



Objectivity in stratification, sampling  
and classification of vegetation

by

ROBERT HOWARD WESTFALL

Submitted in partial fulfilment of the requirements for  
the degree

PHILOSOPHIAE DOCTOR (BOTANY)

in the Faculty of Science  
(Department of Botany)

University of Pretoria  
Pretoria

October 1992

Promoter: Prof. Dr G.K. Theron

Co-promoter: Dr N. van Rooyen

ABSTRACT

Objectivity in stratification, sampling  
and classification of vegetation

by

ROBERT HOWARD WESTFALL

Promoter: Prof. Dr G.K. Theron

Co-promoter: Dr N. van Rooyen

in the

Department of Botany

for the degree

PHILOSOPHIAE DOCTOR (BOTANY)

The aims of this study are to increase objectivity in stratification, sampling and classification of vegetation, thereby, improving repeatability, predictability and relevancy of vegetation classifications.

The aims are achieved by: relating stratification, sampling and classification to scale; improved small-scale vegetation mapping using, satellite imagery; improved plant cover estimations; and vegetation classification by minimum entropy. A comprehensive computer program package was developed to facilitate the aims of this study and reduce time spent on vegetation analyses. It is recommended that the vegetation resource be given the highest national priority because correct vegetation management can also ensure conservation of soil and soil water.



1. INTRODUCTION	1
1.1 AIMS	1
1.2 JUSTIFICATION	2
1.3 HYPOTHESES	6
1.4 THESIS ARRANGEMENT	6
1.5 REFERENCES	7
2. STUDY AREAS	11
2.1 PHYSIOGRAPHY	11
2.2 GEOLOGY	13
2.3 SOILS	13
2.4 CLIMATE	14
2.5 BIOTIC FACTORS	15
2.6 PREVIOUS RESEARCH	17
2.7 REFERENCES	19
3. METHODS	24
3.1 PREPARATORY WORK	25
3.1.1 Scale	25
3.1.1.1 Background	25
3.1.1.2 Methods applied	27
3.1.2 Stand area	28
3.1.2.1 Background	28
3.1.2.2 Methods applied	29
3.1.3 Reconnaissance	30
3.1.3.1 Background	30
3.1.3.2 Methods applied	31



3.1.4 Stratification	32
3.1.4.1 <i>Background</i>	33
3.1.4.2 <i>A method for vegetation stratification using scale-related, vegetation-enhanced satellite imagery</i>	35
3.1.5 Stand location	41
3.1.5.1 <i>Background</i>	41
3.1.5.2 <i>Methods applied</i>	41
3.1.5.3 <i>PHYTOLOC - A random number generator and sample-set location program for stratified random vegetation sampling</i>	44
3.1.6 Sampling unit area	45
3.1.6.1 <i>Background</i>	45
3.1.6.2 <i>Predictive species-area relations and determination of subsample size for vegetation sampling in the Transvaal Waterberg</i>	47
3.2 <i>FIELD SAMPLING</i>	52
3.2.1 Sampling unit location	52
3.2.1.1 <i>Background</i>	52
3.2.1.2 <i>Methods applied</i>	54
3.2.2 Plant identification and verification	55
3.2.2.1 <i>Background</i>	55
3.2.2.2 <i>A new identification aid combining the features of a polyclave and an analytical key</i>	56
3.2.2.2.1 <i>Improvements</i>	66
3.2.3 Species cover	67



3.2.3.1	<i>Background</i>	67
3.2.3.2	<i>The plant number scale - an improved method of cover estimation using variable-sized belt transects</i>	69
3.2.3.2.1	<i>Improvements</i>	73
3.2.4	<i>Floristic data recording</i>	74
3.2.4.1	<i>Background</i>	74
3.2.4.2	<i>Methods applied</i>	74
3.2.4.3	<i>PHYTOCAP. A field-data capture program for the PHYTOTAB program package</i>	77
3.2.4.3.1	<i>Improvements</i>	78
3.2.5	<i>Habitat data</i>	78
3.2.5.1	<i>Background</i>	78
3.2.5.2	<i>Methods applied</i>	80
3.2.5.2 A)	<i>Field data</i>	80
a)	<i>Stand data</i>	80
b)	<i>Sampling unit data</i>	86
3.2.5.2 B)	<i>Derived data</i>	89
a)	<i>Input</i>	90
b)	<i>Output</i>	92
3.3	<i>CLASSIFICATION</i>	94
3.3.1	<i>Background</i>	94
3.3.2	<i>Methods applied</i>	98
3.3.2.1	<i>Relevè sequencing</i>	99
3.3.2.1 A)	<i>Commonality sequence</i>	101
3.3.2.1 B)	<i>Similarity sequence</i>	103
3.3.2.1 C)	<i>Separation unit sequence</i>	104
3.3.2.2	<i>Relevè grouping</i>	105



3.3.2.3 <i>Species sequencing</i>	106
3.4 <i>VERIFICATION</i>	109
3.4.1 Background	109
3.4.2 Methods applied	110
3.4.2.1 <i>Classification efficiency</i>	110
3.4.2.2 <i>Spatial relationships</i>	111
3.4.2.3 <i>Floristic and habitat relationships</i>	113
3.4.2.4 <i>Classification and field relationships</i>	115
3.5 <i>DERIVATIVES</i>	116
3.5.1 Background	117
3.5.2 Methods applied	117
3.5.2.1 <i>Primary derivatives</i>	118
3.5.2.1 A) <i>Plant communities</i>	118
3.5.2.1 B) <i>Community definition</i>	121
3.5.2.1 C) <i>Gradients</i>	123
3.5.2.2 <i>Secondary derivatives</i>	125
3.5.2.2 A) <i>Structure</i>	125
3.5.2.2 B) <i>Community composition analysis</i>	125
3.5.2.2 C) <i>Stand phase analysis</i>	127
3.5.2.2 D) <i>Community cover assessment</i>	128
3.6 <i>PHYTOTAB-PC</i>	129
3.6.1 <i>Classification and data processing</i>	130
3.6.2 <i>Data bank</i>	139
3.6.3 <i>External utility programs</i>	141
3.7 <i>REFERENCES</i>	142
4. <i>RESULTS AND DISCUSSION</i>	149
4.1 <i>PREPARATORY WORK</i>	149



4.1.1 Scale	150
4.1.2 Stand area	152
4.1.3 Reconnaissance	154
4.1.4 Stratification	155
4.1.5 Stand location	158
4.1.6 Sampling unit area	160
4.2 <i>FIELD SAMPLING</i>	162
4.2.1 Sampling unit location	162
4.2.2 Plant identification and verification	164
4.2.3 Species cover	166
4.2.4 Floristic data recording	169
4.2.5 Habitat data	171
4.2.5.1 <i>Field data</i>	173
4.2.5.2 <i>Derived data</i>	181
4.3 <i>CLASSIFICATION</i>	185
4.4 <i>VERIFICATION</i>	222
4.5 <i>DERIVATIVES</i>	225
4.6 <i>PHYTOTAB-PC</i>	228
4.7 <i>REFERENCES</i>	230
5. CRITICAL EVALUATION AND CONCLUSIONS	233
OPSOMMING	245
SUMMARY	247
ACKNOWLEDGEMENTS	249
CURRICULUM VITAE	250
REFERENCES	256
APPENDIX I	269



## CHAPTER 1. INTRODUCTION

Criticisms of the Braun-Blanquet (1928, 1951) approach to vegetation studies have mainly been concerned with two facets, namely,

- i. minimum sampling unit area or quadrat size (Goodall 1961, Werger 1974); and
- ii. observer bias, both in the sampling and classification phases (Goodall 1953, 1961; Poore 1956; Werger 1974).

In the southern African context, doubts on the relevancy of vegetation classifications can also be expressed: Is this the most cost-effective method of inventorizing and describing the variation in the vegetation resource? Is there a demand for classifications and how are they put to use?

Observer bias in the stratification, sampling or classification processes can affect repeatability and hence, the potential for prediction, of the results. This study is an attempt to overcome the criticisms and justify expanded use of the Braun-Blanquet approach in vegetation science, by reducing decision-making, and hence observer bias, through method refinement and increased application of computers. Objectivity in the basic processes of stratification, sampling and classification should thereby be increased.

### 1.1 AIMS

The aims of this study are threefold, namely, to:

- i. reduce subjective decision-making in the stratification, samp-





ling and classification processes for improved repeatability and predictability;

- ii. reduce time spent on the stratification, sampling and classification processes through computer automation, where possible, for greater efficiency; and
- iii. increase relevancy and significance of classifications by emphasizing possible classification derivatives.

## 1.2 JUSTIFICATION

Data collection in many scientific fields is a technical operation. That this is not so in the field of vegetation ecology can be attributed to the reliance placed on subjective decision-making. Reduction of subjective decision-making could ensure that the processes, up to and including classification, become technical. Research personnel could thus be freed to verify the adequacy, as well as improve the relevancy and significance of classifications, and increase areas covered. This should also provide greater scope for the employment of technicians.

In comparison with South Africa, with about 24 000 specific and infra-specific plant taxa (Gibbs Russell *et al.* 1985), the flora of Europe is relatively depauperized with about 15 000 to 17 000 specific and infra-specific plant taxa in an area more than eight times the size of South Africa (Polunin 1969). Application of the Braun-Blanquet methods in the depauperized European flora, in which the methods originated, can be relatively simple compared with South Africa, where natural vegetation still covers over 80%

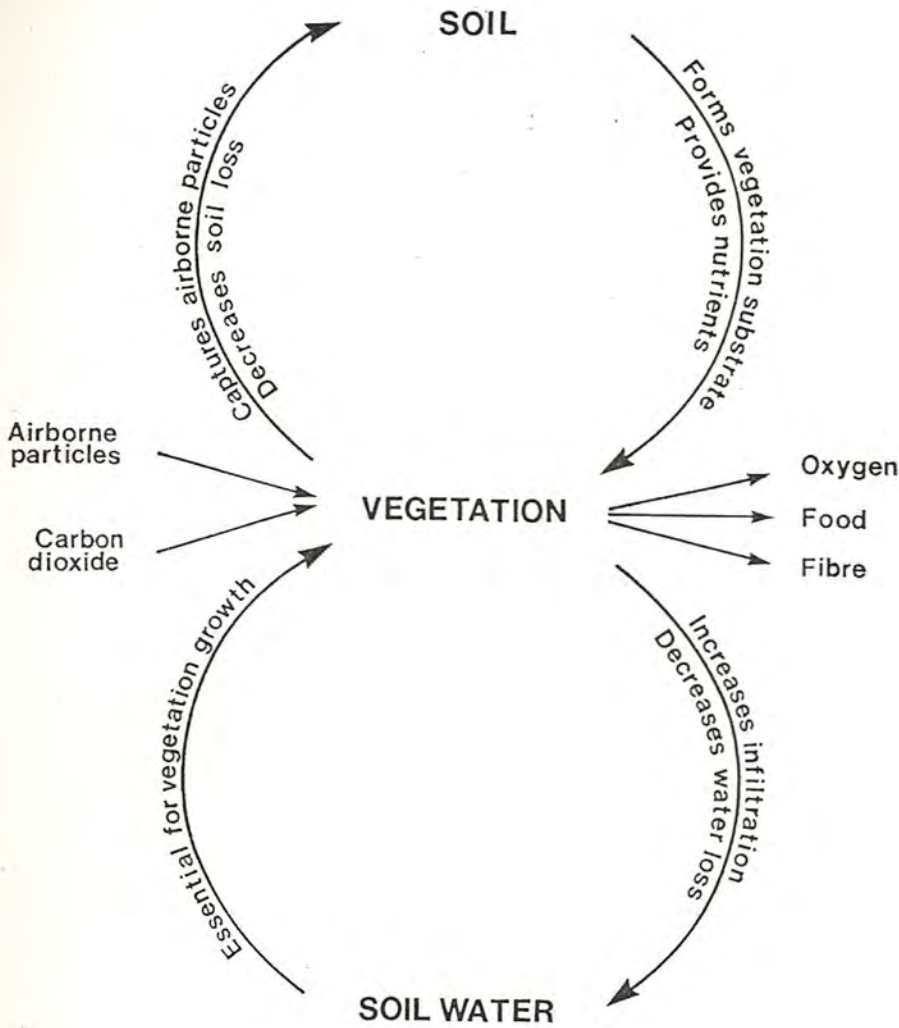


FIGURE 1.1. - A model of vegetation, soil and soil water interactions, illustrating the necessity for vegetation cover for soil and soil water conservation, as well as reducing airborne particles.



of the surface area (Rutherford & Westfall 1986) and consequent need for small-scale work is required. Furthermore, the largest areas of the Savanna, Nama-Karoo and Grassland Biomes occur on the interior plateau (Rutherford & Westfall 1986), with generally weak environmental gradients which can complicate understanding of floristic and environmental relationships, especially at small scales. This environmental and floristic complexity of the vegetation in South Africa, relative to Europe, can complicate subjective decision-making, especially for workers with insufficient experience in Braun-Blanquet methodology.

Why is it necessary to study vegetation at all? Is this a luxury that we cannot afford? Vegetation is essential to life. This resource is primarily responsible for oxygen production and removal of carbon dioxide and pollutants from the atmosphere. It is a primary food and fibre source. Vegetation cover conserves soil without which terrestrial plants would not grow. Vegetation can increase soil water through increased infiltration, decreased runoff and reduced evaporation through shading. Soil water is essential to plant growth and maintenance of the water supply in many streams. These relationships are illustrated in Figure 1.1. It is, furthermore, suggested that with conservation of vegetation, soil and primary water conservation is assured, as can be inferred from Figure 1.1. Therefore, costs can be saved and the total conservation effort could be made more effective by according vegetation conservation the highest priority of the natural resources.

But is classification of vegetation necessary in the study and



conservation of vegetation? Not only vegetation studies and conservation efforts, but any form of land use, from policy determination to the execution thereof, must be related to appropriate areas. For example, areas of interest at national level will generally be larger than those of interest for individual farms. This difference in area can be related to scale where national scale is smaller than farm scale. Such areas should also relate to the purpose for which the area is intended. For example, a catchment area should not be based on a magisterial district.

The assumption that a plant community integrates the effects of its physical environment, thereby indicating natural classes for significant environmental continua in a given area is basic to the Braun-Blanquet method (Poore 1956; Mueller-Dombois & Ellenberg 1974; Werger 1974). This assumption has not been disproved. Plant communities can also be related to scale because they form hierarchies where smaller communities can be grouped to form larger plant communities. This implies that the environmental factors, which differentiate plant communities at different scales, are also related to scale. Areas, based on plant communities should, therefore, be suitable for many purposes and at various scales.

Classification of the vegetation of South Africa into appropriate areas or communities can be advantageous for many land use practices, but is essential for vegetation conservation and utilization management. Reducing the number of decisions required for the stratification, sampling and classification processes can not



only improve the scientific validity of these processes, but can also facilitate these processes.

### 1.3 HYPOTHESES

- i. That adjacent plant communities at a certain scale differ in above-ground phytomass causing variation in the spectral characteristics, of the plant communities. The spectral characteristics are those detectable with a multi-spectral scanner. These differences can be used to objectively delimit plant communities at the given scales.
- ii. That the minimum sampling unit area for objective vegetation sampling is a function of scale, and vegetation structure. Vegetation structure, here, refers to plant size and spacing.
- iii. That more than one solution is possible in the classification of a vegetation data set.
- iv. That "noise" is a quantifiable attribute of a classified data set and is inversely proportional to pattern. "Noise" refers to outliers and the absence of species in a species-group in a classified data set.

### 1.5 THESIS ARRANGEMENT

Page numbering in this work is consecutive according to pages containing text or figures. Published articles are, therefore, renumbered accordingly. Tables and figures are numbered consecutively within each chapter, and are preceded by the chapter number, except in the case of published articles which retain the original



numbering. A wide range of literature was consulted, but is not referred to, unless it contributed materially to the thesis. Literature cited in each chapter, including references of published articles included in the chapter, is referenced at the end of each chapter and a full literature list is referenced at the end of this work. Each computer program, relevant to the aims of this study is dealt with separately, but jointly form the PHYTOTAB-PC program package, except where otherwise indicated. Program listings are available by negotiation with the author.

#### 1.6 REFERENCES

BRAUN-BLANQUET, J. 1928. *Pflanzensoziologie*. 1 Aufl. Springer, Wien.

BRAUN-BLANQUET, J. 1951. *Pflanzensoziologie*. 2 Aufl. Springer, Wien.

GIBBS RUSSELL, G.E., REID, C., VAN ROOY, J. & SMOOK, L. 1985.

List of species of southern African plants. Edition 2.

Recent literature and synonyms. Part 1. Cryptogams,

Gymnosperms, Monocotyledons. *Memoirs of the Botanical*

*Survey of South Africa* 51: 1-152.

GOODALL, D.W. 1953. Objective methods for the classification of vegetation. i. The use of positive interspecific correlation. *Australian Journal of Botany* 1: 39-63.

GOODALL, D.W. 1961. Objective methods for the classification of vegetation. iv. Pattern and minimal area. *Australian Journal of Botany* 9: 162-193.

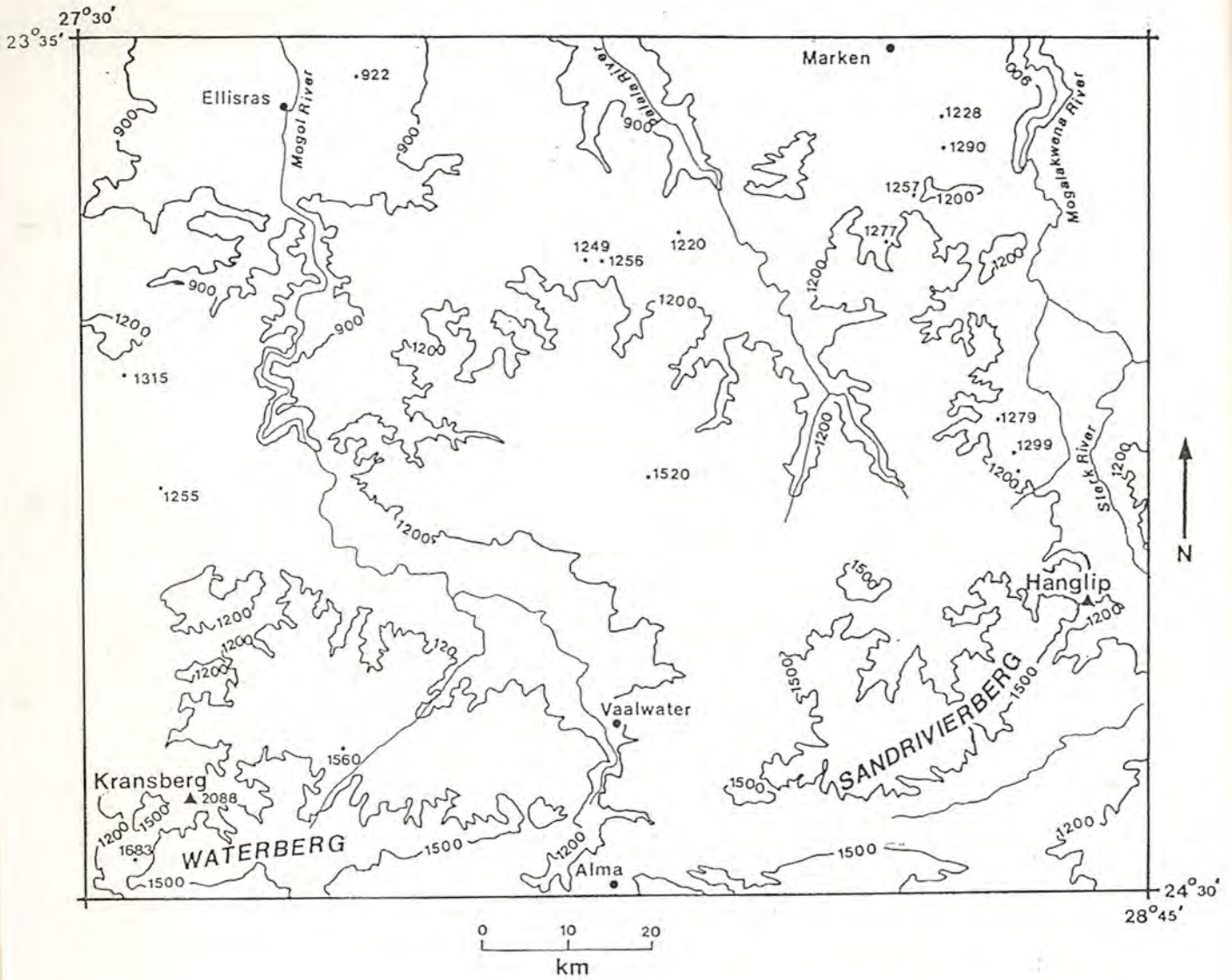


- MUELLER-DOMBOIS, D. & ELLENBERG, H. 1974. *Aims and methods of vegetation ecology*. Wiley, New York.
- POLUNIN, O. 1969. *Flowers of Europe: A field guide*. Oxford University Press, London.
- POORE, M.E.D. 1956. The use of phytosociological methods in ecological investigations. iv. General discussion of phytosociological problems. *Journal of Ecology* 44: 28-50.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1986. Biomes of southern Africa - An objective categorization. *Memoirs of the Botanical Survey of South Africa* 54: 1-98.
- WERGER, M.J.A. 1974. On concepts and techniques applied in the Zürich-Montpellier method of vegetation survey. *Bothalia* 11: 309-323.



FIGURE 2.1 - The Transvaal Province, Republic of South Africa, showing the location of the main study area, in the north western sector of the province.





Legend

- ▲ Mountain peak
- Contour interval: 300 m

FIGURE 2.2 - The physiography of the main study area in the north western Transvaal, showing the main drainage lines.



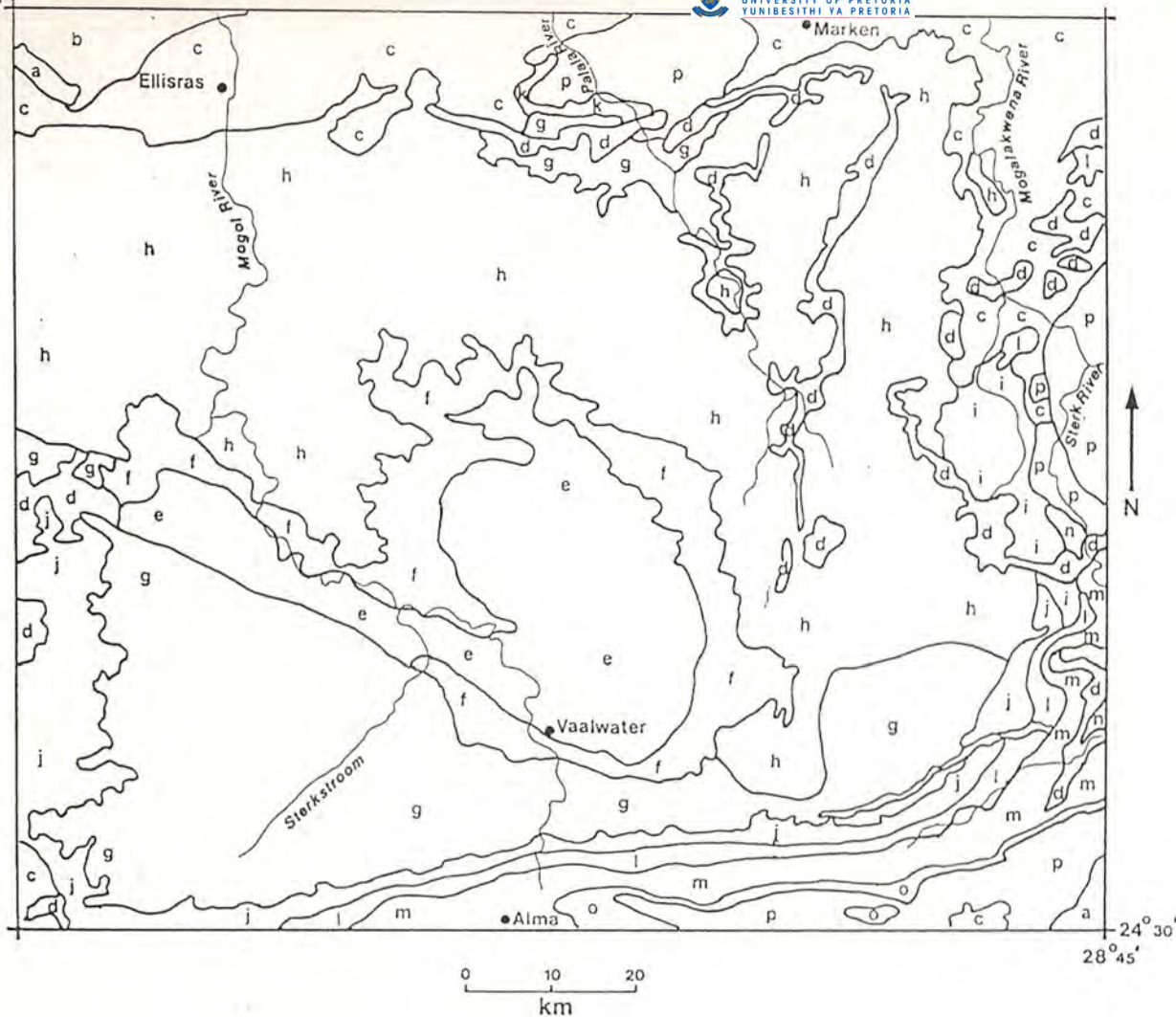
## CHAPTER 2. STUDY AREAS

Two separate study areas were selected. The main study area is situated in the north-western Transvaal (Rutherford & Westfall 1984) between southern latitudes  $23^{\circ} 35'$  and  $24^{\circ} 30'$  and eastern longitudes  $27^{\circ} 30'$  and  $28^{\circ} 45'$  which include the major portion of Sour Bushveld (Acocks 1953, 1975, 1988) of the Transvaal Waterberg, north from Kransberg and Hanglip (Figure 2.1 & 2.2). In contrast to many vegetation studies the area is rectangular in shape which is intended to facilitate later integration of data. The area covers approximately  $13\ 000\ \text{km}^2$  with dimensions of  $100 \times 130\ \text{km}$ . Ellisras is the largest town, situated in the north-west of the study area (Figure 2.1).

The second study area is situated at the CSIR Division of Roads and Transport Technology, Silverton, Pretoria site. The study area comprises two  $10 \times 10\ \text{m}$  plots ten metres apart in a visually homogeneous grassland with no observable variation in physiography, geology, soils, climate or biotic factors.

### 2.1 *PHYSIOGRAPHY*

The physiography of the main study area is extremely irregular, being mainly mountainous with the Waterberg and Sandriviersberg ranges in the south having altitudes of up to  $2\ 088\ \text{m}$  above mean sea level and decreasing in altitude northwards (Figure 2.2). Extensive plains at altitudes between  $800$  and  $1\ 200\ \text{m}$  above mean sea level occur in the far northwest of the study area. The main

27°30'  
23°35'UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

## LEGEND

## Formation and dominant lithology

a	Basalt
b	Clarens Sandstone, siltstone
c	Undifferentiated post-Waterberg
d	Diabase
e	Vaalwater Sandstone, siltstone, shale
f	Clermont Sandstone, grit
h	Mogalakwena Sandstone, grit
g	Sandriviensberg Sandstone, grit
i	Makgabeng Sandstone
j	Aasvoëlkop Siltstone, mudstone, shale
k	Setlaole Sandstone, grit
l	Skilpadkop Sandstone, conglomerate
m	Alma Sandstone, conglomerate
n	Sterksrivier Sandstone, conglomerate
o	Swaershoek Sandstone, conglomerate
p	Undifferentiated pre-Waterberg

FIGURE 2.3. - The geological formations in the main study area showing dominant surface lithology from Jansen (1982).

perennial rivers are the Mogol, Palala and Mogalakwena Rivers which drain northwards into the Limpopo River. Southward incisions by these drainage lines have formed extensive lower-lying floodplains within the mountain area.

## 2.2 GEOLOGY

The geology of the Waterberg area has been described by Jansen (1982). The main geological formations (Figure 2.3) in the east of the study area are the Aasvoëlkop Formation, with mainly siltstone, mudstone, shale and laharite and the Makgabeng Formation, with mainly sandstone. Both these formations are of the Matlabas Subgroup. The main formation in the south, is the Sandriviersberg Formation with mainly grit; in the north and west, the Mogalakwena Formation with mainly sandstone, grit and conglomerate; and in the central part of the study area, the Cleremont Formation with mainly sandstone and grit and the Vaalwater Formation with mainly sandstone, arkose, siltstone and shale. The lastmentioned four formations are of the Kransberg Subgroup. Both the subgroups are of the Waterberg Group and Mokolian Erathem, with ages from approximately 1 700 Ma to 1 300 Ma. The study area has very few known deposits of economic value but coal is mined in the Karoo sediments at the Grootgeluk Mine near Ellisras and iron ore is mined at Thabazimbi, both in close proximity to the study area.

## 2.3 SOILS

One of the most recent soil maps of the Republic of South Africa



is the map of MacVicar (1973) where the entire study area is classified as being a Red-Yellow-Grey latosol plinthic catena with neutral sands/loams, yellow-grey dominant and much rocky land. The lack of differentiation of soil units within the study area can be attributed to the scale of 1: 2 500 000 at which this soil classification is mapped. Although the binomial and taxonomic soil classification systems for South Africa (MacVicar *et al.* 1977, 1991) provides for more detailed mapping of soils, the soils within the study area have not as yet been mapped using these systems. However, a comparison of floristic units with soil forms and series at a mapping scale of 1: 30 000, in a pilot project within a small part of the study area, often showed poor correlation (Westfall 1981). This poor correlation can be attributed to the limits set for the units of the soil classification not being necessarily the same as the limits influencing the vegetation. Furthermore, it was found that the recorded soil chemical properties were far less significant in differentiating natural vegetation, at this scale, than soil physical properties, in the area. Consequently, greater emphasis is placed on physical rather than chemical properties of the soil in the main study area.

#### 2.4 CLIMATE

According to Köppen's classification (Schulze 1947) the main study area is classified as Cwa which describes a warm temperate climate with summer rainfall and a January mean temperature exceeding 22° C. The climate is continental with more than 85% of the mean annual rainfall falling in summer, from October to March (Schulze



1965). Precipitation is variable with mean annual rainfall varying from 650 to 900 mm in the mountainous areas and below 500 mm on the plains in the north of the study area. Precipitation data was obtained from the Weather Bureau, Private Bag X447, Pretoria, 0001, by request. The four hottest months of the year are November to February (Schulze 1965) which, because of predominantly summer rainfall, is the period when optimum plant growth should occur. The irregular topography of the study area would necessitate a very fine grid of climate stations to allow climatic correlation with vegetation at a scale of 1:250 000. The influence of climate on vegetation differentiation is, however, of prime importance in the study area (Westfall 1981). Because insufficient weather data are available from official and published sources, these data could be determined from farm records, where available and by interpolation and other indirect means.

## 2.5 BIOTIC FACTORS

Observations during the course of fieldwork indicate that in terms of area, veld grazing by cattle is the main land use practice in the main study area but the game on many farms also utilize the natural vegetation. Cultivation is mainly confined to the flatter areas where irrigation is possible, such as the floodplains of the Sterk, Palala and Mogol Rivers. Crops include maize, grain sorghum, tobacco and melons.

Observations also indicate that a dramatic human population increase is taking place at Ellisras with its newly developed

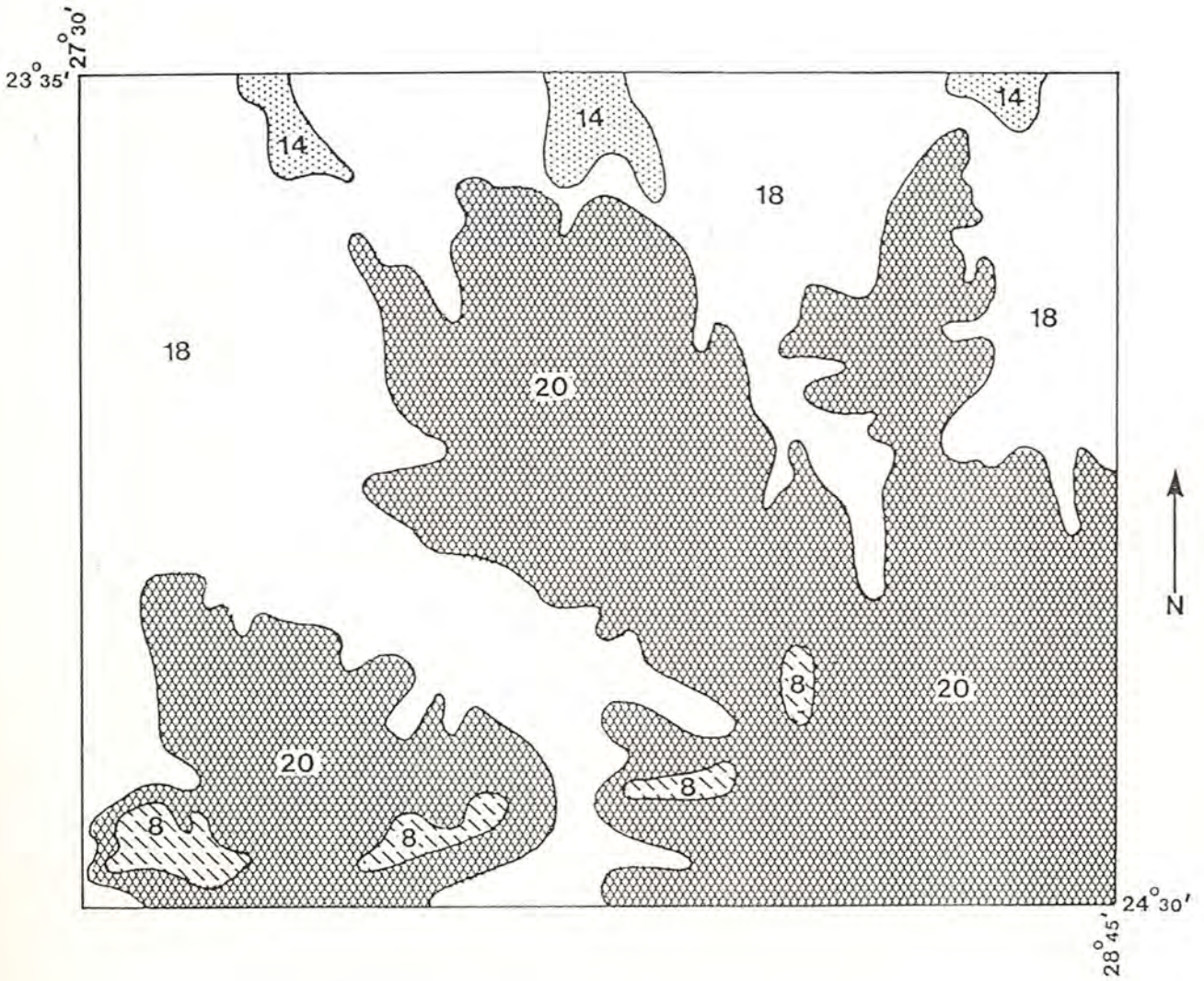


FIGURE 2.4. - The Veld Types (Acocks, 1953, 1975, 1988) of the main study area in the north-western Transvaal. 8, North-Eastern Mountain Sourveld; 14, Arid Sweet Bushveld; 18, Mixed Bushveld; 20, Sour-Bushveld.



coalfields and the construction of the Matimba power station. The relatively unspoilt, scenic grandeur of the Waterberg with the Kransberg massif considered to be amongst the best rock-climbing facilities in the Transvaal, according to members of both the Transvaal and Northern Transvaal Mountain Clubs; the many game and private nature reserves; the new Kransberg National Park and the proximity of the Waterberg to the Pretoria-Witwatersrand complex can combine to lure increasing numbers of visitors and increase development and human exploitation of the environment. Although the study area could be considered underutilized in terms of agricultural production, its potential could be drastically reduced by the effects of a large human population increase.

The following Veld Types (Acocks 1953, 1975, 1988) are represented in the study area and the percentage of the study area covered, is indicated in brackets: Sour Bushveld (53%); Mixed Bushveld (43%); Arid Sweet Bushveld (2%); and North-Eastern Mountain Sourveld (2%) (Figure 2.4). The Sour Bushveld is found in the mountainous areas with Mixed Bushveld on the plains adjacent to the mountains and on the floodplains. Arid Sweet Bushveld is limited to the floodplains of the Mogol, Palala and Sterk Rivers in the north. Outliers of North-Eastern Mountain Sourveld are found only above 1 500 m altitude.

## 2.6 PREVIOUS RESEARCH

Early scientific interest in the vegetation of South Africa emphasized showy plants and plants of nutritional and medicinal value





(Werger 1978). Interest in the geographic distribution of plants, in South Africa and world-wide, developed during the nineteenth century because of the descriptions of travellers such as Lichtenstein (1811, 1812) and Burchell (1822, 1824). In subsequent descriptive phytogeographic divisions of southern Africa, based mainly on observations and experience, the Transvaal Waterberg was placed in the: Regnum Mesembryanthemorum (Schouw 1823); Kalahari Region (Grisebach 1872); Highveld Region (Rehmann 1880); Palaeotropic Dry Region (Engler 1882); Kalahari Region (Bolus 1886); and the Highveld Region (Marloth 1908).

During the first two decades of the twentieth century, botanists such as Schönland, Marloth, Phillips, Burt-Davy and Bews published checklists of plant species and ecological notes for various areas in South Africa (Schönland 1922). Bews, who worked mainly in Natal also studied the succession of plants in South Africa (Bews 1916) and describes the grasslands of South Africa (Bews 1918), where the Transvaal formed part of the Eastern Grassland Region. Guidelines for future surveys, including codes for habitat factors and plant physiognomy, were suggested and checklists and plant succession were emphasized. The works of Bews (1916) and Clements (1916) on plant succession greatly influenced vegetation surveys in South Africa for the next four decades.

During this time the Transvaal Waterberg was described as being part of the: Kalahari Park and Bush Province (Pole Evans 1922); Evergreen and Deciduous Tree and Bush variation of Parkland (Pole Evans 1936) and Small Tree Savanna variation of the Bush Veld



Savanna (Adamson 1938). More recent, small-scale work, describing the vegetation of southern Africa includes that of: Eyre (1963), Meester (1965), Werger (1978), and White (1981). Although the descriptions of these vegetation types indicate a lack of conformity, the major vegetation units of Adamson (1938) correspond most closely with the Biomes of southern Africa (Rutherford & Westfall 1986) where the Waterberg area forms part of the Savanna Biome.

More detailed vegetation studies, with the purpose of improving agriculture subsequently played a greater role. Irvine (1941) classified the vegetation of the northern Transvaal into veld types, based on grazing potential and Acocks (1953, 1975, 1988) relied largely on these data for his classification of veld types, in this area. The vegetation of the Kransberg block in the south west of the study area was described by Coetzee *et al.* (1981) without formal sampling and Westfall (1981) studied the farm Groothoek, also in the south west of the study area, at a detailed scale, using the phytosociological approach.

Although Acocks (1953, 1975, 1988) *inter alia*, adopted a floristic sampling approach it was only with Werger (1973) and subsequent workers that the formal Braun-Blanquet methods of floristic sampling and classification were introduced to South Africa.

## 2.7 REFERENCES

- ACOCKS, J.P.H. 1953. Veld types of South Africa. *Memoirs of the Botanical Survey of South Africa* 28: 1-192.



- ACOCKS, J.P.H. 1975. Veld types of South Africa. 2nd edition.  
*Memoirs of the Botanical Survey of South Africa* 40: 1-128.
- ACOCKS, J.P.H. 1988. Veld types of South Africa. 3rd edition.  
*Memoirs of the Botanical Survey of South Africa* 57: 1-146.
- ADAMSON, R.S. 1938. *The vegetation of South Africa*. Whitefriars, London.
- BEWS, J.W. 1916. Account of the chief types of vegetation in South Africa with notes on plant succession. *Journal of Ecology* 4: 129-159.
- BEWS, J.W. 1918. *The grasses and grasslands of South Africa*. Davis, Pietermaritzburg.
- BOLUS, H. 1886. Sketch of the flora of South Africa. *Official handbook of the Cape of Good Hope* 286-317. Richards, Cape Town.
- BURCHELL, W.J. 1822. *Travels in the interior of southern Africa*. Vol. 1. Paternoster, London.
- BURCHELL, W.J. 1824. *Travels in the interior of southern Africa*. Vol. 2. Paternoster, London.
- CLEMENTS, F.C. 1916. *Plant succession: An analysis of the development of vegetation*. Carnegie Institute, Washington.
- COETZEE, B.J., VAN WYK, P., GERTENBACH, W.P.D., HALL-MARTIN, A. & JOUBERT, S.C.J. 1981. 'n Plantekologiese verkenning van die Waterberggebied in die Noord-Transvaalse bosveld. *Koedoe* 24: 1-23.
- ENGLER, A. 1882. *Versuch einer Entwicklungsgeschichte der Pflanzwelt insbesondere der Florengebiete seit der Tertiärperiode*. Engelmann, Leipzig.



- EYRE, S.R. 1963. *Vegetation and soils. A world picture.* Arnold, London.
- GRISEBACH, A. 1872. *Die vegetation der Erde nach ihrer klimatischen Anordnung.* Engelmann, Leipzig.
- IRVINE, L.O.F. 1941. *The major veld types of the northern Transvaal and their grazing management.* D.Sc. (Agric.) thesis, University of Pretoria, Pretoria.
- JANSEN, H. 1982. The geology of the Waterberg Basins in the Transvaal, Republic of South Africa. *Memoirs of the Geological Survey of South Africa* 71: 1-98.
- LICHTENSTEIN, H. 1811. *Reisen im südlichen Africa in der Jahren 1803, 1804, 1805, 1806.* Vol. 1. Salfeld, Berlin.
- LICHTENSTEIN, H. 1812. *Reisen im südlichen Africa in der Jahren 1803, 1804, 1805, 1806.* Vol. 2. Salfeld, Berlin.
- MACVICAR, C.N. 1973. *Soil map, Republic of South Africa.* 1: 2 500 000. Soil and Irrigation Research Institute, Pretoria.
- MACVICAR, C.N., DE VILLIERS, J.M., LOXTON, R.F., VERSTER, E., LAMBRECHTS, J.J.N., MERRYWEATHER, F.R., LE ROUX, J., VAN ROOYEN, T.H. & HARMSE, H.J.VON M. 1977. *Soil classification: A binomial system for South Africa.* Department of Agricultural Technical Services, Pretoria.
- MACVICAR, C.N., LAMBRECHTS, J.J.N., BENNIE, A.T.P., BRUCE, R.W., DE VILLIERS, J.M., DOHSE, T.E., ELLIS, F., ELOFF, J.F., FEY, M.V., GREY, D.C., HARMSE, H.J.VON M., HARTMANN, M.O., HENSLEY, M., IDEMA, S.W.J., LAKER, M.C., LOXTON, R.F., MERRYWEATHER, F.R., MEYER, J.H., MICHAEL, D., PATERSON, D.G., SCHLOMS, B.H.A., SCHOEMAN, J.L., SCHONAU, A.P.G., SCOTNEY, D.M.,



- SNYMAN, K., TURNER, D.P., VAN NIEKERK, B.J., VAN ROOYEN, T.H.,  
VERSTER, E. & YAGER, T.U. 1991. Soil classification: A  
taxonomic system for South Africa. *Memoirs on the Agricultural  
Natural Resources of South Africa* 15: 1-257.
- MARLOTH, R. 1908. *Das Kapland, insonderheit der Reich der  
Kapflora, das Waldgebiet und die Karoo pflanzengeografisch  
dargestellt*. Fischer, Jena.
- MEESTER, J. 1965. The origins of the southern African mammal  
fauna. *Zoologica Africana* 1: 87-93.
- POLE EVANS, I.B. 1922. The main botanical regions of South  
Africa. *Memoirs of the Botanical Survey of South Africa*  
4: 49-53.
- POLE EVANS, I.B. 1936. A vegetation map of South Africa.  
*Memoirs of the Botanical Survey of South Africa* 15: 1-23.
- REHMANN, A. 1880. Geo-botanische Verhältnisse von Süd-Afrika.  
*Botanische Centralblad* 1: 1119-1128.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1984. Sectors of the  
Transvaal Province of South Africa. *Bothalia* 15: 294-295.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1986. Biomes of southern  
Africa - An objective categorization. *Memoirs of the  
Botanical Survey of South Africa* 54: 1-98.
- SCHÖNLAND, S. 1922. South African botanical literature. *Memoirs  
of the Botanical Survey of South Africa* 4: 69-85.
- SCHOUW, J.F. 1823. *Grundzüge einer allgemeinen Pflanzen-  
geographie und pflanzengeographischer Atlas*. Reimer,  
Berlin.



- SCHULZE, B.R. 1947. The climates of South Africa according to the classifications of Köppen and Thornthwaite. *South African Geographic Journal* 29: 32-42.
- SCHULZE, B.R. 1965. *Climate of South Africa. Part 8. General Survey*. Government Printer and Weather Bureau, Pretoria.
- WERGER, M.J.A. 1973. *Phytosociology of the Upper Orange River Valley, South Africa*. D.Sc. thesis, University of Nijmegen, Nijmegen.
- WERGER, M.J.A. 1978. Biogeographical division of southern Africa. In Werger, M.J.A. (ed.). *Biogeography and ecology of southern Africa*. pp.145-170. Junk, The Hague.
- WESTFALL, R.H. 1981. *The plant ecology of the farm Groothoek, Thabazimbi District*. M.Sc. thesis, University of Pretoria, Pretoria.
- WHITE, F. 1981. *Vegetation map of Africa*. Scale 1:5 000 000. UNESCO, Paris.



