<u>MULTITEMPORAL</u>: Viewed as a combination of colour-codes, this V-L is well-defined as a unit but there is considerable confusion with V-L's 1 and "artificial" 4, and the V-L group consisting of V-L's 2 and 9.

Although the overall distribution of V-L 3 in the multitemporal classification is more accurate, the single date classification is preferred for its unique isolation of part of V-L 3.

V-L 5:

SINGLE DATE: V-L 5 has an eastern and western distribution in the NW quadrat. Both the distributions are relatively well classified, but they are dissimilar in terms of their colour-code combinations. Thus while they are separated as units, they are also separated from each other.

MULTITEMPORAL: V-L 5 is poorly classified with a combination of colour-codes and is confused with parts of V-L 1.

The single date classification is more successful.

V-L 18:

SINGLE DATE: The classification of this V-L is most unsatisfactory. It consists of small patches of colour-codes typical of other V-L's.

MULTITEMPORAL: This V-L is completly confused with the colour-code combination of V-L 5 and cannot be extracted as a unit at all. It is not confused with a large number of V-L's as in the single date classification.

Neither the single date nor the multitemporal classification is acceptable.

1.2 SOUTH-EAST QUADRAT (GHAAP PLATEAU)

V-L 4:

A small area of V-L 4 occurs in the north-west corner of the classification area.

SINGLE DATE: V-L 4 is almost indistinguishable from V-L 17d. V-L's 4, 17d, 17a and 16b are similar in colour-coding.

MULTITEMPORAL: There is a large overlap with the western portion of V-L 17d, but no confusion with V-L's 17a and 16b.

While neither classification is really acceptable, the multitemporal classification is preferred.

V-L 17d:

SINGLE DATE : V-L 17d is confused with V-L's 4 and 17a and the boundaries between these three V-L's are indistinguishable.

MULTITEMPORAL: Separated from V-L 17a but there is an overlap with V-L 4. The classification of the western distribution is better than that of the south-eastern distribution. In the south-east there is some confusion between V-L's 17d, 17a and 36a.

While neither classification is good, the multitemporal classification is preferred for its separation of the major western portion of V-L 17d.

V-L 17a:

SINGLE DATE: There is a fair separation of this V-L, but there is confusion with 17d and 36a.

MULTITEMPORAL: The southern portion of this V-L is accurate, but the northern portion is confused with 36a. There is a small amount of confusion with 17d.

In the final analysis the multitemporal classification is preferred because V-L 17a is confused largely with V-L 36a, and the confusion between the two grasslands, V-L's 17a and 17d is low.

V-L 36a:

SINGLE DATE: The northern portion of this V-L is a relatively separate and distinct unit; otherwise there is considerable overlap and confusion with V-L's 17d and 17a.

MULTITEMPORAL: Central and southern fingerlike areas have been classified as V-L 33 and the northern part has been classified simalarly to 17a. However, even though the individual colour-codes may be more typical of other V-L's, the unique combination found in V-L 36a results in a unit with distinct boundaries.

Neither classification is satisfactory, but the clear definition of boundaries and a more accurate distribution contribute to a preference for the multitemporal classification.

V-L 33:

SINGLE DATE: Although V-L 33 has been classified as a distinct patch with a single colour-code, there is confusion with V-L's 36a and 16b.

MULTITEMPORAL: The major portion of V-L 33 is confused with V-L 16b. The remaining portion is confused with parts of V-L 36a and V-L 17a.

The single date classification is preferred.