### CHAPTER 3. METHODS

In order to achieve the stated aims, the Braun-Blanquet method was viewed from a basic perspective, in terms of the components of the processes of stratification, sampling and classification, rather than to view only selected components. The methods that follow describe, in application sequence, techniques that build on the basic elements of the Braun-Blanquet method, within the flexibility of the Braun-Blanquet method (Coetzee 1975). The techniques were developed and refined over a ten year period and improvements have, therefore, been added to published versions where applicable.

The basic elements of the Braun-Blanquet method are vegetation familiarization leading to a formal or an informal stratification, species presence, sampled in quadrats, cover estimation and classification of species (rows) and samples (columns) in matrix form with cover-abundance values forming the matrix. Accordingly, the two 10 x 10 m plots in the second study area were each divided into 25 contiguous 2 x 2 m sampling units, wherein species presence only were recorded.

The components of the stratification and sampling processes can, for convenience, be grouped according to preparatory work and field sampling.



### 3.1 PREPARATORY WORK

Preparatory work is that work required before field sampling can commence. This includes a clearly stated, well-defined aim which is probably the most important aspect of a study because selection of appropriate methods is dependent on the aim. Results and the interpretation thereof, will be dependent on the methods selected, to achieve the aim. In this study, preparatory work also included determining linear regressions of rainfall data, obtained from the Weather Bureau, and a combination of altitude and latitude, so that interpolations of mean annual rainfall could be estimated for each sampling site within the study area.

#### 3.1.1 Scale

Scale determines the maximum detail that can be obtained with a particular study and is dependent on the aims of the study.

##### 3.1.1.1 Background

The plant community is a "group of plants sharing a common environment and is distinguished by a particular floristic composition" (Mueller-Dombois & Ellenberg 1974) at any scale smaller than that of the individual plant.

It is generally agreed that plant communities form hierarchies (Poore 1956; Mueller-Dombois & Ellenberg 1974) in that smaller plant communities can be included in larger plant communities.





However, this inclusion is often partial rather than whole. Du Reitz (in Poore 1956) states that "a stand may have parts of other ecological groups or vegetation units". This can often complicate plant community border recognition and the level of heterogeneity to include in a plant community. Which plant community to recognize depends on scale. Scale, in turn, is determined by the purpose of the study. In other words, the amount of detail required in a study determines the scale. Scale should, therefore, have an explicit role in the stratification, sampling and classification of vegetation.

According to Rutherford & Westfall (1986) the detail required in a study determines the working or associated sampling scale which in turn determines the smallest mappable unit area (SMUA). The SMUA is also related to the practical considerations of cartographic accuracy and user requirements of a linear precision seldom less than a millimetre (Rutherford & Westfall 1986). To these may be added the humidity-related, stretch and shrinkage of printed maps. At a given working scale, the area under study is subdivided by lines separating stratified units each containing a number of SMUA's which can include potential sampling sites (Rutherford & Westfall 1986). Each SMUA at working scale should, according to Rutherford & Westfall (1986):

- i. not be closer than 1 mm to a stratified unit's border; and
- ii. not be closer than 2 mm to another sampling site.

It follows that the width of an unmappable ecotone or transition is scale-dependent, being less than 2 mm on a map at a given work-



ing scale. A unit equal to or broader than 2 mm could, therefore, have community status at the relevant scale. The relationship of working scale to the smallest mappable unit area (SMUA) is derived from the distance from sampling unit area to stratified unit border and is given by:

$$y = e^{(-12,67 - \ln x)}$$

where  $y$  = SMUA in  $m^2$ , and  $x$  = scale, as a representative fraction. In their account, Rutherford & Westfall (1986) give examples of MSDB (minimum sampling distance from a stratified unit border) and SCD (shortest cross distance of a mappable unit). Because working scale corresponds with detail, any map enlargement beyond the working scale will not improve the detail.

#### 3.1.1.2 *Methods applied*

If the smallest mappable unit area (SMUA) is taken as a circle then the relationship between SMUA and working scale can be simplified, as follows:

$$r = x/1\ 000$$

where  $r$  = radius of SMUA in metres, and  $x$  = denominator of the scale as a representative fraction. For example, the radius, in metres of the SMUA for a working scale of 1:50 000, is 50 m. This is also the minimum distance on the ground between a stratified unit border and the SMUA, at this working scale. The radius of the SMUA and minimum distance between SMUA and stratified unit border is, therefore, always 1 mm on a map at any working scale.



The scale used in this study is 1: 250 000 which corresponds with the detail required for regional planning by the Department of Agricultural Development (see section 4.1.1).

### 3.1.2 Stand area

The New Collins Concise Dictionary of the English Language (1985) defines a stand as "a growth of plants in a particular area, especially trees in a forest or a crop in a field". The effect of taking scale into account in sampling at 1:250 000 scale (small scale sampling) is that it becomes increasingly more difficult to recognize vegetation stands because of increasing heterogeneity with decreasing scale.

#### 3.1.2.1 *Background*

Gabriel & Talbot (1984) define a stand as follows: "A concrete (vs. abstract) aggregation of plants of more or less similar uniformity in physiognomy, species composition, spatial arrangement and condition to distinguish it from adjacent communities. Concrete stands which we sample or measure are aggregated into abstract communities and communities are further abstracted into a general vegetation". Werger (1974) states that "Stands for sampling should be selected in such a manner that each is representative of the vegetation of which it is part and that each plant sampled therein should yield a more or less typical description of that vegetation in terms of both floristic composition and structure". Werger (1974) further states that "Stands which are obviously heteroge-



neous in habitat, structure or floristic composition, and might, therefore, logically be expected to represent parts of two or more units or associations, should be avoided because they do not contribute information which can be used to describe the two or more community types that they represent". From these definitions it can be inferred that a stand is generally a relatively homogeneous unit of vegetation, smaller than the community of which it is a part, but not smaller than the sampling plot with which it is sampled. Furthermore, a stand should be representative of a community and not of an ecotone.

#### 3.1.2.2 *Methods applied*

Stand area is not explicit in the stand definitions given in the previous section but a stand should be sufficiently large to be representative of the community it is to represent. On the other hand, a stand should not be smaller than the smallest mappable unit area (SMUA) because of scale limitations, where it is intended to map communities.

Werger (1972) says that a vegetation stand can fully manifest itself, in terms of species composition and structure, in an area of about one half to one hectare but recommends a much smaller plot area. However, if stand area were to be equated with the SMUA then the following would apply:

- 1: 5 000 scale represents a stand area of 78 m<sup>2</sup>;
- 1: 8 000 scale represents a stand area of 201 m<sup>2</sup>;
- 1: 10 000 scale represents a stand area of 314 m<sup>2</sup>;



1: 20 000 scale represents a stand area of 1257 m<sup>2</sup>;

1: 50 000 scale represents a stand area of 7854 m<sup>2</sup>; and

1: 250 000 scale represents a stand area of 20 ha.

Equating stand area with the SMUA permits scale to be taken into account when sampling vegetation, providing that the sampling plot is then representative of the stand. Furthermore, the criteria defined in section 3.1.2.1 can still be applicable.

The stand is, therefore, defined as a circular area within a stratified unit with radius equal to the denominator of the scale, as a representative fraction. Thus, the minimum area for a community is a function of scale and stand heterogeneity is likely to increase with decreasing scale. The stand radius for the main study area is 250 m, representing a stand area of about 20 ha.

### 3.1.3 Reconnaissance

A reconnaissance is the preliminary inspection of the study area prior to sampling and can have various objectives.

#### 3.1.3.1 Background

According to Werger (1974) "The area of investigation must be well known in all its variety before the study is started". This viewpoint is supported by Mueller-Dombois & Ellenberg (1974) who go even further in stating that "Once the entitation or subdivisioning of the vegetation cover has been clarified, the communities are essentially established". In other words, tentative communities



should be mapped prior to field sampling. The principle of separate mapping and classification processes, in which the classification confirms the mapping units, has been adhered to in this study. However, what is required for the inexperienced researcher are more explicit details of the actions necessary during reconnaissance. Although plant empathy (Tinley 1977) can give added perspective, systematically recorded observations can only increase the scientific value of a reconnaissance.

#### 3.1.3.2 *Methods applied*

The main purpose of reconnoitring the study area is terrain familiarization. This entails travelling throughout the study area to estimate floristic and environmental variation, familiarization with the flora and to obtain permission from landowners for later work and camping sites, if required. Routes followed during the reconnaissance should correspond to the main environmental gradients such as rainfall, topography, geology and soil, where possible, using available roads, in order to detect the range of floristic variation over these gradients. Data recorded for the main study area include:

- i. estimation of the number of structural/floristic units as an indication of the variation in plant communities in the study area. This is required for stratification and need not be precise, but an overestimate is preferable to an underestimate for colour allocation (section 3.1.4). The number of structural/floristic units, can be estimated by counting the different structural/floristic units which intercept the



reconnaissance route. The structural component is according to Edwards (1983) and the floristic component is the dominant or co-dominant species of each structural stratum. Borders of these units can be indicated on a route map, where they intercept the route for stratification validation, if necessary. Scale can be taken into account by only recognizing those structural/floristic units that exceed 4 times the stand radius on the route. This permits a minimum of one stand, with a buffer zone equal to the stand radius, on either side of the stand, to be included in the structural/floristic unit (see section 3.1.4). Aerial photographs could also be used for these estimations;

- ii. estimation of species richness, in terms of species per  $m^2$ , in each structural/floristic unit. This is required for determination of sampling unit area (section 3.1.6). This is done by counting the number of plant species in an area commensurate with plant height and spacing such that any area in the structural/floristic unit would result in a similar count;
- iii. voucher specimen collection of dominant and other common plant species. These are the plants most likely to be encountered during field sampling. The ability to identify these plant species effectively in the field reduces the time required for field sampling (see section 3.2.2); and,
- iv. photographic records and notes on any features commensurate with the aims of the study can also be taken.



### 3.1.4 Stratification

Stratification is the process of preliminary vegetation mapping prior to vegetation sampling and classification. Generally, the sampling intensity in a community, using the Braun-Blanquet approach, is far too low for the community borders to be defined according to the sampling units, without some form of stratification.

#### 3.1.4.1 Background

It can be argued that to map plant communities using the Braun-Blanquet approach is not necessary because each plant community can be recognized in the field using the community diagnosis. This, however, ignores the spatial relationships between communities, is difficult for those not familiar with the techniques to apply, and some sort of vegetation unit map is required, in any case, for sampling the vegetation.

In stratified sampling the vegetation is divided into preliminary vegetation units to which are allocated specific portions of the total sample. Such sampling can actually improve the accuracy of estimates since it is ensured that the various vegetation units that make up the total vegetation are represented in the sample (Freund & Williams 1958).

Such divisions prior to sampling also reduce the problem of redundancy at vegetation unit borders (Werger 1974), and ensure an even





distribution of sampling units in each vegetation unit, irrespective of vegetation unit area.

Stratification for random or systematic sampling can be based on aerial photo interpretation of a study area, as is often the case, as well as mappable non-vegetation factors, such as geology, geomorphology and soil types. Visual integration of different factors, to detect discontinuities in vegetation, is subject to the observer's experience, for repeatability. Furthermore, correspondence between vegetation and non-vegetation (environmental) factors can usually only be shown after the classification process, although correspondence may be implicit, which can limit the usefulness of aerial photo interpretation for stratification purposes. Non-vegetation factors can only be reliably used for stratification, where their limits correspond with those of the vegetation units.

Stratification of vegetation can also be based on satellite imagery. The advantages of satellite imagery over aerial photography include improved geographic fidelity, data in a digitized form, suitable for processing, and suitability for small scale work. The following article describes the methods applied in this study.

# A method for vegetation stratification using scale-related, vegetation-enhanced satellite imagery

R. H. WESTFALL\* and O. G. MALAN\*\*

**Keywords:** Landsat, Munsell parameters, South Africa, stratification, Transvaal, vegetation

## ABSTRACT

A method for visual vegetation stratification and pattern refinement, using scale-related, vegetation-enhanced satellite imagery, is described. The method simplifies colour assignment, facilitates accurate vegetation mapping and could lead to balanced floristic classifications.

## UITTREKSEL

'n Metode vir visuele plantegroei-stratifikasie en patroon verbetering wat van skaalverwante, plantegroei-versterkte satellietbeelde gebruik maak, word beskryf. Die metode vereenvoudig kleurtoekenning en vergemaklik akkurate plantegroei-kartering en kan tot meer gebalanseerde floristiese klassifikasies lei.

## INTRODUCTION

In vegetation sampling (Werger 1974) the subjective selection of sample sites, based on vegetation homogeneity, can only be done effectively by an operator with considerable experience. In addition, there is often a lack of repeatability in the methods and a tendency to ignore vegetation dynamics by only sampling those areas representative of 'good' vegetation.

Stratified random sampling overcomes these problems and increases sampling efficiency by ensuring adequate representation of subdivisions (Elliott 1983). Furthermore, in contrast to random or systematic sampling, the heterogeneity of vegetation, in terms of possible number of communities, can be related to the number of stratified units. Stratification also facilitates the avoidance of transitions which generally do not contribute more information than the adjacent communities (Werger 1974).

Stratification of vegetation prior to floristic sampling entails primarily the categorization of vegetation according to structural characteristics. The categories can be further refined according to the factors apparently responsible for differentiating the strata. The most important factor is usually taken to be topography but others such as geology, pedology, climate or combinations of the four may also be decisive.

Problems encountered with vegetation stratification when using aerial photographs include radial distortions, altitude-related scale differences and often inconvenient scales, which do not facilitate precise vegetation mapping. The excessive detail present in aerial photographs can be potentially confusing and time consuming for stratification, especially

for small-scale work. The use of small-scale, almost orthographic satellite imagery for stratification overcomes these problems but introduces the problems of pattern interference by factors such as soil, and colour assignment where patterns are formed by unfamiliar colours and textures. This paper describes a method for vegetation stratification, using satellite imagery that can overcome the problems of pattern interference and colour assignment and facilitates later pattern refinement.

## CONVENTIONAL USE OF LANDSAT DATA IN VEGETATION STRATIFICATION

### *False colour images*

The Landsat multi-spectral scanner (MSS) records radiance from the earth's surface in four spectral bands: 500-600, 600-700, 700-800 and 800-1100 nm, usually referred to respectively as bands 4, 5, 6 and 7. These data are obtainable in image or digital form. The ground resolution of the system is nominally  $79 \times 56$  m which corresponds to a picture element of about  $0,30 \times 0,22$  mm at a scale of 1:250 000.

The MSS collects six lines of data simultaneously using a set of six detectors for each spectral interval. The mismatch in these sets is one of the major causes of noise in MSS data, which is apparent as striping with a periodic cycle of six lines at extreme radiometric enhancements.

Traditionally Landsat MSS data have been used in mapping at scales between 1:1 000 000 and 1:250 000 (under special circumstances up to 1:100 000 or even 1:50 000 scale) in the form of false colour images, i.e. bands 4, 5 and 7 displayed respectively as the colour primaries, blue, green and red.

This type of display suffers from two disadvantages: firstly only three of the four bands can be displayed, with a possible loss of crucial information in band 6. Secondly, because of the high degree of correlation between bands (Table 1), a very limited re-

\* Botanical Research Institute, Department of Agriculture and Water Supply, Private Bag X101, Pretoria 0001.

\*\* National Physical Research Laboratories, CSIR, P.O. Box 395, Pretoria 0001.

gion of the available three-dimensional colour space is utilized: live vegetation appears exclusively in various shades of red, depending on its structure and vigour.

TABLE 1. — Correlation matrix of the radiance values of the Landsat MSS data, in the four Landsat MSS bands, for the study area

Band	4	5	6	7
4	1,0			
5	0,88	1,00		
6	0,71	0,62	1,00	
7	0,59	0,47	0,94	1,00

### Digital multispectral classification

Digital multispectral classification methods have been used successfully in crop mapping. The methods have proved to be a problem in the stratification of natural vegetation, particularly under southern African conditions. The reasons are mainly: rugged topography (crops are normally grown on level fields), heterogeneity of stands with the consequent problems of selecting 'typical' training sites of sufficient size (i.e. several hectares in size) for the extraction of spectral signatures and the interference of soil reflectance caused by incomplete canopy cover.

Furthermore, this method relies exclusively on spectral characteristics and cannot make use of the contextual or contextual information consciously or subconsciously available to the human interpreter.

### DIGITAL ENHANCEMENT OF LANDSAT DATA

In the alternative approach of enhancement of the Landsat MSS data by digital image processing followed by visual interpretation, the versatility of the human analyst is assisted by means of quantitative enhancement of vegetation differences in stratification but particularly in pattern refinement.

### Principal Component Analysis

Principal component analysis (PCA) provides a convenient method of data compression and removal of redundant correlation between bands (Lasserre *et al.* 1983). Experience has shown that typically more than 99% of the total variance in the data is retained in the first three principal components (Table 2), the maximum which can be accommodated in any colour display.

TABLE 2. — Variance in terms of proportional eigenvalues for the first four principal components of the Landsat MSS data for the study area

Principal Component	First	Second	Third	Fourth
	0,872	0,106	0,016	0,005

The first component represents shadow-enhanced topography and overall terrain brightness differences, whereas the second and third reflect the spec-

tral differences in surface cover. Although the third component normally contains considerably less variance than the second, this may be crucial information for stratification. The fourth component contains predominantly noise.

For optimum results PCA must be based on the statistics of a (composite) subscene, approximately equally representative of the relevant floristic subdivisions, of the area to be stratified. The three principal components may be displayed in any combination of the colour primaries. However, particularly in regions of large variations in overall brightness, such as in areas of rough topography, this results in multicoloured imagery which is difficult to interpret.

A much more informative product can be obtained by displaying the first three components respectively as the Munsell colour parameters, brightness, hue and saturation.

### Display in the Munsell colour space

In order to generate a practicable colour display, the components representing the Munsell colour parameters must be converted into the colour primaries, blue, green and red by computation, special care being taken that visual hue differences in the display truly reflect numerical differences in the data (Malan & Lamb 1985). The first and third components are first contrast-stretched to about 1–2% of the data in maximum and minimum values. The distribution of the first component is approximately Gaussian (Figure 1) while a histogram equalization stretch is applied to the second component (Figure 2). These stretch lookup tables must be based on the statistics of the representative subscene used for PCA.

In the final product the overall impression of terrain brightness of the original image is retained, thus facilitating registration with overlays and later pattern refinement. The effect of the histogram equalization stretch of the hue component is to spread the spectral differences of the cover over the complete hue gamut (as modified in saturation by the third component) as opposed to the limited range of hues in the conventional false colour representation (compare Figures 3 & 4).

### Filtering

When vegetation mapping is to be done at a scale where the shortest cross distance of a mappable unit is greater than a ground resolution of  $79 \times 56$  m, the original resolution of the Landsat MSS data could lead to a product with too much small detail. Prior smoothing of the data produces a scale-related product which facilitates stratification and pattern refinement.

The minimum colour-stratified area is given by  $12,56 n \text{ mm}^2$  with a shortest cross distance of 4 mm (Rutherford & Westfall 1986). The value of  $n$  is determined by the minimum number of samples required for possible floristic subdivision of a colour-stratified area which is dependent on scale as related to sample area and spacing as well as the area of the colour-stratified unit. For example, with  $n=4$ , full resolution Landsat MSS data could therefore theo-

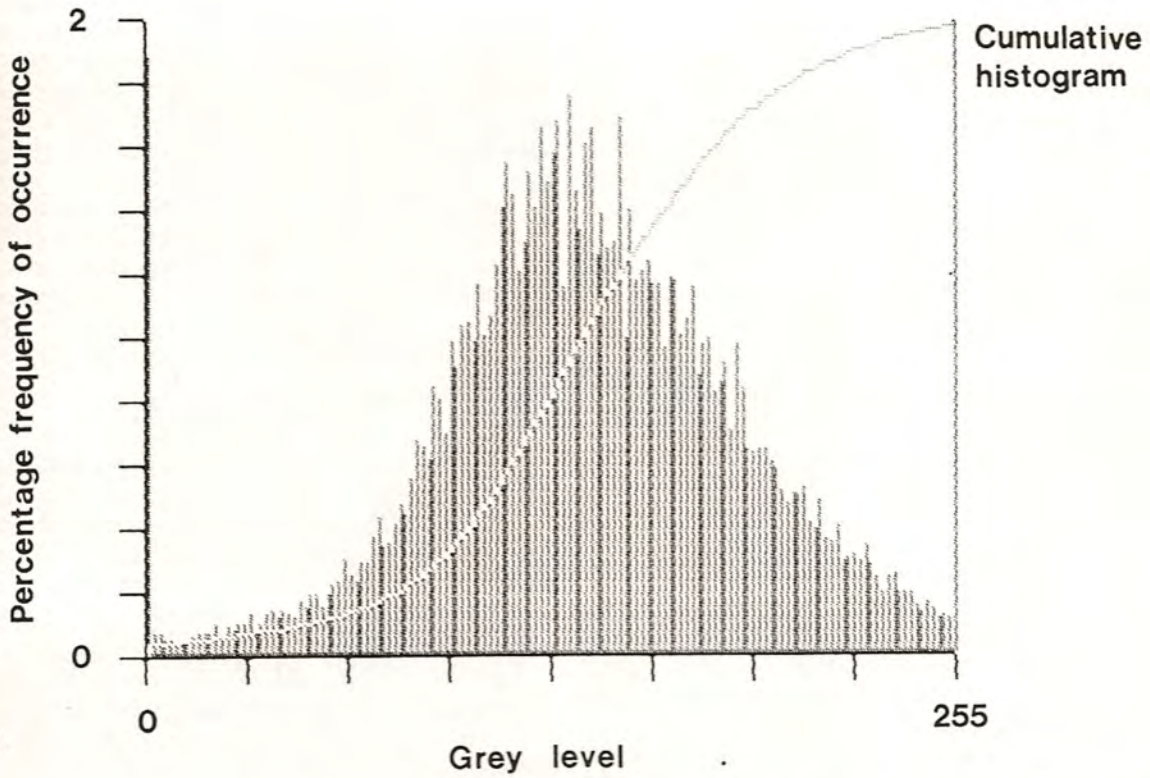


FIGURE 1. — Histogram showing the approximately Gaussian distribution of the first principal component, displayed as the Munsell brightness parameter.

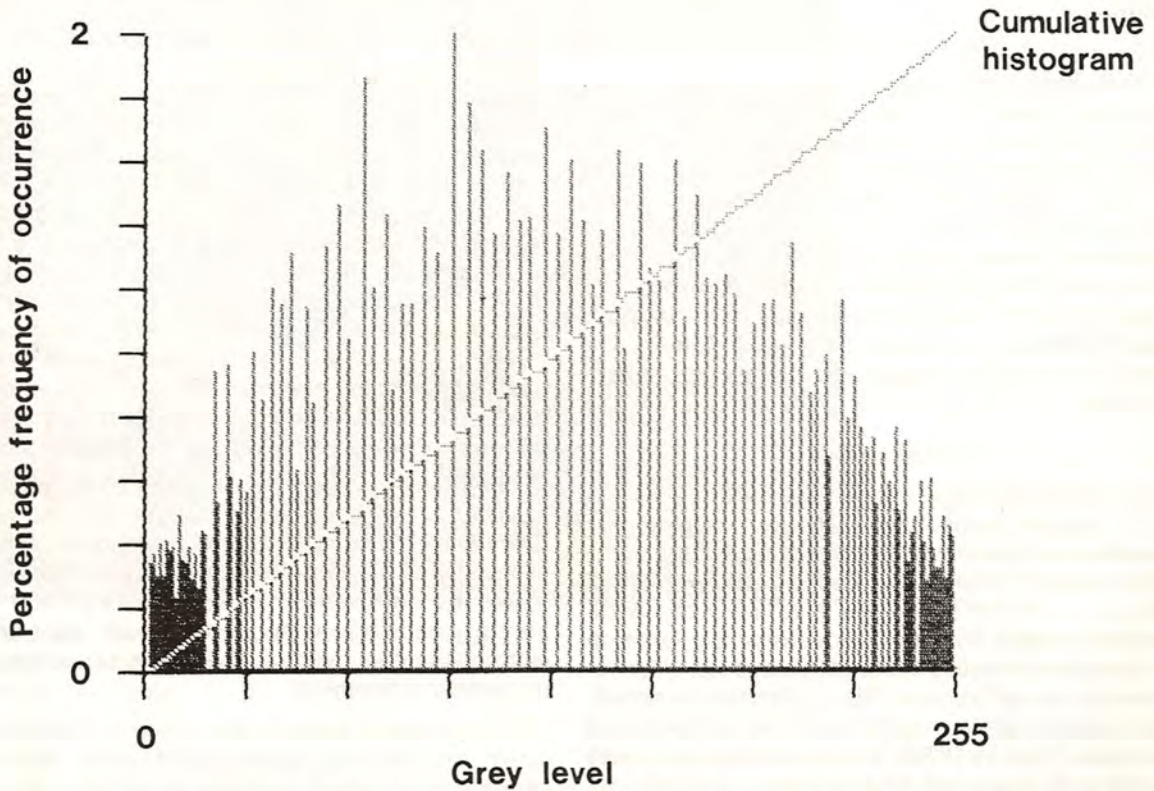


FIGURE 2. — Equalized histogram of the second principal component, displayed as the Munsell hue parameter.



retically be used for vegetation stratification up to a scale of 1:10 000. However, PCA dramatically enhances noise in the data in the higher principal components, particularly the striping in Landsat MSS data. The application of a median filter with a kernel at least six lines wide smoothes the data and removes striping effectively. Consequently, in practice, a median filter with a minimum kernel size of  $6 \times 9$  picture elements corresponding to a square ground resolution of about 24 ha, is used. This corresponds to a maximum useful working scale of almost 1:50 000.

The (first) brightness component is excluded from the filtering process, to retain fine detail in the representation of topography, which assists in pattern refinement as well as in the identification of ground control points for registration with map overlays.

#### PATTERN REFINEMENT

Pattern refinement refers to the process of modifying the stratified units, usually after sampling and classification of the vegetation and prior to mapping floristic units. This process includes the grouping together of similar, smaller, discreet areas and subdivision of larger, uniform colour-stratified areas by contextual comparison if required, with suitable, simplified topographical, geological, pedological or meteorological overlays and sample-set classification at the given working scale. The choice of any or all of these overlays is determined by the range of variation exhibited by the overlays that can relate to the structural variation. For example, smaller units of differing colour can be grouped together on the basis of floristics and soils or topography while larger, uniform units could be subdivided on the basis of floristics and geology or climate in order to refine the patterns.

Minor inaccuracies can occur in the registration of Landsat hard copy images and other maps at the same scale used for overlays which are attributed to differential stretch and shrinkage caused by fluctuating humidity as well as some obvious inaccuracies in the maps. These errors can be compensated for by shifting local fit to achieve the maximum number of registration points rather than be compounded by maintaining fixed registration points. However, the effect of stretch and shrinkage can be further reduced by the use of dimensionally stable transparencies of both satellite images and other maps, where available.

#### RESULTS

The results of colour-stratification using scale-related vegetation-enhanced satellite imagery are given for a portion of the Transvaal Waterberg in the north-western Transvaal (Rutherford & Westfall 1984) at 1:250 000 scale. The area has a highly diverse topography with a consequently high variation in vegetation structure which is inferred from the variation in colour pattern. The vegetation is mainly representative of Sour and Sourish Mixed Bushveld veld types (Acocks 1975). Structural heterogeneities related to topographic diversity were indicative of potentially small stratification units and the value of  $n=4$  was accordingly assigned. This corresponds to a

map area of 50 mm<sup>2</sup> which is equivalent to 320 ha. The appropriate filter kernel size was therefore  $32 \times 22$  picture elements. A contrast-stretched false-colour MSS image (acquisition date March 1981) is shown in Figure 3. A vegetation-enhanced image, at the original resolution of the same scene, is shown in Figure 4, which illustrates the complexities of colour assignment exacerbated by noise, particularly six line striping. A scale-related, vegetation-enhanced image at the hard copy scale of 1:250 000 of the same scene is shown in Figure 5 with resultant simplification of colour pattern to form stratified units.

#### INTERPRETATION

A critical prerequisite for successful stratification is the choice of the optimum data acquisition date for maximum differentiation between the structural subdivisions. Because the four wettest months in the study area are from November to February, the data acquisition date of March ensured high vegetation cover with minimal cloud interference.

The inputs required from the user are location of study area, working scale and data acquisition dates of scenes required. Primary colour-stratification is automated after these inputs and is, therefore, objective and time and labour saving. Many researchers use topographic, geologic, pedologic or meteorologic maps for comparison with floristic units or simply to show the variation in these factors. Unless working scale is standardized at the outset of a project, comparisons are difficult. Overlays at a standardized scale are used for pattern refinement and can also be used for later comparisons with floristic units in addition to the enhanced Landsat MSS image. They can, therefore, serve dual purposes and their effective use is increased.

Training sites would often be required for unsimplified images (Westfall & Malan in press), whereas the use of scale-simplified images largely overcomes the need for this training. Furthermore, scale-related stratified units should improve floristic classifications by providing a balanced distribution of sample sites commensurate with vegetation heterogeneity and the amount of detail required for a given working scale. It should be pointed out that, like PCA, the colour stratification by vegetation enhancement is highly scene-dependent. In practice, however, this is not a great disadvantage because one Landsat scene covers 34 000 km<sup>2</sup>. Also multiples of this size can be treated identically if they are contiguous images on one north-south Landsat swath.

The use of satellite images, which have better geometric fidelity than aerial photographs, also simplifies the process of accurate vegetation mapping, especially where first-order stereo-restoration instruments are not available. This ensures greater mapping precision when compared to base maps and facilitates the effective use of overlays for comparison or pattern refinement.

The proposed methods also ensure objectivity because the primary colour-stratification process is computerized, which produces repeatable stratification units. It is doubtful that subjectively stratified units could be repeated by different workers. This

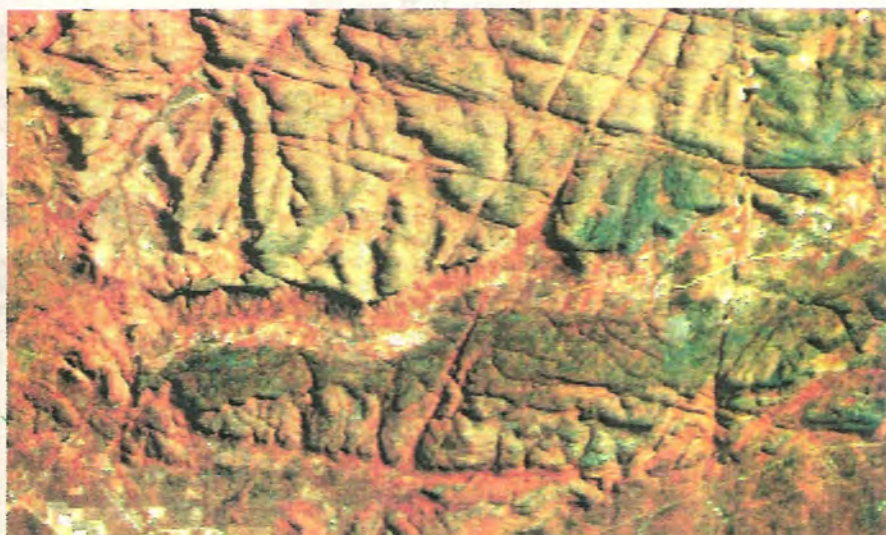


FIGURE 3. — A contrast-stretched, false-colour, multispectral-scanner (MSS) image of a portion of the Transvaal Waterberg with limited range of hues. Scale 1:250 000.

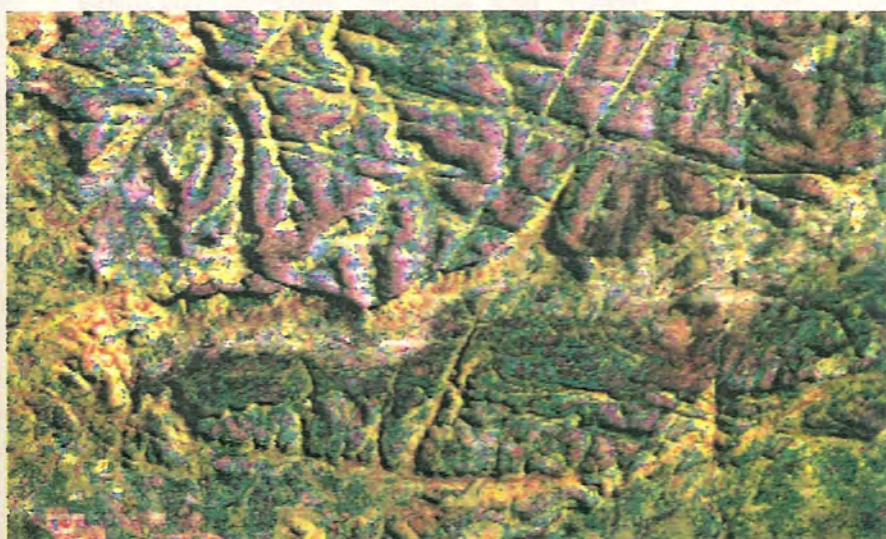


FIGURE 4. — A vegetation-enhanced image at the original resolution of the same scene as shown in Figure 1 using the complete hue gamut.

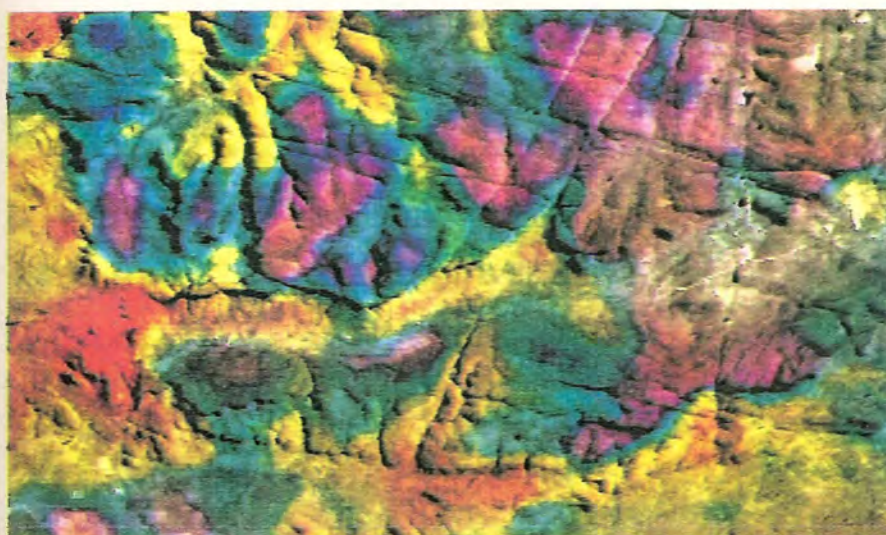


FIGURE 5. — A scale-related vegetation-enhanced image of the same scene as shown in Figure 2 showing simplification of colour pattern to form primary stratified units.

could affect the balance of the resultant floristic classifications and could be especially significant in comparisons over time.

#### CONCLUSIONS

The proposed methods of vegetation stratification prior to floristic sampling are objective, and time- and labour-saving. The process of colour assignment

to stratified units is simplified. The stratified units are related to working scale and the structural heterogeneity present in the vegetation. Overlay comparison and accurate vegetation mapping are facilitated and more balanced floristic classifications can be expected. Landsat data are generally more detailed than required for vegetation stratification at scales smaller than 1:50 000 and hence filtering is necessary.



ACKNOWLEDGEMENTS

The authors thank Dr J. C. Scheepers for comments and suggestions.

REFERENCES

ACOCKS, J. P. H. 1975. Veld types of South Africa. *Memoirs of the Botanical Survey of South Africa* No. 40.

ELLIOTT, J. M. 1983. *Some methods for the statistical analysis of samples of benthic invertebrates*. Kendall, Wilson.

LASSERRE, M., MALAN, O. G. & TURNER, B. 1983. *The application of principal component analysis to Landsat MSS data*. Proceedings of seminar on Principal component analysis in the atmospheric and earth sciences, Pretoria 7-8 February 1983.

MALAN, O. G. & LAMB, A. D. 1985. *Display of digital image data; quantitative and optimal*. Proceedings of the third

South African symposium on Digital image processing, Durban 22-23 July 1985.

RUTHERFORD, M. C. & WESTFALL, R. H. 1984. Sectors of the Transvaal Province of South Africa. *Bothalia* 15: 294-295.

RUTHERFORD, M. C. & WESTFALL, R. H. 1986. Biomes of southern Africa — an objective categorization. *Memoirs of the Botanical Survey of South Africa* No. 54.

WERGER, M. J. A. 1974. On concepts and techniques applied in the Zürich-Montpellier method of vegetation survey. *Bothalia* 11: 309-323.

WESTFALL, R. H. & MALAN, O. G. in press. *A comparison of vegetation units derived from vegetation-enhanced satellite imagery with the vegetation units derived from the floristic classification of the farm Groothoek, Thabazimbi District*. Proceedings of the symposium on Pattern recognition in remote sensing and geophysics.



### 3.1.5 Stand location

The main disadvantage of stratified systematic sampling is the possibility of hidden periodicities or recurring phenomena in the data, which could influence results (Freund & Williams 1958) so stratified random sampling is preferred. In stratified random sampling the location of stands within each stratified vegetation unit is random, to reduce observer bias. Randomization of stand localities does not necessarily infer statistical validity on the sampling process because the number of repetitions in vegetation sampling within a vegetation unit is generally too few for statistical tests that are based on normal distributions.

#### 3.1.5.1 *Background*

Randomization of stand localities is generally achieved by a grid co-ordinate system in which co-ordinates are selected using random tables or a random number generator. Whatever system of co-ordinate selection is used, the appropriate minimum spacing between stand sites, and hence grid size, should be known, because Werger (1974) states that contiguous stands are not to be recommended as much unnecessary work would be done. Apart from non-contiguous sampling, the literature does not appear to address the problem of minimum spacing of stands.

#### 3.1.5.2 *Methods applied*

From the geometric considerations given in section 3.1.1, a grid



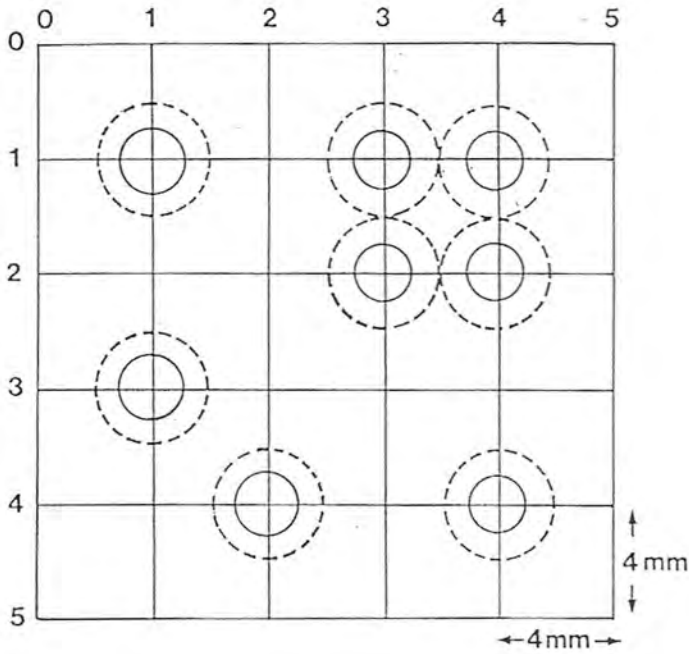


FIGURE 3.1. - A 4 x 4 mm numbered grid for map overlay to determine possible stand sites at the grid intersections. Solid circles represent stands; and dotted circles represent buffer zones around stands. Buffer zone radius = 2x stand radius = maximum ecotone width = minimum distance between stands (from Rutherford & Westfall 1986).



size of 4 x 4 mm should be used, at any given working scale, as a map overlay to locate stand sites from random co-ordinates. This ensures a minimum spacing of one stand diameter between stands. Stand area in relation to the grid is illustrated in Figure 3.1. The use of a standard grid, for any working scale, can simplify computer generation of random numbers for selecting grid co-ordinates and their conversion to the co-ordinates of degrees, minutes and seconds, to facilitate stand location in the field. In the field, if a randomly selected stand includes obvious heterogeneity such as disturbance not commensurate with the working scale, then it is ignored and the following stand is selected. Stands can also be ignored in the field if they are inaccessible. Furthermore, stands may not be used if the circumference is less than one stand diameter from a stratified unit border. It is for these reasons that more stand localities are generated than are actually required. Observer bias is, therefore, reduced by decision-making regarding stand location, although some decisions may be necessary regarding stand suitability. The following article describes a program for the generation of random stand co-ordinates to be used with a 4 mm grid and grid co-ordinate conversion to degrees, minutes and seconds for field location of stands.



## PHYTOLOC — A RANDOM-NUMBER GENERATOR AND SAMPLE-SET LOCATION PROGRAM FOR STRATIFIED RANDOM VEGETATION SAMPLING

The stratified random method of vegetation sampling is objective and efficient, in terms of sample-set distributions, for floristic classifications (Westfall & Malan 1986). However, the commonly used random number tables and calculator-generated random numbers often require number abbreviation and manual recording of the numbers, which can be time-consuming. But these inconveniences are insignificant when compared with the time taken to measure the location of random sample sets and express their location in terms of the latitude and longitude co-ordinates of degrees, minutes and seconds. The PHYTOLOC program was developed to generate random numbers for sample set location in terms of random co-ordinates and, in addition, to express these co-ordinates as latitudes and longitudes in degrees, minutes and seconds, thereby saving considerable time and effort.

The program is written in Basic and runs on a Sharp PC 1500 computer. Use is made of a consecutively numbered 4 mm transparent grid map overlay which is related to any working scale, in terms of sample set spacing and size (Rutherford & Westfall 1986). Grid overlay registration with a base map is according to the zero co-ordinates of the overlay with the intersection of minimum latitude and longitude of the study area on the base map, as well as with the zero x-axis of the overlay with the minimum latitude of the base map. Inputs required for the program are:

- (i) maximum latitude of the study area in decimal degrees;
- (ii) minimum latitude of the study area in decimal degrees;
- (iii) minimum longitude of the study area in decimal degrees;
- (iv) difference in millimetres, between minimum and maximum latitudes, at the given working scale;
- (v) mean distance in millimetres, between 1 minute longitudes at the minimum latitude;
- (vi) mean distance in millimetres, between 1 minute longitudes at the maximum latitude;
- (vii) number of sample sets required, estimated by  $10 \text{ SU} + (0,25 \times 10 \text{ SU})$  where SU is the number of stratified units. This should generally allow for omissions due to transitions and proportionality;
- (viii) upper limit (integer), within the study area, of the x-axis of the grid overlay, and
- (ix) upper limit (integer), within the study area, of the y-axis of the grid overlay.

The program generates and prints random numbers for the x- and y-axes of the grid overlay and computes and prints the equivalent values in degrees, minutes, seconds and decimal fractions of seconds for longitude and latitude respectively. Convergence of longitude is also taken into account. Each set of co-ordinates, representing a potential sample site, is numbered consecutively. In addition the means and standard deviations of the x- and y-arrays are computed to show the statistical distribution of potential sample sites. Co-ordinates are transferred to the base map using the grid overlay and the print-out can be used for field allocation of latitude and longitude to the field data sheets. These co-ordinates

can also be used on larger-scale maps for more precise field location of sample sites.

For valid categorization and analysis of floristic units, based on multivariate data, a minimum of four sample sets are required, although a single sample set is mappable at the given working scale. Consequently for a single floristic division of a stratified unit a minimum of eight sample sets would be required. However, for statistical comparisons of univariate data such as biomass or number of taxa, sample-set number should be proportional to area (Elliott 1983). It is, therefore, suggested that a sampling intensity of 2,5% of the potential sampling sites (i.e. total number of co-ordinate interceptions) within a stratified unit should be maintained to ensure proportionality. This is approximately commensurate with the relationship of study area to sample number (Rutherford & Westfall 1986) but modified by vegetation heterogeneity in terms of number of stratified units.

A non-random set of sample sites could, therefore, be required to fulfil the categorization and analysis requirements for stratified units with less than a total of 320 interception points. The additional non-random sample sites can be selected objectively by:

- (i) using best fit of additional sample sets for areas equal to those of eight or less interception points (i.e. 100% sampling intensity), and
- (ii) using additional random sample sets, to ensure representation of vegetation variation, for areas equal to those with between 8 and 320 interception points. (i.e. > 2,5% but < 100% sampling intensity).

Additional random sample sets can be obtained together with the relevant co-ordinates, if required, by the same procedure, but with each relevant stratified unit registered separately on the 4 mm grid.

## ACKNOWLEDGEMENTS

The author thanks Drs J. C. Scheepers and H. van Ark for comments and suggestions.

## REFERENCES

- ELLIOT, J. M. 1983. *Some methods for the statistical analysis of samples of benthic invertebrates*. Kendall, Wilson.
- RUTHERFORD, M. C. & WESTFALL, R. H. 1986. Biomes of southern Africa — an objective categorization. *Memoirs of the Botanical Survey of South Africa* No. 54.
- WESTFALL, R. H. & MALAN, O. G. 1986. A method for vegetation stratification using scale-related, vegetation-enhanced satellite imagery. *Bothalia* 16: 263-268.

R. H. WESTFALL



### 3.1.6 Sampling unit area

The sampling unit is the vegetation sample plot used to sample a vegetation stand and in the classification phase is referred to as a relevè (Mueller-Dombois & Ellenberg 1974). Where more than one plot is used to sample a vegetation stand then the sampling unit is a sub-plot.

#### 3.1.6.1 Background

Werger (1972), in his review of species-area relationships, concludes that no convincing minimal-area for sampling unit size has been formulated and regards the concept unsuitable for ascertaining a suitable sampling unit size for sampling vegetation. He further suggests a sampling unit size much smaller than half a hectare which depends on species richness and structure of stands. This may be an extreme statement but in southern Africa, sampling unit area can vary considerably. For example Coetzee (1975) used 10 x 10 m; Coetzee *et al.* (1976) used 10 x 20 m for herbaceous plants and variable sized units for woody plants; Bredenkamp & Theron (1978) used 16 m<sup>2</sup> for grasslands and 100 m<sup>2</sup> for woodlands; Van der Meulen (1979) used 10 x 20 m; Van Rooyen *et al.* (1981) used 10 x 20 m and variable size for the woody component; Westfall *et al.* (1985) used 10 x 20 m for all plants and Le Roux *et al.* (1988) used 10 x 20 m for herbaceous plants and 100 x 100 m for trees and shrubs.

Such variation in sampling unit area can be expected to complicate



comparisons within and between communities from different studies. Furthermore, difficulty was experienced in selecting a single 10 x 20 m sampling plot to represent a 20 ha stand. A sub-sampling approach was, therefore, adopted to sample stands. It must be emphasized that the sub-samples are based on minimum area for recording species presence only and although the total sample area for a stand is less than 200 m<sup>2</sup> the species number recorded was often more than that which could be expected from a 10 x 20 m plot.

The following article describes the methods applied in this study to determine minimum sub-sample area.

# Predictive species – area relations and determination of subsample size for vegetation sampling in the Transvaal Waterberg

R.H. Westfall\*, J.M. van Staden and M.D. Panagos

Botanical Research Institute, Department of Agriculture and Water Supply, Private Bag X101, Pretoria, 0001 Republic of South Africa

Accepted 20 August 1986

An expression for predicting the number of species in a given area is described. Derivatives of this expression, to increase sampling efficiency in a vegetation stand, include a minimum of four separate subsamples; a maximum number of subsamples when less than 10% increment in new species is achieved; and subsample size. It is also suggested that species diversity in terms of species per unit area can be more consistent when derived from this expression.

'n Uitdrukking om die aantal spesies vir 'n bepaalde gebied te voorspel word beskryf. Afleidings van hierdie uitdrukking, om doeltreffendheid by die monsterneming van 'n plantegroei stand te verbeter, sluit die volgende in: 'n minimum van vier afsonderlike submonsterpersele; 'n maksimum aantal submonsterpersele wanneer die nuwe spesiesaanwas minder as 10% is; en submonsterperseelgrootte. Dit word ook voorgestel dat spesies diversiteit, ooreenkomstig spesies per eenheidoppervlakte, meer konsekwent kan wees wanneer dit van hierdie uitdrukking afgelei word.

**Keywords:** sample size, savanna, species number, Transvaal, vegetation sampling

\*To whom correspondence should be addressed

## Introduction

The vegetation ecology of the Transvaal Waterberg is currently being investigated at a scale of 1:250 000 (Westfall, in prep.). The smallest mappable unit area or vegetation stand, of which a sample should be representative, is determined by scale (Rutherford & Westfall 1986) and is 20 ha for this study. A single sample of this dimension would be prohibitive both in terms of cost and time. Random subsamples within the stand offer an objective method of obtaining a representative sample. The efficiency of the subsamples, in terms of plant species recorded, can be determined if the total number of plant species in the stand can be estimated.

The object of this study is to determine the optimum subsample size, in terms of reduced effort and improved efficiency, for sampling the vegetation of the Transvaal Waterberg.

## Methods

### Vegetation analysis

Two observers, A and B, counted species in eight samples representing four very different vegetation types, namely closed grassland (sample 1); open woodland (sample 2); closed woodland (samples 3 and 4) and unidominant forest (samples 5 and 6). In samples 1 and 2 each observer repeated the countings of the other observer for the same area, each without knowledge of the other's results. In each sample cumulative plant species totals were recorded in 14 nested rectangular subsamples, each with a width-to-length ratio of 1:2 (Figure 1). Subsamples size increased from 62,5 mm × 125 mm (0,0078125 m<sup>2</sup>) to 10 m × 20 m (200 m<sup>2</sup>) and the length of each subsample was perpendicular to the contour.

In addition, both observers jointly counted the observable species present in a 20-ha stand of which sample 2 was representative. In this stand nested circular subsamples were also used with radii of 6,3 m (128 m<sup>2</sup>), 12,7 m (512 m<sup>2</sup>), 25,5 m (2 048 m<sup>2</sup>), 51 m (8 192 m<sup>2</sup>), 102 m (32 768 m<sup>2</sup>), 204 m (131 070 m<sup>2</sup>) and 252 m (200 000 m<sup>2</sup>).

## Synthesis

The data in the form of species number for a given area were tested for best fit with various curves (Parton & Innis 1972) as well as linear relationships. The curve which best fitted the data is described by the function:

$$f(x, a, b, c, d) = \frac{a}{(1 + b/x)} + \frac{c}{(1 + d/x)}$$

where a and c = the parameters which control the maximum value of the function and b and d = the parameters which control the rate at which the function approaches its maximum value (Parton & Innis 1972).

No parameters could be found with actual or derived values similar to the computed values for parameters a, b, c & d. This reduced the application potential of the function and it was accordingly not further applied.

The linear regression which best fitted the data is described by:

$$y = e^{(m \ln x + \ln c)}$$

where y = number of species for a given area (x), m = slope (rate of species increase for increasing area), x = area (m<sup>2</sup>), c = number of species in 1 m<sup>2</sup>.

This linear regression was then used to derive the number of species in 1 m<sup>2</sup> from a given area and to synthesize data for illustrating species – area relationships.

The validity of the relationships was tested by using data pertaining to the first three relevés recorded in the Transvaal Waterberg (Westfall, in prep.) which utilized these relationships.

## Results

The number of species recorded by observers A and B for the nested samples 1 to 6, with predicted values, are given in Table 1 and Figure 2. The linear regressions of these samples are illustrated in Figure 3. With the exception of sample A6, the slopes are generally similar. Sample A6 represents a small stand of unidominant *Podocarpus latifolius*

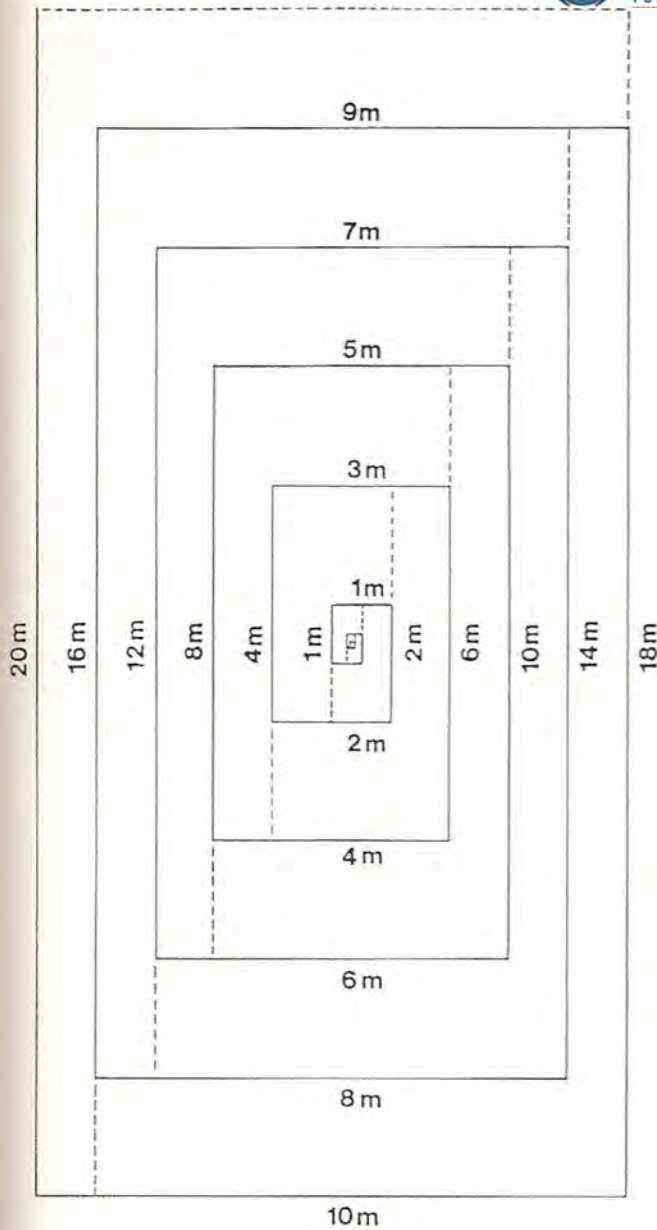


Figure 1 Layout of rectangular nested subsamples. Subsample sizes are: 62,5 mm  $\times$  125 mm; 125 mm  $\times$  250 mm; 250 mm  $\times$  500 mm; 500 mm  $\times$  1000 mm; 1 m  $\times$  2 m; 2 m  $\times$  4 m; 3 m  $\times$  6 m; 4 m  $\times$  8 m; 5 m  $\times$  10 m; 6 m  $\times$  12 m; 7 m  $\times$  14 m; 8 m  $\times$  16 m; 9 m  $\times$  18 m; and 10 m  $\times$  20 m.

(Thunb.) R. Br. Ex Mirb. in which forest-margin species were recorded before the maximum sample size of 200 m<sup>2</sup> was reached and can therefore be considered atypical. The mean slope of the regressions, excluding sample A6, is 0,34 but is taken as 0,3 which was used for calculating the predicted values in Tables 1 and 2. The mean correlation co-efficient for the recorded and predicted values for the seven samples (excluding sample A6) is  $R = 0,9860$ , whereas the mean correlation co-efficient for the corresponding recorded and predicted values using the curve function is  $R = 0,9819$ . The difference between these correlation co-efficients is not significant. The correlation co-efficient for recorded and predicted values in the 20-ha sample is 0,9725 (Table 2).

The value of  $c$  is derived from  $y = e^{(mlnx + lny)}$  and is given by

$$c = e^{(-mlnx + lny)}$$

which is dependent on the species diversity of the vegetation

type concerned. The values of  $c$ , derived from different areas and corresponding derived species number are given in Figure 4.

The derived relationship between number of subsamples and percentage new species increment is given in Figure 5 and the relationship between subsample number and percentage accumulated new species total is given in Figure 6.

The results of the 20-ha sample together with predicted values is given in Table 2.

### Discussion

The prediction of number of species for a given area has been tested for a wide range of vegetation types with satisfactory results. The larger difference between recorded and predicted number of species for areas greater than 2 048 m<sup>2</sup> (Table 2) can be attributed to the difficulty in observing all species in these areas. Destructive sampling would be necessary to ensure the recording of all species. Both observers encountered increasing difficulty in observing all species with increasing sample size where sample size was greater than 20 m<sup>2</sup>. With nested subsamples, however, species omitted in a particular subsample were often recorded in a larger subsample. This is illustrated by the similarity in totals obtained by observers A and B for 200 m<sup>2</sup> and the differences in totals for 18 m<sup>2</sup> to 128 m<sup>2</sup> in samples 1 and 2 (Table 1).

The effect of sampling two highly distinct vegetation types in a small sample is shown in Table 1 (A6) where a greater proportion of new species was recorded from 8 m<sup>2</sup> than in the other samples. This is also illustrated in Figure 3 where the slopes ( $m$ ) of all samples except A6 are similar regardless of vegetation type. Local heterogeneity in the 20-ha sample did not influence the results as much as with the smaller sample. This can be attributed to the logarithmic increase in area used in the prediction expression where slope is not significantly affected by a greater proportion of new species in a large area. It is, therefore, essential for predictive purposes that the vegetation should be relatively homogeneous for a sampling area of at least 200 m<sup>2</sup>.

For a given vegetation type, the value of  $c$  in the expression should be more or less constant when derived from different areas and a horizontal line could, therefore, be expected when these values are plotted against the area from which they are derived. However, in Figure 4 the values of  $c$  increase to an area equivalent to 2 m<sup>2</sup> before remaining more or less constant. This increase can be attributed to the edge effect of small subsample size where the proportion of species intercepted by the subsample border to species within the subsample is greater than for larger subsamples. It is difficult to record fractions of species present so that in practice intercepted species are often ignored. Subsample size should, therefore, be greater than or equal to 2 m<sup>2</sup> in the vegetation concerned to reduce this edge effect. Similarly the value of  $c$  should also be derived from an area greater than or equal to 2 m<sup>2</sup> in the vegetation concerned which will also ensure greater precision in comparability of species diversity in terms of species per unit area.

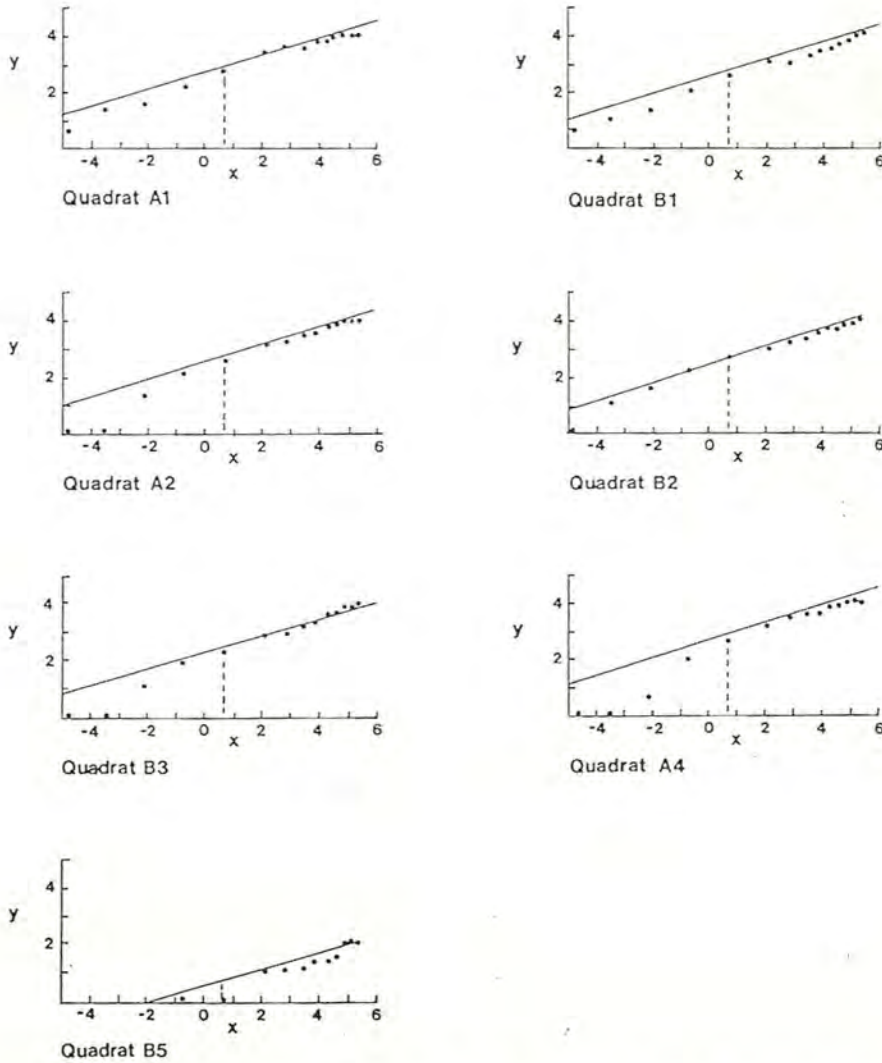
The predicted increment of new species as a percentage of the total number of species for increasing number of subsamples (Figure 5) has maximum inflexion for four subsamples or less or an increment of 8,3%, or greater for given  $c$  values and subsample sizes. This indicates an optimum efficiency of four subsamples for vegetation sampling or an increment of 8,3%. With additional subsamples only relatively small increments decreasing from 6,5% could be expected (Figure 5). In practice this could be taken as an increment

**Table 1** The number of species recorded by observers A and B for nested quadrats 1 to 6 with the predicted values\* in brackets and correlations with observed and predicted values

Area (m <sup>2</sup> )	A1	B1	A2	B2	B3	A4	B5	A6
0,00781	2 (4)	2 (3)	1 (3)	1 (3)	0 (2)	0 (3)	0 (0)	0 (0)
0,03125	4 (5)	3 (4)	1 (4)	3 (4)	0 (3)	1 (5)	0 (1)	0 (1)
0,12500	5 (8)	4 (6)	4 (6)	5 (6)	3 (5)	2 (8)	0 (1)	0 (1)
0,50000	9 (12)	8 (10)	8 (10)	9 (9)	7 (7)	8 (12)	1 (1)	1 (1)
2,0	17 (18)	15 (15)	13 (15)	14 (13)	10 (11)	14 (18)	1 (2)	2 (2)
8,0	28 (28)	22 (22)	22 (22)	20 (20)	17 (17)	27 (27)	3 (3)	3 (3)
18,0	37 (36)	22 (28)	26 (28)	24 (26)	19 (22)	34 (34)	3 (4)	8 (4)
32,0	38 (42)	27 (33)	34 (33)	26 (30)	25 (26)	36 (41)	3 (5)	8 (5)
50,0	44 (49)	34 (38)	38 (38)	35 (35)	28 (29)	40 (47)	4 (5)	12 (5)
72,0	46 (54)	35 (43)	46 (43)	39 (39)	35 (33)	49 (52)	4 (6)	16 (6)
98,0	50 (59)	40 (47)	49 (47)	41 (42)	39 (36)	52 (57)	6 (6)	24 (6)
128,0	53 (64)	48 (51)	52 (51)	43 (46)	47 (39)	57 (62)	8 (7)	26 (7)
162,0	55 (69)	55 (54)	54 (54)	49 (49)	51 (42)	60 (67)	9 (7)	30 (7)
200,0	57 (74)	57 (58)	55 (58)	52 (53)	55 (45)	61 (71)	9 (8)	33 (8)
Slope	1,199	1,022	0,945	1,002	0,791	1,044	0,773	0,207
Correlation co-efficient (R)	0,9891	0,9894	0,9970	0,9971	0,9910	0,9948	0,9434	0,9365
Value of lnC**	2,7085	2,4673	2,4673	2,3720	2,2095	2,6721	0,4749	0,4749

\*Prediction according to:  $y = e^{(m \ln x + \ln C)}$  where  $y$  = predicted number of species,  $m = 0,3$  (slope),  $x$  = area (m<sup>2</sup>)

\*\*C = species number for 1 m<sup>2</sup> derived from 8 m<sup>2</sup>



**Figure 2** The relationship between ln area (m<sup>2</sup>) (x-axis) and ln number of species (y-axis) showing observed numbers (dots) and predicted values (regression lines). Generally, fewer species are observed in quadrats smaller than 2 m<sup>2</sup> (broken lines) than predicted, which is attributed to edge effect. Observer difficulty in recording all species is indicated by the trend of fewer observed species than predicted for the larger quadrat sizes (greater than 20 m<sup>2</sup>).



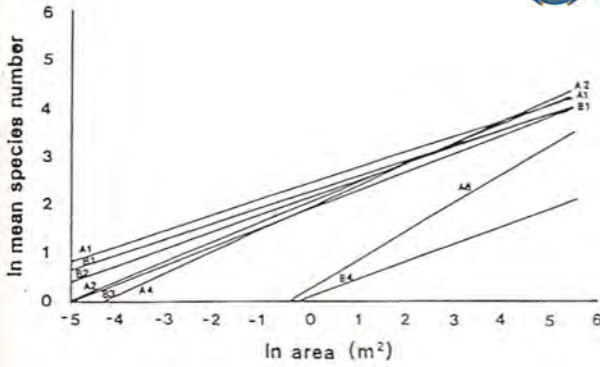


Figure 3 The relationship between ln species number and ln area for eight samples, showing the similarity in slope (*m*). Sample A6 is atypical.

Table 2 The number of species recorded in nested subsamples with predicted values in brackets ( $C = 2,5147$  derived from  $128 \text{ m}^2$ )

Area ( $\text{m}^2$ )	Number of species
128	53 (53)
512	82 (80)
2 048	120 (122)
8 192	165 (184)
32 768	199 (280)
131 072	213 (424)
200 000	274 (481)

Correlation co-efficient (*R*) between recorded and predicted values:  $R = 0,9725$

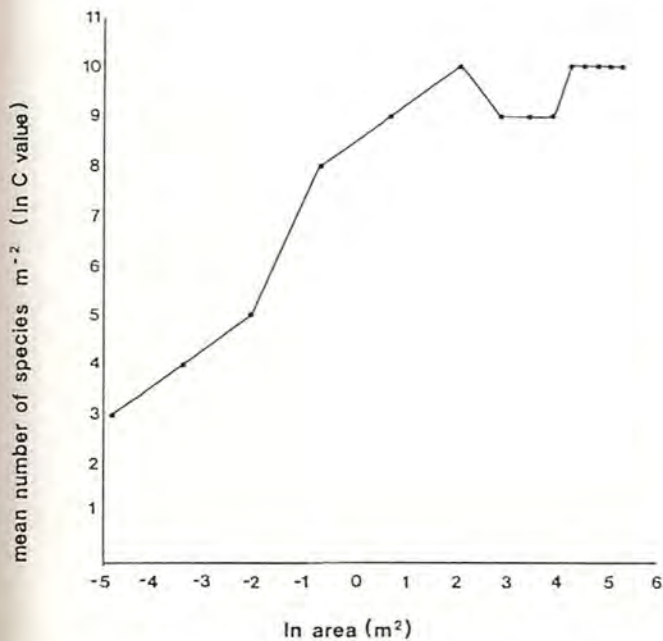


Figure 4 The relationship between ln *c* values, derived from different areas and corresponding ln area showing the increase in *c* values to  $2 \text{ m}^2$  ( $\ln 2 = 0,693$ ) representing sample edge effect.

of less than 10% to allow for local heterogeneity. A minimum of four subsamples is recommended because if an increment of less than 10% is achieved with fewer subsamples local disturbance could be indicated.

Species distribution in a stand varies from closely spaced species (high frequency) to 1 species in the stand (low

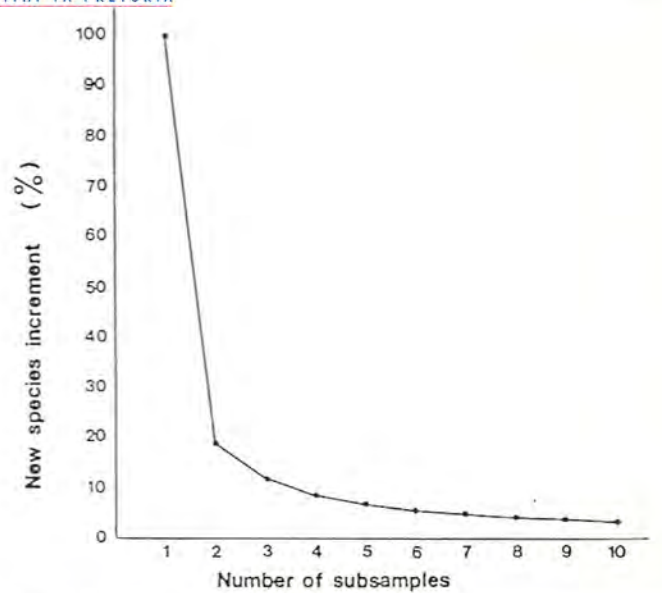


Figure 5 The relationship between number of subsamples and new species increment expressed as a percentage of the accumulated total. Subsample increments: 1 = 100%; 2 = 18,8%; 3 = 11,5%; 4 = 8,3%; 5 = 6,5%; 6 = 5,3%; 7 = 4,5%; 8 = 3,9%; 9 = 3,5% & 10 = 3,1%.

frequency). Too small a subsample size would not include sufficient high frequency species for the relationship in Figure 5 to be valid as each subsample would include only a fraction of the high frequency species resulting in a greater number of new species for successive subsamples. In contrast, too large a subsample size would include too many of the less frequent species with the same result. The percentage difference in the increment between subsamples 3 and 4 is  $11,5 - 8,3 = 3,2\%$ . This difference should be reflected in a limited number of new species. If the number is only a fraction, then the subsample size is too small and conversely, if too large then the subsample is too large. For convenience and to allow for some local vegetation variation 3,2% difference can be taken to represent 3 new species. This difference can also be expressed as percentage accumulated totals between subsamples 4 and 3 where the first subsample is taken as 100%:  $151,6 - 139 = 12,6\%$  (Figure 6). The number of species (*y*) in the first subsample with difference of 12,6% = 3 new species between subsamples 3 and 4 is, therefore:

$$\frac{139}{100} y + 3 = \frac{151,6}{100} y$$

$$y = \frac{3}{0,126}$$

$$= 23,81 \text{ species}$$

If  $c = 12 \text{ species m}^{-2}$  then the area (*x*) corresponding to 23,81 species is:

$$x = e^{\frac{(\ln y - \ln c)}{m}}$$

$$= 9,82 \text{ m}^2$$

Subsample size is, therefore,  $9,82 \text{ m}^2$ . This relationship also takes the diversity of the vegetation into account because subsample size increases as diversity (*c* value) decreases. The subsample size is also within the range of: greater or equal to  $2 \text{ m}^2$ , for reduced edge effect; and less than  $20 \text{ m}^2$  for observer efficiency, in the vegetation concerned in this study. For less diverse vegetation, maximum subsample size could be increased if required. To ensure a width to length ratio

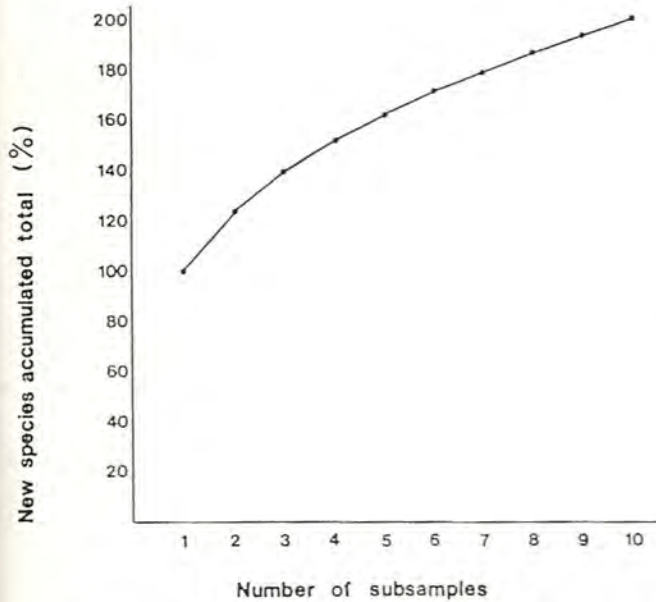


Figure 6 The relationship between number of subsamples and accumulated total of new species expressed as a percentage. Subsample accumulated totals: 1 = 100,0%; 2 = 123,2%; 3 = 139,0%; 4 = 151,6%; 5 = 162,1%; 6 = 171,1%; 7 = 179,2%; 8 = 186,5%; 9 = 193,3% & 10 = 199,5%.

of 1:2 for rectangularity of subsamples (Mueller-Dombois & Ellenberg 1974) the following convenient sizes can be used:

- 1 m × 2 m (2 m<sup>2</sup>)
- 1,5 m × 3 m (4,5 m<sup>2</sup>)
- 2 m × 4 m (8 m<sup>2</sup>)
- 2,5 m × 5 m (12,5 m<sup>2</sup>)
- 3 m × 6 m (18 m<sup>2</sup>)
- 3,5 m × 7 m (24,5 m<sup>2</sup>)

The nearest convenient size which is not less than 9,82 m<sup>2</sup> is, therefore, 2,5 m × 5 m (12,5 m<sup>2</sup>). In the first three relevés recorded in the investigation of the vegetation ecology of the Transvaal Waterberg (Westfall, in prep.) using this subsample, size increments of less than 10% were reached with four subsamples in two cases and with five subsamples in one case. These results substantiate the validity of the subsample number and area relationships. Species recorded in each relevé varied from 38 to 60 and sample areas covered varied from 50 to 62,5 m<sup>2</sup>. This is a considerably higher number of species than that recorded by Westfall (1981) in vegetation with similar diversity using single 200 m<sup>2</sup> sample areas. Separate random subsamples, therefore, appear to increase efficiency. The smaller area also facilitates cover estimation of the herbaceous stratum. It is noteworthy that in all three relevés (Westfall, in prep.) only approximately 10% of the total number of

species estimated to be present in each 20-ha stand was recorded. Recording of all species within the stand would have been prohibitive in terms of time and effort. It can, therefore, be assumed that the remaining 90% of the species have a very low frequency and that the so-called 'infrequent species' in phytosociological classifications are often not as infrequent as assumed. The method proposed here is also objective in distinguishing species, which on a frequency basis, should be more relevant in classifications. In randomly selected subsamples each species chance for inclusion in the subsamples is proportionate to its frequency.

The value of 0,3 for slope ( $m$ ) can be refined with more data for improved predictions especially for areas greater than 20 ha.

The expressions described in this paper, as well as those describing area, sample number and scale relations (Rutherford & Westfall 1986) have, for convenience, been programmed for the SHARP PC 1500 computer. This program, titled VEGFORM, is available from the first author on request.

### Conclusions

The advantages of expressing species diversity in terms of the  $c$  value include:

- (1) overcoming the edge effects of small quadrat size in certain vegetation types; thereby
- (2) improving comparability of species diversity when expressed as species per m<sup>2</sup>;
- (3) associating subsample size with species diversity;
- (4) determining number of subsamples required for sampling a vegetation stand; and
- (5) predicting the approximate number of plant species which could occur within a vegetation stand.

### Acknowledgements

The authors thank Dr J.C. Scheepers, Dr M.C. Rutherford, Dr H. van Ark and Mrs J. Schaap for assistance.

### References

- MUELLER-DOMBOIS, D. & ELLENBERG, H. 1974. Aims and methods of vegetation ecology. Wiley, New York.
- PARTON, W.J. & INNIS, G.S. 1972. Some graphs and their functional forms. U.S. International Biological Program, Grassland Biome Project Technical Report No. 153: 1-41.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1986. The Biomes of southern Africa — An objective categorization. *Mem. bot. Surv. S. Afr.* No. 54.
- WESTFALL, R.H. 1981. The plant ecology of the farm Groothoek, Thabazimbi District. M. Sc. thesis, Univ. of Pretoria.
- WESTFALL, R.H. (in prep.) The vegetation ecology of Sour Bushveld in the Transvaal Waterberg.



### 3.2 FIELD SAMPLING

The purpose of field sampling is to acquire data, that with processing and logical interpretation will achieve the aims of the study. Because field sampling often incurs additional transport and subsistence costs, savings can be effected by reducing time spent in the field, through improvements in methodology.

#### 3.2.1 Sampling unit location and number

This refers to sampling unit location (sample plot) within a vegetation stand and not stand location, as described in section 3.1.5.

##### 3.2.1.1 Background

Where working scale is such that the scale-related vegetation stand can be represented by a single quadrat then selection of quadrat location, based on representativeness, is probably the most efficient method of quadrat location. However, where the area of the scale-related vegetation stand is far larger than the area of a single quadrat used for sampling the stand, then randomization of sub-quadrat locations will reduce the number of observer decisions to be made. This is confirmed by Werger (1974) who states that "sampling plots (sampling units) should be representative of the stands they represent unless random or systematic sampling is carried out when the question of representativeness is by-passed."

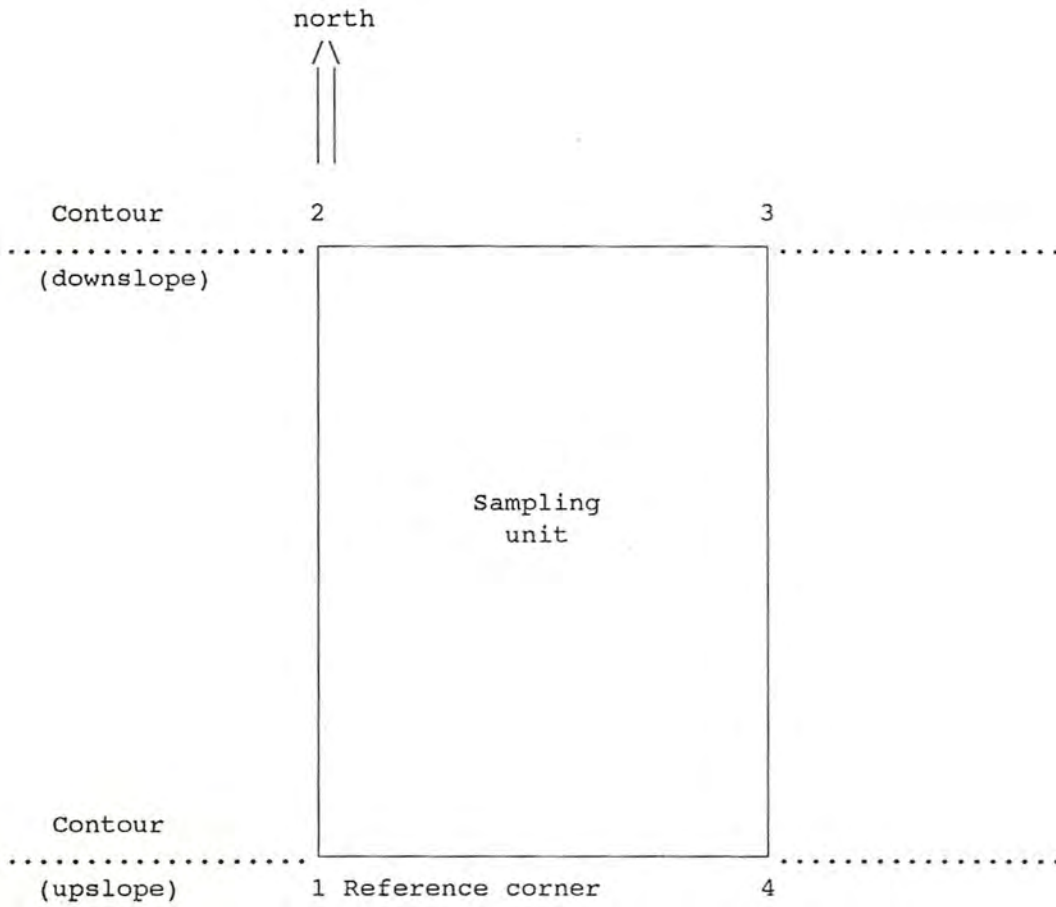


FIGURE 3.2. - Reference corner (1) of sampling unit in terms of direction, for level terrain; and slope, for sloping terrain.



In this study, where 2,5 x 5 m sampling units are used (section 3.1.6) to sample a 20 ha stand, representativeness is extremely difficult to determine and random sampling unit placement is required.

#### 3.2.1.2 *Methods applied*

The stand centre is a point in the field represented by grid intercepts and co-ordinates (section 3.1.5). Placement of sampling units within the stand can be simplest with the stand centre as reference because the stand is circular. A co-ordinate system based on compass bearing and distance from the stand centre is a convenient method of locating a sampling unit.

The program PHYTOLOC (section 3.1.5.3) can be used to generate random co-ordinates for this purpose, where the x co-ordinate is a compass bearing and the y co-ordinate is the distance from the stand centre in metres, or *vice versa*. A list of these co-ordinates can be printed prior to field sampling, to save time.

Decision-making for sampling unit orientation can be further reduced by standardizing on the sampling unit reference corner. This can also facilitate vegetation monitoring in terms of sampling unit re-location. The reference corner is that point on the ground represented by the sampling unit co-ordinates and is the left upslope corner of the sampling unit, when facing downslope, with the longer sides orientated downslope (Figure 3.2). If the ground is level, then the reference corner is the south west corner, when facing



north, with the longer sides orientated due north.

### 3.2.2 Plant identification and verification

Plant identification aids are often used by researchers in the field. These can be portable field herbaria, photographic records, photostats of herbarium specimens and notes on identifying characteristics.

#### 3.2.2.1 *Background*

Although it is common practice to collect voucher specimens of specific and infra-specific taxa recorded in a study area for reference purposes, the criteria for field identification of these plants are rarely made known. Comparison with a voucher specimen is insufficient if the criteria used for comparison are not given. The problem is exacerbated by the reliance of most identification aids, such as floras, on fertile material, whereas, field identifications are often on sterile material. This can complicate efforts to repeat a given study, for example, for monitoring purposes, and inhibit the transfer of species identification knowledge, which is unique to each field researcher. A lack of standardization is also evident in the character and character states used for identifying plants in published works. Can vegetation studies be repeatable if the criteria used by researchers for plant identifications in the field, are not made known?



# A new identification aid combining features of a polyclave and an analytical key

ROBERT H. WESTFALL, HUGH F. GLEN AND MICHAEL D. PANAGOS

*Botanical Research Institute, Department of Agriculture and Water Supply, Private Bag X101, Pretoria 0001, South Africa*

*Received December 1984, accepted for publication February 1985*

WESTFALL, R. H., GLEN, H. F. & PANAGOS, M. D., 1986. **A new identification aid combining features of a polyclave and an analytical key.** A new identification aid combining features of a polyclave and an analytical key is described. It is based on presence-or-absence characters and is presented in the form of a matrix, with the characters in rows and the taxa in columns. The PHYTOTAB program package is used to order the matrix, in order to facilitate identification. The method was used to construct an identification aid for vegetative material and has wider taxonomic and teaching implications.

ADDITIONAL KEY WORDS:—Computer-aided-identification – matrix – phytosociology – PHYTOTAB

## CONTENTS

Introduction . . . . .	65
Existing identification methods . . . . .	66
Analytical keys . . . . .	66
Polyclaves . . . . .	67
Methods . . . . .	68
Results . . . . .	70
Discussion . . . . .	70
Conclusions . . . . .	72
Acknowledgements . . . . .	72
References . . . . .	73

## INTRODUCTION

During the course of work on a pilot study of the Transvaal Waterberg (Westfall, 1981) the need for an objective method of site identification of the numerous plant species arose. The use of field herbaria proved time-consuming and cumbersome. Existing keys would have been of limited assistance because, first, they often use evanescent characters at an early stage (e.g. requiring flowers out of season), and secondly, they are only available for a minority of the

taxa present in this study area. When a specimen is collected for later identification by a reputable herbarium, the onus of identifying correctly the same taxon at other sites rests entirely on the knowledge, memory and experience of the field ecologist. During phytosociological work, vegetative characters are necessarily used for identification, whereas vegetative keys are rarely available.

Bearing these problems in mind, we decided to investigate alternative methods of identification, a subject which had been pursued by one of us in some detail (Glen, 1974). Traditionally, taxonomists have used one or another form of dichotomous key for identification, and the alternative methods have been given very little practical use. A selection of identification strategies is discussed below.

#### EXISTING IDENTIFICATION METHODS

A wide variety of different manually operated identification systems has been proposed for use in biology (Leenhouts, 1966), as well as a small but growing number of computer-aided systems (Morse, 1971; Pankhurst, 1978). These include such apparently diverse systems as the 17th-century precursors of the indented dichotomous key described by Voss (1952) and the computer-aided system described by Morse (1971), in which the computer interrogates the operator. All these techniques belong to one or other of two groups: analytical keys, which have only one entry point, and polyclaves (this term, coined by Duke (1969), is used here in the wider sense given to it by Morse, 1974), which have many entry points. The method described later in this paper is unique in being usable as both an analytical key and as a polyclave.

#### *Analytical keys*

Analytical keys are easily reproduced, being printed documents, and relatively easily optimized, the theory of optimization having been first worked out by Lamarck (1778). Possibly the most accessible review of these is by Osborne (1963). A dichotomous key can be used with maximum efficiency if each step splits the remaining group of taxa into two equal subgroups. In general, a polychotomous step with  $N$  branches at any point in an analytical key operates at maximum efficiency if it splits the group into  $N$  equal subgroups. This means that the mean path length from the start of the key to the level at which an identification is achieved is minimized.

Attempts have been made (Morse, 1971, 1974) to weight different taxa according to their relative commonness or rarity, so that common taxa are keyed out with the shortest average path length, and to weight characters so that the easiest to observe are used preferentially. Taxon weighting would tend to minimize the number of errors per unit time on the assumption that the key will be used more often to identify common taxa than rarities. Character weighting tends to minimize the chance of using a difficult character near the start of the process, and so it minimizes the possibility of using the wrong major section of the key, thus minimizing the number of errors per identification.

Analytical keys do have practical disadvantages, despite their clear theory and apparent simplicity. These are largely related to their single entry point and





limited number of paths to the correct identification. The most frequently encountered problem is that it is very difficult to construct a key that will enable one to identify a fragmentary (sterile) specimen unambiguously if it belongs to a large group, for example the Mesembryanthemaceae. Keys to the genera of this family have been constructed both by Bolus (1958) and by Herre & Volk in Herre (1971). Both of these require flowers and ripe fruit for certainty in identification, yet both parts are seldom available at the same time.

Variable taxa present a problem in the construction of analytical keys. If one allows for the variation by allowing the taxon in question to key out in more than one place, then one degrades the overall efficiency of the key by lengthening it. If one allows for variation by describing it where it occurs in the key, then efficiency is degraded by basing a dichotomy on a property which is less than desirably clear, or by lowering the ease of observation.

The addition of taxa to an existing analytical key is a relatively severe problem involving the rewriting of at least part of the key. If the additions are more than minor, it is probably most efficient to rewrite the key from the beginning, if optimization is to be retained. Computer-constructed keys can achieve this end more easily, for example those of Morse (1974) and Pankhurst (1971).

Ease of observation optimization requires that a minimum of characters be used at each step. In fact, the use of more than one character in a step implies that either the characters used are not fully independent, or that one or more of those used are variable. Rypka *et al.* (1967) have shown that it is theoretically possible to construct a key to any number ( $T$ ) of taxa using not more than  $N$  characters, where  $N = \log_2 T$ , rounded up to the nearest integer. In fact, considerably more characters than this minimum number are usually used, but the number of characters rarely exceeds the number of taxa. Therefore, it may happen that a highly distinctive and potentially useful character, for example flower colour or an accurate locality label, is present on the specimen but is not used in the key.

### *Polyclaves*

A polyclave has as many entry points as there are characters, and an effectively infinite number of pathways from the start to identification of the specimen. Very little skill is necessary to identify a plant using a polyclave, but the speed and accuracy of identifications improve significantly with increasing skill. In general, it appears that the most efficient strategy for using a polyclave is at each step to use the rarest character that is available and still unused.

Sneath & Sokal (1973) divide polyclave algorithms into simultaneous and sequential. The basic strategy of a simultaneous polyclave is to calculate the value of one or another similarity function between the unknown specimen and all taxa in the identification matrix. The value of this function should be high with only one taxon, to which the unknown may be assigned if the value is above a predetermined threshold. Alternatively, taxa are considered to be regions in character hyperspace, and the locality of the unknown in this space may be calculated. The unknown is assigned to a particular taxon if it falls into that region of the space. Typical simultaneous methods are those of Gyllenberg (1965) and Lapage *et al.* (1970, 1973).

In sequential polyclaves, characters are considered one at a time. In each step, a subset of taxa having the character in question in the same state as the unknown is extracted from the complete set, and compared with the subset resulting from the previous step. Those taxa common to both subsets form the set of possible identifications for the next step. When this step contains only one taxon, the unknown has been identified. A wide variety of such polyclaves has been described by Leenhouts (1966). Essentially, they fall into three groups: printed, card- and computer-operated, although the same polyclave may be converted from one form to another.

Printed polyclaves may be even cheaper to reproduce than the equivalent dichotomous key, because they can be compressed into less space. For example, the key of Meyer (1969) to the families of flowering plants of SW Africa/Namibia, a polyclave, occupies about 220 lines (excluding explanatory figures), while the dichotomous key by Merxmüller (1972) to the same taxa plus 15 families of Pteridophyta occupies about 1500 lines. However, a polyclave is frequently more expensive to reproduce.

A form of polyclave on plastic cards has been described (Leenhouts, 1966), but polyclaves on paper cards are far commoner. Most punch-card polyclaves fall into two groups: centre-punched one card per character state, and edge-punched one card per taxon. The former kind is exemplified by the keys of Bianchi (1931) (the first punch-card polyclave), Hansen & Rahn (1969) and Weber & Nelson (1972); the latter by Baker (1970), a key to the species of the genus *Erica*.

With the possible exception of computer-generated punch-card polyclaves, this kind of key is very expensive to reproduce, as accurate punching dies must be made for each card in the key. With computer-generated keys, standard 80-column cards are punched with a standard card punch, and the key can be produced for a realistic price. With the almost total extinction of the computer card reader-punch as a result of improved magnetic storage media, it is not certain whether computer-generated punch-card keys will remain viable for any length of time.

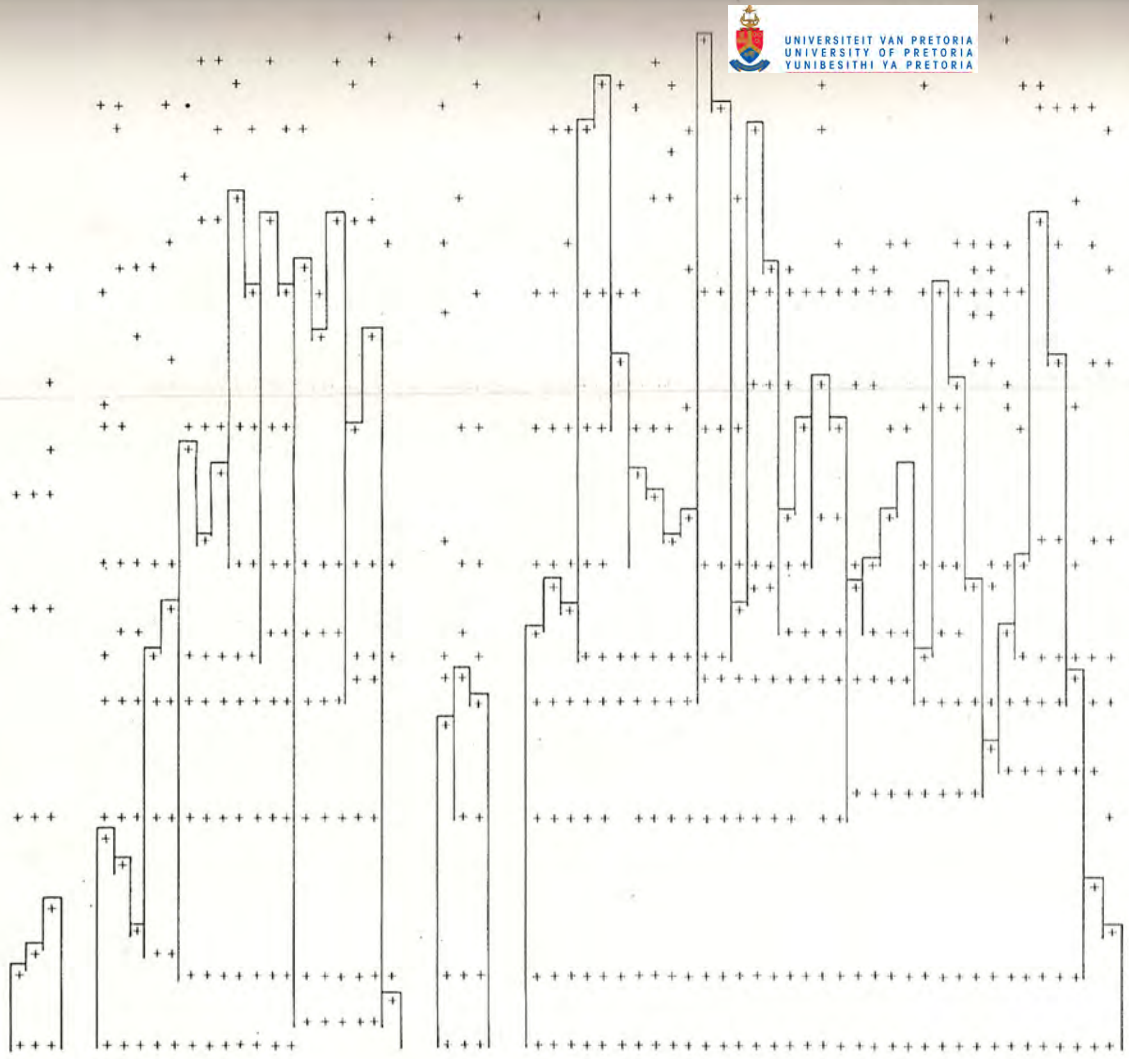
Edge-punched card keys can be operated with a mechanical card sorter or with a skewer, but are more subject to damage than centre-punched cards. On the other hand, the latter do not usually have any space for annotation.

Computer-aided identification systems (automated polyclaves) are many, for example those of Boughey, Bridges & Ikeda (1968), Goodall (1968), Morse (1974), Pankhurst (1978) and the simultaneous methods mentioned above. Each has its own advantages and disadvantages, but at present none can be used in the field. It is possible that in the future one or more may be made to work on a portable personal computer. Morse (1975) has given a particularly lucid account of the computer-aided systems available at that time. We plan to add to the system described here, routines enabling it to be used in the field on a micro-computer.

#### METHODS

Eighty plant specimens were collected in the Transvaal Waterberg during March 1983. The specimens were numbered and sent to the National Herbarium, Pretoria (PRE) for identification.





Cladodromous			Venation
Semicraspedodromous			
Reticulodromous			
Hyphodromous			
Brochidodromous			
Craspedodromous			
Eucamptodromous			
Asymmetrical base			
Asymmetrical lamina			
Oblong			
Narrow /oblong	Leaf form		
Obovate			
Ovate			
Elliptic			
Cuneate			
Lobate			
Decurrent	Leaf base		
Cordate			
Obtuse			
Acute			
Even pinnate			
Multifoliate			
Twice pinnate	Leaf type		
Odd pinnate			
Digitately trifoliate			
Simple			
Opposite			
Whorled	Leaf arrangement		
Decussate			
Alternate/scattered			
Very different	Colour	Leaf surfaces	
Similar	Texture		
Truncate			
Rounded			
Mucronate	Leaf apex		
Obtuse			
Acute			
Erose			
Serrate			
Dentate	Leaf margins		
Revolute			
Crenate			
Lobed			
Entire			
Amplexicaul			
Sessile	Leaf attachment		
Petiolate			

al Waterberg: Trees, shrubs, climbers, creepers, forbs, ferns

Characters relevant to each specimen were observed in the field and recorded on a field data sheet on which provision is made for recording the collector's number, provisional name and locality or relév . If more than 60 characters are recorded additional page(s) could be used.

The morphological characters used in the preliminary key were taken from various sources such as Henkel (1934), Metcalfe & Chalk (1979) and Tainton, Bransby & Booysen (1976) and condensed into an illustrated photocopied booklet for field use. This proved useful in maintaining consistency in the allocation of characters.

The PHYTOTAB program package (Westfall *et al.*, 1982) was used to construct the matrix. The data from the field data sheets were transferred to encoding forms using the PHYTOTAB input format where each character was regarded as a rel v . Taxa were grouped into trees and shrubs, climbers and creepers, forbs, ferns and graminoids for convenience. This would also allow for each group to be shown on a separate table when more data are collected. Characters can also form groups such as leaf-margin patterns and venation patterns. The character groups are sequenced in descending order of character (group) frequency from right to left. Characters not in groups are on the left of the table in frequency order (Tables 1 & 2). Spaces were inserted between characters to facilitate vertical reading.

Taxa are sequenced within groups so that dichotomies may be formed. That is, taxa with similar characters are grouped together. The dichotomies are continued until they consist of a single line for each taxon, and are ended with a vertical blocking line after the first unique character for each taxon when compared with the two adjacent characters. Blocking was done by hand. If a taxon has two or more character states that express possible variation within a character group falling within the blocked area, the species may be repeated to facilitate use of the key. This is illustrated in our example by *Combretum zeyheri* Sond., in which the leaf apices may be either obtuse or mucronate, and by *Syzygium cordatum* Hochst., in which the leaf variability is so great as to require that the same taxon appears four times (Table 1).

To use the table, the specimen to be identified is compared with the characters given for the group to which the specimen belongs, starting from the right-hand side of the table. The user determines whether the character is present or absent in the specimen. If the character is present in the state described, then the user follows the 'present' pathway indicated by +. If the character is sometimes absent it is still shown as +. If it is always absent then the 'absent' pathway, indicated by blanks, is followed. The next character is found to the left of the newly completed one. This procedure is repeated until the end of a block, indicated by a vertical blocking line, is reached. The user should now be able to look across to only one name on the left-hand side of the table. This will give him the desired identification. The blocked pathways indicate the minimum characters needed for identification. Additional characters outside the blocks may be used for corroboration.

Two other possible methods of using the table suggest themselves. However, these methods should only be necessary as a teaching aid, or if the specimen lacks important characters or in a preliminary version of the key. It must be stressed that these methods are by no means obligatory.

First, it is possible to use the table as a form of tabular key (Newell, 1970).



A profile of the unknown specimen is obtained by noting the characters present, in the order of the table, and this profile is compared with those of the taxa in the table, one by one, until a matching profile is found. This would be the name of the unknown specimen. This method is probably too cumbersome to be of practical value.

Another method of using the table would be as a classical polyclave. The unknown specimen is examined for any conspicuous (or unusual, if one is familiar with the table) character, and the taxa which display this character in this state are determined. Another character is searched for, to see which of the taxa found in the previous step display this character in the required state. The specimen has been identified when only one taxon remains.

#### RESULTS

The problems discussed in the introduction were overcome by devising a method for constructing vegetative keys during the reconnaissance phase of phytosociological work. The method had to allow for increasing the number of taxa as the work progressed. It had also to be simple to apply in the field. The key presented here will operate satisfactorily only in the geographical area concerned; it is used in this paper solely as an example.

In Table 2, *Schizachyrium sanguineum* (Retz.) Alst. in the Graminoid group has an open pathway indicating that this species does not key out with the available characters. However, its identity can be inferred from the absence of flattened stem bases which are present in *Hyparrhenia filipendula* (Hochst.) Stapf var. *pilosa* (Hochst.) Stapf.

#### DISCUSSION

The matrix method of key construction described in this paper allows for biological variation where more than one character state, such as different leaf forms, may be present in the same taxon, without detracting from the effectiveness of the key. Characters which are periodically available for which the population is variable may be used, as well as corroborative characters. The right-to-left method of use described above in the Methods may seem at first sight to be idiosyncratic. This method is used because the phytosociological example makes use of the species name file for both vegetation and taxonomic classification. The PHYTOTAB program package allows the user to print species names to the right of the table as well as printing the matrix in mirror image, thereby reading from left to right. Thus, in ordinary taxonomic usage the table would be rearranged to read 'correctly'. A further advantage of this package is that in the printout form, both horizontal and vertical lines may be employed to improve legibility.

This form of key is unique in that it can be used as both a polyclave and an analytical key. It combines the flexible strategy of the former with the ease of reproduction of the latter to produce a very efficient identification system. One advantage of polyclaves is that either taxa or characters can be added at will, with minimal effort required, although the addition of taxa may result in ambiguities in identifications, which may only become apparent after long use. By using PHYTOTAB, both taxa and characters can be added to existing keys

Table 2. A preliminary vegetative key for some species of the Transvaal Waterberg: Graminoid

Species	Purple below nodes	Taste	Ligules	Hairy	Base/sheath ratio	Leaves	Life form
	Culm bent at nodes	Caraway Turpentine					
<i>Astrida aequiglumis</i>							+
<i>Hyparrhenia filipendula</i> var. <i>pilosa</i>		+					+
<i>Schizachyrium sanguineum</i>							+
<i>Perotis patens</i>							+
<i>Aristida spectabilis</i>			+				+
<i>Trachyandra spicata</i>							+
<i>Rhynchelytrum setifolium</i>							+
<i>Tristachya biseriata</i>							+
<i>Aristida canescens</i> subsp. <i>canescens</i>		+					+
<i>Brachiaria nigropedata</i>							+
<i>Loudetia simplex</i>							+
<i>Themeda triandra</i>							+
<i>Digitaria eriantha</i> subsp. <i>eriantha</i>							+
<i>Setaria perennis</i>							+
<i>Andropogon schirensis</i> var. <i>angustifolius</i>							+
<i>Eragrostis racemosa</i>							+
<i>Digitaria eriantha</i> subsp. <i>stolonifera</i>	+						+
<i>Cymbopogon excavatus</i>							+
<i>Hyperthelia dissoluta</i>							+
<i>Diheteropogon amplexens</i>							+
<i>Cymbopogon prolixus</i>							+
<i>Panicum maximum</i>							+
<i>Setaria sphacelata</i> var. <i>sphacelata</i>							+
<i>Pogonarthria squarrosa</i>							+
<i>Setaria ustilata</i>							+

at will, and ambiguities will be revealed when the tables are blocked. Correcting a key is also a straightforward task.

It will be noted that the horizontal length of the block for each taxon is roughly proportional to the path length that would be required in an analytical key to identify that taxon. Two uses for this proportionality (which, it should be stressed, is no more than a rough approximation) suggest themselves. First, it offers a check on the discriminatory efficiency of the characters chosen. If all the blocks are long, then redundant characters which do not contribute to any identifications are present. These can be removed, or placed where they are only used as corroborative characters. Secondly, the proportionality can be used to demonstrate Lamarck's principles of efficient key construction to students elegantly and graphically (Lamarck, 1778; Osborne, 1963). This form of key seems to be the only one available which demonstrates its effectiveness with field data.

A further teaching application of the 'PHYTOTAB-key' may be mentioned. With the falling cost of computer usage (at least in terms of constant units), and with the relative ease and accessibility of use of the PHYTOTAB package, a method is presented of allowing students to learn the mysteries of good key construction in the most effective possible manner, by making their own mistakes in the course of constructing their own keys to taxa they think they know.

The PHYTOTAB program package (Westfall *et al.*, 1982) allows the user to construct a key to which taxa and characters may be added as and when more specimens are collected. The keys can be subdivided easily to form tables of convenient size for field use. An advantage with computer printout is that keys may be duplicated conveniently and cheaply.

#### CONCLUSIONS

Our identification aid has peculiar and significant advantages in three areas: in teaching, in reconnaissance and in the field. Its suitability as a teaching tool stems from the way in which the printed table, while being usable as a polyclave, displays explicitly the structure of an analytical key. Because the PHYTOTAB program package is relatively simple, quick and inexpensive to use, students can be given practical experience of key construction using this package. For the same reasons, 'PHYTOTAB-keys' can be produced repeatedly during the course of a taxonomic investigation. Ambiguities in the key at each stage may be taken to indicate areas requiring further research, and the keys produced will help in field-work while the group is under revision. In the field, use of keys made by this method will mean that characters present only in fresh material can be used to improve the accuracy of determinations. It may be expected that the facility with which both taxa and characters may be added to the table will be found useful in both taxonomic and ecological work.

#### ACKNOWLEDGEMENTS

We thank A. V. Hall, O. A. Leistner, R. A. Lubke, E. J. Pankhurst and J. C. Scheepers for helpful suggestions. Miss A. P. Backer and Mr J. F. van Blerk are thanked for assistance in the early stages of this work. The staff of the National Herbarium, Pretoria identified the specimens collected.





REFERENCES

- BAKER, H. A., 1970. A key for the genus *Erica* using edge-punched cards. *Journal of South African Botany*, 36: 151-156.
- BIANCHI, A. T. J., 1931. Een nieuwe determinatie-methode. *Tectona* 24: 884-893.
- BOLUS, H. M. L., 1958. Key to the genera of Mesembryanthemaceae. *Notes on Mesembryanthemum and allied Genera*, 2: 392-403.
- BOUGHEY, A. S., BRIDGES, K. W. & IKEDA, A. G., 1968. An automated biological identification key. *Museum of Systematic Biology, University of California, Irvine, Research Series*, 2: i-xix, 1-35.
- DUKE, J. A., 1969. On tropical tree seedlings. I. Seeds, seedling systems and systematics. *Annals of the Missouri Botanical Gardens*, 56: 125-161.
- GLEN, H. F., 1974. 'A taxonomic revision of the subtribe Gibbaeinae (Mesembryanthemaceae)'. Unpublished M.Sc. Thesis, University of Cape Town, Cape Town.
- GOODALL, D. W., 1968. Identification by computer. *BioScience*, 18: 485-488.
- GYLLENBERG, H. G., 1965. A model for the computer identification of micro-organisms. *Journal of General Microbiology*, 39: 401-405.
- HANSEN, B. & RAHN, K., 1969. Determination of angiosperm families by means of a punched-card system. *Dansk botanisk Archiv*, 26: 1-46 & 172 punched cards.
- HENKEL, J. S., 1934. *A field book of the woody plants of Natal and Zululand*. Durban: Robinson.
- HERRE, A. G. J., 1971. *Genera of Mesembryanthemaceae*: Cape Town: Tafelberg.
- LAMARCK, J. B. A. P. M. de, 1778. *Flore Francaise*, 3 vols. Paris: Imprimerie Royale.
- LAPAGE, S. P., BASCOMB, S., WILCOX, W. R. & CURTIS, M. A., 1970. Computer identification of bacteria. In A. Baillie & R. J. Gilbert (Eds), *Automation, Mechanisation and Data Handling in Microbiology*: 1-22. London: Academic Press.
- LAPAGE, S. P., BASCOMB, S., WILCOX, W. R. & CURTIS, M. A., 1973. Identification of bacteria by computer: general aspects and perspectives. *Journal of General Microbiology*, 77: 273-290.
- LEENHOUTS, P. W., 1966. Keys in biology: a survey and a proposal of a new kind. *Proceedings of the Koninklijke Nederlandsche Akademie van Wetenschappen Amsterdam, Series C*, 69: 571-596.
- MERXMÜLLER, H., 1972. *Prodromus einer Flora von Suedwest Afrika*. Lehre: Cramer.
- METCALFE, C. R. & CHALK, L., 1979. *Anatomy of the Dicotyledons*, Vol. 1, 2nd edition. Oxford: Oxford University Press.
- MEYER, P. G., 1969. *Eintuehrung in die Pflanzenwelt Suedwestafrikas*. Windhoek: S.W.A. Wissenschaftliches Cesellschaft.
- MORSE, L. E., 1971. Specimen identification and key construction with time-sharing computers. *Taxon*, 20: 269-282.
- MORSE, L. E., 1974. Computer programs for specimen identification, key construction and description printing using taxonomic data matrices. *Publications of the Museum, Michigan State University, Biological Series*, 5: 1-128.
- MORSE, L. E., 1975. Recent advances in the theory and practice of specimen identification. In R. J. Pankhurst (Ed.), *Biological Identification with Computers*: 11-52. London: Academic Press (Systematics Association special volume, 7).
- NEWELL, I. M., 1970. Construction and use of tabular keys. *Pacific Insects*, 12: 25-37.
- OSBORNE, D. V., 1963. Some aspects of the theory of dichotomous keys. *New Phytologist*, 62: 144-160.
- PANKHURST, R. J., 1971. Botanical keys generated by computer. *Watsonia*, 8: 357-368.
- PANKHURST, R. J., 1978. *Biological Identification*. London: Edward Arnold.
- RYPKA, B. W., CLAPPER, W. E., BOWEN, I. G. & BABB, R., 1967. A model for the identification of bacteria. *Journal of General Microbiology*, 46: 407-424.
- SNEATH, P. H. A. & SOKAL, R. R., 1973. *Numerical Taxonomy*. San Francisco: Freeman.
- TAINTON, N. M., BRANSBY, D. I. & BOOYSEN, P. de V., 1976. *Common Veld and Pasture Grasses of Natal*. Pietermaritzburg: Shuter & Shooter.
- VOSS, L. G., 1952. The history of keys and phylogenetic trees in biology. *Journal of the Science Laboratory, Denison University*, 43: 1-25.
- WEBER, W. A. & NELSON, P. P., 1972. *Random access key to the genera of Colorado mosses*. Boulder: University of Colorado Museum.
- WESTFALL, R. H., 1981. 'The plant ecology of the farm Groothoek, Thabazimbi district'. Unpublished M.Sc. Thesis, University of Pretoria, Pretoria.
- WESTFALL, R. H., DEDNAM, G., VAN ROOYEN, N. & THERON, G. K., 1982. PHYTOTAB—a program package for Braun-Blanquet tables. *Vegetatio*, 49: 35-37.



### 3.2.2.2.1 Improvements

Subsequent improvements to the techniques of plant identification are:

- i. an illustrated booklet of plant characters used for identification in this study (Appendix I);
- ii. a program was written, as an option for the PHYTOTAB-PC program package (section 3.6), for sequencing species for plant identification purposes. This program groups species according to the required dichotomies for key production. It was found that the character sequence is easier to apply if characters are standardized in their sequence, namely, from large to small attributes. This sequence is also used in the illustrated booklet and encoding forms. Using the same characters it is also possible to produce a family key; and,
- iii. visual verification of plant identification by means of photographic slide/plant comparisons. A photographic colour slide is taken of the plant when the voucher specimen is collected. The slide incorporates the specimen number. After development of the slide film, the slides are not mounted but returned to the film cassette.

After a plant is identified by means of the matrix key, the cassette with the relevant specimen number is inserted into a field viewer for comparison. The field viewer used here is an old motor-drive Konica FS 1 camera with a) a translucent perspex back for light transmission and b) a magnifier lens for viewing which replaces the standard lens. The mirror system is also removed. The motor-drive facilitates slide location.



### 3.2.3 Species cover

In this study species cover refers to total projected canopy cover for a plant species and cover estimations are required with a precision suitable for monitoring cover change and determining relative species composition.

Generally species cover is visually estimated and not sampled. However, where greater precision than a visual estimate is required, as in this study, and cover estimation is achieved through sampling, then the sampling unit area for cover should relate to cover. Therefore, the sample attributes of presence and cover can require different sampling unit areas.

#### 3.2.3.1 Background

Mueller-Dombois & Ellenberg (1974) list several scales for estimating species cover-abundance such as the Domin-Krajina scale with eleven classes, the Braun-Blanquet scale with seven classes and the Daubenmire scale with six classes. Werger (1974) advocates a nine class scale.

Londo (1976), however, states that a cover scale should meet the following requirements:

- i. the scale should be sufficiently detailed;
- ii. its points should be related to the actual cover values so that these can be treated arithmetically;
- iii. cover and abundance should be strictly separated. It



- is not logical to combine such dissimilar features in a single quantitative scale; and,
- iv. the symbols should be as simple as possible.

The criterion of precision, commensurate with repeatability, could also be added to these requirements for monitoring cover change. However, the aim of cover estimations should first be considered. If the aim is to detect cover pattern in a matrix then, possibly, an eleven class scale, such as the Domin-Krajina scale, could confuse pattern and the ideal might be a simplified five class scale. On the other hand if arithmetic processing of cover data is required, such as for monitoring cover change or to determine relative species composition, then even an eleven class scale is likely to be too crude. Precision in cover determination is inversely proportional to cover class size and directly proportional to effort required for such determinations. A balance should, therefore, be achieved between class size and effort required, for cover determinations, to meet for example, monitoring or species composition requirements.



## Miscellaneous notes

VARIOUS AUTHORS

### THE PLANT NUMBER SCALE — AN IMPROVED METHOD OF COVER ESTIMATION USING VARIABLE-SIZED BELT TRANSECTS

#### INTRODUCTION

The vegetation ecology of the Transvaal Waterberg is currently being investigated at a scale of 1:250 000 (Westfall in prep.). Vegetation structure is being analysed according to Edwards (1983) using the cover meter (Westfall & Panagos 1984) for cover determinations in each height class. In the floristic analysis, individual species cover is estimated by using the Domin-Krajina cover-abundance scale (Mueller-Dombois & Ellenberg 1974).

A comparison of recorded species cover, being the sum of the class midpoints according to the Domin-Krajina scale (Mueller-Dombois & Ellenberg 1974), with the structural cover should result in the summed species cover for a stand being: 1, greater than the cover of the height class with the greatest cover, and 2, less than the summed cover of all the height classes, provided that 3, quadrat size is such as to include those species contributing significantly to the total cover of the stand.

In the vegetation being investigated, quadrat size is generally commensurate with species richness (Westfall *et al.* 1987). However, the summed estimated species cover was often considerably less than the mean cover for the height classes with the greatest cover for the same stands. Overcompensation for the underestimation of species cover often led to the summed estimated species cover being considerably greater than the summed cover of all the height classes for the same stands (Westfall in prep.). Clearly, improved species cover estimations are required for species cover to have any relevance other than an approximate indication of relative abundance.

In estimating species cover, according to Edwards (1983), the observer is often inclined to ignore grasses without inflorescences and to estimate from a static position without taking plant size and distribution into account. For example, a larger plant should require a larger area to be observed than a smaller plant. Furthermore, although mean canopy diameter can be readily estimated it is often far more difficult to estimate mean distance apart in terms of mean canopy diameter because of often highly irregular plant distribution.

To overcome these problems, the approach suggested here is based on a simple estimate of area and a count of the individuals of a species within the area.

#### METHODS

The cover of a species is given by Edwards (1983) for hexagonal packing by

$$c = \frac{90,7}{(n+1)^2}$$

where  $c$  = percentage crown cover and  $n$  = the mean number of crown diameters by which the plant crowns are separated.

Assuming hexagonal packing, the transect area, of which the percentage crown cover is a proportion, is given by:

$$\sin 60^\circ (n+1) 30D$$

where  $D$  = mean crown diameter and  $30$  = the value for obtaining a minimum of 0,1% cover. Cover of less than 0,1% is not considered significant. Transect length is, therefore,  $30D$  and transect width is  $\sin 60^\circ (n+1)$ . In practice, transect width was taken as slightly less than the average gap between plants within or nearest to the sample quadrat plus the mean crown diameter. The number of individuals of a species was then counted within the transect. Only species occurring within the sample quadrat were recorded and for each a count of individuals within a transect commensurate with each species spacing and size was made. Counts of individuals did not include the first individual as the transect was started adjacent to the first individual. This permitted a cover of less than 0,1% where no individuals were counted. Transect width was never greater than the length as this could have resulted in actual cover values of less than 0,1% being given higher cover values.

The mean number of crown diameters ( $n$ ) by which the plant crowns are separated within each transect is given by  $n = \frac{30-I}{I}$  where  $I$  = number of individuals counted. Percentage crown cover can then be calculated according to Edwards (1983). Table 1 shows number of individuals counted, representation by a single character symbol and percentage cover for recording purposes in the field.

Vegetation structure was analysed using the cover meter and the summed species cover was estimated with both the Domin-Krajina scale and the plant number scale as outlined above for five vegetation stands represented by 21 quadrats. The mean of the shortest and longest cross distances of each crown was taken as the crown diameter for each species and these distances were noted in four categories (Edwards 1983) namely, forbs (herbs), grasses, shrubs and trees. Class intervals were selected on a basis of trial and error to give an approximately normal distribution of occurrences within crown diameter class intervals. All estimations were done by an independent observer.

#### RESULTS

The results of the crown cover determinations are given in Table 2.



TABLE 1.—Number of plant individuals counted with single character symbol and percentage crown cover

No.	Symbol	% Crown cover
0	+	0,00
1	1	0,10
2	2	0,40
3	3	0,91
4	4	1,61
5	5	2,52
6	6	3,63
7	7	4,94
8	8	6,45
9	9	8,18
10	A	10,08
11	B	12,20
12	C	14,51
13	D	17,03
14	E	19,75
15	F	22,68
16	G	25,80
17	H	29,12
18	I	32,65
19	J	36,38
20	K	40,31
21	L	44,44
22	M	48,78
23	N	53,31
24	O	58,05
25	P	62,99
26	Q	68,13
27	R	73,47
28	S	79,10
29	T	84,76
30	U	90,70
31	V	96,85
>31	W	100,00

The Fibonacci sequence, where each number is the sum of the preceding two numbers, provided the closest resemblance to a normal distribution of occurrences within crown diameter class intervals. This is illustrated in Figure 1 with a frequency polygon with the class intervals on a natural logarithmic scale to reduce the effect of increasingly larger class intervals.

The class intervals used for mean crown diameters according to the Fibonacci sequence, are shown in Table 3. Transect lengths were determined by the midpoints of each class interval.

#### INTERPRETATION

In Table 2 estimations according to the Domin-Krajina cover-abundance scale are considerably lower than those for the single height classes with the most cover. The plant number scale, in contrast, yielded higher values than that of the single height class with the most cover and lower values than the cover of the combined height classes for each relevé, except relevé 37. This indicates a greater precision in estimating cover when using the plant number scale as opposed to the Domin-Krajina scale.

According to Westfall *et al.* (1987), quadrat size should have been larger for the vegetation type represented by relevé 37, but this was not apparent using the Domin-Krajina scale at the time of sampling. However, simple summation of the values obtained by the plant number scale in the field indicated inadequate quadrat size. It is far too time-consuming to verify quadrat size

for each quadrat according to Westfall *et al.* (1987). The plant number scale together with a structural analysis of the vegetation provides a simple means of verifying adequacy of quadrat size.

In the frequency polygons (Figure 1) the peaks to the left of the central troughs for forbs, grasses and shrubs are caused by a relatively higher proportion of 0,2 m diameter crowns. This can be attributed to the observer rounding off crown diameters to 0,2 m some of which should have fallen into the 0,211 to 0,34 class. If class intervals had been known at the time of recording, it can be expected that greater care would have been exercised in measurements where crown diameters were close to class borders. The troughs mentioned are, therefore, considered to be a result of measurement inaccuracies which could be overcome by using class intervals.

The use of standard transect lengths as illustrated in Table 3 should simplify transect length determination and provide for variability in crown size. A further advantage could be the simultaneous counting of individuals of different species with similar crown diameters and spacing to save time. It is also suggested that a simple counter be used for recording number of individuals for each species as marking paper for this purpose requires stopping at each individual recorded.

It must be emphasized that the parameter determined here is projected crown cover and not projected foliage cover which is more species and age-dependent.

The class 'r' on the Braun-Blanquet scale and '+' on the Domin-Krajina scale (Mueller-Dombois & Ellenberg 1974) both with 'solitary, insignificant cover' are difficult to determine. Species with a single occurrence in a sample quadrat often have significant cover outside the quadrat. If a stand is defined, as in this study, as 20 ha (Westfall *et al.* 1987), it is impracticable to determine whether a species is 'solitary' within that area. The concept of 'solitary' is relative to the area defined. In the plant number scale used here the lowest cover class is less than 0,1%, which seems better defined than 'solitary'.

In contrast to the Domin-Krajina scale, the plant number scale has proportionately finer subdivisions at the lower cover values of the scale. This is of significance in the South African context with often high species richness characterized by many dominant species with generally lower cover in contrast to the few dominant species with higher cover often found in the relatively impoverished European vegetation.

TABLE 2.—Percentage crown cover in five vegetation stands represented by relevés 33 to 37

Relevé no.	Single height class	Percentage crown cover		
		Combined height classes	Domin-Krajina scale	Plant number scale
33	55	100	27	60
34	31	104	25	74
35	44	103	19	79
36	41	86	21	53
37	63	146	17	35

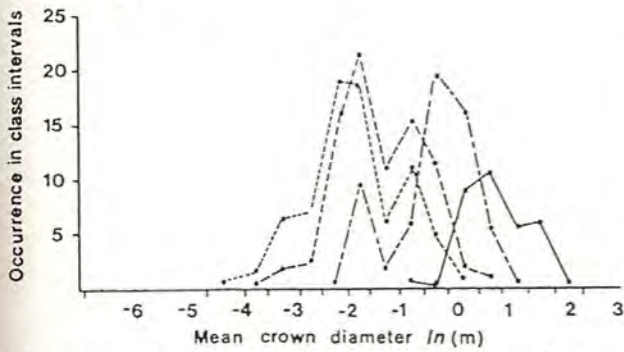


FIGURE 1.—Frequency polygons of occurrence of plants in mean crown diameter class intervals for forbs (....), grasses (---), shrubs (-·-·-) and trees (—). Class intervals are according to the Fibonacci sequence on a ln scale.

CONCLUSIONS

The projected crown cover determinations based on area estimations and counting of individuals shows improved precision compared to the Domin-Krajina cover-abundance scale. Although this method is more time-consuming than a purely visual estimation of cover, the use of standard class intervals and a counter should decrease the time required for cover determinations. The method appears more suitable for the species-rich South African vegetation than the traditional European cover-abundance estimation scales. The method also provides a means of verifying quadrat size adequacy.

ACKNOWLEDGEMENTS

The authors thank Dr J.C. Scheepers for comments and suggestions.

REFERENCES

EDWARDS, D. 1983. A broad-scale structural classification of vegetation for practical purposes. *Bothalia* 14: 705–712.

TABLE 3.—Class intervals of crown diameters according to the Fibonacci sequence for determining standard transect lengths

Crown diameter class interval (m)	Transect length (m)
0,001–0,01	0,15
0,011–0,02	0,45
0,021–0,03	0,75
0,031–0,05	1,20
0,051–0,08	1,95
0,081–0,13	3,15
0,131–0,21	5,10
0,211–0,34	8,25
0,341–0,55	13,35
0,551–0,89	21,60
0,891–1,44	34,95
1,441–2,33	56,55
2,331–3,77	91,50
3,771–6,10	148,05
6,101–9,87	239,55

MUELLER-DOMBOIS, D. & ELLENBERG, H. 1974. *Aims and methods of vegetation ecology*. Wiley, London.  
WESTFALL, R.H. & PANAGOS, M.D. 1984. A cover meter for canopy and basal cover estimations. *Bothalia* 15: 241–244.  
WESTFALL, R.H., VAN STADEN, J.M. & PANAGOS, M.D. 1987. Predictive species area relations and determination of subsample size for vegetation sampling in the Transvaal Waterberg. *South African Journal of Botany* 53: 44–48.  
WESTFALL, R.H. in prep. *The vegetation ecology of Sour Bushveld in the Transvaal Waterberg.*\*

R.H. WESTFALL and M.D. PANAGOS

MS. received: 1987.07.31.



TABLE 3.1. - Symbols used for crown diameter classes

Crown diameter (m)	Symbol
0,001 - 0,010	A
0,011 - 0,020	B
0,021 - 0,030	C
0,031 - 0,050	D
0,051 - 0,080	E
0,081 - 0,130	F
0,131 - 0,210	G
0,211 - 0,340	H
0,341 - 0,550	I
0,551 - 0,890	J
0,891 - 1,440	K
1,441 - 2,330	L
2,331 - 3,770	M
3,771 - 6,100	N
6,101 - 9,870	O





### 3.2.3.2.1. Improvements

Subsequent improvements to the methods described in section 3.2.3.2 include the allocation of symbols (Table 3.1) to the crown diameter classes and recording of these class symbols together with the cover symbols for each species. The following data can be derived from the relationship between mean cover and mean canopy diameter:

- i . mean cover (%);
- ii . mean crown diameter (m);
- iii. individuals per hectare;
- iv. square metres per individual;
- v. mean spacing, plant centre to centre (m);
- vi. mean canopy radius (m); and,
- vii. mean canopy to canopy gap (m).

It was found, during field testing of these methods, that the precision of species cover estimations could be improved considerably, for the higher strata, if rooting within a sampling unit was substituted for canopy overhang in the sampling unit.

The following is, therefore, applied:

- i. forbs, grasses and dwarf shrubs: species presence recorded if half or more than half of the stem or tuft base occurs within the sampling unit; and,
- ii. shrubs and trees: species presence recorded if any part of the canopy overhangs the sampling unit.

In this way vegetation structure is taken into account for sampling presence and cover.



### 3.2.4 Floristic data recording

Floristic data is taken here to mean the data relating to terrestrial plant life and its distribution, in a given area. The recording of plant identification data is described in section 3.2.2. Taxonomic nomenclature used in this work is according to Gibbs Russell *et al.* (1985, 1987).

#### 3.2.4.1 *Background*

Mueller-Dombois & Ellenberg (1974) suggest that apart from species presence and cover-abundance other aspects are often worth recording such as species quantity, dispersion, morphological structure or layering. To this is added vitality or vigour and sociability. Although the purpose of the study will often dictate what should be recorded, the minimum requirements for optimum results should be considered, taking into account the aspects that can be derived from those that are recorded.

#### 3.2.4.2 *Methods applied*

It follows from the previous sections that the minimum floristic data which should be recorded in the field, for this study, are the following:

- i. species name abbreviation or voucher specimen reference number. The number is preferred because it requires fewer keystrokes than a name abbreviation if direct computer data capture is used; the plant identification key can serve as a num-



ber control whereas abbreviations would require a separate control to ensure that the same abbreviation is not used for a different species; the number necessitates at least some reference to the plant identification keys which can improve identification accuracy; and programmatic inclusion of voucher specimen reference numbers in a checklist is then possible. Although use of name abbreviations can be quicker than numbers, initially, if the species are known, it is suggested that overall time can be saved by using numbers, especially if data correction, where for example, the same abbreviation is used for different species, is taken into account;

- ii. percentage canopy cover class symbol (section 3.2.3); and,
- iii. crown diameter class symbol (section 3.2.3).

In addition, the growth form symbol (from Edwards 1983 and Rutherford & Westfall 1986) is also recorded, as follows:

- T: Trees are rooted, woody, single-stemmed plants over 2 m high or multi-stemmed over 5 m high;
- S: Shrubs are rooted, woody, multi-stemmed plants from 1 to 5 m high, or single-stemmed when less than or equal to 2 m high;
- D: Dwarf shrubs are rooted, woody or partly woody, plants less than 1 m high;
- G: Grasses are rooted, herbaceous plants belonging to the family Poaceae or graminoid plants such as Cyperaceae and Restionaceae which resemble grasses; and
- F: Forbs are rooted, non-graminoid herbaceous plants.

Not all plant species occurring in a sampling unit can be recorded. Those occurring in seed form and those occurring underground at



the time of sampling, for example, are difficult to identify. Similarly, sterile plant material with a low cover is also often difficult to identify. Such plants would generally have a low ecological significance within the plant community concerned although their taxonomic significance as, for example, endemic species, might be high. Furthermore, unidentifiable plants with a low cover and occurrence, are of doubtful value as diagnostic species. It is seldom likely that identifiable material for plants with high cover and occurrence cannot be obtained. The criterion for the recording of a plant species occurrence within a sampling unit is, therefore, the availability of identifiable material at the time of sampling. It is suggested that this is a more realistic criterion than the often-quoted "total floristic recording" which seldom can be true.

The minimum floristic data requirements for classification, however, are initially, relevè number, species name abbreviation or voucher specimen reference number and cover percentage symbol. The following article describes a field data capture program for these data.



## PHYTOCAP. A FIELD-DATA CAPTURE PROGRAM FOR THE PHYTOTAB PROGRAM PACKAGE

Manual field recording of floristic data, for phytosociological studies, entails re-encoding data on computer encoding forms, transfer to magnetic tape and finally loading onto the mainframe computer for multi-variate analyses. The time taken from re-encoding to access on the mainframe computer can be from one to six weeks. Classification of data while fieldwork is in progress, therefore, becomes impracticable in the summer-rainfall areas because fieldwork generally takes place in the relatively short growth period and the delays in computer access would limit fieldwork considerably if data were classified during this period. A serious disadvantage of classifying vegetation only after completion of fieldwork is that vegetation units are often either under-sampled, resulting in invalid syntaxa, or over-sampled resulting in wasted labour and expense. Furthermore, the potential for errors is increased by re-encoding the data.

These problems were overcome by using computerized field-data capture where data is recorded directly onto a hand-held computer and loaded onto the mainframe computer on return from the field. Multivariate analysis techniques used are the PHYTOTAB program (Westfall *et al.*, 1982), which are compatible with both the DECORANA (Hill, 1979a) and TWINSPAN (Hill, 1979b) programs. Preliminary classification using either PHYTO20 or TWINSPAN can then be available within a day of returning from fieldwork.

The system used for field-data capture is the Sharp PC 1 500 computer with an additional 8K expansion memory module, printer/cassette interface, cassette recorder and programmable RS232C interface. The program PHYTOCAP, written in BASIC, is used for recording floristic data in the PHYTOTAB (Westfall *et al.*, 1982) format. The program features include automatic line number allocation, sample and/or subsample numbers, alphanumeric species codes, cover-abundance values, data pertinent to individual species, data display, data printing in two formats, halting and continuing program execution, line editing, saving data to tape and loading data from tape. Furthermore, data input is verified for errors such as sample number length, species omission and cover-abundance omission. The user is also informed when five lines of memory are left. The capacity of the computer is 70 lines which is approximately 14 relevés or samples with 40 species per sample, which is generally more than adequate for the floristic data recorded in one day.

Field procedure includes quadrat location, quadrat demarcation, floristic sampling, voucher specimen collection and environmental parameter sampling. The Sharp PC 1 500 computer is used for floristic sampling. Species for which voucher specimens exist are input as a four-letter genus code and a three-letter species code. Species for which voucher specimens are required are tagged with pre-numbered, specimen number, adhesive address labels

and input as a left-justified specimen number. Specimen collection is effected after completion of floristic data recording. This process ensures a smooth flow of data input and reduces the possibility of species being overlooked. Environmental parameters are recorded directly on a field data sheet. It is envisaged that a second Sharp PC 1 500 be used for recording environmental parameters and as a standby machine. Memory capacity currently precludes the use of a single machine for both floristic and environmental data.

After a day's recording the computer is attached to the printer/cassette interface, which remains in the vehicle. A printout of the floristic data is obtained of each sample for stapling to the field data sheets which form the hardcopy for eventual permanent safe-keeping at the Botanical Research Institute, Pretoria. The data are then transferred to C 15 cassettes, when the computer can be cleared for the following data set.

Loading data to the Burroughs B 7 900 mainframe is effected by means of a Burroughs ET 1 100 terminal, RS232C interface and a transfer program called DATATRAN written by S. J. Crafford. Data are read from the cassette and simultaneously transmitted at 300 Baud.

The advantages of this system of data recording include the cost-saving production of preliminary classifications for optimum sampling as well as the labour- and cost-saving of not having to re-encode data. The potential for errors is also reduced by the reduction in data handling. Documentation and taped copies of the program are available from the author. Please forward a blank C 15 cassette for copying.

## ACKNOWLEDGEMENTS

The author thanks Dr J. C. Scheepers for comments and suggestions and Mr S. J. Crafford for writing the program DATATRAN.

## REFERENCES

- HILL, M. O., 1979a. *DECORANA*. A FORTRAN program for detrended correspondence analysis and reciprocal averaging. Unpublished report. Ecology and Systematics, Cornell University, Ithaca, New York. pp. 1-30.
- HILL, M. O., 1979b. *TWINSPAN*. A FORTRAN program for arranging multi-variate data in an ordered two-way table by classification of the individuals and attributes. Unpublished report. Ecology and Systematics, Cornell University, Ithaca, New York. pp. 1-48.
- WESTFALL, R. H., DEDNAM, G., VAN ROOYEN, N. & THERON, G. K., 1982. PHYTOTAB. A program package for Braun-Blanquet tables. *Vegetatio* 49: 35-37.

R. H. WESTFALL\*

\* Botanical Research Institute, Department of Agriculture & Water Supply, Private Bag X101, Pretoria 0001.



#### 3.2.4.3.1 Improvements

A program PHYTOFORM written for the Sharp PC 1500 in BASIC converts the sampling unit data (sub-plots), for each stand to a single relevè representing each stand. The conversion is based on the mean cover for each species recorded for the stand, thereby reducing the sampling unit species lists to a single species list for each stand. Transfer to the host computer is the same as for the original data. Thus the host computer can have two data sets, namely, one with relevè data and the other with sampling unit data.

#### 3.2.5 Habitat data

Habitat data are the environmental data, both biotic and abiotic, which describe the environment of a plant community and which can explain the floristic differences between plant communities. These data are usually recorded in the field.

##### 3.2.5.1 Background

Werger (1974) states that "In the Zürich-Montpellier approach it is empirically determined that patterns in floristic composition correspond with patterns in the environment". The National Working Group for Vegetation Ecology in South Africa, for example, suggested minimum environmental parameters for vegetation and environment correlation and prediction of floristic potential. These are:

- i. sample number;



- ii. sample co-ordinates (latitude, longitude);
- iii. date;
- iv. altitude (above sea level) in metres;
- v. aspect of sample in degrees;
- vi. slope inclination in degrees;
- vii. lithostratigraphy;
- viii. percentage rock and outcrop cover;
- ix. geomorphology;
- x. vegetation structure according to Edwards (1983); and
- xi. land use and management.

The first three items are extremely important, especially if the site is to be re-sampled to monitor change. However, certain questions relating to the aims of a study need clarification. If an environmental description of the vegetation is all that is required then the environment could be described in descriptive terms. A descriptive account of the environment could be improved by standardization of terms to facilitate comparisons with other studies. However, if plant communities are to be correlated with environmental factors, to determine differentiating causes, then sampling of environmental factors is necessary.

Is sample size the same for both vegetation sampling and environmental factor sampling? If the sample size for vegetation sampling does not meet statistical criteria then how can a sample of, for example, soil depth be adequate using the same criteria? Mappable habitat factors, such as lithostratigraphy and soil type should not require statistical sampling if these have already been



mapped. However, factors such as aspect, slope and soil depth are sampled by what are essentially point samples, which cannot be equated with the samples used for vegetation. Cognizance should also be taken of scale when considering environmental factors. A hill could be regarded as a mountain when scale is increased. What environmental factors should be sampled and what can be derived from other sources?

These questions could affect possible community and environment correlations and hence influence the value of a classification.

#### 3.2.5.2 *Methods applied*

Habitat data were determined from two sources, namely, direct measurements in the field at the time of sampling and indirectly, as derived data from topographic and geological maps.

##### 3.2.5.2 A) *Field data*

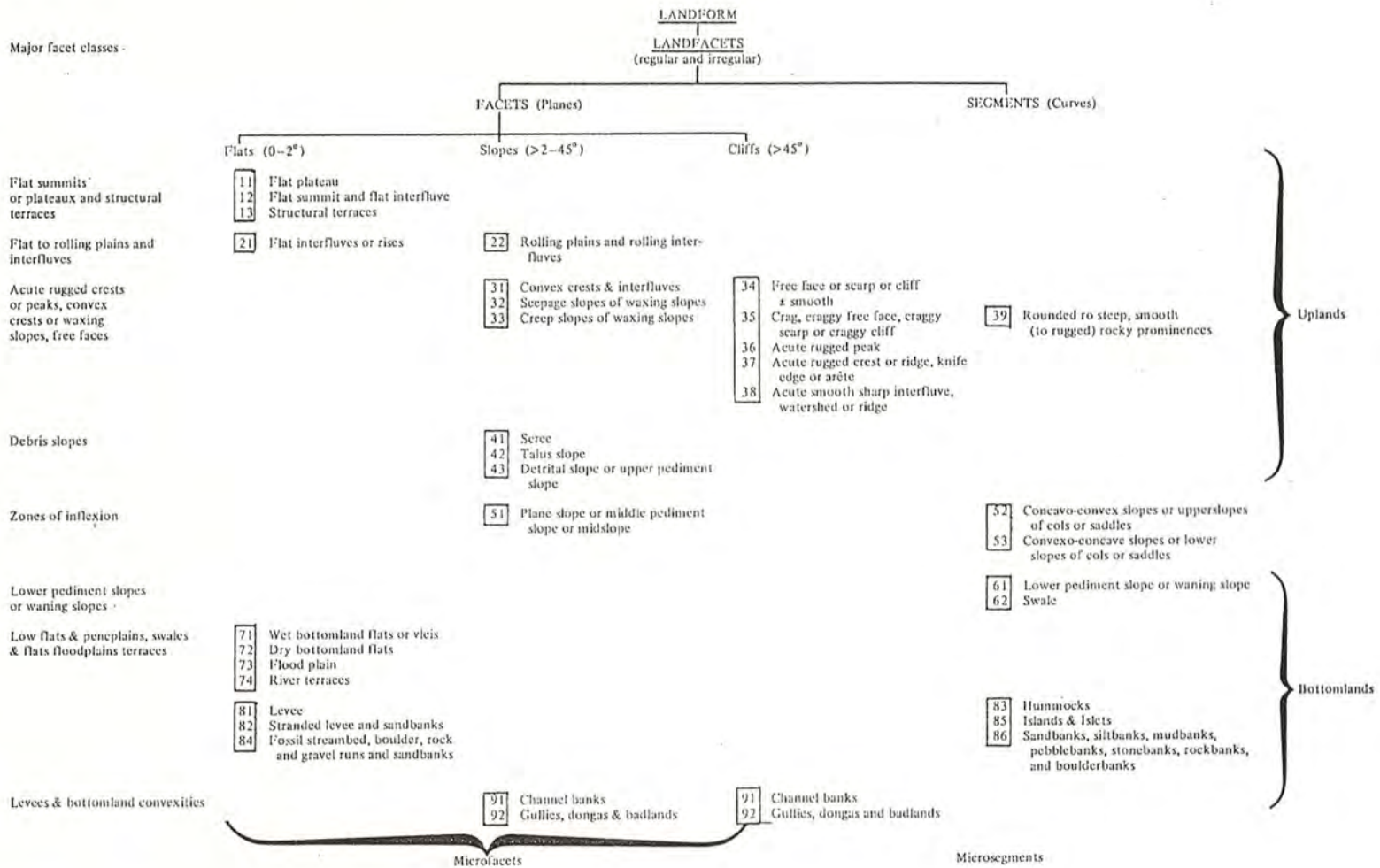
The disparity in area between stand and sampling unit for this study, the extent of a particular variable's influence on the plant community and sampling considerations necessitated the categorization of habitat data into two groups for recording purposes, namely, those suitable for stand characterization and those suitable for sampling unit characterization.

##### a) Stand data

- i. Lithostratigraphy, being formation and surface rock type



TABLE 3.2. - Geomorphology classes in terms of facets and segments from Scheepers<sup>1</sup> (unpubl.)



<sup>1</sup> J.C. Scheepers, Private Bag X05, Lynn East, 0039

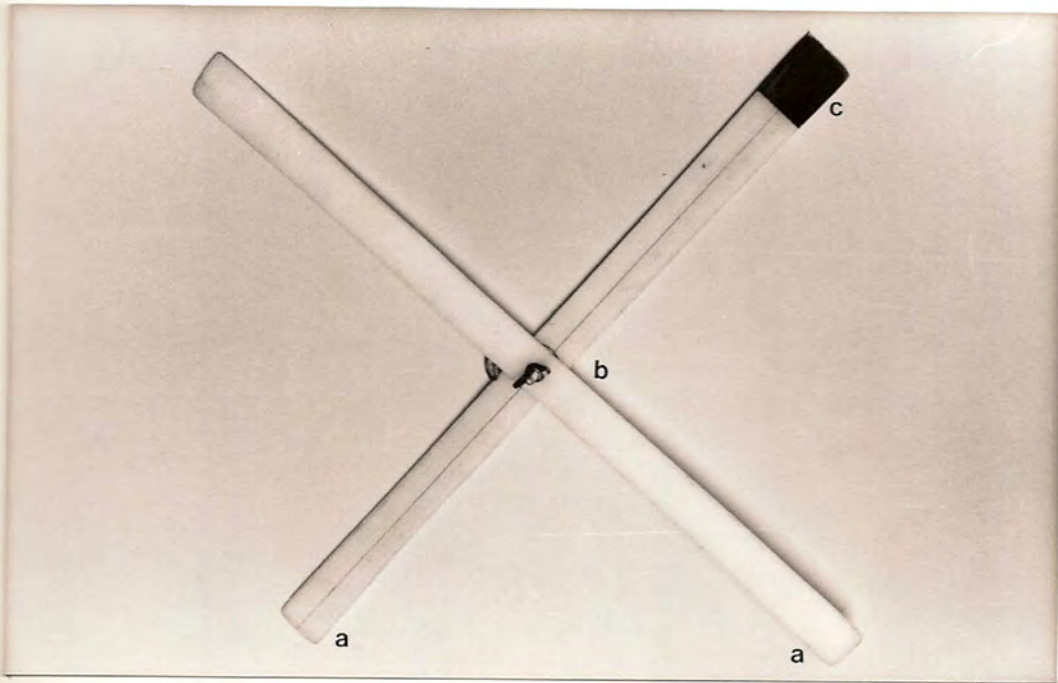


FIGURE 3.3. - A flexible plastic cross for selecting random pairs of plants in ground layer vegetation. a) intersecting arms, b) wing nut pivot - the first component of the plant pair is that which is nearest to the pivot, and c) marked - the second component of the plant pair is that which first intercepts the arm or continuation thereof.



FIGURE 3.4. - A tripod-mounted, manual spinner for selecting consecutive plants in pairs, at random, in vegetation above the ground layer. a) spinner, b) sight, and c) telescopic tripod.



TABLE 3.3. - A dual level system of land use categorization

<u>Land use:</u> (level 1)	<u>Primary</u> <u>symbol</u>	<u>Land use:</u> (level 2)	<u>Secondary</u> <u>symbol</u>
<i>Vegetation utilization</i>		<i>Production:</i>	
None observable	A	Cattle	A
Component mostly utilized:		Game	B
Grass	B	Goats	C
Herbs	C	Horses	D
Shrubs	D	Ostriches	E
Trees	E	Sheep	F
All	F	Other stock	G
<i>Vegetation replaced by:</i>		Bark	H
Pasture - dry land	I	Hay	I
Pasture - irrigated	J	Flowers	J
Crops - dry land	K	Fruits	K
Crops - irrigated	L	Leaves	L
Plantation	M	Root	M
Exotics (other)	N	Seeds	N
<i>Construction:</i>		Wood	O
Dam	Q	Cereals	P
Road - unsurfaced	R	Vegetables	Q
Road - surfaced	S	Alien	R
Footpath/track	T	<i>Waste:</i>	
Railway	U	Solid	W
Single storey	V	Liquid	X
Multi-storey	W	Gas	Y
Cleared	X		



according to Jansen (1982) with field verification of out-crops.

- ii. Altitude, being the altitude of the stand centre in metres, taken from the relevant 1:50 000 SA topo series map, to the nearest half-contour i.e. 10 m or 25 feet.
- iii. Geomorphology class, as visually categorized according to Table 3.2 .
- iv. Vegetation structure, being the total canopy cover for each of the height classes suggested by Edwards (1983) as determined by a cover meter (Westfall & Panagos 1984). Five random samples were taken in each height class. For the height classes up to one metre, a plastic cross (Figure 3.3) was arbitrarily thrown and for the height classes above one metre, a tripod-mounted spinner (Figure 3.4) was used to select random pairs.
- v. Land use, being a visual assessment according to the categories in Table 3.3 .
- vi. Insolation exposure, being 180 degrees minus the sum of the angles of the eastern and western horizons from the horizontal, at the stand centre and measured with an inclinometer. The directions of the eastern and western horizons, for the study area, were taken as due east and due west less the difference in degrees between the stand centre latitude and the Tropic of Capricorn. Maximum solar insolation is thus taken as being at the summer solstice.
- vii. Temperature of bare ground, ground layer vegetation and vegetation at chest height, being the mean of the measurements at five arbitrarily selected sites for each category around the stand centre. Temperatures were measured with an Instatherm



non-contact, infra-red thermometer. The temperature of a standard grey scale card was also measured at each stand for comparative purposes. These data are intended to assess the effect of vegetation cover on soil temperature.

viii. Grazing, as visually assessed according to the following four-class scale:

- 0, ground layer grazed short (less than 100 mm);
- 1, ground layer grazed short and patchy;
- 2, ground layer grazed evenly; and
- 3, no grazing evident.

ix. Browsing, as visually assessed according to a two-class scale:

- 0, browse line evident; and
- 1, browse line not evident.

x. Erosion, as visually assessed according to the following four-class scale:

- 0, donga erosion evident, by donga or gully formation;
- 1, sheet erosion evident, by general soil accumulation in a similar direction against plants;
- 2, donga and sheet erosion evident; and
- 3, no erosion evident.

xi. Fire, as visually assessed according to a two-class scale:

- 0, evidence of fire present; and
- 1, no evidence of fire present.

b) Sampling unit data

- i. Slope inclination, from the horizontal in degrees as determined with an inclinometer.
- ii. Aspect of inclination, in degrees as determined with a magne-



TABLE 3.4. - The "sausage" method of determining the clay percentage or texture of soil at field capacity (from F.S.S.A. 1974)

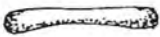
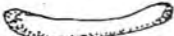



No sausage	Sand	<10% clay
	Loamy sand	10 - 15% clay
	Sandy loam	>15 - 20% clay
	Sandy clay loam	>20 - 35% clay
	Sandy clay	>35 - 55% clay
	Clay	>55% clay

TABLE 3.5. - The "finger test" method of determining the clay percentage or texture of soil at field capacity (from F.S.S.A. 1974)

Fingers stay clean	0 - 6% clay
Fingers slightly dirty	>6 - 12% clay
Fingers slightly sticky	>12 - 20% clay
Slight glossy smear and concretions	>20 - 35% clay
Highly glossy smear	>35% clay



- tic compass and allowing for magnetic deviation.
- iii. Percentage surface rock and outcrop cover, as a visual estimate.
  - iv. Soil surface compaction, being the mean of five measurements at the sampling unit centre and corners, using a pocket penetrometer, calibrated in  $\text{kg.cm}^{-2}$ .
  - v. Soil depth, as determined with a soil auger at the sampling unit centre, if the soil depth was greater than 300 mm or as the mean of five depth spike measurements at the sampling unit centre and corners, if the soil depth was less than 300 mm. These recordings are to the nearest 10 mm.
  - vi. Soil texture, being the mean of the class midpoints of the percentage clay in the A horizon, as estimated using both the "sausage method" (Table 3.4) and the "finger test" method (Table 3.5).
  - vii. Soil colour of the A and B horizons, determined at field capacity with a soil colour chart (Munsell Soil Color Charts 1954).
  - viii. Soil form, as determined by diagnostic horizon combinations according to MacVicar *et al.* (1977).
  - ix. Litter cover, as an estimated percentage of soil surface covered.
  - x. Litter depth, being the mean of five depths measured in mm to the nearest 10 mm at the sampling unit centre and the four corners.
  - xi. Relative biomass of the ground layer, being the mean of five measurements at the sampling unit centre and corners, using a disc pasture meter (Trollope & Potgieter 1983), calibrated





in cm. Conversion to  $\text{kg}\cdot\text{ha}^{-2}$  requires correlations with clipped, dry material. However, relative comparisons can be made between stands directly, although less precise.

A program HABIMEAN written in BASIC for the Sharp PC 1500 converts the sampling unit habitat data to stand data. The conversion is generally based on mean values which are transformed to SI units, where applicable. In the case of soil form, the percentage occurrence of the soil forms in the stands is calculated. Mean aspect is calculated as the mean of the vectors of slope and aspect for each sampling unit. Mean soil colour is taken as the means of the Munsell values for brightness, saturation and hue, to the nearest Munsell class. This program does not form part of the PHYTOTAB-PC program because of the difficulty in standardizing environmental variables.

#### 3.2.5.2 B) *Derived data*

Derived data are those habitat data for each stand which, in this case, were derived from topographic maps. The program SIDA is not included in the PHYTOTAB-PC package because current developments in extracting such information from geographic information systems will make this approach obsolete. The program description is, nevertheless, included here to show the type of information that can be derived from topographic maps. Some of this information is not usually recorded in the field because of difficulties in field determinations, such as stand position relative to watershed and drainage line.



The co-ordinates, in degrees, minutes and seconds generated with the program PHYTOLOC (section 3.1.6) for stand centre location were plotted on the relevant 1:50 000 SA topo series sheets. The circumference of each stand (5 mm radius at 1:50 000 scale) was circumscribed around the stand centres and the relevè number for each stand was noted on the sheets. The following was then determined from the topo sheets for input to the program SIDA written in BASIC for the Sharp PC 1500 computer:

a) Input

- i. Stand or relevè number.
- ii. Contour interval, in feet or metres, depending on the contour interval used on the relevant topo series sheet.
- iii. Sheet grid number, for reference purposes.
- iv. Latitude, in degrees, minutes and seconds of the stand centre.
- v. Altitude of the stand centre, in feet or metres (map dependent) to the nearest half-contour.
- vi. Aspect, in degrees from north of the stand inclination being the direction from the stand centre to the nearest point on the lowest contour, most adjacent to the stand circumference. This forms the aspect line.
- vii. Lowest contour altitude, in feet or metres of the lowest contour most adjacent to the stand circumference.
- viii. Lowest contour distance, in mm from the stand centre to the nearest point on the lowest contour, most adjacent to the stand circumference.
- ix. Highest contour altitude, in feet or metres of the highest contour most adjacent to the stand circumference.



- x. Highest contour distance, in mm from the stand centre to the nearest point on the highest contour, most adjacent to the stand circumference.
- xi. Watershed distance in mm, to the nearest half-contour from the stand centre along the highest contour distance line to the nearest crest or watershed. This is the watershed line. If the stand is situated on a crest then the watershed line is to the nearest drainage line approximately opposite the first drainage line (see xiii) so that the stand position can be related to two drainage lines.
- xii. Watershed altitude, in feet or metres, of the watershed or drainage line.
- xiii. Drainage distance, in mm, to the nearest half-contour, from the stand centre along the lowest contour distance line to the nearest drainage line. This is the drainage line. If the stand is situated in a depression then the drainage line is to the nearest watershed approximately opposite the first watershed line (see xi) so that the stand position can be related to two watersheds.
- xiv. Drainage altitude, in feet or metres of the drainage line or watershed.
- xv. Angle horizon east or west. The directions of the eastern and western horizons are due east and west plus the difference between the stand centre latitude and  $23\frac{1}{2}^{\circ}$  south (as in section 3.2.5.2 A) (a) vi). The horizon angles are determined by a subroutine of the program SIDA when contour line altitudes and distances from the stand centre, in mm, along the directions of the eastern and



western horizons, are input. The subroutine displays extrapolated altitudes so that the user can determine the number of contour lines to be input. Contour lines lower than the stand are ignored.

- xvi. Cumulative distance, in mm, being the distance from drainage line to watershed or if watershed altitude is less than the stand centre then to the highest point along the watershed line.

#### b) Output

The output is based mainly on simple trigonometric functions of the input variables. The following are printed for each stand:

- i. stand or relevè number, as input;
- ii. grid number, as input;
- iii. altitude, in metres;
- iv. aspect, in degrees;
- v. mean soil depth, in mm, based on a linear regression of the combined slope and position between watershed and drainage line and measured soil depth;
- vi. mean annual rainfall, in mm, based on a linear regression of the combined altitude and latitude and the recorded rainfall for officially recognized rainfall stations in the study area where recordings generally exceeded 20 years;
- vii. insolation exposure, in degrees;
- viii. relative available moisture, in mm, as a function of mean annual rainfall, slopes, aspect, position in the landscape, altitude, latitude, insolation exposure and mean soil depth;

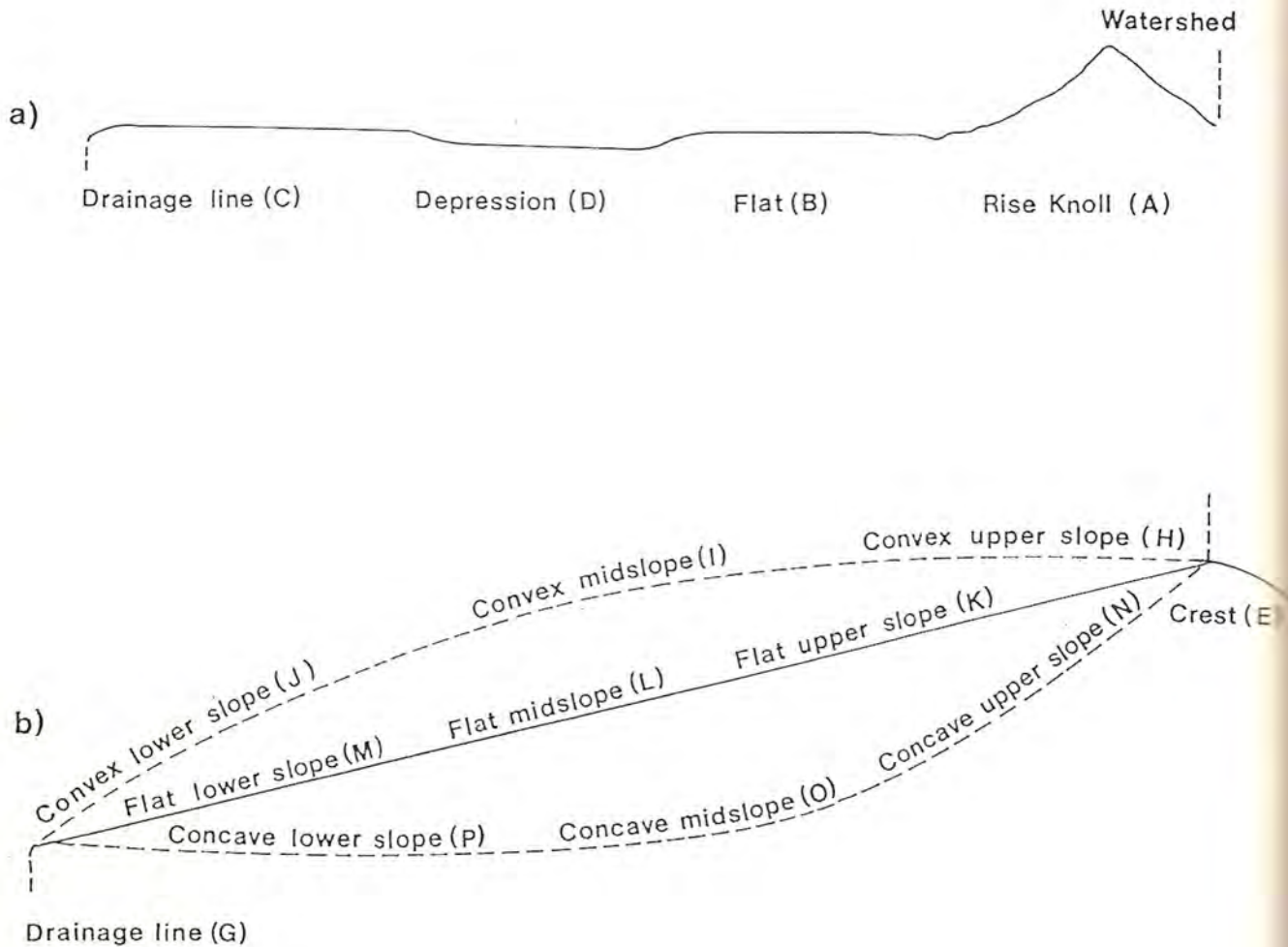


FIGURE 3.5. - Geomorphology classes determined with the SIDA program with class codes in brackets. a) Landscape slope  $\leq 2^\circ$  b) Landscape slope  $> 2^\circ$ . Slope categorization is restricted if the slope is less than 12x stand radius (R) to: crest/koppie if slope  $< 4R$ ; midslope if slope  $< 8R$ ; lower slope and upper slope if slope  $< 12R$ .



- ix. topographic profile, from watershed to drainage line with the stand diameter proportionally indicated according to position;
- x. topographic profile slope (mean), in degrees;
- xi. stand slope, in degrees;
- xii. scale of topographic profile;
- xiii. geomorphology, in classes as listed in Figure 3.5; and,
- xiv. stand drainage class according to the topographic profile and stand slopes where:
  - A Outflow: Run-off > Run-in
  - B Throughflow: Run-off = Run-in
  - C No flow: Run-off = 0; Run-in = 0
  - D Inflow: Run-off < Run-in.

### 3.3 CLASSIFICATION

Classification in this sense refers to the ordering of species presence/absence data in a matrix form where columns represent relevès (sampled stands) and rows represent species. The matrix values at the intercepts of columns and rows indicate presence and quantity by a cover or cover-abundance symbol.

#### 3.3.1 Background

Gauch (1982) defines classification as "grouping similar entities together in clusters". The New Collins Concise Dictionary of the English Language (1985) defines classification as "the placing of animals and plants in a series of increasingly specialized groups because of similarities in structure, origin, etc. that indicate



common relationships". The first definition seems more suited to a description of scatter diagrams, whereas, the second definition appears more appropriate to the needs for a vegetation classification, such as data reduction, identification and diagnosis of hierarchically grouped vegetation units and to show the relationships between and variation within, such units.

The procedure proposed by Mueller-Dombois & Ellenberg (1974) entails identification of "common species groups as a key to grouping relevès". The concatenation of relevès in which selected species occur, to achieve a relevè sequence, gives rise to the problem of diagnostic and non-diagnostic species recognition. In this sense, diagnostic refers to those species whose distributions are influenced by environmental gradients within the study area, at the scale concerned. They, therefore, have limited but definite distributions in the study area and can be used to delimit plant communities at any level in the hierarchy. Non-diagnostic species are not thus influenced and their limits within the study area are not clear. Clearly, reliable recognition of these two species groups can only be achieved after the classification process. If a set of species were chosen, for example, from the non-diagnostic species-group for initial relevè concatenation, a totally different classification could emerge, compared to that based on a set of diagnostic species. In practice, however, relevès are often grouped according to noticeable environmental gradients as a first step in sequencing relevès. Programs such as TWINSPAN (Hill 1979), PHYTO 20 (Westfall *et al.* 1982) and PHYTO 21 (Westfall & De Wet 1988) can also be used for initial classification but these programs do

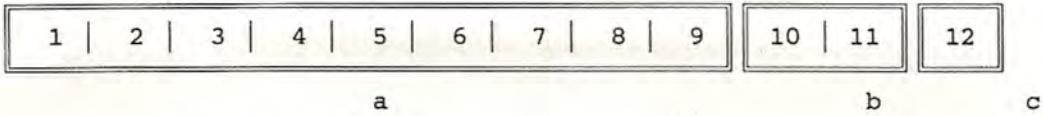


FIGURE 3.6. - Communities (numbered) sequenced according to the main environmental gradient responsible for their differentiation (a); subsidiary gradients (b and c) in which communities are sequenced similarly but inserted after the main environmental gradient.

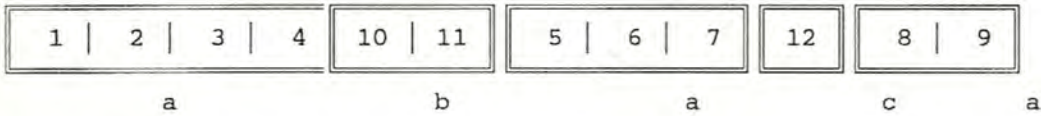


FIGURE 3.7. - Communities (numbered) sequenced according to the main environmental gradient responsible for their differentiation (a); subsidiary gradients (b and c) in which communities are sequenced similarly but inserted in the positions to which they correspond on the main environmental gradient.





not meet the needs for a final classification and further sequencing is generally required.

In a matrix classification the relevè number sequence is linear. This can pose problems where multiple environmental gradients are responsible for community differentiation. Should communities differentiated by subsidiary gradients be placed at the end of the communities differentiated by the main gradient or inserted within the communities of the main gradient at their respective positions on the main gradient? The last-mentioned possibility is preferred, as will be shown later. However, it is often likely that the two relevès adjacent to the point of insertion, of a community differentiated by a subsidiary gradient, are more similar to each other in terms of floristics than they are to the relevès to be inserted. These relationships are illustrated schematically in Figures 3.6 and 3.7.

In sequencing relevès according to similarity, an additional problem was found, namely, the selection of the first relevè with which the other relevès are to be compared. If all the relevès are tested in the first position and each relevè is compared with the first and subsequent relevès, then the number of tests is a factorial of the total number of relevès. The magnitude of this number is prohibitive even if mirror-image sequences are excluded. Furthermore, how can the "best" sequence be determined if only two relevès are compared at a time and not the entire sequence? What is required is a value relating to the effectiveness of a particular sequence of relevès.

The questions that arise from an objective approach to a Braun-Blanquet classification are:

- i. given that no two plant species have exactly the same distribution, can a particular classification be regarded as unique and, if not, what are the alternatives?
- ii. are the community relationships in the classified matrix data linear, as suggested by Poore (1956)?
- iii. what is the basis of the classification? Should species be sorted before relevès according to similarity in distribution as suggested by Werger (1974)?
- iv. is it possible in a data set with little or no observable discontinuities for two independent workers to achieve the same classification? It should be possible where discontinuities are obvious; and,
- v. can the classification process be automated thereby reducing decision-making in the classification process?

Subjective sampling followed by automated classification will not make the classification objective. For a classification to be objective and hence repeatable, the processes of stratification, sampling and classification should not have observer bias.

### 3.3.2 Methods applied

The approach adopted in this study, is in accordance with the needs for classification and has the following aims:

- i. to obtain a relevè sequence where (a) relevè-groups can be formed, based on floristic similarity; and (b) the relevè-



groups thus formed are also sequenced according to floristic similarity. This is not necessarily the same as simply sequencing the relevès according to similarity because of the possible insertions of subsidiary gradients;

- ii. to delimit relevè-groups; and,
- iii. to obtain a species sequence where the relevè-groups and their relationships are emphasized.

#### 3.3.2.1 *Relevè sequencing*

The refinement of a preliminarily sequenced, relevè sequence often depends on an intuitive assessment and reduction of gaps in the matrix. The gaps are included blanks, between the first and last occurrence of a species in the matrix, caused by the absence of that species in the relevant relevès. The gaps before and after the first and last occurrence, respectively, of a species in the matrix are not taken into consideration, so that the sequence of relevès can affect the total. If all gaps were considered then the total would be the same irrespective of relevè sequence. The gaps can be caused by sampling omissions, outliers and, more often, the irregular spacing of many plant species. Visual sequencing of relevès according to gap minimization from the outset is, however, so difficult as to be virtually impossible.

Redundancy is the occurrence of species in two or more relevès to form pattern (Gauch 1982). Pattern can not be adequately quantified because value judgements have to be taken on pattern formation. Judgements can be based on a combination of number of



species and relevès and the constancy of species occurrence in the relevès forming the pattern. These judgements are subjective.

Gauch (1982) defines pattern as the co-ordinated occurrence and "noise", the unco-ordinated occurrence of species in the matrix. He regards "noise" and pattern as opposites, so that if "noise" is reduced in a matrix, then pattern is enhanced. Although Gauch (1982) does not regard "noise" as quantifiable, the effect of noise can be attributed to gaps in the matrix in terms of included blanks. The included blanks are species absences between the first and last occurrence of each species in the matrix, for a given relevè sequence. This can be quantified by addition.

For a given relevè sequence the included gaps for a single species are referred to as "separation units" and for all the species as "total separation units". The total separation units are inversely proportional to redundancy in the matrix and hence are related to classification adequacy. The optimum classification is that where the total separation units are a minimum. This presupposes that each relevè be tested in each possible position. The resulting possibilities are:  $n!$ , where  $n$  is the total number of relevès. A program was written to determine the time it would take to test these possibilities on a Burroughs B7900 mainframe computer, for a small matrix of 25 relevès x 33 species. The answer was in excess of 2 000 million, million years. Even excluding mirror-image sequences this is impracticable.

A heuristic approach, where not all, but only the best possibili-



ties are tested, was adopted. Clearly, a similarity index could also play a role in excluding many possible permutations. However, the initial relevè sequence could still affect the final sequence. It was found, for example, that the final sequence using TWINSPAN (Hill 1979) could change when the initial sequence was altered (see Chapter 4). A solution to this problem could be to standardize the initial sequence in as unique a manner as possible.

#### 3.3.2.1 A) *Commonality sequence*

The sampling sequence, and hence relevè numbering, is often highly arbitrary. The aim of standardizing the relevè sequence is, therefore, to obtain a sequence in which the first relevè would be the same, irrespective of the sampling sequence. This implies that each relevè should have some sort of unique attribute independent of any relevè sequence. A method of achieving this aim, is to count the occurrences in the matrix, of all the species in a relevè. This is done for each relevè in the matrix. The relevès can then be sequenced, either ascending or descending, according to the totals for each relevè. The extremes of this sequence, have in tests, proved to be unique although some similarity can exist between some intermediate relevès. This sequence is called the commonality sequence because the one extreme of the sequence represents relevès which have the most generally occurring species in common and the other extreme the least generally occurring species in common.



TABLE 3.6.- Illustration of the possible permutations of species presence and absence (rows) in three relevès (columns). Presence is indicated by a "+" and absence by a "0".

		Columns		
		1	2	3
Rows	1	+	+	+
	2	+	+	0
	3	0	+	+
	4	+	0	+
	5	+	0	0
	6	0	+	0
	7	0	0	+
	8	0	0	0



### 3.3.2.1 B) *Similarity sequence*

A problem encountered with similarity co-efficients such as those of Jaccard, Sorensen, Ellenberg and Gleason (Mueller-Dombois & Ellenberg 1974) is that relevès are compared in pairs in isolation, with no cognizance taken of the matrix of which they form a part. For example two adjacent relevès in a matrix, representing two different communities could have an extremely low similarity co-efficient. This is because of the weight attached to negative association, generally found in similarity co-efficients.

If relevès are considered in three's (Table 3.6), then in terms of the central relevès position in the matrix, the following can be stated:

- i. positive associations exist in the case of the first three rows;
- ii. a possible positive association could exist if single absence resulted from sampling omission or irregular spacing, in the fourth row;
- iii. no positive association exists, but negative association can not be shown because of sampling omissions or irregular spacings in the fifth to seventh rows; and,
- iv. absence in all three columns could be considered a similarity but not a negative association, in the eighth row.

Therefore, if only the first two relevès are considered (Table 3.6) then positive associations exist for rows 1 and 2 (joint occurrences), possible positive associations exist for rows 3, 4, 5 and 6 (absences could be because of sampling omissions), and no



negative association exists in rows 7 and 8 (joint absence could indicate similarity).

Similarity weighting for a pair of relevès is, therefore, based on positive and possible positive associations with no negative values given, as negative associations cannot be shown for two relevès. The weightings given are as follows:

Positive association 2

Possible positive association 1.

The similarity co-efficient (C) for comparing two relevès in a matrix is, therefore:

$$C\% = \frac{1/2 S + J}{S + J} \times 100 \quad \text{where}$$

S = number of single occurrences in both relevès and

J = number of joint occurrences.

### 3.3.2.1 C) *Separation unit sequence*

After standardization of the sampling sequence of relevès and sequencing according to similarity, the total separation units for the matrix is calculated. The first relevè is then moved, successively, one position in the relevè sequence and the total separation units are calculated after each move. The relevè sequence with the lowest total separation units is retained and the process is repeated with the new relevè sequence. This results in  $n^2 - n + 1$  calculations of total separation units, where n is the total number of relevès, after all the relevès have been moved in this manner.





The relevè sequence with the lowest total separation units is then reversed and the entire process is repeated successively until no decrease in total separation units is achieved. The reversal of the relevè sequence tends to average the position of a relevè where more than one position for that relevè exists.

The last retained sequence is then used to determine new similarity and separation unit sequences and these processes are iterated until the total separation units increase when the sequence with the lowest total separation units is reversed and the iteration then continues. When the lowest total separation units have been achieved in this manner, the iterations are terminated and the relevant relevè sequence is retained. This relevè sequence should then represent a matrix with a minimum of included blanks. Species occurring in all or a single relevè do not affect the position of the relevès in which they occur, in the relevè sequence.

#### 3.3.2.2 *Relevè grouping*

After the relevès have been sequenced the next step in the classification process is to delimit relevè-groups, some of which could represent plant communities. If the optimum relevè sequence has been achieved then it could be expected that the relevès that best represent a relevè-group would be situated towards the middle of the relevant relevè-group, because of a sine curve effect of the increment and decrement of species. Furthermore, if the relevè-groups are differentiated by different species then the first and last occurrence of each species distribution in the matrix could be



significant, in terms of relevè-group delimitation. Thus a sine curve, or approximation thereof, could be expected from the values obtained by the difference between the number of species representing first and last occurrences in the matrix, for each relevè. Species occurring in all the relevès or in a single relevè are ignored.

The number of relevè-groups to be recognized depends on the degree of difference permitted between relevès. Clearly, this will relate to the scale of the study, and the number of relevè-groups should be closest to the original number of vegetation units stratified, for objectivity. Thus, scale, based on the stratification process, is taken into account in the classification process. However, the programs allow flexibility in the number of relevè-groups that can be recognized so that data can be classified, even if the stratification and sampling processes are not according to scale. Relevè-group delimiters are inserted into the relevè sequence automatically, at the appropriate positions.

#### 3.3.2.3 *Species sequencing*

The final step in vegetation classification is to sequence species so that the relevè-groups and their relationships are emphasized. This is achieved by increasing pattern through species sequencing whereby species with similar distributions according to relevè-groups or combinations of relevè-groups are placed together in species-groups. A balance must also be achieved between outliers and redundancy or pattern. Various combinations of outliers to re-



dundancy were tested without definite ratios being evident. In terms of total separation units, however, single occurrences of species in the matrix are not significant as they cannot contribute to pattern. Similarly a single occurrence of a species in a relevè-group or combination of relevè-groups cannot be considered significant for the group or groups of relevès because no pattern is formed. Therefore, minimum species occurrence is taken as two, for a relevè-group or combination of relevè-groups.

Outliers are also limited to single occurrences in any relevè-group, provided that the total number of outliers does not exceed 38% of the total for that species in the matrix, because occurrences of more than one in a relevè-group could contribute to pattern. The 38% limit for outliers is based on tests and also corresponds with the Fibonacci fraction. The relationship between outliers and pattern can be considered to be based on minimum outliers. It must be emphasized that "noise" in terms of included blanks is present regardless of species sequence and that sequencing species does not affect relevè grouping or the relationships between relevè-groups. Furthermore, the maximum number of outliers that can be allowed in a given matrix is dependent on an appreciation of the pattern improvement that can be effected. This is obviously subjective, but does not affect the relevè grouping or the relationships between relevè-groups, as has been stated.

Species are also sequenced in descending order of occurrence in each species-group in the diagnostic species portion of the matrix. Species-group order is according to presence in:

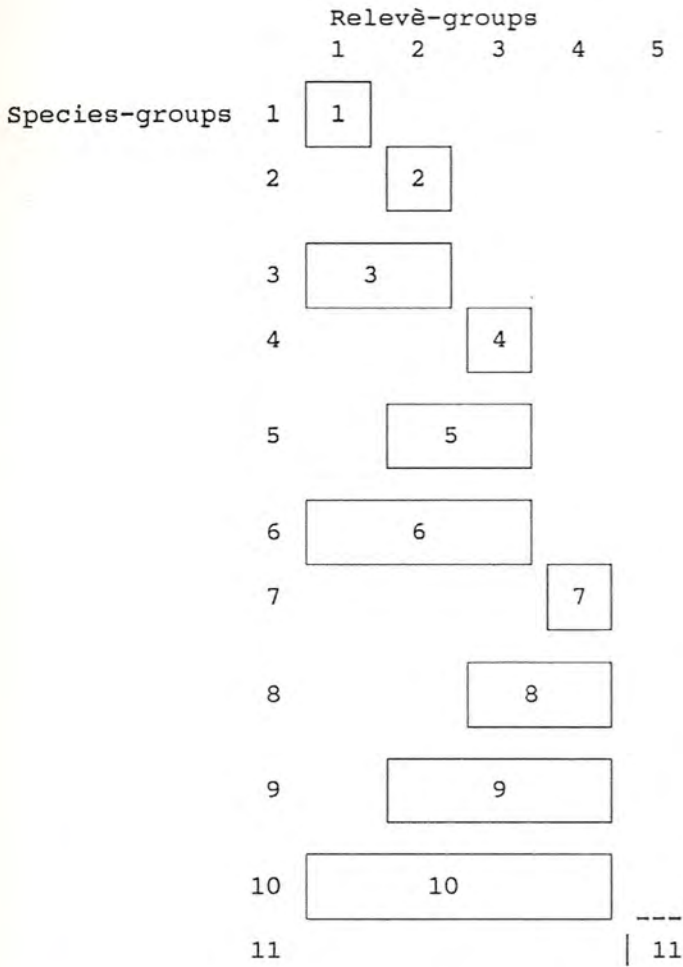


FIGURE 3.8. - Simplified schematic diagram showing the sequence of species-groups construction, according to the relevè group sequence.



- i. first relevè-group;
- ii. second relevè-group;
- iii. first and second relevè-groups;
- iv. third relevè-group;
- v. second and third relevè-groups; and,
- vi. first, second and third relevè-groups and so forth as illustrated in Figure 3.8.

Species-group delimiters are then inserted into the species sequence.

Species-groups with a single species can occur. The user should determine whether or not this is significant and either retain or combine such species-groups. The flexibility that is possible with species sequencing can be carried even further by combining species-groups, consisting of two or more relevè-groups, to correspond with the limits of environmental gradients (section 3.5.2.1 C). This simplifies the matrix, in terms of species-groups, without loss of information.

#### 3.4 VERIFICATION

A classification should be tested to ensure its validity, because more than one solution to a classification could be possible. This testing procedure can be referred to as a verification process.

##### 3.4.1 Background



Wenger (1974) stresses the need for confirmation of "coincidence between the pattern on the phytosociological table and specific habitat conditions". Apart from gaining insight into floristic and habitat relationships, the degree of "co-incidence" can also confirm a classification. However, if it is accepted that each species in the classified table has its own particular distribution according to factors which may limit it, then it is feasible that any combination of relevés could produce some or other correspondence with one or other environmental variable. Correspondence between community and habitat is, therefore, not necessarily verification of the validity of a classification, but can support a classification. The total separation units is a rating of the adequacy of a classification but the rating obtained is not necessarily the minimum total separation units because the relevés are not tested in every possible position. Other means should, therefore, be sought to jointly verify a classification.

#### 3.4.2 Methods applied

The following sections describe some methods in which a classification can be verified.

##### 3.4.2.1 *Classification efficiency*

A measure of the efficiency of a classification is the relationship between all the blanks (not just included blanks) in the matrix and the total separation units. This is a relative value, with which classifications can be compared, unlike the total separation



units, which is an absolute value. The classification efficiency (E) is calculated as follows:

$$E\% = 100 - \frac{(TSU \times 100)}{AG}$$

where TSU = total separation units and AG = all gaps in the matrix. A matrix with no noise and correctly classified would, therefore, have a classification efficiency of 100%. It appears that, in tests, adequately classified matrices have classification efficiency values of above 60%. Values markedly lower than this could indicate shortcomings in the classification process or data.

#### 3.4.2.2 *Spatial relationships*

The next step suggested in verifying the adequacy of a classification is to determine the spatial relationships between relevè-groups. This can be done by using an overlay technique. In general relevè-groups should exhibit a relatively high degree of spatial integrity, except where the scale is such that mosaics can be formed by relevè-groups. A low degree of spatial integrity could indicate shortcomings in the classification process or data.

A measure of correspondence in the spatial relationships between the stratification process and the classification process can be indicative of classification adequacy. The degree of correspondence can be quantified by comparing the two number sequences; the one representing the relevès grouped according to the stratification; the other representing the relevès grouped according to the classi-



a)

Habitat variable ranges

1 Altitude (m <sup>10</sup> )	40-45	35-42	25-42	31-47	55-68	53-70	52-65	60-83
2 Soil depth (cm)	10-14	11-13	15-20	16-18	10-13	12-16	17-18	17-19
3 Soil texture	Sand	Clay	Sand	Clay	Sand	Clay	Sand	Clay
4 Communities:	1	2	3	4	5	6	7	8

b)

Habitat variable ranges

2 Soil depth (cm)	10-14	15-20	11-13	16-18	10-13	17-18	12-16	17-19
1 Altitude (m <sup>10</sup> )	40-45	25-42	35-42	31-47	55-68	52-65	53-70	60-83
3 Soil texture	Sand	Sand	Clay	Clay	Sand	Sand	Clay	Clay
4 Communities:	1	3	2	4	5	7	6	8

c)

Soil depth listed according to ascending range

10-13 10-14 11-13 12-16 15-20 16-18 17-18 17-19

FIGURE 3.9. - Hypothetical model illustrating the sequence of three environmental factors, namely, altitude, soil depth and soil texture to hierarchically differentiate eight communities (a). If soil depth is first in the sequence (b) then differentiation of only four groups of two communities each, occurs. This is because in (b) the ranges overlap so that no groups can be formed, as can be seen in (c), whereas, the division of soil depth by altitude in (a) creates non-intersecting sub-groups of soil depth.





ication. A program to determine correspondence between two sets of grouped number sequences is included in the PHYTOTAB-PC package. The degree to which the relevè-groups in both sequences intersect indicates the amount of correspondence. Furthermore, the stratification can form a basic hypothesis which is tested with the classification. A low degree of correspondence between the two processes can invalidate the hypothesis, whereas, the converse can also provide the basis for mapping.

#### 3.4.2.3 *Floristic and habitat relationships*

Not only can an examination of floristic and habitat relationships aid in verifying the classification process but can also lead to an understanding of the causes of community differentiation. Obviously, the correspondence process is limited to those habitat factors which are available, such as those that have been recorded or derived. Factors not available, could cause problems in the correspondence process as the tested relationships between the habitat factors and floristics would then be incomplete.

Correspondence between floristics and habitat can be assessed with ordination techniques using programs such as DECORANA (Hill 1989b) and CANOCO (Ter Braak 1987). However, the former program requires overlaying selected environmental variables over the floristic scattergram to assess correspondence, and the latter program combines environment and floristics in a single classification process so that the classification is not based on floristics alone.

The following method is based on a simple hierarchical model which



tests direct correspondence between relevè-groups and the associated habitat factors available on a relevè basis. The first step is to construct a hypothetical model, as in Figure 3.9a. The hypothetical model is merely to determine the sequence of habitat factors, because the sequence, being hierarchical, is necessary in determining correspondence. This is illustrated in Figure 3.9b where fewer communities are differentiated because the sequence has changed. It is not possible because of permutation limitations to test all possible sequences programmatically, hence the necessity to construct the hypothetical model. Such a hypothetical model could also be of benefit, in terms of relevancy and economy, in determining which habitat factors to record in the field. The sequence of habitat factors should usually be from the general to the particular. For example, soil texture should not precede mean annual rainfall.

Starting from the first, most general, habitat factor, as determined with the model, the ranges of that habitat factor are determined for each relevè-group. Each unique set of ranges for one or more relevè-groups is allocated a different class symbol. If all the range sets have interceptions with other sets then a single class symbol is allocated and no correspondence for that habitat factor and the relevè-group is implied.

The next habitat factor is then tested in the same manner, but separately within each of the unique sets, determined with the previous habitat factor. The process is repeated for all the



habitat factors. Individual relevè-groups can be differentiated at various levels in the habitat factor hierarchy. If a poor differentiation of relevè-groups is achieved then the following conditions could apply:

- i. the classification is inadequate;
- ii. the habitat factor selection is inadequate;
- iii. the model is inadequate; or,
- iv. the limits for a habitat factor do not correspond precisely with the limits of a relevè-group. In other words, adjacent relevès, representing two communities have habitat factor values which do not correspond with the border between the two communities.

In the case of the last condition the process can be repeated using either nodal relevès or omitting relevès adjacent to relevè-group delimiters.

#### 3.4.2.4 *Classification and field relationships*

The final step in the verification process is comparing the classified relevè-groups with the units to be mapped, on the ground. This step is not to confirm the occurrence, on the ground, of the plant species in the relevant relevè-groups, which should clearly be the case, but the following:

- i. to assess the degree to which the plant species of the relevant relevè-group are representative of the unit to be mapped;
- ii. to assess the variation represented by the relevès of the relevant relevè-group and the degree to which this corresponds



with the variation of the unit to be mapped, at the relevant scale;

- iii. to assess the value of plant species which can be used for plant community diagnosis as a means of identifying the plant community in the field. Here plant spacing, size and regularity of occurrence is important so that the community can be easily recognized;
- iv. to assess the reliability of the hypothesized floristic and habitat relationships;
- v. to assess the validity of the borders of the units to be mapped; and,
- vi. to assess which of the relevè-groups are mappable at the relevant scale.

If the classification has a classification efficiency of about 60% or more, does not invalidate the stratification, and corresponds well with the habitat data and the field assessment, then stratification, habitat correspondence and classification can all be deemed satisfactory. However, lack of correspondence in any of these processes does not necessarily invalidate the classification, but rather the process concerned could be questioned. The relevancy of the classification will ultimately depend on the uses derived therefrom.

### 3.5 DERIVATIVES

The term derivative is taken here, to mean information that can be derived directly (primary derivatives) or indirectly (secondary



derivatives) from the classification. Spatial relationships such as the locality of vegetation units or their co-occurrence with mappable factors such as geology or soils can be illustrated with maps, map overlays or integrated maps. Mapped vegetation units, as proposed in this study, are the product of the stratification process, which can be modified by the classification, rather than being derivatives of the classification.

### 3.5.1 Background

If a classification and a map illustrating spatial relationships of vegetation units are the end-products of a study then the value of such a study will be limited to those workers who are able to interpret and derive additional information from the classification. Those best able to interpret and derive information from a classification should be the compilers thereof. Information that can be derived from a classification will largely determine the uses for classifications and hence the need for such work. It is, therefore, of great importance to workers in this field to derive as much information as possible from classifications, so that the demand for such work can be increased.

### 3.5.2 Methods applied

The classified matrix can include a summary of coded variables for each habitat factor, above each relevè, to form a passive classification of habitat variables, (Deall & Westfall 1989) which can be visually correlated with the relevè-groups.



There is usually much redundancy in a classification and the information contained therein can be considerably reduced by constructing a synoptic matrix (Werger 1974) where relevè-groups are reduced to a single column according to constancy based on species presence. A five class scale is generally used (Mueller-Dombois & Ellenberg 1974). This tends to visually over-emphasize the importance of single occurrences which can be overcome by using:

- i. a "+" symbol for single occurrences; or,
- ii. the plant number scale described in section 3.2.3.

The classified habitat data and synoptic matrix, although dependent on the classification are not considered derivatives, in this study.

#### 3.5.2.1 *Primary derivatives*

The primary derivatives that follow, can be derived from the synoptic matrix.

##### 3.5.2.1 A) *Plant communities*

Braun-Blanquet (1928) defined a plant community as being characterized by "its own species". As more data were gathered, syntheses of matrices took place and community distinctions became less clear. Terms such as exclusive, selective and preferential character species as well as differential, territorial and regional species were introduced to clarify the situation (Werger 1974). Clearly, unless a plant community is restricted to a particular study area, its composition, diagnosis and description for



the study area, could differ considerably from that of the community of which it is a part. This relates to the proportion of the community variation present in the study area, which is usually not known where the community is not entirely included in a study area. The following is, therefore, suggested for individual study areas where no synthesis with other study areas is involved:

- i. a relevè-group represents a plant community only where the plant community is completely included in the study area concerned. A relevè-group cannot adequately represent a plant community that is only partially included in a study area, because the proportion of the variation represented by the relevè-group, is not known;
- ii. the species of the species-group that is restricted to the relevè-group representing the plant community, is the diagnosis of that plant community, in the study area, and are referred to as community diagnostic species;
- iii. formal ranking of plant communities should not be applied for individual study areas, but only with syntheses, because there is usually insufficient information to elucidate relationships;
- iv. informal ranking of plant communities within a study area is based on combined relevè-groups, where conditions i. and ii. apply for the combined group;
- v. relevè-groups without community diagnostic species can represent ecotones, extremes in variation for a particular community, or communities for which the particular scale or extent in the study area is inappropriate; or,

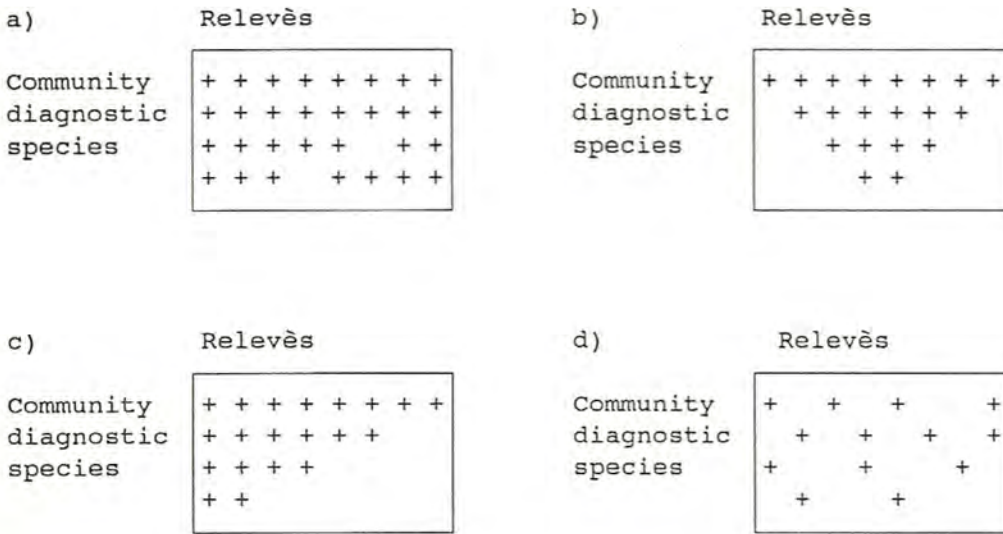


FIGURE 3.10. - Species presence in community diagnostic species-groups: a) few blanks indicating a well-defined community corresponding to an abrupt environmental change at the community limits; b) species presence approximates a normal curve, indicating a well-defined community but with a less abrupt change at the community limits than in (a); c) a variation of (b) where the relevès have been re-arranged so that the strongest community expression, in terms of species presence, is at the left of the group, decreasing towards the right. This pattern does not adequately show spatial relationships because the relevès on the extreme right can often be adjacent to different communities; and d) a poorly-defined community.





vi. plant community composition by all the species, both diagnostic and non-diagnostic, present in the relevè-group/s representing the plant community. This is, however, only a sample of the species present in the community.

This information as well as plant community names can be obtained directly from the synoptic matrix.

#### 3.5.2.1. B) *Community definition*

Community diagnostic species pattern can be used to infer the degree to which the community is defined, if the relevè sequence is according to minimum total separation units. For example, community diagnostic species with no or few gaps, as illustrated in Figure 3.10a, should indicate a well-defined community which is differentiated by an abrupt environmental change at the community limits and can, therefore, be mapped with a high degree of precision.

A community where the community diagnostic species pattern approximates a normal curve (Figure 3.10b) should indicate a less abrupt environmental change at the community limits but the community is still well-defined. Mapping precision will be according to the width of the ecotone. Depending on sample skewness, the relevès best expressing community composition are those in the central part of the relevè-group. The relevès can be re-arranged within the relevè-group as illustrated in Figure 3.10c without affecting community relationships. However, the total separation units will then increase because the original relevè

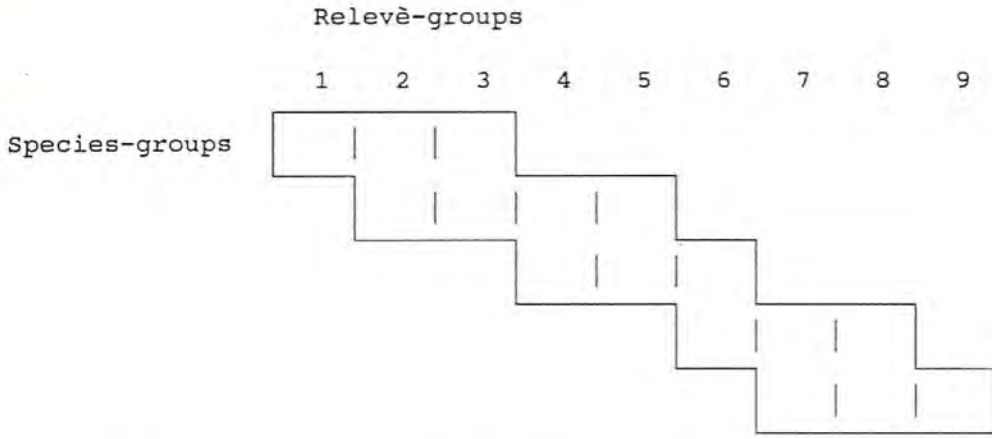


FIGURE 3.11. - Species-groups representing two or more communities, sequenced to correspond to an environmental gradient of which both the upper and lower limits for differentiated communities can be ascertained. This also applies to the horizontal mirror-image of the illustrated pattern.

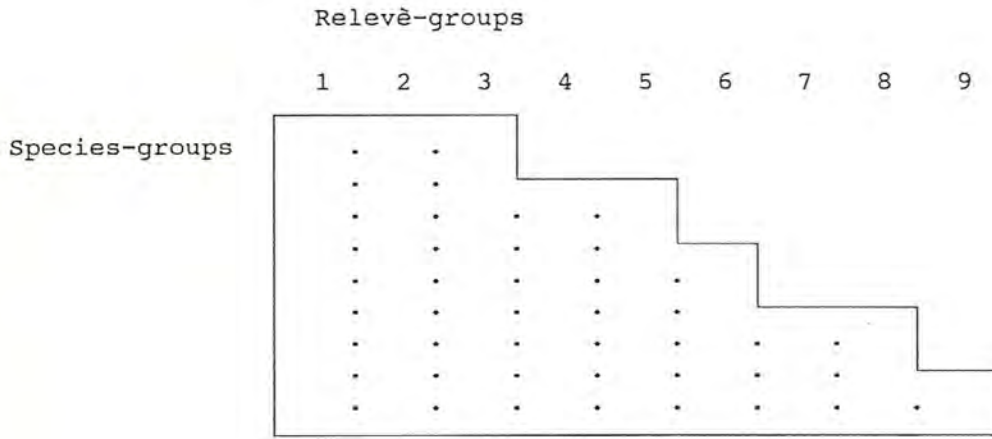


FIGURE 3.12. - Species-groups representing two or more communities, sequenced to correspond to an environmental gradient of which only the upper or lower limits for differentiated communities can be ascertained. This also applies to the horizontal mirror-image of the illustrated pattern.



sequence is based on minimum total separation units.

A community diagnostic species pattern with many blanks and irregular spacing of occurrences (Figure 3.10d) can be attributed to any of the following:

- i. a poorly defined community;
- ii. a gradual environmental change over the community limits;
- iii. heterogeneity caused by widespread disturbance over the community; or,
- iv. inadequate sample size.

Mapping precision of such a community would be low.

#### 3.5.2.1 C) *Gradients*

The matrix can be further simplified by grouping overlapping species-groups, representing two or more relevè-groups, to correspond with environmental gradients. Two main patterns can emerge, namely, a) a gradient with both upper and lower environmental limits (Figure 3.11), or b) a gradient with either upper or lower environmental limits only (Figure 3.12). Both these types can be mirror-images of those illustrated in Figures 3.11 & 3.12. The limits should be evident after correlation of communities with habitat (section 3.4.2.3).

The simplification of the species-groups, in this way, will not affect the total separation units or the relationships between relevès or relevè-groups, so that information is not lost from the matrix but rather additional pattern becomes evident. The matrix

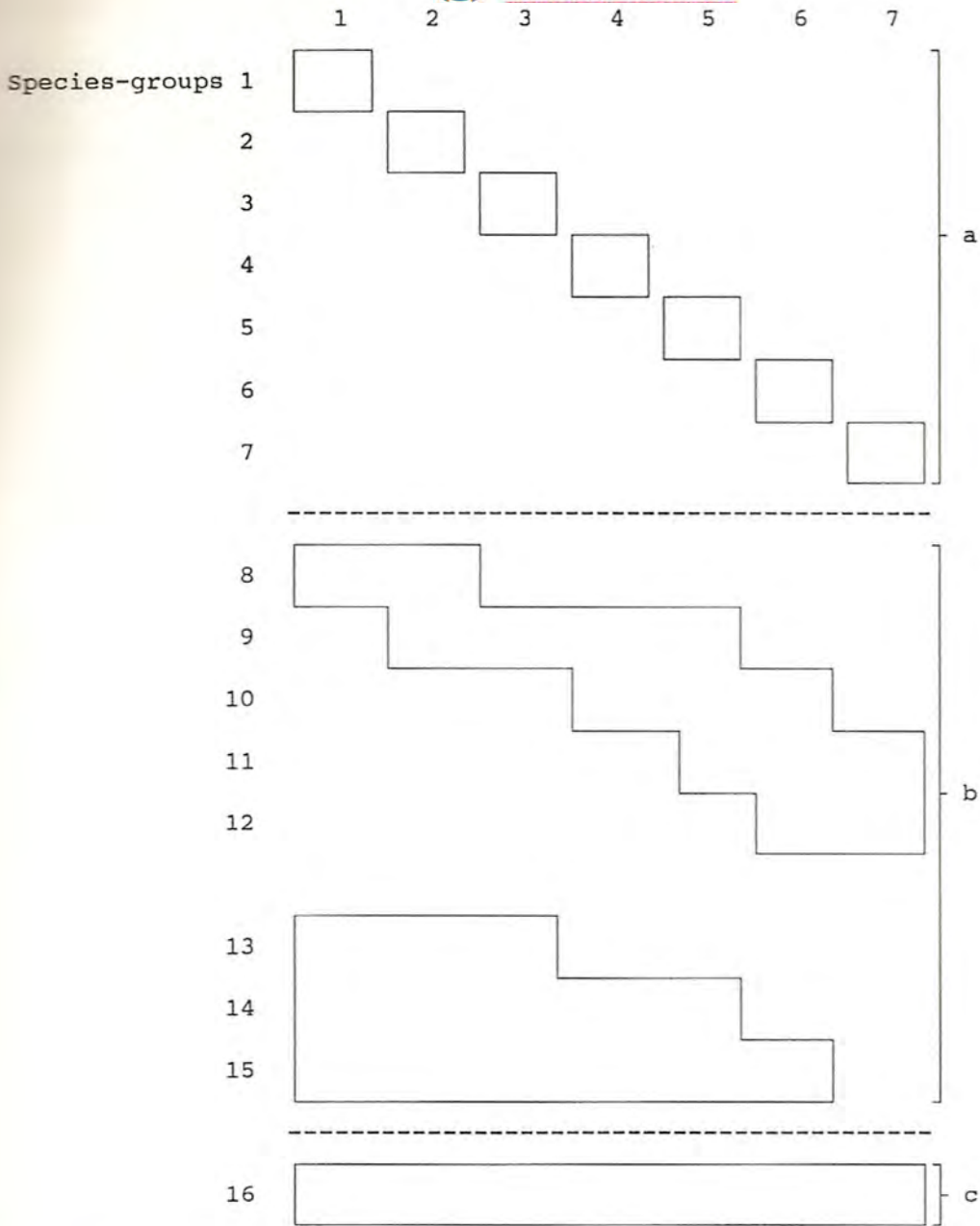


FIGURE 3.13. - Schematic illustration of the simplification of a phytosociological table into 3 sections: a) community diagnostic species-groups; b) species-groups corresponding to environmental gradients. Horizontal mirror-images of the illustrated pattern also apply; and c) non-diagnostic species.



will now consist of three parts, (Figure 3.13) namely, a) community diagnostic species-groups; b) species-groups indicating environmental gradients and c) non-diagnostic species.

#### 3.5.2.2 *Secondary derivatives*

The secondary derivatives are those results computed from the recorded data but are dependent on the classification for grouping.

##### 3.5.2.2 A) *Structure*

The general correspondence between the recorded growth forms and height (section 3.2.4), permit layer diagrams of vegetation structure, according to total recorded canopy cover for all species within each growth form class (after Ito 1979) for each community, to be made. This also allows differentiation of the ground layer into grasses and forbs which is not possible where structure is recorded as total cover per height class.

##### 3.5.2.2 B) *Community composition analysis*

A plant species which utilizes a large resource-space per individual plant is likely to:

- i. have a correspondingly high crown cover;
- ii. have a high resource-space requirement, in the mature phase;
- and,
- iii. have a high cover in relation to its frequency.



The converse could also be valid, namely, that a plant species with a low cover in relation to its frequency of occurrence, requires a correspondingly small resource-space per individual plant in the mature phase. It is obvious that in comparisons of cover-to-frequency ratios, grasses for example, cannot be compared with trees as the disparity in resource-space requirements is too great in the mature phase. These statements are generalizations as individual plants can differ considerably in their reactions to external factors. However, the elucidation of intra-community dynamics in this study is dependent on species reactions, not that of individual plants.

Plant species within a defined plant community are, therefore, categorized according to the growth form classes defined in section 3.2.4. Cover-to-frequency ratios for each species, are then calculated and a linear regression for the cover and frequencies in each growth form class is determined. A linear relation between cover and frequency is assumed. The expected cover for the actual frequency of each species, according to the regressions is also calculated. The species are then arranged in decreasing order of the differences between actual and predicted cover.

Species outside the standard error of the mean, for the cover-to-frequency regressions, form two distinct groups, namely, those with a higher cover, and those with a lower cover than the standard error of the mean. The first group is referred to as strong competitor species because of their resource-space



requirements and those in the second group are referred to as weak competitor species. Generally most species fall within a range between these two groups forming a third group which can be referred to as the normal competition range. Thus, species within a community can be categorized into 15 different classes (i.e. three competitor groups in each of five growth form classes) which relate to the species present in the community and the effect of their differing resource-space requirements.

An extension of these programmatic calculations is the determination of cover proportions, of the growth form classes, within a community and the total cover within each growth form class.

#### 3.5.2.2 C) *Stand phase analysis*

After the community composition analysis, it is then possible to compare the cover of any stand of vegetation within the community, at the scale at which the community was sampled, with that obtained for the community with the community composition analysis. For practical purposes the community frequency can be used and only cover differences taken into account. Thus the differences in cover between the stand and the community are emphasized.

Five distinct phases, in the comparison between stand and community, can be identified, namely:

Phase 0: all the species fall into the normal competition range.

i.e. there are no strong or weak competitor classes and,



therefore, no class is advantaged or disadvantaged;

Phase 1: weak competitors occur in the grass growth form class,  
indicating possible selective utilization;

Phase 2: strong competitors occur in the grass growth form class,  
indicating excessive grass utilization;

Phase 3: weak competitors occur in the woody growth form classes.

The class could be dependent on the biome or vegetation formation. In the case of savanna a progression through the classes from dwarf shrubs to trees is possible. Some of these species in the juvenile stage are potential strong competitors; and

Phase 4: Strong competitors occur in the woody growth form classes. The classes, as in Phase 3, could be biome or formation dependent. A tendency towards unidominance and/or moribundity could exist.

These main phases indicate the stage of vegetation development of the stand relative to the community of which the stand is a part. Dynamics is thus inferred, assuming progression through the stages. Reversal of the sequence is probably a rare occurrence in southern Africa.

#### 3.5.2.2 D) *Community cover assessment*

The ideal vegetation composition for utilization is dependent on the type of utilization applied. For example, the vegetation composition ideally suited to cattle grazing is not necessarily the ideal for goats. It would appear far more efficient to adapt the type of animal utilization and degree of utilization to the





specific vegetation resource than to try to manage the resource to suit the form of utilization.

Vegetation condition can, therefore, be assessed from two perspectives, namely, a) suitability for a particular type of utilization and b) sustained production and resilience or ecological condition. In terms of the justification (section 1.3) ecological condition should primarily relate to the ability of the vegetation to protect the soil. This implies an adequate ground cover with increasing root penetration as slope increases, for soil binding. The following, accordingly, describes an elementary method for assessing tree cover adequacy.

Tree cover values, determined in the community composition analysis, and mean community slope are used as input. The relationship between tree cover and slope was taken after trial and error to be :

$$c = \tan s \times 100$$

where  $c$  = tree cover (%) and  $s$  = slope ( $^{\circ}$ ).

This is merely a provisional indication of the minimum tree cover required for root soil binding which varies from 0% on level ground to 100% on a  $45^{\circ}$  slope. No assessment is made of the adequacy of ground cover.

### 3.6 PHYTOTAB-PC

PHYTOTAB-PC is a program package to facilitate the methods des-



cribed in this chapter and also facilitate handling of phytosociological data. The program package consists of 147 program and text files written in TURBO BASIC, a compiled version of BASIC, for an AT or 286 IBM-PC with 640 KB RAM, 20 MB hard disk and a 360 KB, 5¼" floppy disk drive, minimum configuration. The package is divided into two parts, both menu driven, namely,

- a) directory C:\PT for classification and processing, and
- b) sub-directory C:\PT\PD for a floristic data bank.

### 3.6.1 Classification and processing

The programs are executed from the C:\PT directory, or if ";C:\PT" is included in the path of the AUTOEXEC.BAT file, from any directory, by typing "PT". The main menu selection is then invoked. Menu items are executed by entering the symbol, which is bound by a single bracket, preceding the item. Application sequence is normally according to the numeric main menu items. Alphabetic main menu items are used for documentation and utility programs which can be executed when necessary. The functions and procedures for the program package are described, according to the menus, as follows:

#### D) Documentation

This details setup of directories, program transfer and program initiation as well as program usage by displaying or printing the on-line manual.

#### 1) Matrix Data Input

This initiates keyboard input of the matrix in either PHYTOTAB format: relevè (species; cover-abundance) or TABIN format: species



relevè; cover-abundance). Provision is made for a new data set or continuation of input. Species are abbreviated to an eight digit or eight character code.

## 2) File Handling

This option can be used if PC-WRITE (Version 3.00) is resident, to view and edit sequential files, as well as to back up and restore files to diskette or, for file transfer.

## 3) Transform Matrix to QMAT

QMAT is the random access file used for further processing. This process must be completed before any further processing can be done. Outputs are: relevè sequence file, species sequence file, codename file for species and relevè computer code numbers, TABIN format file (if PHYTOTAB format is the input) and QMAT file. The last file is a random access file for speed of computation. Corrections should be done to either the PHYTOTAB or TABIN format files and transformed again. It should be noted that the relevè and species sequence files contain the computer allocated numbers for the species and relevès. Relevè numbers need not, therefore, be in sequence or range from 1 to n. However, relevè data should not be split. For example, relevè number 2 cannot occur in part before and after another relevè number.

## 4) Relevè and Species Sequencing

Eleven options are available in two submenus. On the first submenu, the first option is used for programmatically sequencing relevès. Matrix size limitation for relevè sequencing, using this option, is the product of species and relevès not greater than 186 000. This option can also utilize considerable processing time, in the order of hours to days, depending on matrix size,



noise (see section 3.3.1). and processor.

The second option, for sequencing species, is executed after the relevè sequence has been determined, either by the user or using the previous option. The species sequencing option is also used to group relevès where required (see sections 3.3.2 & 3.3.3). The third option sequences the species abbreviations alphabetically and the relevè numbers numerically and the codename file can be updated accordingly to facilitate abbreviation checking or to match full species names when downloading. The fourth option is for inserting or deleting numbers in a sequence. Separation characters for species and relevè-group delimiters can be inserted using this option or an editor can be used. Deleting a number decreases the value of subsequent numbers by one. It is important to remember that the numbers on the relevè file are computer codes which must be converted to the original numbers, processed and then converted back. The fifth option invokes the second submenu.

On the second submenu the first four options permit keyboard entry of a relevè or species sequence, corrections to an existing sequence (or an editor can be used), and generation of a number sequence from 1 to n; reversal of an existing sequence (for a mirror-image sequence), which is not affected by the presence of delimiters; and a random sequence from 1 to n, without replicates. The fifth and sixth options are sequence number to sequence code conversions and *vice versa*. These options must be used if any of the previous four options are executed. The final option is a permutation of relevès to obtain the optimum relevè sequence, exclud-



ing mirror images, based on minimum total separation units. This option should not be used for more than about 20 relevès, as it is extremely time-consuming.

#### 5) Peripheral data input

This option invokes a submenu for input of species names, community names, table title, habitat data descriptions and habitat data classified codes. Other data which can be input are growth form codes for community composition analysis and check lists; cover codes and midpoint values used in the matrix and the actual values of environmental variables which can be either descriptive or numerical.

#### 6) Print files

Options 1 to 7 are used to print files that have been created, for checking or reference purposes. Option 8 prepares an alphabetic sequence of species names with, optionally, growth forms, species frequency and separation units per species, in PC-WRITE format for annotating as a checklist. This option can also be used if an alphabetic species listing is required.

#### 7) Print Working Table

This option permits printing of a matrix before habitat data and full species names have been added. The output includes total occurrences and separation units for each species; total species for each relevè; total separation units for all species; and percentage classification efficiency (see section 3.4.1).

#### 8) Print Final Table

This option permits printing of a matrix with full species names and other peripheral data.



## 9) Process Data

This option invokes a submenu for the following:

1. Synoptic table production in which the matrix values can be according to a five class scale; a five class scale with single occurrences indicated with a "+" or the 33 class plant number scale (see section 3.5). All options available to the full matrix can be applied here.
2. Community composition analysis which includes community structure (see sections 3.5.2.1 & 3.5.2.2).
3. Stand phase analysis (see section 3.5.2.3).
4. Community cover assessment (see section 3.5.2.4)
5. Community and habitat correlation with relationships indicated by means of a dendrogram (see section 3.4.3).
6. Species cover relationships (see section 3.2.3.7).

## U) Utility programs

The utility programs, being executable from the main menu, are referred to as internal utility programs. External utility programs do not form part of the PHYTOTAB-PC program package. The first three options with the invoked submenu are for correcting format errors that can occur after downloading matrix data or species names from a mainframe or data bank or to change the parameters for the species name file.

The fourth option is used to generate species sequences for plant identification keys (see section 3.2.2.8) and to split matrices for key or other purposes.

The fifth option converts the matrix data to the CONDENSED format



for the Cornell programs TWINSPAN (Hill 1979a) and DECORANA (Hill 1979b) or CANOCO (Ter Braak 1986, 1987). Should habitat data be required for the CANOCO programs these can be loaded, according to the necessary format, using an editor. Other format options available are those for Affinity Analysis (Scheiner & Istock 1987) and for the PHYTOTAB mainframe format (Westfall *et al.* 1982).

The sixth option is used for basic statistical applications on paired or unpaired variables. The applications include various means, range, standard deviation, variance as well as linear regressions. Optional output can be used as input for Harvard Graphics. Paired variables can also be sorted according to x or y values ascending.

The seventh option allows for fixed raster mapping according to grids, using the printer.

The eighth option is used to merge large sequential files, such as could be required for raster mapping.

The ninth option is used for inputting point co-ordinates for the ARC/INFO-PC geographic information system. This is required for mapping relevè localities.

The tenth option (C) converts cover codes to the plant number scale. This does not improve the precision of the original estimates.

The eleventh option (L) is used for generating and locating random

TABLE 3.7. - A model of the PHYTOTAB-PC program package, primary options with square brackets indicating keystroke sequences

Main output

Processes

A. Instructions

On-line manual

Documentation [D]

B. Classification

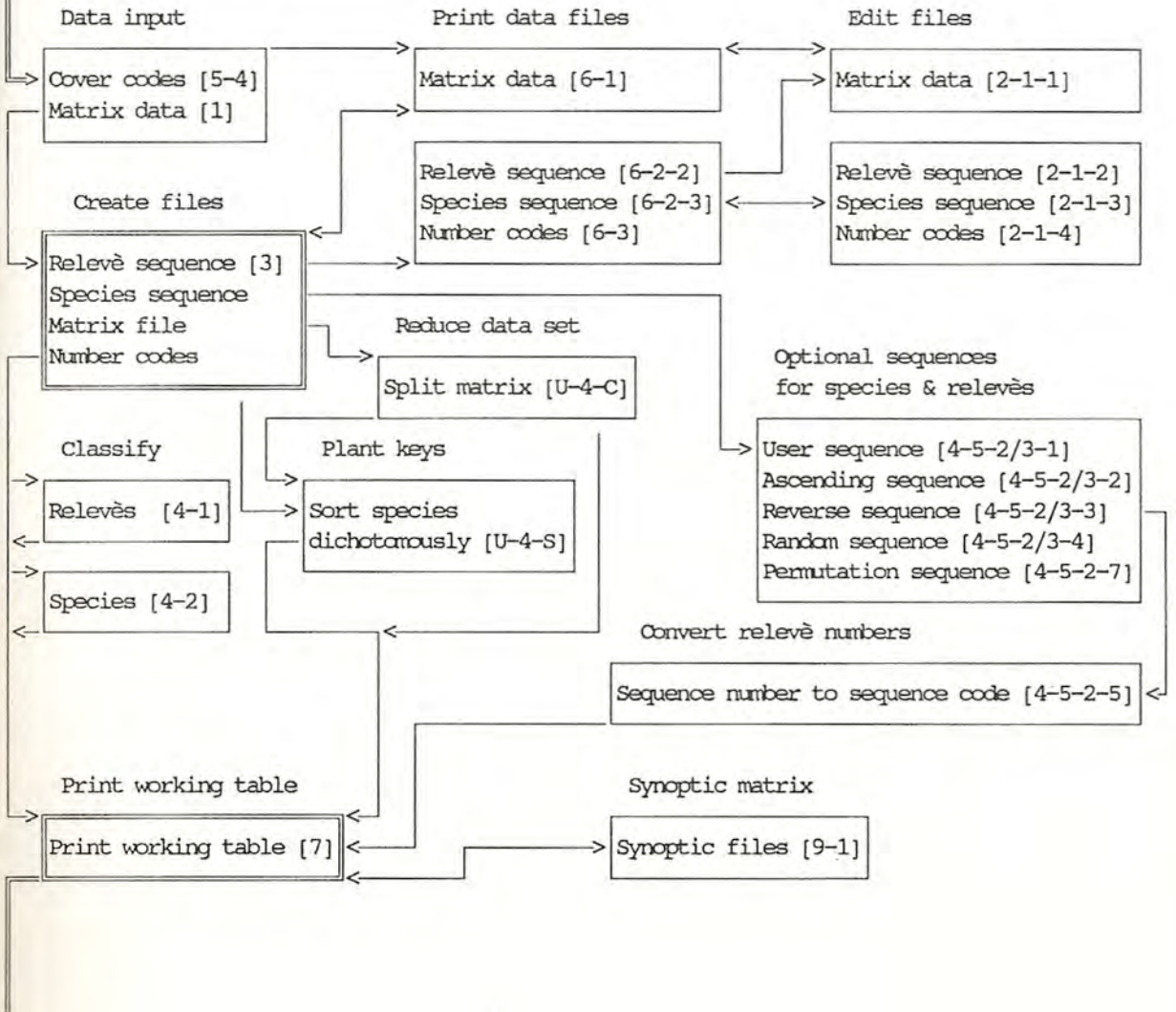






TABLE 3.7. (continued) - A model of the PHYTOTAB-PC program, package primary options with square brackets indicating keystroke sequences

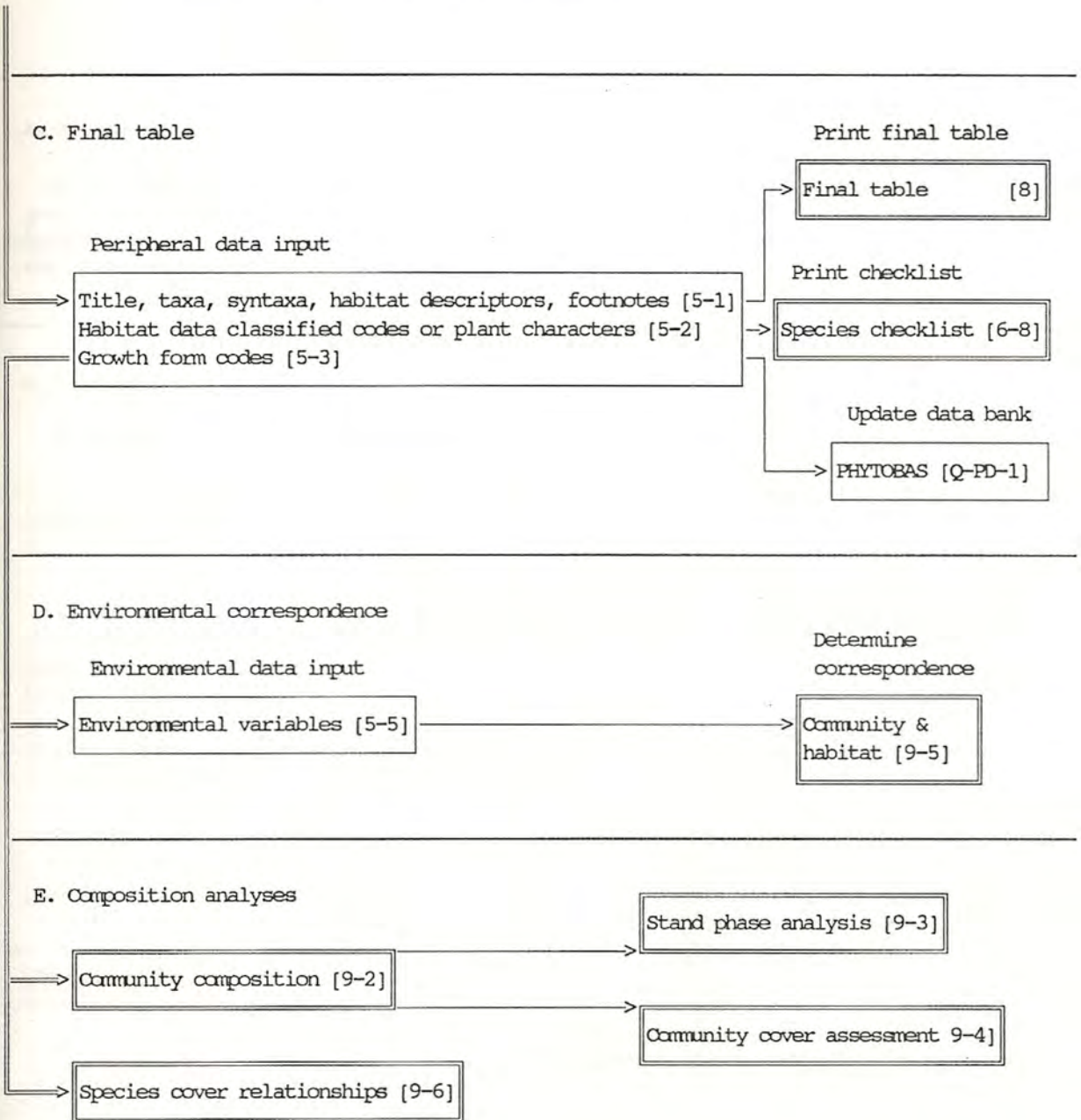




TABLE 3.7. (continued) - A model of the PHYTOTAB-PC program,  
package primary options with square brackets  
indicating keystroke sequences

F. Internal utility programs

Remove blanks from matrix file  
(usually when transferring data) [U-1]

Convert PHYTOBAS SPPNAM to \PT\DESCRI  
(for transfer from data bank) [U-2]

Redimension or remove names or blanks  
from file \PT\DESCRI [U-3]

Convert to other formats (CANOCO,  
AA, PHYTOTAB-mainframe ) [U-5]

Statistics / Linear regressions [U-6]

Digital mapping [U-7]

Merge large sequential files [U-8]

Co-ordinate conversions or generate  
random points [U-L]

Sample/stand dimensions [U-S]

Cover code conversions [U-C]

Number set comparisons [U-N]

G. External utility programs

3-D Ordination output -  
Plotter is required [CAN3D]

Species comparisons with PRE data  
set, search or add authors [SP]



points according to latitude, longitude co-ordinates and a 4 mm grid. Provision is also made for converting fractional degrees to or from decimal degrees and degrees, minutes and seconds.

The twelfth option (N) is for the comparison of two sets of grouped number sequences such as comparing two relevè sequences of the same data set.

The thirteenth option (S) is used for determining the following:

- i. stand dimensions from scale, stand area and stand radius;
- ii. sampling unit dimensions from species number and area; and,
- iii. species number estimation for a given area from species number per unit area.

Two function keys are operative for main menu items. The F1 key invokes the relevant section of the on-line manual and the F2 key invokes a problem solving guide.

All files generated with PT are saved to separate diskettes. This simplifies file maintenance with multiple users and if backup file copies are made both data integrity and security are ensured. File identification tags are printed automatically for each diskette. A flow chart of primary program options is shown in Table 3.7.

### 3.6.2 Data bank

This is a separate set of programs, on a subdirectory C:\PT\PD, functioning as a floristic data bank. The data bank programs are



executed from the C:\PT directory, or if ";C:\PT" is included in the path of the AUTOEXEC.BAT file, from any directory, by typing "PD". Four files are resident on a hard disk, namely, matrix data, species names, community names and an index. Data are loaded from diskettes generated with the classification programs. The main menu selections are the following:

1. Load data

This is from diskette in the PHYTOTAB-PC format, only after the final table has been printed and deemed satisfactory, if the programs are to be used for a data bank. The programs can also be used for analysing a single data set and splitting matrices.

2. List data

The following data can be listed: species names, both according to code or alphabetically; an index of data sets resident; relevè identifiers, being computer codes and original relevè numbers; species sequences being the order in which species should occur for a particular data set and community names, both numerically or alphabetically.

3. Search routines

Searches can be made starting with relevè identifiers, relevè numbers, data sets or multiple groups of relevè numbers. This permits subdivision and merging of data sets. The necessary files for classification are generated on diskette.

Searches can also be made starting from species codes, species names or truncated species names. The Boolean "and" logic applies when more than one name is entered for a search, i.e. all the species should be present in a relevè for the relevè to be select-



ed. Apart from creating diskettes for use with the classification programs, a list of species associated with selected species, on a relevè basis, can be obtained.

#### 4. Correct random files

This applies to the species and community name files. The index and floristic data can be corrected with an editor.

#### 5. Check species names

Species names on diskette, for table production, can be compared with those on the data base, for errors.

The data bank is limited by the capacity of the hard disk drive. Search and loading time is proportional to the amount of data present. For large data sets, an 80386 or 80486 processor is an obvious advantage.

### 3.6.3 External utility programs

Apart from an ASCII file editor such as PC-WRITE (version 3.00 or higher), the following programs can further facilitate vegetation analyses:

DECORANA: The CANOCO (Ter Braak 1986; 1987) version is preferred for ordination because of the options available and the inclusion of sample and species numbers in the scatter diagrams.

AA: Affinity analysis (Scheiner & Istock 1987) for analysing the variation of degrees of compositional relatedness among the communities of a data set.

CAN3D: This program is used for three-dimensional plotting of DECORANA output. A plotter is required.



SPECOM: (Westfall in prep.). This program is used for checking species names against the computerized National Herbarium synonym list and adding author names for checklists.

### 3.7 REFERENCES

- ACOCKS, J.P.H. 1975. Veld Types of South Africa. *Memoirs of the Botanical Survey of South Africa* 40: 1-128.
- BAKER, H.A. 1970. A key for the genus *Erica* using edge-punched cards. *Journal of South African Botany* 36: 151-156.
- BIANCHI, A.T.J. 1931. Een nieuwe determinatie-methode. *Tectona* 24: 884-893.
- BOLUS, H.M.L. 1958. Key to the genera of Mesembryanthemaceae. *Notes on Mesembryanthemum and allied Genera* 2: 392-403.
- BOUGHEY, A.S., BRIDGES, K.W. & IKEDA, A.G. 1968. An automated biological identification key. *Museum of Systematic Biology, University of California, Irvine, Research Series*, 2: i-xix, 1-35.
- BRAUN-BLANQUET, J. 1928. *Pflanzensoziologie*. 1 Aufl. Springer, Wien.
- COETZEE, B.J. 1975. A phytosociological classification of the Rustenburg Nature Reserve. *Bothalia* 11: 561-580.
- DEALL, G.B. & WESTFALL, R.H. 1989. Improving the resolution of floristic/habitat pattern correlations on phytosociological tables. *Bothalia* 19: 263-266.
- DUKE, J.A. 1969. On tropical tree seedlings. I. Seeds, seedling systems and systematics. *Annals of the Missouri Botanical Gardens* 56: 125-161.



- EDWARDS, D. 1983. A broad-scale structural classification of vegetation for practical purposes. *Bothalia* 14: 705-712.
- ELLIOTT, J.M. 1983. *Some methods for the statistical analysis of samples of benthic invertebrates*. Kendall, Wilson.
- F.S.S.A., 1974. Manual of soil analysis methods. *The Fertilizer Society of South Africa* 37: 1-65.
- GAUCH, H.G. 1982. *Multivariate analysis in community ecology*. Cambridge, New York.
- GIBBS RUSSELL, G.E., REID, C., VAN ROOY, J. & SMOOK, L. 1985. List of species of southern African plants. Edition 2. Recent literature and synonyms. Part 1. Cryptogams, Gymnosperms, Monocotyledons. *Memoirs of the Botanical Survey of South Africa* 51: 1-152.
- GIBBS RUSSELL, G.E., WELMAN, W.G., RETIEF, E., IMMELMAN, K.L., GERMISHUIZEN, G. PIENAAR, B.J., VAN WYK, M. & NICHOLAS, A. 1987. List of species of southern African plants. Edition 2. Recent literature and synonyms. Part 2. Dicotyledons. *Memoirs of the Botanical Survey of South Africa* 56: 1-270
- GLEN, H.F. 1974. *A taxonomic revision of the subtribe Gibbaeinae (Mesembryanthemaceae)*. M.Sc. thesis, University of Cape Town.
- GOODALL, D.W. 1968. Identification by computer. *BioScience* 18: 485-488.
- GYLLENBERG. H.G. 1965. A model for the computer identification of micro-organisms. *Journal of General Microbiology* 39: 401-405.
- HANSEN, B. & RAHN, K. 1969. Determination of Angiosperm families by means of a punched-card system. *Dansk botanisk Archiv* 26: 1-46 & 172 punched cards.



- HENKEL, J.S. 1934. *A field book of the woody plants of Natal and Zululand*. Robinson, Durban.
- HERRE, A.G.J. 1971. *Genera of Mesembryanthemaceae*. Tafelberg, Cape Town.
- HILL, M.O. 1979a. *TWINSpan - A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes*. Unpublished report. Cornell University, Ithaca, New York.
- HILL, M.O. 1979b. *DECORANA - A FORTRAN program for detrended correspondence analysis and reciprocal averaging*. Unpublished report. Cornell University, Ithaca, New York.
- ITO, K. 1979. A tentative study of stratification diagrams. In Miyawaki, A. & Okuda, S. (eds). *Vegetation und Landschaft Japans*. pp.357-368. The Yokohama Phytosociological Society, Yokohama.
- JANSEN, H. 1982. The geology of the Waterberg Basins in the Transvaal, Republic of South Africa. *Memoirs of the Geological Survey of South Africa* 71: 1-98.
- LAMARCK, J.B.A.P.M. de. 1778. *Flore Francaise* 3 vols. Imprimerie Royale, Paris.
- LAPAGE, S.P., BASCOMB, S. WILCOX, W.R. & CURTIS, M.A. 1970. Computer identification of bacteria. In A. Baillie & R.J. Gilbert (eds), *Automation, Mechanisation and Data Handling in Microbiology* 1-22. Academic Press, London.
- LAPAGE, S.P., BASCOMB, S. WILCOX, W.R. & CURTIS, M.A. 1973. Identification of bacteria by computer: general aspects and perspectives. *Journal of General Microbiology* 77: 273-290.
- LASSERRE, M., MALAN, O.G. & TURNER, B. 1983. *The application of*





*principal component analysis to Landsat MSS data. Proceedings of seminar on Principal component analysis in the atmosphere and earth sciences, Pretoria 7-8 February 1983.*

- LEENHOUTS, P.W. 1966. Keys in biology: a survey and a proposal of a new kind. *Proceedings of the Koninklijke Nederlandsche Akademie van Wetenschappen Amsterdam, Series C.* 69: 571-596.
- MACVICAR, C.N., DE VILLIERS, J.M., LOXTON, R.F., VERSTER, E., LAMBRECHTS, J.J.N., MERRYWEATHER, F.R., LE ROUX, J. VAN ROOYEN, T.H. & HARMSE, H.J.VON M. 1977. *Soil classification: A binomial system for South Africa.* Department of Agricultural Technical Services, Pretoria.
- MALAN, O.G. & LAMB, A.D. 1985. *Display of digital image data; quantitative and optimal.* Proceedings of the third South African symposium on Digital image processing, Durban 22-23 July 1985.
- MERXMUELLER, H. 1972. *Prodromus einer Flora von Suedwest Afrika.* Cramer, Lehre.
- METCALFE, C.R. & CHALK, L. 1979. *Anatomy of the Dicotyledons.* Vol. 1, 2nd edition. Oxford University Press, Oxford.
- MEYER, P.G. 1969. *Eintuehrung in die Pflanzenwelt Suedwestafrikas.* S.W.A. Wissenschaftliches Gesellschaft, Windhoek.
- MORSE, L.E. 1971. Specimen identification and key construction with time-sharing computers. *Taxon* 20: 269-282.
- MORSE, L.E. 1974. Computer programs for specimen identification, key construction and description printing using taxonomic data matrices. *Publications of the Museum, Michigan State University,, Biological Series* 5: 1-128.
- MORSE, L.E. 1975. Recent advances in the theory and practice of



- specimen identification. In R.J. Pankhurst (Ed.), *Biological identification with Computers*: 11-52. Academic Press, London.
- MUELLER-DOMBOIS, D. & ELLENBERG, H. 1974. *Aims and methods of vegetation ecology*. Wiley, New York.
- MUNSELL SOIL COLOR CHARTS, 1954. Munsell Color co., Baltimore.
- NEWELL, I.M. 1970. Construction and use of tabular keys. *Pacific Insects* 12: 25-37.
- OSBORNE, D.V. 1963. Some aspects of the theory of dichotomous keys. *New Phytologist* 62: 144-160.
- PANKHURST, R.J. 1971. Botanical keys generated by computer. *Watsonia* 8: 357-368.
- PANKHURST, R.J. 1978. *Biological identification*. Edward Arnold, London.
- PARTON, W.J. & INNIS, G.S. 1972. *Some graphs and their functional forms*. U.S. International Biological Program, Grassland Biome Project Technical Report 153: 1-41.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1984. Sectors of the Transvaal Province of South Africa. *Bothalia* 15: 294-295.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1986. Biomes of southern Africa - An objective categorization. *Memoirs of the Botanical Survey of South Africa* 54: 1-98.
- RYPKA, B.W., CLAPPER, W.E., BOWEN, I.G. & BABB, R. 1967. A model for the identification of bacteria. *Journal of General Microbiology* 46: 407-424.
- SCHEEPERS, J.C. (Unpubl.) *Geomorphology classes*. Grassland Research Centre, Pretoria.
- SCHEINER, S.M., & ISTOCK, C.A. 1987. Affinity analysis: methodologies and statistical inference. *Vegetatio* 72: 89-93.



- SNEATH, P.H.A. & SOKAL, R.R. 1973. *Numerical Taxonomy*. Freeman, San Francisco.
- TAINTON, N.M., BRANSBY, D.I. & BOOYSEN, P. de V. 1976. *Common Veld and Pasture Grasses of Natal*. Shuter & Shooter, Pietermaritzburg.
- TER BRAAK, C.J.F. 1986. Canonical correspondence analysis: A new Eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.
- TER BRAAK, C.J.F. 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio* 69: 69-77.
- TROLLOPE, W.S.W. & POTGIETER, A.L.F. 1983. *Characteristics of fire behaviour in the Kruger National Park*. Report. Department of Agronomy, University of Fort Hare, Alice.
- VOSS, L.G. 1952. The history of keys and phylogenetic trees in biology. *Journal of the Science Laboratory, Denison University* 43: 1-25.
- WEBER, W.A. & NELSON, P.P. 1972. *Random access key to the genera of Colorado mosses*. University of Colorado Museum, Boulder.
- WERGER, M.J.A. 1972. Species-area relationships and plot size: with some examples from South African vegetation. *Bothalia* 10: 583-594.
- WERGER, M.J.A. 1974. On concepts and techniques applied in the Zürich-Montpellier method of vegetation survey. *Bothalia* 11: 309-323.
- WESTFALL, R.H. 1981. *The plant ecology of the farm Groothoek, Thabazimbi District*. M.Sc. thesis, University of Pretoria.



- WESTFALL, R.H. (in prep.). SPECOM: A program for interrogating and comparing the PRECIS synonym list with species name data sets.
- WESTFALL, R.H., DEDNAM, G., VAN ROOYEN, N., & THERON, G.K. 1982. PHYTOTAB - A program package for Braun-Blanquet tables. *Vegetatio* 49: 35-37.
- WESTFALL, R.H. & DE WET, B.C. 1988. New programs for preliminary sequencing of relevès and species in phytosociological data sets. *Bothalia* 18: 122-123.
- WESTFALL, R.H. & MALAN, O.G. 1984. A comparison of vegetation units derived from vegetation-enhanced satellite imagery with the vegetation units derived from the floristic classification of the farm Groothoek, Thabazimbi District. Proceedings of the symposium on Pattern recognition in remote sensing and geophysics, Johannesburg, 2-3 July 1984.
- WESTFALL, R.H. & MALAN, O.G. 1986. A method for vegetation stratification using scale-related, vegetation-enhanced satellite imagery. *Bothalia* 16: 263-268.
- WESTFALL, R.H. & PANAGOS, M.D. 1984. A cover meter for canopy and basal cover estimations. *Bothalia* 15: 241-244.
- WESTFALL, R.H., VAN STADEN, J.M. & PANAGOS, M.D. 1987. Predictive species-area relations and determination of subsample size for vegetation sampling in the Transvaal Waterberg. *South African Journal of Botany* 53: 241-244.



## CHAPTER 4. RESULTS AND DISCUSSION

The methods developed and described in this study have been applied successfully by Van Staden (1992), whose methods were based on this study, and his results are, therefore, used to substantiate this work, where necessary.

Field work in the main study area is ongoing and 63 stands with a total of 270 sampling units have been sampled. Much of the work in this area has been developmental; for example the reconnaissance methods advocated were not applied comprehensively from the outset and several cover estimation techniques were used before the plant number scale (Westfall & Panagos 1988) was developed.

### 4.1 PREPARATORY WORK

Analysis of computerized rainfall data supplied by the Weather Bureau, on request, for 21 stations within the main study area, with records for periods longer than 20 years, gave the following regression, according to the formula  $y=ax+c$ :

$$y = 154 (x) + 379$$

where  $y$  = mean annual rainfall; and

$$x = \frac{\text{altitude (m)}}{1\ 000} \times (\text{latitude (decimal degrees)} - 23,5)$$

A correlation coefficient of  $r=0,90$  between the  $x$  and  $y$  values,



was obtained using the statistics utility in the PHYTOTAB-PC program package. A poor correlation between the variables was, however, obtained when stations outside the study area were included, where a rain shadow effect seems to occur. This implies that mean annual rainfall, for any stand within the main study area, can be estimated from the altitude and latitude of the stand. The technique illustrates the feasibility of using regressions to predict rainfall on a stand basis, in studies where topography can influence rainfall.

#### 4.1.1 Scale

The scale used in this study and that of Van Staden (1992) is 1:250 000 which is considered commensurate with the detail required by the Department of Agricultural Development for vegetation resource inventory at a regional level. Stand radius is, therefore, 250 m and minimum stand spacing is 500 m (section 3.1.5). The smallest area for a stratified unit, in which four stands with minimum stand spacing can be included, is approximately 314 ha. This in effect, relates to a stand and spacing radius of 500 m per stand or approximately 78,5 ha per stand. The minimum of four stands per stratified unit is based on the same argument as the minimum number of sampling units per stand (section 3.1.6) and relates to stratified unit variation sampled. The stratified unit is that area mapped for sampling the vegetation. The vegetation unit is the area classified as a relevè or group of relevès. Correspondence between the stratified unit and vegetation unit can



assist in verifying a classification.

The high degree of correspondence between stratification, classification and environment, achieved by Van Staden (1992), is attributed mainly to cognizance being taken of scale. Scale determines the minimum area of a vegetation unit or potential community which is taken to be a stand. Scale can also influence environmental correspondence with vegetation units. For example, at biome scale (1:10 000 000) the vegetation units, or biomes, correspond to seasonality and summer aridity (Rutherford & Westfall 1986). Within a biome, rainfall, geology and soils can vary considerably, such as rainfall in the Fynbos Biome, which varies from about 300 mm to over 2 400 mm per annum (Rutherford & Westfall 1986). These factors can be expected to differentiate vegetation units at scales larger than biome scale, hence the differentiating environmental factors for a vegetation unit can be regarded as being scale-related. If scale is not taken into account, especially in small-scale work, then mixed scales could result, thereby complicating environmental correspondence and syntheses of studies.

On the other hand, cognizance of scale does not appear to be as critical in large scale work. This can be attributed to sampling unit area either, being similar to stand area or, being of such an area that it can be representative of a stand, without the stand being defined in terms of scale.

Scale also relates to heterogeneity, where stands at large scales are likely to be relatively homogeneous, whereas, as scale de-



creases, heterogeneity is likely to increase. This is also likely to apply to the vegetation units of which the stands are a part, so that the heterogeneity of a vegetation unit is inversely proportional to scale. Thus, cognizance of scale can also aid the sampling of mosaics in vegetation, where scale determines whether the components of the mosaic should be sampled separately or jointly.

Scale can also affect sampling intensity, for example, at biome scale (1:10 000 000) sampling intensity in terms of stands per km<sup>2</sup> would be considerably less than at 1:50 000 scale. However, in order to adequately sample the increased vegetation variation at biome scale, stand radius at biome scale is considerably larger, being 10 km, than the 50 m required at at 1:50 000 scale. Thus, increasing stand area for decreasing scale compensates for increasing heterogeneity with decreasing scale.

#### 4.1.2 Stand area

The stand area, for the main study area as defined in section 3.1.2, corresponds to a scale of 1: 250 000. The stand radius is, therefore, 250 m with stand area being approximately 20 ha.

Decisions required for stand definition are inversely proportional to scale. In other words the smaller the scale the more difficult it is to identify a stand of vegetation at the required scale. The direct association advocated in this study between scale and stand area is the means by which scale is taken into account during





sampling and is also easy to apply. The stand is integral to the Braun-Blanquet approach to vegetation sampling (Mueller-Dombois & Ellenberg 1974). Its relevance can be considerably increased for small scale work by the definition proposed in this study without invalidating large scale work.

It is interesting to note that the stand area of the Veld Types (Acocks 1975, 1988), according to scale, would have been about  $7 \text{ km}^{-2}$ . This is approximately the area covered by Acocks during sampling (J.C. Scheepers pers. comm.)

A practical application of the stand, as defined in this study, is the determination of the minimum area required for conserving and sustaining a particular vegetation unit. The minimum sustainable area for a vegetation unit is the stand at the largest scale at which the vegetation unit can be recognized. However, the periphery of such a conserved unit is likely to change with time, depending on practices adjacent to the border of the conserved unit. A buffer area would, therefore, be required. This buffer area can be equated with the ecotone or transitional area, between two communities, which is also related to scale. The minimum sustainable area (s) for a vegetation unit is, therefore, given by:

$$s = \pi 2r^2$$

where s=minimum sustainable area ( $\text{m}^2$ ) and r is equal to the denominator, in metres, of the largest scale, as a representative fraction, at which the vegetation unit can be recognized.



#### 4.1.3 Reconnaissance

Approximately 60 structural/floristic units were estimated present in the main study area at the relevant scale during reconnaissance (see section 3.1.3). The structural/floristic unit total was a deliberate over-estimate for adequate colour allocation in the stratification process (section 3.1.4). Species richness averaged 11 species per  $m^2$  as determined by regressions of species counts in 4  $m^2$  quadrats, selected to represent the range in species richness variation. These results were obtained using the sample/stand dimension utility in the PHYTOTAB-PC program package and are required for determining minimum sampling unit dimensions (sub-plots) for stand sampling.

The minimum requirement of a reconnaissance of a study area is an estimate of the variation in vegetation expressed as the approximate number of vegetation units at the scale required. This is essential for both stratification and classification. The criterion of species richness to determine sampling unit dimensions, is only necessary where stands are subsampled. The particular aim of a study will determine whether other variables require preliminary assessment of variation. However, observations on vegetation/environment relations should be made for initial model construction (section 3.4.2.3). Voucher specimen collection of the common and dominant plant species, can also be made during reconnaissance, for initial plant identification key purposes, as these are the species most likely to be encountered during sampling.



#### 4.1.4 Stratification

A total of 44 units were stratified in the study area at the relevant scale. The smallest stratified unit is such that four stands, each with buffer zones equal to the stand radius, could be included (section 4.1.1). The buffer zones provide for possible ecotones, at the relevant scale, should community borders fall between two stands and also ensure non-contiguity of stand placement. The area represented by the smallest stratified unit is, therefore, approximately 314 ha.

The successful use of LANDSAT MSS data to differentiate vegetation units corresponding to plant communities is shown in Van Staden (1992). The ground resolution of LANDSAT MSS data, of approximately 79 x 79 m, is such that individual plant differences are unlikely to affect spectral characteristics for small-scale work. The textural and contextual detail of plant communities at this scale are such, however, that a fairly wide range of structural and cover variation can be tolerated, locally, within a plant community. In this sense, structure refers to the horizontal zonation of the vegetation. Van Staden (1992) has shown a high correspondence between vegetation units, classified on the basis of floristics, and stratified units, determined by means of satellite imagery, despite variations in cover because of severe overgrazing. It is likely that utilization of a particular layer of vegetation, can be offset over time, by an increase in phytomass in a different layer so that the effect of grazing on aboveground phytomass, is reduced. Furthermore, Van Staden (1992)



has shown that the environmental gradients differentiating plant communities relate primarily to available moisture, in terms of effective soil depth. This implies that adjacent plant communities differ in the availability of moisture, which is probably the main factor limiting phytomass, in the study concerned. Although plant communities with similar structure can be floristically different, and vice versa, it is suggested that the combination of structure and cover, which is a function of phytomass, can differ in adjacent plant communities. This leads to the assumption that total aboveground phytomass is more likely to be responsible for differences in plant community spectral characteristics than either structure or cover alone, in the study concerned. Hypothesis i (section 1.3) is, therefore, not disproved.

It is often highly arbitrary to delimit vegetation units primarily according to an environmental factor, such as altitude, which mostly forms a continuum in its range. The decision as to what particular altitude constitutes a discontinuity in the vegetation must, therefore, take the vegetation into account. The methods used in this study for stratification of vegetation suggest improvements, for small scale work, to the usual method of aerial photograph interpretation where topography and vegetation are visually integrated. Rather than visually integrate topography and vegetation, the primary delimitation could be done according to pattern analysis of vegetation. Topographic features would then play a secondary role and scale could be taken into account.

This would entail, firstly, constructing a transparent overlay



TABLE 4.1. - PHYTOLOC output for random location of stands, showing random point numbers, x-y co-ordinates for overlay grid and corresponding latitude and longitude, in degrees (to the left of the decimal), minutes (first two decimal places) and seconds (third and fourth decimal places with decimal fraction following). Stand numbers are consecutive and are not necessarily those of the random points.

---

COORD NO. X-AXIS, Y-AXIS		
1	X-AX= 53	Y-AX= 123
Lat	24.41390147	Lon 28.00590064
2	X-AX= 83	Y-AX= 89
Lat	24.2313596	Lon 28.1900061
3	X-AX= 73	Y-AX= 52
Lat	24.03106403	Lon 28.13339644
4	X-AX= 25	Y-AX= 90
Lat	24.23461083	Lon 27.44453031
5	X-AX= 89	Y-AX= 25
Lat	23.48328078	Lon 28.2332387
6	X-AX= 30	Y-AX= 107
Lat	24.32588177	Lon 27.47371387
7	X-AX= 101	Y-AX= 102
Lat	24.30162561	Lon 28.29241894
8	X-AX= 13	Y-AX= 111
Lat	24.3508867	Lon 27.37375639
9	X-AX= 65	Y-AX= 78
Lat	24.17159605	Lon 28.08298468
10	X-AX= 75	Y-AX= 3
Lat	23.36375369	Lon 28.15248322
11	X-AX= 52	Y-AX= 106
Lat	24.32263054	Lon 28.0032904
12	X-AX= 75	Y-AX= 118
Lat	24.38564532	Lon 28.13544628
13	X-AX= 98	Y-AX= 20
Lat	23.45502462	Lon 28.2902484
14	X-AX= 55	Y-AX= 64
Lat	24.09407881	Lon 28.02425018
15	X-AX= 37	Y-AX= 43
Lat	23.58180295	Lon 27.52084009
16	X-AX= 52	Y-AX= 106
Lat	24.32263054	Lon 28.0032904
17	X-AX= 90	Y-AX= 63
Lat	24.09082758	Lon 28.23323076
18	X-AX= 9	Y-AX= 111
Lat	24.3508867	Lon 27.3516775
19	X-AX= 63	Y-AX= 26
Lat	23.49053202	Lon 28.07532632
20	X-AX= 55	Y-AX= 52
Lat	24.03106403	Lon 28.02494252
21	X-AX= 17	Y-AX= 98
Lat	24.28062068	Lon 27.40006091
22	X-AX= 32	Y-AX= 21
Lat	23.46227586	Lon 27.49163862

---



with the dimensions of the smallest stratified unit, at the required scale, adjusted to that of the aerial photographs, as a guide. Then mount the aerial photographs on a wall to form a single image. Observation distance is that at which the smallest homogeneous pattern, represented by textural detail on the aerial photographs, is no smaller than that of the overlay guide, at the same distance. Confusing detail at larger scales than required is thus avoided. Finally, trace the borders of all homogeneous pattern units, greater than or equal to the area of the guide, and recognizable at the observation distance, onto the aerial photographs. Vegetation as represented by pattern, is thus the sole criterion of primary delimitation. Secondary delimitation of the primary units can be based on topographic features or other suitable criteria, as deemed necessary.

It must be emphasized that the sampling intensity for vegetation classification is generally too low to accurately map vegetation units according to stands because as few as four stands can represent a vegetation unit. The stratification process is, in practice, the process by which vegetation units are primarily demarcated and mapped, but can be modified and supported by the classification.

#### 4.1.5 Stand location

An example of output from the program PHYTOLOC is given in Table 4.1. The stand location reference, taken as the stand



centre, is given in degrees, minutes and seconds. The x-y coordinates refer to the grid overlay intersection points for transfer to the relevant map. The PHYTOLOC program has been included in the PHYTOTAB-PC program package.

Stand location in the field need not be precise for classification purposes, when working at small scales, because of the random location of stands. However, comparisons with the program SIDA (section 4.2.5.2) showed that greater precision in locating the stand centre in the field was required, than could be obtained with simple visual estimation, if predicted habitat data were to correspond with that of the stand. Stand location was improved by using an altimeter to verify the contour position in the field. An optical rangefinder with a 2 km limit, used to improve stand location, proved inadequate in terms of accuracy and range.

Geographic positioning by satellite systems (GPS) should, in the future, overcome the problems of stand location. The price of such systems have decreased considerably to below R10 000 and the size has also decreased, being somewhat larger than a pocket calculator. Such a system will not only show the co-ordinates of the current position but can indicate the direction and distance to the required position.

Precise stand location will also be required, in the future, when use is made of Geographic Information Systems (GIS) for obtaining habitat data and incorporating vegetation sample sites in such systems.



#### 4.1.6. Sampling unit area

Sampling unit area commensurate with a species richness of 12 species per  $m^2$  is  $2,5 \times 5$  m (section 3.1.6), determined with the sample/stand dimension utility included in the PHYTOTAB-PC program package. Sample area, for each stand in the main study area, is, therefore, a minimum of  $50 m^2$  (four  $12,5 m^2$  sampling units) or 0,025% of the 20 ha stand area. Species sampled, however, represent only about 13% (40 species out of 274) of the species present in the stand based on a species count in a representative 20 ha stand. The total canopy cover, however, represented by the species sampled is estimated to be about 95% of the total canopy cover of all the species present, based on the estimated canopy cover of each species during the species count. It must be emphasized that the sampling unit dimensions illustrated here refer to the minimum dimensions for a particular scale and for a particular species richness. Increase in sampling unit area could improve the constancy of species in a matrix, but decrease in area is likely to affect the classification.

Sampling unit area is not sample size because area is not what is being sampled. Species presence is the attribute of the samples. A minimum of four stands each with four sampling units, could, in statistical terms, be considered insufficient replication for a normal distribution (Freund & Williams 1958) within a





TABLE 4.2. - The range of sampling units per stand in the main study area

	Number of sampling units per stand			
	4	5	6	7
Number of stands	50	9	3	1
Percentage of total	79,4	14,3	4,7	1,6

TABLE 4.3. - PHYTOLOC output for location of sampling units where the x- co-ordinate represents direction (degrees); and the y- co-ordinate represents distance from the stand centre (m)

COORD NO., X-AXIS, Y-AXIS
1 X-AX= 121 Y-AX= 181
2 X-AX= 227 Y-AX= 114
3 X-AX= 168 Y-AX= 176
4 X-AX= 53 Y-AX= 88
5 X-AX= 5 Y-AX= 71
6 X-AX= 166 Y-AX= 138
7 X-AX= 230 Y-AX= 159
8 X-AX= 209 Y-AX= 81
9 X-AX= 149 Y-AX= 118
10 X-AX= 294 Y-AX= 194
11 X-AX= 285 Y-AX= 46
12 X-AX= 80 Y-AX= 21
13 X-AX= 320 Y-AX= 106
14 X-AX= 252 Y-AX= 24
15 X-AX= 69 Y-AX= 96
16 X-AX= 281 Y-AX= 235
17 X-AX= 203 Y-AX= 236
18 X-AX= 242 Y-AX= 102
19 X-AX= 106 Y-AX= 193
20 X-AX= 252 Y-AX= 17
21 X-AX= 196 Y-AX= 120
22 X-AX= 2 Y-AX= 23



plant community. This would often preclude statistical tests based on normal distributions. The sample, however, is not designed for statistical comparisons but to determine those species which best characterize the community and those responsible for the most significant portion of the vegetation cover. The test for the adequacy of the sample should, therefore, be scientific validity and not necessarily statistical validity. By scientific validity is meant that the processes should be repeatable by independent observers.

#### 4.2 FIELD SAMPLING

Because sampling is ongoing the results pertaining to field sampling are examples rather than a synopsis of all the processes. The number and percentage of total, of the sampling units for all stands is given in Table 4.2.

The methods used in this study for field sampling are extremely rigorous. This confines thought and observation in the field chiefly to sampling units. The opposite extreme, is descriptive accounts of vegetation with no explicit methodology, such as Edwards (1967). In such accounts observations are not restricted and many sound conclusions have been derived. The ideal would be to utilize the best of both approaches.

##### 4.2.1 Sampling unit location

Table 4.3 shows an example of the output from the program PHYTOLOC



for random sampling unit location. Position of the reference corner, for each sampling unit within the stand, is in relation to the stand centre. The greatest disadvantage in subsampling the stand, using quadrats, is in the time taken to locate, mark and record such subsamples. It was seldom that two stands could be completed in a single day.

The stand as defined in this study raises interesting possibilities regarding sampling unit location. Apart from subsampling, as described for small scale work, representative samples can be used for large scale work. Furthermore, it is conceivable that point methods could also be used within the defined stand context, thereby permitting comparison between point and quadrat data. Obviously, the number of points within a stand should result in a comparable number of species with that obtained with quadrats. A plotless method could also be used to determine the species representative of a stand. Such a method could make use of the plant number scale, whereby, species with cover greater than, say 0,05% or less than 40 crown diameters apart could be recorded within the stand. Thus the ideal of an informal sampling approach coupled with objective stand dimensions and location could be achieved. In this case the location of sampling units would not be relevant and considerable field time could be saved. The problem of minimum species area would also fall away. Hypothesis ii (section 1.3) would, therefore, be irrelevant as minimum sampling area would not be applicable.



#### 4.2.2 Plant identification and verification

The criteria for plant identification in the field according to specimens collected are given in Tables 4.4 to 4.8 (back pocket). The same characters and character states used for plant identification have been used for identification of the families represented by the specimens collected (Table 4.9, back pocket).

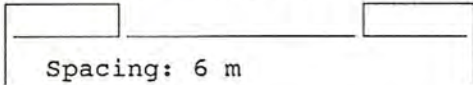
The initial input required for the construction of the plant identification key is time-consuming. It could be argued that the benefits of validating identification in this manner do not warrant the additional effort required. However, the benefits are increased by direct knowledge transfer. Van Staden (1992) used that portion of the key relevant to his study area for plant identifications, thereby reducing his own input. It is unfortunate that knowledge gained in field identification of plants cannot be easily transferred. This method goes some way towards solving that problem. The method is also very useful for training in plant identification. This is probably the first study in vegetation ecology where the criteria for plant identification in the field are given explicitly and comprehensively. The key to the families in the study area, although useful for field use, is unlikely to be of any taxonomic significance. This, nevertheless, illustrates the potential of the programs.

A camera with a data back, for imprinting reference numbers on slides, would be a decided advantage for numbering slides of plant specimens and stands, in the field. Use of a blackboard with



Acacia caffra

Community/relevè: 1  
Recorded cover: 25.8%  
Derived cover 25%  
Mean crown diameter: 3.05 m  
Individual/ha: 353  
m sq/individual: 28.31 m sq  
Spacing:centre-centre: 6 m  
Canopy radius: 1.52 m  
Canopy-canopy gap: 2.96 m



Acacia caffra

Community/relevè: 2  
Recorded cover: 0.91%  
Derived cover 1%  
Mean crown diameter: 3.05 m  
Individual/ha: 12  
m sq/individual: 802.87 m sq  
Spacing:centre-centre: 31.97 m  
Canopy radius: 1.52 m  
Canopy-canopy gap: 28.93 m

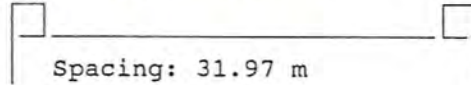


FIGURE 4.1. - Example of SPECODA output showing cover data for one crown and two cover classes.

#### 4.2.3 Species cover

Examples of illustrated species cover using the program SPECODA for selected species are given in Figure 4.1. The calculations used in this program are included in the processing facility in the PHYTOTAB-PC program package. A comparison of total species cover with cover determined according to vegetation height classes for a stand should result in the following:

- a) total species cover should be greater than the cover of the height class with the least cover; and
- b) total species cover should be less than the sum of the cover for all height classes.

Cover estimations using the Domin-Krajina cover-abundance scale (Mueller-Dombois & Ellenberg 1974) often resulted in total species cover greatly exceeding these limits. Estimations with the plant number scale (Westfall & Panagos 1988) are within these limits.

Although fewer class intervals could enhance cover pattern on a matrix, the effectiveness of the plant number scale for estimating cover is demonstrated by Van Staden (1992). Precision is according to a whole plant. It is unlikely that greater precision would be required. However, less precision could influence results of the community composition analysis. This method also has potential for monitoring individual species change. Permanently marking transect corners would allow monitoring by merely counting individuals at required periods. In this case, the transect refers to the area



required for counting species numbers (section 3.2.3.2). Changes caused by defoliation or increases due to plant growth on the other hand, could be detected by determining the increase or decrease in transect area required for a particular species.

The over- and under-estimates obtained using the Domin-Krajina cover-abundance scale are attributed to the static position often adopted by the observer, where:

- a) an adequate sample for cover estimation was often obscured by vegetation leading to under-estimation; and
- b) local clumping often resulted in an over-estimation.

The plant number scale in contrast necessitates that the observer moves along a transect related to the crown diameter of a plant species. It is also doubtful whether or not summation of midpoints of large class intervals, such as found in the Domin-Krajina cover-abundance scale, can approximate actual cover.

The plant number scale method is ideally suited to informal stand sampling which would also allow for observations on species selection for monitoring to be made subjectively at time of sampling.

The potential of the species spacing illustrations and species density (Figure 4.1) was not explored. It is conceivable that these derivatives of the plant number scale could be used to examine plant spatial requirements, competition between species as well as species reactions to environmental gradients. The plant number scale also complies fully with the requirements of a scale



TABLE 4.10. - PHYTOCAP output for a selected stand showing species present (collector's number) and cover (symbol) in each sub-quadrat. Numbers preceded by "0" are line numbers followed by a number to which the first digit refers to the sub-quadrat and the last three to the stand. Cover symbols are those of the plant number scale

001 1047	004 2047	007 3047	010 4047
2149 C	2149 1	2297 +	2297 D
2297 A	2240 3	2053 F	2240 2
2053 3	2192 +	2154 1	2149 8
2157 1	2131 2	2131 7	2066 1
2192 +	2132 +	2132 3	2131 5
2039 3	2045 1	2192 1	2039 2
2131 1	2043 +	2229 +	2132 3
2135 A	2250 +	2194 4	2192 1
002 1047	005 2047	008 3047	011 4047
2073 +	2078 7	2039 1	2241 3
2005 +	2077 3	2041 2	2004 2
2194 1	2167 6	2077 2	2078 5
2165 A	2004 2	2008 8	2165 5
2241 B	2241 1	2165 2	2008 1
2167 5	2012 7	2078 5	2003 +
2077 5	2243 2	2116 1	2167 3
2200 2	2074 3	2167 6	
003 1047	006 2047	009 3047	
2012 4	2128 1	2223 1	
	2144 1	2004 3	

TABLE 4.11. - PHYTOFORM output for the stand in the previous Table showing conversion of data to PHYTOTAB mainframe format. Number of species = 35. Cumulative cover = 62.37%

001 00472149	72297	82053	82157	+2192	12039	22131	42135	5
002 00472073	+2005	+2194	22165	62241	62167	52077	32200	1
003 00472012	42240	22132	32045	+2043	+2250	+2078	52004	2
004 00472243	12874	12128	+2144	+2154	+2229	+2008	42116	+
005 00472223	+2066	+2003	+					





according to Londo (1976) (section 3.2.3).

Crown cover estimations are preferred to basal cover estimations because, apart from the requirements of the Braun-Blanquet approach, more factors can be related to crown cover than basal cover which inherently contains no more information than that which can be derived from species frequency. For example, crown cover can relate to a) degree of soil protection; b) utilizable material within height classes; and c) competitiveness between similar size species.

#### 4.2.4 Floristic data recording

Examples of sampling unit data using the PHYTOCAP program are given in Table 4.10. Conversion of these data to stand data using the PHYTOFORM program are shown in Table 4.11. It is doubtful whether computerized field data capture is advantageous for floristic data input. The transfer programs are slow and the computer is an additional burden in the field. Additionally it is generally quicker to use pencil and paper in the field for these recordings than single finger typing on the miniaturized keyboards. For these reasons neither of these programs are included in the PHYTOTAB-PC program package. It is, furthermore, suggested that greater efficiency can be obtained in the field by recording floristic data on field sheets, with provision for relevant casual observations, and that the data be captured on a laptop computer, loaded with the PHYTOTAB-PC programs, daily after fieldwork.



If a classification is the aim of a study then species presence only is the minimum floristic data required. PHYTOTAB-PC does not require cover values for classification or for synoptic table generation. Ordinations of communities based on synoptic tables also do not require cover values.

A quick visual assessment of dominance from a phytosociological table requires a scale with few cover classes. Too much detail would require numerical treatment of the data for understanding. Such treatment could imply a precision greater than that possible with cover estimation techniques used, such as the Domin-Krajina cover-abundance scale. This may not be too important if a visual assessment of dominance only, is required. If plant spacing and density in terms of individuals per unit area, are required then the plant number scale is highly convenient because only the recording of the crown diameter class is additionally necessary. This value is, in any case, necessary for determining transect length.

Any comparison of plant species should require categorization because of the disparity in plant species. A simple categorization such as growth form classes can be adequate. The system of growth form categorization used in this study has the advantage of being related to height classes. This permits a separate structural analysis of vegetation to be replaced by a growth form/cover analysis with little loss of information. Additional input is the growth form code.



Any additional information recorded would be dependent on the aims of a particular study. However, it must be emphasized that descriptive data such as vitality and sociability are difficult to treat numerically and their value is often lost. It is far easier to incorporate descriptive data into a small data set than a large data set which can only be reduced numerically.

If an informal method of stand sampling is adopted then casual observations could be relevant. A danger can exist of an intermittent phenomenon being recorded when first observed but its occurrence in prior situations could have been overlooked. This problem is overcome with a pre-determined list of considerations for observation.

A full numerical analysis of data using the PHYTOTAB-PC programs requires species presence together with cover code, crown diameter code and growth form code to be recorded in the field.

#### 4.2.5 Habitat data

In the main study area, eleven parameters relating to stands, eleven parameters relating to sampling units (quadrats) and twelve parameters derived from 1:50 000 Topo series maps (section 3.2.5.2) were recorded, apart from the general requirements of stand description.

The general requirements for describing a stand sample are:



TABLE 4.12. - HABIMEAN output for stand 47 showing conversion of recorded subquadrat habitat data to stand data

---

Stand number: 47  
Aspect vector: 294 degrees  
Slope (mean): 0 degrees  
Litter cover: 2%  
Litter depth: 2 mm  
Soil depth (min): 1000 mm  
Soil depth (mean): 1200 mm  
Soil depth (max): >1200 mm  
Soil colour (mean): 10YR 4/2  
Soil texture (% clay): 4%  
Soil form: SP-100%  
Surface rock cover: 0%  
Surface compaction: 132.38 kPa  
Relative herbaceous biomass: 35 mm\*

---

\*measured as disc pasture meter drop height



- a) stand number, starting at 1 and increasing consecutively for each stand sampled within a study area or data set. This is the unique number for each stand required by the PHYTOTAB-PC programs. In the abstraction of communities in the phytosociological tables, stand or sample numbers are referred to as relevè numbers (Mueller-Dombois & Ellenberg 1974);
- b) date of sampling the stand, being the temporal reference of the stand; and
- c) sample co-ordinates of latitude and longitude in degrees, minutes and seconds, being the spatial reference of the stand. The program PHYTOLOC (section 3.1.5) determines location and co-ordinates to fractions of a second so that recording of this information is simple. Locating the point on the ground, however, could cause problems (section 4.1.5).

#### 4.2.5.1 *Field data*

Numerical data describing the stand include altitude, exposure and temperature. Mappable or spatial data include lithostratigraphy, geomorphology, vegetation structure and land use. Non-spatial, descriptive data include grazing, browsing, erosion and fire. With the exception of soil form, sampling unit data recorded are numerical. Examples of the averaging of sampling unit data using the HABIMEAN program for a selected stand are given in Table 4.12. This program is not included in the PHYTOTAB-PC program package because its use is restricted to the habitat parameters used in this study, whereas, the decision as to what parameters to record in a study, can differ widely.



The results of the program HABIMEAN (Table 4.12) are mainly arithmetic means, except for aspect, soil form and soil colour. Aspect is the vector of aspects with magnitude slopes. The main influence of aspect modified by slope is thus taken into account. Combination of a north and south aspect, for example, where the slopes are  $10^\circ$  and  $2^\circ$  respectively would result in a north aspect with a mean slope of  $6^\circ$ . This is a possible solution to the problem of averaging aspects where a single numerical value is required for correlation purposes and the effect of slope must be included. However, the effectiveness of such vectors has not been shown and other factors such as latitude could influence the function appreciably. Soil forms are not averaged but the percentage occurrence of a soil form for the sampling units of a stand are calculated. For soil colour the individual components of hue, value and chroma (Munsell Soil Color Charts 1954) are averaged to reflect an average colour for the stand. Soil colour is, therefore, a numerical value, according to the Munsell notations, which can be treated accordingly.

The argument that descriptive data is difficult to include in a numerical treatment of a large data set (section 4.2.4) partially applies to habitat data. The exception is spatial descriptive data such as geology and soil type which can be visually correlated with vegetation units as overlays or manipulated as coverages in a Geographic Information System. Descriptive, non-spatial habitat information, such as the occurrence of fire and estimation of grazing intensity, would, therefore, have a low relevance in large data sets.



<p>(a)</p> <p style="text-align: center;">HEADER</p> <p>Stand no: ...      Grid no: .....</p> <table border="0" style="width: 100%;"> <tr> <td style="width: 50%;">Altitude (m)</td> <td style="width: 50%; border-left: 1px solid black;">MAR (mm)</td> </tr> <tr> <td>Aspect &lt;-----</td> <td style="border-left: 1px solid black;">Exposure</td> </tr> <tr> <td>(deg) Mean soil depth (mm)</td> <td style="border-left: 1px solid black;">(deg) RAM (mm)</td> </tr> </table> <p>Topo profile: (from drainage to watershed)</p> <p style="text-align: center;">Landscape slope (deg)</p> <p>Stand slope (deg)</p> <p>Scale (in relation to stand diameter)</p> <p>Geomorphology:</p> <p>Drainage:</p>	Altitude (m)	MAR (mm)	Aspect <-----	Exposure	(deg) Mean soil depth (mm)	(deg) RAM (mm)	<p>(b)</p> <table border="0" style="width: 100%;"> <tr> <td style="width: 50%;">Stand no: 47</td> <td style="width: 50%;">Grid no: 2428AA</td> </tr> <tr> <td style="text-align: center;">1200</td> <td style="border-left: 1px solid black; text-align: center;">607</td> </tr> <tr> <td style="text-align: center;">0 &lt;-----</td> <td style="border-left: 1px solid black; text-align: center;">180</td> </tr> <tr> <td style="text-align: center;">419</td> <td style="border-left: 1px solid black; text-align: center;">200</td> </tr> </table> <p>Topo profile:-</p> <div style="text-align: right; margin-right: 50px;"> <p>_____ 0</p> <p>_____</p> <p>Scale: 500 m</p> </div> <p>Geomorphology: Flat (B)</p> <p>Drainage: No flow -C</p>	Stand no: 47	Grid no: 2428AA	1200	607	0 <-----	180	419	200
Altitude (m)	MAR (mm)														
Aspect <-----	Exposure														
(deg) Mean soil depth (mm)	(deg) RAM (mm)														
Stand no: 47	Grid no: 2428AA														
1200	607														
0 <-----	180														
419	200														

FIGURE. 4.2. - SIDA output for stand 47 showing header form with explanations (a) and data (b).



TABLE 4.13. - Statistical output for all relevès showing correlations between SIDA data (x-axis) and the corresponding field data (y-axis) for a) altitude

				<u>Linear regression (y=xb+c)</u>
a) Minimum	X= 890	Y= 890		Slope (b )= 1.005
Maximum	X= 1760	Y= 1775		Angle of slope= 45.150 degrees
Range	X= 870	Y= 885		Y-axis interception (c )=-9.671
Total	X= 81260	Y= 81080		Correlation coefficient (r )= .995
Mean	X= 1289.84	Y= 1286.98		Regression variance= 463.843
Median	X= 1300	Y= 1280		Standard error of the estimate= 21.537
Midrange	X= 1325	Y= 1332.5		
Harmonic mean	X= 1248.31	Y= 1244.65		
Mean deviation	X= 4.759	Y= 4.714		
Variance	X= 53495.125	Y= 54518.179		
Standard deviation	X= 231.290	Y= 233.491		
Coefficient of variation	X= 17.931%	Y= 18.142%		
Standard error of the mean	X= 29.139	Y= 29.417		

Scatter diagram

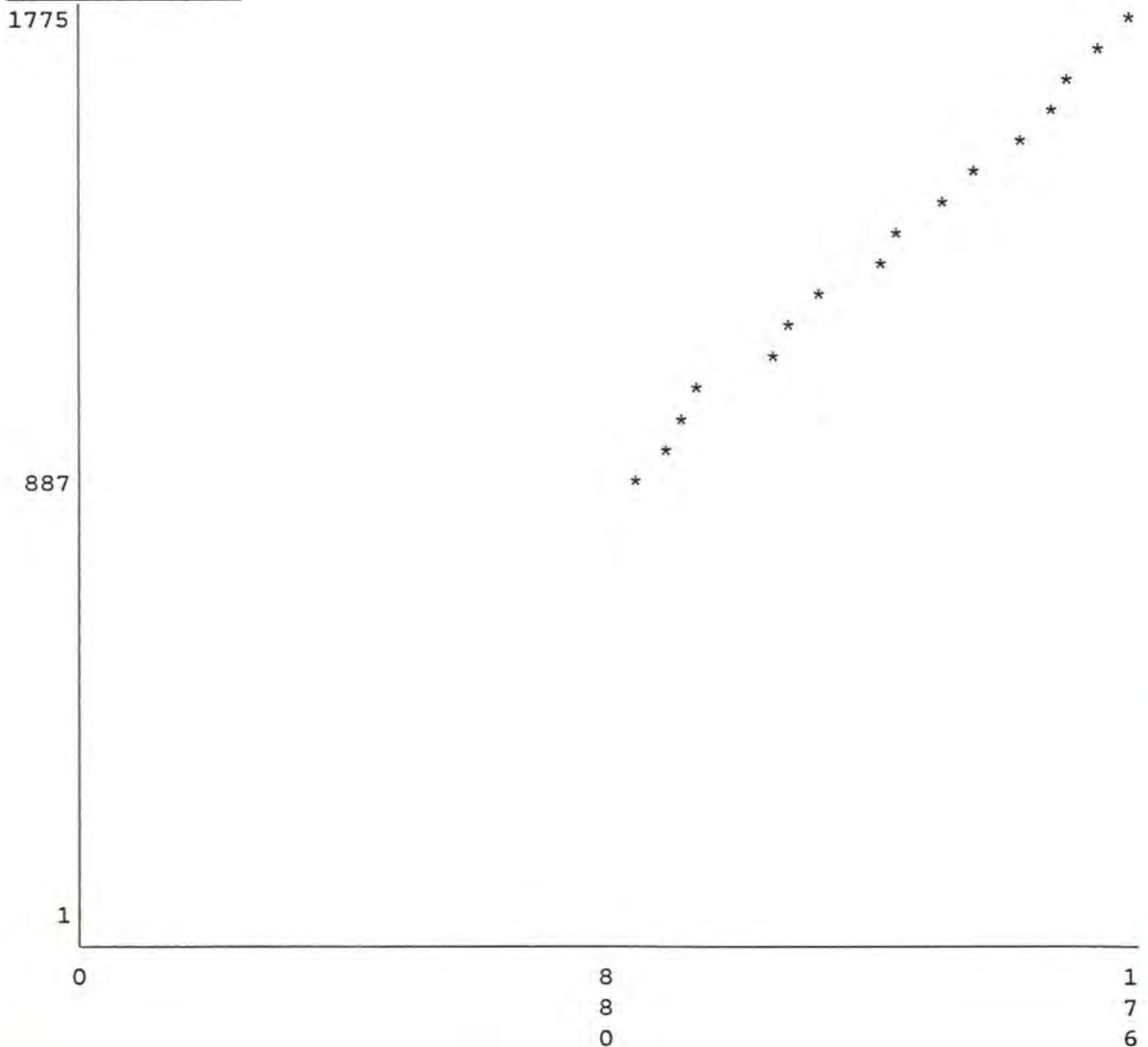






TABLE 4.13 (continued). - Statistical output for all relevès showing correlations between SIDA data (x-axis) and the corresponding field data (y-axis) for b) aspect

b)		Minimum X= 0	Y= 0	Linear regression (y=xb+c)
	Maximum X= 444	Y= 448		Slope (b )= .801
	Range X= 444	Y= 448		Angle of slope= 38.724 degrees
	Total X= 12875	Y= 14755		Y-axis interception (c )= 70.33
	Mean X= 204.36	Y= 234.20		Correlation coefficient (r )= .801
	Median X= 231	Y= 274		Regression variance= 6706.697
	Midrange X= 222	Y= 224		Standard error of the estimate= 81.894
	Geometric mean X= 0	Y= 0		
	Mean deviation X= 3.243	Y= 3.717		
	Variance X= 18412.912	Y= 18437.716		
	Standard deviation X= 135.694	Y= 135.785		
	Coefficient of variation X= 66.397 %	Y= 57.976 %		
	Standard error of the mean X= 17.095	Y= 17.107		

Scatter diagram

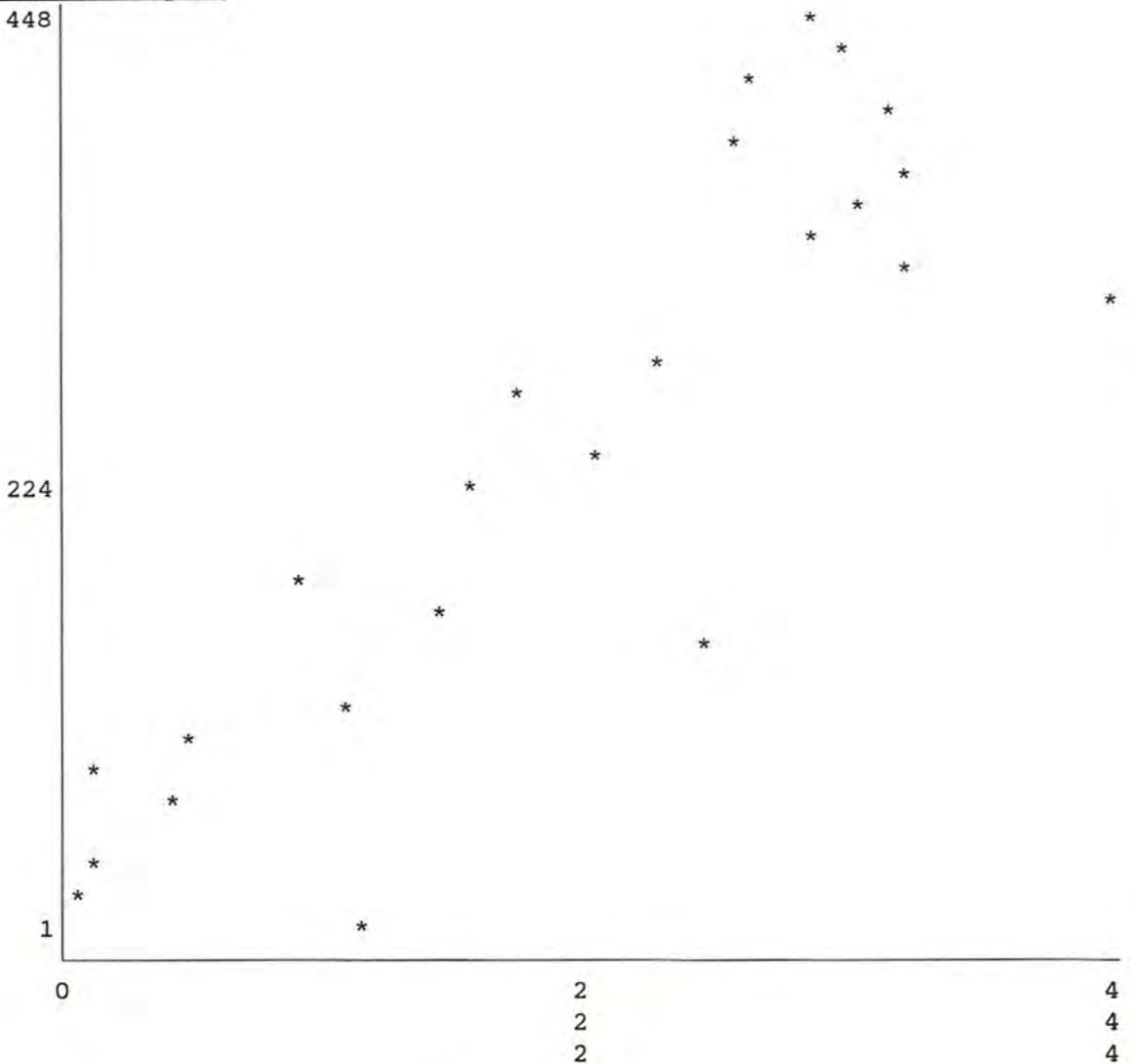




TABLE 4.13 (continued). - Statistical output for all relevès showing correlations between SIDA data (x-axis) and the corresponding field data (y-axis) for c) slope

		Linear regression (y=xb+c)
c) Minimum X= 0	Y= 0	Slope (b)= .707
Maximum X= 180	Y= 180	Angle of slope= 35.281 degrees
Range X= 23	Y= 29	Y-axis interception (c)= 49.523
Total X= 11127	Y= 10993	Correlation coefficient (r)= .567
Mean X= 176.61	Y= 174.49	Regression variance= 23.368
Median X= 178	Y= 176	Standard error of the estimate= 4.834
Midrange X= 168.5	Y= 165.5	
Harmonic mean X= 176.489	Y= 174.284	
Mean deviation X= .0006	Y= .0039	
Variance X= 21.788	Y= 33.899	
Standard deviation X= 4.667	Y= 5.822	
Coefficient of variation X= 2.642 %	Y= 3.336 %	
Standard error of the mean X= .588	Y= .733	

Scatter diagram

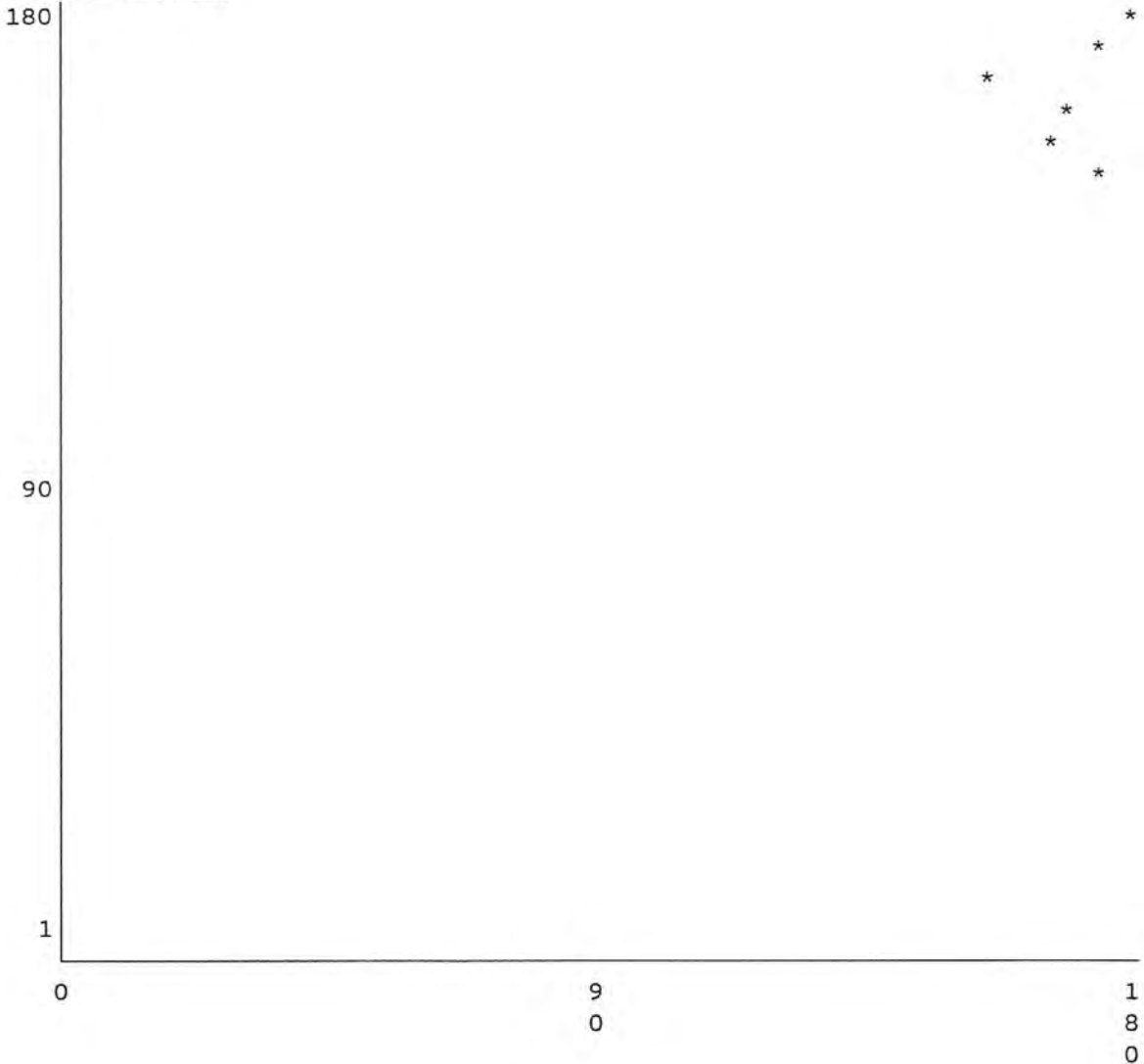




TABLE 4.13 (continued). - Statistical output for all relevès showing correlations between SIDA data (x-axis) and the corresponding field data (y-axis) for d) exposure

		Linear regression (y=xb+c)
d) Minimum X= 157	Y= 151	Slope (b)= .707
Maximum X= 180	Y= 180	Angle of slope= 35.281 degrees
Range X= 23	Y= 29	Y-axis interception (c)= 49.523
Total X= 11127	Y= 10993	Correlation coefficient (r)= .567
Mean X= 176.61	Y= 174.49	Regression variance= 23.368
Median X= 178	Y= 176	Standard error of the estimate= 4.834
Midrange X= 168.5	Y= 165.5	
Harmonic mean X= 176.489	Y= 174.284	
Mean deviation X= .0006	Y= .0039	
Variance X= 21.788	Y= 33.899	
Standard deviation X= 4.667	Y= 5.822	
Coefficient of variation X= 2.642 %	Y= 3.336 %	
Standard error of the mean X= .588	Y= .733	

Scatter diagram

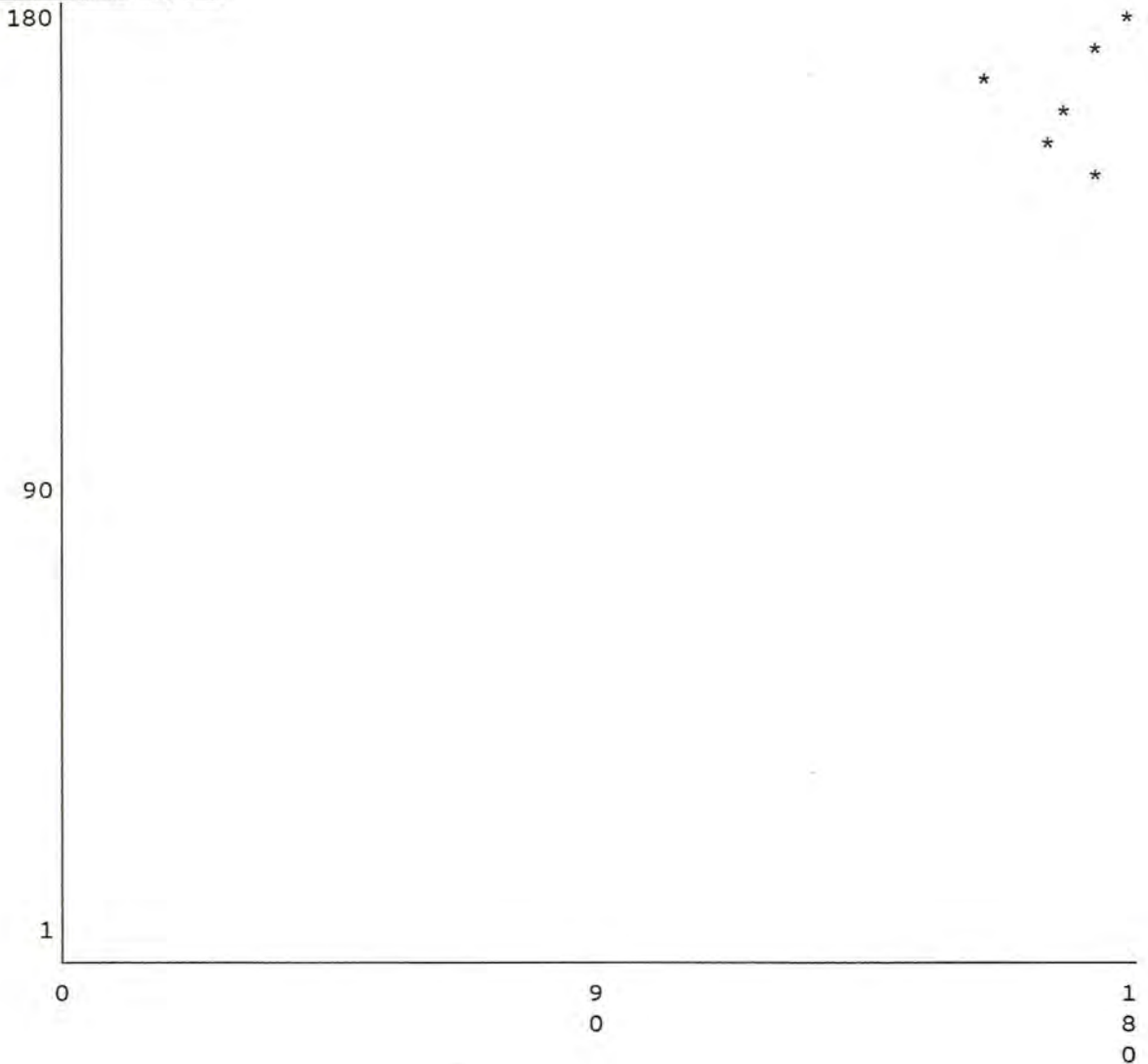
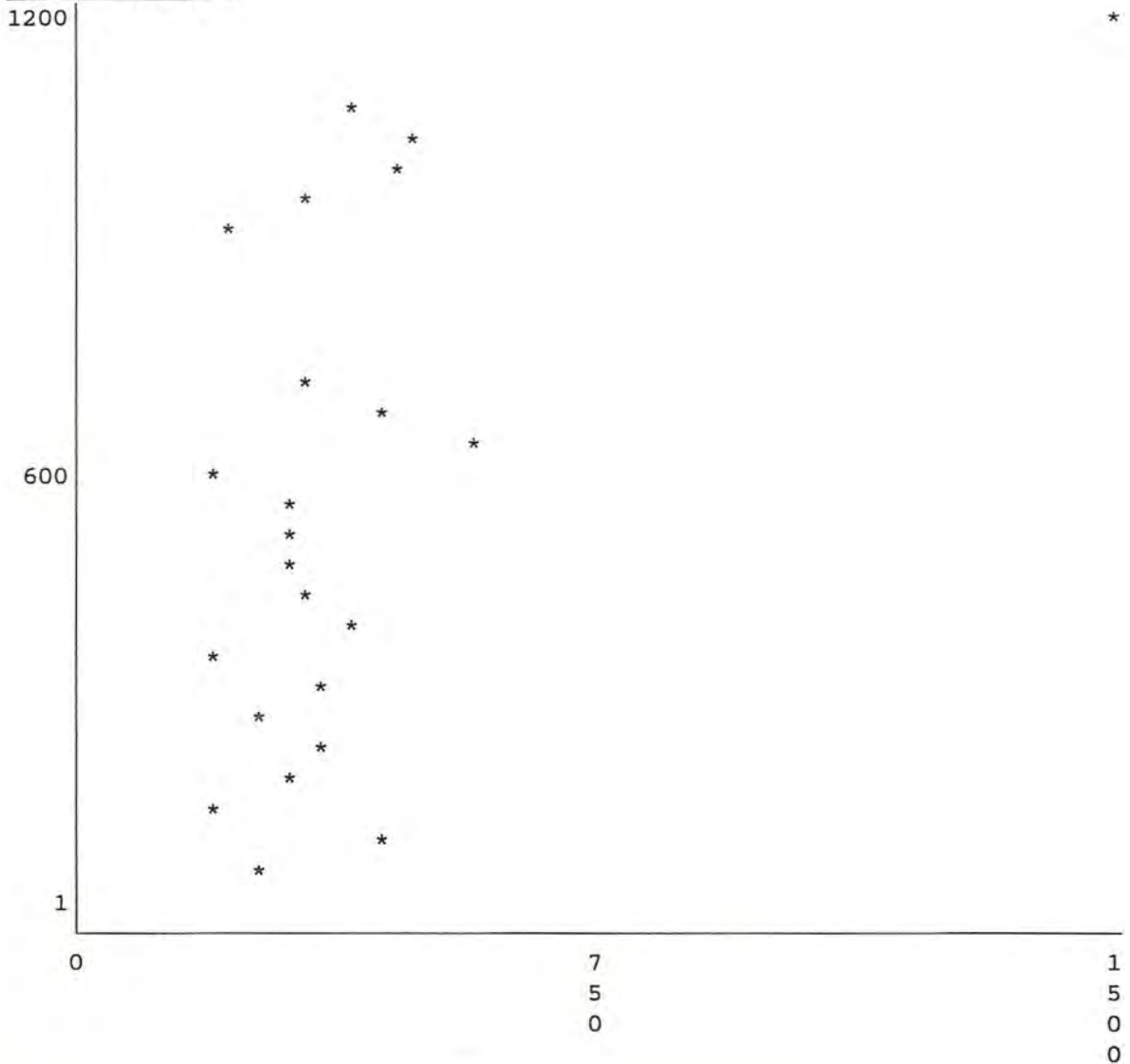




TABLE 4.13 (continued). - Statistical output for all relevès showing correlations between SIDA data (x-axis) and the corresponding field data (y-axis) for e) soil depth

				<u>Linear regression (y=xb+c)</u>
e)	Minimum	X= 75	Y= 90	Slope (b)= .678
	Maximum	X= 1500	Y= 1200	Angle of slope= 34.159 degrees
	Range	X= 1425	Y= 1110	Y-axis interception (c)= 288.967
	Total	X= 23915	Y= 34433	Correlation coefficient (r)= .381
	Mean	X= 379.60	Y= 546.55	Regression variance= 154199.5
	Median	X= 349	Y= 413	Standard error of the estimate= 392.682
	Midrange	X= 787.5	Y= 645	
	Harmonic mean	X= 296.425	Y= 261.602	
	Mean deviation	X= 1.942	Y= 8.197	
	Variance	X= 56192.152	Y= 177586.609	
	Standard deviation	X= 237.048	Y= 421.410	
	Coefficient of variation	X= 62.446 %	Y= 77.102 %	
	Standard error of the mean	X= 29.865	Y= 53.092	

Scatter diagram





Based on observations in this study as well as Westfall (1981) and Van Staden (1992) the two main factors responsible for vegetation unit differentiation, in areas where moisture is a limiting factor and data are not readily obtainable elsewhere (section 4.2.5.2), are soil texture and soil depth. Determination of the former using both the "sausage" and "finger test" methods (section 3.2.5.2) appears adequate. It is, however, unreasonable to expect that a random sample of four auger holes could always be representative of the soil depth of a 20 ha stand. Increasing the number of auger holes could improve correlations but the number required to satisfy statistical requirements is impracticable. A more efficient method of soil depth determination is required.

#### 4.2.5.2 *Derived data*

Examples of habitat data derived from 1:50 000 Topo series maps using the SIDA program are given in Figure 4.2. Comparisons of SIDA output with the corresponding variables recorded in the field using the program MINISTAT are given in Table 4.13. Correlation of 355° and 5° aspect, although numerically disparate are not very dissimilar in terms of direction, being only 10° apart. The following transformations were applied, for comparison to overcome this problem:

- a) for SIDA values between 360° and 90° and relevè values between 270° and 360° add 360° to SIDA value;
- b) for relevè values between 360° and 90° and SIDA values between 270° and 360° add 360° to relevè value; and



c) if SIDA or relevè values are between  $90^\circ$  and  $270^\circ$  then  $360^\circ$  is taken as  $0^\circ$ .

The best correlation co-efficient ( $r=0,99$ ) was obtained with altitude because the same source was used in both cases. Aspect, slope, exposure and soil depth gave correlation co-efficients of 0,80; 0,67; 0,56 and 0,38 respectively. That better correlation co-efficients were not obtained, could relate to the shortcomings in the field sampling methods, where sample size was inadequate, or incorrect stand location, in the field. For example, the discrepancies between SIDA data and field habitat data for altitude can be attributed to imprecise stand location, because both recordings were derived from the same source. It is also likely that the incorrect location of certain stands could have affected the correlations of other parameters (Table 4.13). Specific problems relating to field data measurement include: a) local variation in aspect, slope and soil depth; b) vegetation obscuring the horizon for exposure determinations; and c) subsurface rocks and gravel layers for soil depth determinations.

The soil depth determinations using the SIDA program are, furthermore, based mainly on degree of slope. This implies that all level ground, not on summits, would be greater than 1 500 mm in depth. Clearly this is not the case, hence it is suggested that soil depth be recorded in the field despite the limitations associated with this measurement. The good correlation between mean annual rainfall and altitude and latitude is applicable only to the main study area. Van Staden (1992) found that altitude alone was suf-



ficient and the SIDA program was adapted accordingly for that study area. Rainfall correlations should, therefore, be done on a local basis for stand interpolation.

Another source for derived rainfall data using a similar rationale namely, altitude and topography, are the data available from the Computing Centre for Water Research, University of Natal, Pietermaritzburg (Dent et al. 1987). These data, however, may not be sufficiently precise for stand purposes because they are only available on a one minute grid basis.

An advantage of the method of determining exposure described in this study is the effect of kloof orientation on exposure. Although such a situation was not sampled in this study, a kloof with a north-south orientation would be exposed to insolation for a far shorter duration than a kloof with an east-west orientation. These differences can be detected with the method described. The significance of such differences has yet to be assessed. It, furthermore, appears reasonable to consider exposure in terms of its components because the components are easier to quantify. If, by exposure is meant exposure to sun and wind, then these two components should be considered separately. Hence, exposure in this account is taken to mean exposure to insolation only.

The SIDA parameter of relative available moisture (RAM) is unlikely to be of much significance because of the large number of components used to derive the values. The algorithms are



TABLE 4.14. - PHYTOTAB-PC classification of a synthetic data set using both the heuristic and permutation methods. Noise is absent. Total separation units=0. Classification efficiency=100%.

Relevè number:	00 00 00 00 42 81 53 76
Species 1	++
Species 2	++
Species 3	++
Species 4	++
Species 5	++ ++
Species 6	++ ++
Species 7	++ ++
Species 8	++ ++
Species 9	++
Species 10	++
Species 11	++ ++
Species 12	++ ++
Species 14	++ ++ ++
Species 13	++ ++ ++
Species 15	++ ++ ++ ++
Species 16	++ ++ ++ ++
Species 19	+ +
Species 18	+ +
Species 17	+ +
Species 20	+

TABLE 4.15. - TWINSPAN classification of the synthetic data set in Table 4.14. Total separation units=21. Classification efficiency=74%.

Relevè number:	00000000 35671248
Species 9	++
Species 20	+
Species 19	+ +
Species 12	++++
Species 11	++++
Species 10	++
Species 14	+++++ +
Species 13	+++++ +
Species 8	++ + +
Species 7	++ + +
Species 18	+ +
Species 16	+++++++
Species 15	+++++++
Species 6	++++
Species 5	++++
Species 4	+ +
Species 3	+ +
Species 2	++
Species 17	+ +
Species 1	++





somewhat complex. A far simpler approach is effective soil wetting depth, used with good results by Van Staden (1992).

It is conceivable that for many of the parameters recorded in the field, more precise data could be obtained by using the SIDA program. This would also save field work time and allow information to be utilized which would otherwise be impracticable to record in the field, such as stand position in terms of drainage line and watershed, which is often impracticable to determine in the field. Although the SIDA program illustrates the potential of an indirect approach to recording certain habitat data, manual data input from topographic maps is extremely time-consuming. The approach is likely to be better adapted to digital terrain data and GIS processing, when such data becomes available. The SIDA programs have, accordingly, not been included in the PHYTOTAB-PC program package. However, the MINISTAT program is included in the PHYTOTAB-PC program package under the statistics utility.

#### 4.3 CLASSIFICATION

The PHYTOTAB-PC classification results of a synthetic data set are given in Table 4.14. The data set was constructed in such a manner that noise, in the form of separation units, would be absent if correctly classified. The initial relevè sequence was random. Identical results were achieved using both the heuristic approach which includes commonality, similarity and separation unit sequencing and the permutation approach which can only be used on small data sets. In the case of the last-mentioned



TABLE 4.16. - Number sequence comparison of the grouped relevè sequences in Tables 4.14 & 4.15

a) Correspondence of relevè-groups

b) Percentage correspondence of relevè-groups

		Table 4.14			
		1	2	3	4
Table 4.15	1	0	0	2	0
	2	0	0	0	2
	3	1	1	0	0
	4	1	1	0	0

		Table 4.14			
		1	2	3	4
Table 4.15	1	0	0	100	0
	2	0	0	0	100
	3	50	50	0	0
	4	50	50	0	0

Mean correspondence = 67%



approach, mirror-images are excluded. Both approaches are available on the PHYTOTAB-PC program package. Noise is absent, as shown by zero separation units and classification efficiency is 100%.

These results show that the algorithms used for relevè sequencing in the PHYTOTAB-PC program package (heuristic approach) can produce results comparable with relevès being tested in each possible position (permutation approach) for small data sets. Processing time limitations, even on a mainframe computer, preclude similar tests on large data sets but it is assumed that the results would be similar. However, the user should be aware that the relevè sequence obtained using the heuristic approach does not test relevès in each possible position and that the total separation units need not necessarily be the lowest obtainable for a particular data set. Classification efficiency values, however, indicate that the heuristic approach is an improvement on all classifications tested.

Table 4.15 gives the result of a TWINSPAN (Hill 1979a) classification of the same data set used in Table 4.14, with the initial relevè sequence ranging from 1 to 8 ascending, as required. Noise is present, as shown by the 21 separation units and classification efficiency is reduced to 74%. The differences, between the two methods, are illustrated in Table 4.16 in which the grouped number sequences of both results are compared in both absolute and percentage terms, with a mean correspondence of 67%.



TABLE 4.17.- TWINSpan classification of the synthetic data set in Table 4.14, using a random relevè input sequence. Total separation units=14. Classification efficiency=82%.

Relevè            00000000  
 number:         18243567

Species 3	++
Species 4	++
Species 5	++++
Species 6	++++
Species 7	++ ++
Species 8	++ ++
Species 13	++ ++++
Species 14	++ ++++
Species 15	++++++
Species 16	++++++
Species 18	+ +
Species 1	++
Species 2	++
Species 17	++
Species 11	++++
Species 12	++++
Species 19	+ +
Species 20	+
Species 9	++
Species 10	++



The TWINSPAN results are not satisfactory in that, groups formed by relevès 4;2 and 8;1 (Table 4.14) are contraposed in the TWINSPAN classification. Furthermore, the extreme relevès (3;8) in the TWINSPAN classification are centrally positioned in Table 4.14.

Table 4.17 gives the result of a TWINSPAN classification of the same data set used in Tables 4.14 & 4.15, but with the initial relevè sequence in a random order. However, the original relevè numbers have been retained. The results differ from the previous TWINSPAN results (Table 4.15) as shown by noise present which is represented by 14 separation units and a classification efficiency of 82%. The results in Tables 4.14 & 4.17 correspond in terms of grouping of relevès because a number sequence comparison between the two relevè sequences gives a mean correspondence of 100%. However, there is lack of correspondence in the position of the relevè-groups, where the first two releve-groups are contraposed. It is noteworthy that a random initial relevè sequence for TWINSPAN gives a better result, in this case, than the required input sequence.

The data set used in these examples contains no outliers and relevè-group pattern is consistent. Worse results could be expected with TWINSPAN where outliers are present and relevè-groups contain gaps. These results show that TWINSPAN is inadequate for final classifications in terms of relevè-group definition or sequence. Care should, furthermore, be exercised when using TWINSPAN to select species upon which to group relevès.

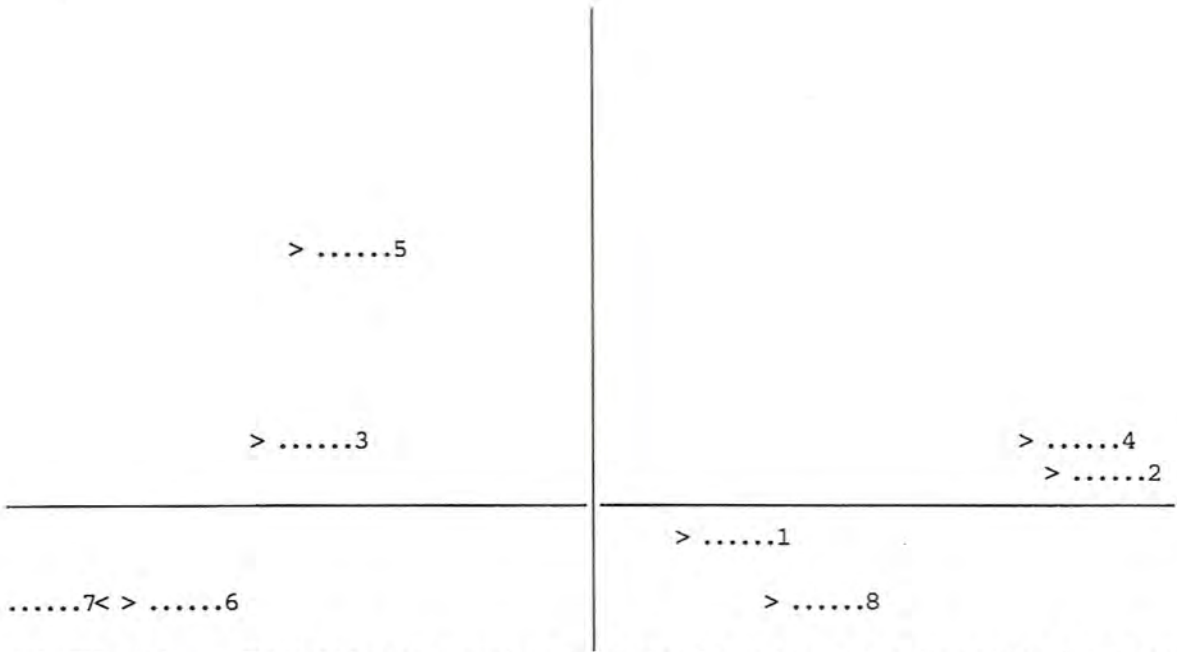


FIGURE 4.3. - DECORANA ordination of relevés using data in Table 4.14 and the CANOCO program. Arrowheads indicate position of each relevè. The horizontal axis is axis 1 and the vertical axis is axis 2.



On the other hand, the result of a DECORANA (Hill 1979b) ordination of the same data set (Table 4.14) using CANOCO (Ter Braak 1986), (Figure 4.3) shows complete correspondence in terms of grouping relevès with the groups formed in Table 4.14, as the mean correspondence between the two grouped relevè sequences is 100%. Furthermore, the relevè-groups thus formed, also occur in similar sequences in both the ordination (first axis) and the PHYTOTAB-PC classification, with the latter in reverse order. For this data set, the results of the DECORANA ordination, therefore, correspond far better with the PHYTOTAB-PC classification than with either of the TWINSpan classifications. Changing the input sequence did not affect the DECORANA results. DECORANA, therefore, supports the PHYTOTAB-PC classification but is not suitable for classification by itself, because of the difficulty in determining relevè-groups from cluster diagrams in large data sets.

Van Staden (1992) has also shown the relevance of DECORANA ordination in analysing the components of the main environmental gradient within his study area. However, the aforementioned application of ordination was not a species by relevè ordination but a diagnostic species by community ordination. This is an ordination of the diagnostic species in the matrix, that is, those species contributing to relevè-group pattern and using the relevè-groups as relevès, which are obtained with the synoptic table. This effectively reduces noise in the form of non-diagnostic species which can often be from one third to a half of the species total and reduces redundancy in the form of relevès belonging to the same relevè-groups. Gauch (1982) mentions that excessive noise can



TABLE 4.18.- PHYTOTAB-PC classification of the first data set in the second study area. Total separation units=127. Classification efficiency=70%.

Relevè-group number:	1	2	3	4	5
Relevè number:	2122121	012	0110011100	010	02
number:	3921800	554	9378212667	341	45

*Helichrysum rugulosum*  
*Lippia scaberrima*  
*Convolvulus sagittatus*

1	11
11	
	11

1

*Scabiosa columbaria*

111
-----

*Tephrosia capensis*  
*Salvia runcinata*  
*Tragus berteronianus*  
*Phyllanthus maderaspatensis*  
*Aristida congesta* subsp. *barbicollis*  
*Eragrostis pseudosclerantha*  
*Helichrysum nudifolium*

1

1+ 11 1
1111
111 1
11 11
11 1
1 11
11

1

*Tagetes minuta*

11	11 111	1
----	--------	---

*Chaetacanthus* sp.

111
-----

*Aristida congesta* subsp. *congesta*

21	12	1	2
----	----	---	---

*Sutera* sp.

111 1	1	1	11
-------	---	---	----

*Anthospermum pumilum*  
*Heteropogon contortus*

11	1	1111	1	11	111	
22	22	22	22	1	1	23

*Cymbopogon excavatus*

3

33	2	3	33	232	23
----	---	---	----	-----	----

*Eragrostis chloromelas*  
*Hyparrhenia anamesa*  
*Elionurus muticus*  
*Conyza podocephala*  
*Vernonia oligocephala*  
*Hermannia* cf. *grandifolia*  
*Crabbea angustifolia*  
*Verbena brasiliensis*  
*Protasparagus suaveolens*  
*Melinis repens*  
*Hibiscus microcarpus*  
*Verbena tenuisecta*  
*Athrixia elata*

3333223	333	3333433233	324	33			
33333	3	333	3333433333	333	33		
2222	22	222	22	2222	23	221	22
1	11111	111	1111111111	1	11		
	11	11	1	11	1	11	11
	1	1	1		111	11	11
	11		11	1			
11			1	1			
		1					
		1					
			1				
							+
		1					





TABLE 4.19.- PHYTOTAB-PC classification of the second data set in the second study area. Total separation units=99. Classification efficiency=72%.

Relevè-group number:	1	2	3	4	5
Relevè number:	01100	1121	21	2200	0102101120
	62143	7349	18	2352	7690580451

*Chaetacanthus* sp.

11 11 1 1

*Hermannia* cf. *grandifolia*

111

*Hibiscus microcarpus*

11

*Helichrysum nudifolium*

11

*Phyllanthus maderaspatensis*

11

*Aristida congesta* subsp. *barbicollis*

11 111

*Indigofera zeyheri*

11 11

*Sutera* sp.

11

*Themeda triandra*

1 1 1 1111 11

*Heteropogon contortus*

12 21 2 2 2121 2 1 2

*Eragrostis chloromelas*

33323 3333 33 3323 3333233333

*Elionurus muticus*

22222 1221 21 2222 2322 22223

*Hyparrhenia anamesa*

33333 3333 33 3333 333333 333

*Coryza podocephala*

111 1 1111 11 1111 11111 1111

*Trachypogon spicatus*

111 1111 11 11 1 111111111

*Vernonia oligocephala*

11111 1 1 11 111 11111111 1

*Scabiosa columbaria*

1111 11 11 1111 111111 1

*Crabbea angustifolia*

11111 11 1 111 111 111

*Cymbopogon excavatus*

3333 32 2 223 1 22

*Brachiaria serrata*

11 1111 11 11 1

*Anthospermum pumilum*

11 1 1 1 1 1 1

*Tephrosia capensis*

11 1 1 1 11

*Melinis repens*

1 1 1 11

*Eragrostis capensis*

11 1 1

*Tagetes minuta*

1 1 11

*Helichrysum rugulosum*

1

*Aristida congesta* subsp. *congesta*

1

*Gazania krebsiana*

1

*Dicoma zeyheri*

1

*Eragrostis racemosa*

1

*Sonchus wilmsii*

+

1	1	1		2		5		
-----				-----		-----		
3	3		1	1	1			
-----			-----		-----			
3	3		3		4		2	
-----				-----		-----	-----	
3	3		3	3		1		
-----			-----		-----	-----		
4		3		4		5		2

FIGURE 4.4. - Position of 2 x 2 m sampling units in the first 10 x 10 m quadrat showing relevè grouping according to the classification (Table 4.18), by relevè-group numbers in the position of each relevè. Each of the relevè-groups has a common border with the 10 x 10 m quadrat. Only one relevè-group (3), exhibits spatial integrity.

3		4	4		2		5	
-----			-----		-----	-----		
5		2		3		2		5
-----			-----		-----	-----	-----	
1		1		2		5	5	
-----			-----		-----	-----	-----	
1		5	5	5	5			
-----			-----		-----	-----	-----	
5		4		1	1		4	

FIGURE 4.5. - Position of 2 x 2 m sampling units in the second 10 x 10 m quadrat showing relevè grouping according to the classification (Table 4.19), by relevè-group numbers in the position of each relevè. Each of the relevè-groups has a common border with the 10 x 10 m quadrat. No relevè-groups exhibit complete spatial integrity.



influence DECORANA results and that noise and pattern are opposites. The argument that ordination can determine discontinuities in vegetation and hence the need for classification (Gauch 1982) is no longer relevant because the PHYTOTAB-PC programs determine such discontinuities as part of the classification process. If no discontinuities are present, only one relevè-group will be obtained. The CANOCO version of DECORANA is preferred because a) greater flexibility is possible with axis comparisons; b) point data is printed on the scattergrams; and c) other ordination options are available such as principal component analysis, reciprocal averaging and canonical correspondence analysis. The last-mentioned option was not tested because vegetation is classified using combined floristics and environmental variables, whereas the aim of this study is to classify purely according to floristics. A combined classification is also likely to be influenced by the selection of environmental variables.

The results of the PHYTOTAB-PC classification of the two 10 x 10 m quadrats are given in Tables 4.18 and 4.19 with diagrams of the sampling unit positions shown in Figures 4.4 and 4.5 respectively. In Table 4.18 all the relevè-groups except the last (relevè-group 5) are characterized by community diagnostic species. Table 4.19 is similar but relevè-group 3, without community diagnostic species, is situated between relevè-groups 2 and 4. It could be argued that relevè-group 3 should be sequenced at the end of the relevè-groups to eliminate the gap in the middle of the matrix. However, this would increase the total separation units to 147 and reduce the classification efficiency to 59%. The reason for this



TABLE 4.20. - PHYTOTAB-PC classification of the combined data sets in the second study area. Total separation units=375. Classification efficiency=68%

Relevé-group number:	1	2	3	4	5	6	7	8	9	10	11	12	13
Relevé number:	02	121221	010011110	02	4434434	13	5	3344	04	33	0014	334	223222
number:	45	009218	618237629	54	9408346	58	0	5950	37	23	7142	761	891763
<i>Helichrysum rugulosum</i>			11 1			1							
<i>Lippia scaberrima</i>			11		1								
<i>Convolvulus sagittatus</i>			11										
<i>Salvia runcinata</i>			11	11									
<i>Tragus berteronianus</i>				11	11								
<i>Phyllanthus waderaspatensis</i>				1111		1	1						
<i>Eragrostis pseudosclerantha</i>			1	111									
<i>Verbena brasiliensis</i>			-11	1	1								
<i>Eragrostis capensis</i>								1111					
<i>Hibiscus microcarpus</i>			1					11					
<i>Melinis repens</i>								1111	11				
<i>Indigofera zeyheri</i>										1111			
<i>Trachypogon spicatus</i>						11	1111	1	1	1111	1	11	11
<i>Scabiosa columbaria</i>					11	11111	11			111	1	11	11
<i>Heteropogon contortus</i>			22	2	11	22	2	2	122	21	11	2	22
											32	1	222
<i>Eragrostis chloromelas</i>	33	323332	333433233	33	3323333	33	3	3332	33	33	3423	333	323333
<i>Hyparrhenia anaxesa</i>	33	3	3333	333433333	33	3333333	33	3	333	33	3333	333	333333
<i>Elionurus muticus</i>	22	22222	22222	22	22	2122122	22	2	222	22	22	3121	223
<i>Conyza podocephala</i>	11	11	111	111111111	11	11111111	11	1	1111	11	1	1	1111
<i>Vernonia oligocephala</i>	11	11		1	11	1	1	111111	1	1111	1	11	111
<i>Cymbopogon excavatus</i>	23	3	3	23	33	2	322	3	22	22	323	331	33
<i>Anthospermum pumilum</i>				11	1	1111	1	1	1	11	111	1	1
<i>Crabbea angustifolia</i>	11			1	1	1	1	1	1	1	1	111	111
<i>Hermannia cf. grandifolia</i>	11	1	1	1	1	1	1	1			1111		
<i>Tagetes minuta</i>		11	11	111	1	1	1	1		11			
<i>Tephrosia capensis</i>			1+	11	1	1	1	1	1	1	1	11	1
<i>Brachiaria serrata</i>					11	1	1	1	1	1	1	1	111
<i>Sutera sp.</i>					11	11	1	1	1	1	1	1	1
<i>Themeda triandra</i>						1	1	1	111	1	11		
<i>Chaetacanthus sp.</i>						1	1	1	1	1	11	1	111
<i>Aristida congesta subsp. barbicollis</i>				111		1	1	111					
<i>Aristida congesta subsp. congesta</i>				1	21		1		1		22		
<i>Helichrysum nudifolium</i>				1	1		11						
<i>Verbena tenuisecta</i>											+		
<i>Athrixia elata</i>					1								
<i>Protasparagus suaveolens</i>								1					
<i>Gazania krebsiana</i>												1	
<i>Dicoma zeyheri</i>									1				
<i>Eragrostis racemosa</i>												1	
<i>Sonchus wilmsii</i>													+

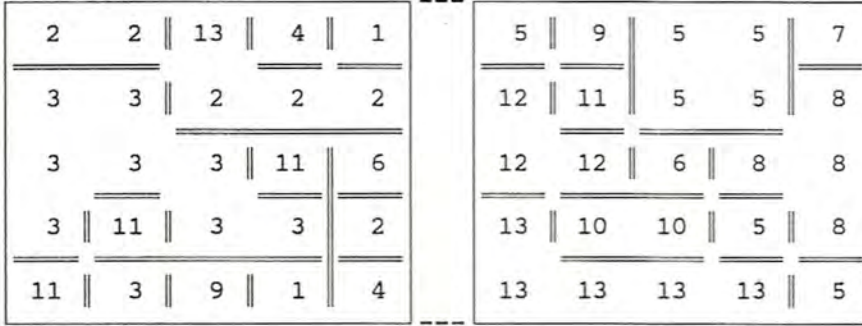


FIGURE 4.6. - Position of 2 x 2 m sampling units in both 10 x 10 m quadrats showing relevè grouping according to the combined classification (Table 4.20), by relevè-group numbers in the position of each relevè. Each of the relevè-groups has a common border with the 10 x 10 m quadrats except for relevès 32 & 33 (relevè-group 10), which are completely included within other relevè-groups. Only three relevè-groups, (6, 9 & 11) comprising relevès: 15, 38; 3, 47; and 7, 1, 14 & 42, respectively, intersect both data sets.



is the distribution of species such as *Melinis repens*, *Eragrostis capensis*, *Tagetes minuta* and *Helichrysum rugulosum* which, if the relevè-group were moved to the right of the matrix would introduce more gaps than those obtained. This relevè-group is, therefore, regarded as intermediate, floristically, between relevè-groups 2 and 3.

All the relevè-groups have common borders with the 10 x 10 m quadrats. The proportion of the variation of the vegetation units represented by each relevè-group is, therefore, unknown. The unknown proportions, which lie outside the 10 x 10 m quadrats, could obviously be larger in area than the proportions which are represented. Although spatial integrity is restricted to one relevè-group (Figure 4.4), in that all the relevès of that relevè-group have common borders, most of the relevès have common borders with the relevès in each relevè-group. This supports the classifications, especially if the detailed scale is taken into account.

The results of combining the two 10 x 10 m quadrats and producing a single classification are given in Table 4.20. Mean correspondence for the grouped relevè sequences for the separate and combined classifications are 69% for the first quadrat and 35% for the second quadrat. The positions of the classified units are shown in Figure 4.6.

As in Table 4.19, the position of the relevè-groups without community diagnostic species in Table 4.20, are related to species in the non-diagnostic section of the matrix, such as *Hermannia* cf.



*grandifolia*, *Sutera* sp., *Themeda triandra* and *Aristida congesta* subsp. *congesta* which, would increase separation units and decrease classification efficiency if these relevè-groups were sequenced at the right of the matrix. Based on these species the relevè-groups are regarded as floristically intermediate between the adjacent relevè-groups. This emphasizes the importance of considering all species in a matrix, and not merely the species in the diagnostic section of the matrix, for grouping and sequencing relevè-groups. Furthermore, all species must be considered before a logical grouping into diagnostic and non-diagnostic species can be made. This approach appears to correspond with what occurs in the field where a species is influenced by all the species in its vicinity and not only diagnostic species.

The relevè sequence is of prime importance in classification because this determines relevè-groups, floristic relationships with environment and mapping units. The species sequence is of secondary importance because it only shows species distribution over the determined relevè-groups. Species sequences can, therefore be modified by the user, allowing for more outliers where deemed necessary. In the Tables illustrated in this work, the criterion for species sequences has been minimum outliers, as programmatically sequenced. Relevè-group delimitation, on the other hand, can be changed without affecting the classification efficiency or total separation units i.e relevè-group delimiters can be moved, inserted or removed. However, Van Staden (1992) has shown that moving relevè-group delimiters can affect a community



ordination by reducing the number of SD units obtained. This has the effect of explaining less variation in the data.

Only three relevè-groups intersect both quadrats to form relevè-groups which did not occur in the separate classifications (Figure 4.6). No relevè-group retained its original relevès in the combined classification and only one species-group remained intact in the combined classification. A single vegetation unit, therefore, could be represented by two relevè-groups, each from a different 10 x 10 m quadrat because each relevè-group represents different variation within the vegetation unit, at the scale concerned. Although all plants within the study area were sampled with the contiguous sampling units, this nevertheless, illustrates that the sampling was inadequate in terms of the variation of the vegetation units, included within the study area. In other words, the vegetation units included in the study area are too small to adequately represent the variation of the vegetation units of which each is a part. Hence, the inconsistency in synthesizing the two data sets. It is, therefore, suggested that syntheses can only be reliably done where the vegetation units synthesized are entirely included within the respective study areas. In practical terms this means that vegetation units which are intersected by the borders of a study area can not be reliably synthesized because the proportion of variation included within the study area is unknown. Thus vegetation unit sampling adequacy relates not only to the proportion of species sampled within the study area but also to the degree that the vegetation unit variation is represented within the study area. This should also be considered



TABLE 4.21.- Synoptic version  
 PHYTOTAB-PC. Total separation units=5.  
 Classification efficiency=82%

Community:	1	2	3	4	5
<i>Helichrysum rugulosum</i>	3				
<i>Lippia scaberrima</i>	2	+			
<i>Convolvulus sagittatus</i>	2				
<i>Scabiosa columbaria</i>			5		
<i>Tephrosia capensis</i>		+	3		
<i>Salvia runcinata</i>			2		
<i>Tragus berteronianus</i>			2		
<i>Phyllanthus maderaspatensis</i>		+	2		
<i>Aristida congesta</i> subsp. <i>barbicollis</i>			2		
<i>Eragrostis pseudosclerantha</i>	+		2		
<i>Helichrysum nudifolium</i>			1		
<i>Tagetes minuta</i>	2		3	+	
<i>Chaetacanthus</i> sp.					5
<i>Aristida congesta</i> subsp. <i>congesta</i>			2	4	
<i>Sutera</i> sp.			5	2	4
<i>Anthospermum pumilum</i>	2	+	4	5	
<i>Heteropogon contortus</i>	3	4	2	4	
<i>Cymbopogon excavatus</i>	+		3	5	5
<i>Eragrostis chloromelas</i>	5	5	5	5	5
<i>Hyparrhenia anamesa</i>	5	5	5	5	5
<i>Elionurus muticus</i>	5	5	4	5	5
<i>Conyza podocephala</i>	5	5	5	+	5
<i>Vernonia oligocephala</i>	2	4	2	4	5
<i>Hermannia</i> cf. <i>grandifolia</i>	2	+	2	4	5
<i>Crabbea angustifolia</i>	2		2		
<i>Verbena brasiliensis</i>	2		1		
<i>Protasparagus suaveolens</i>		+			
<i>Melinis repens</i>		+			
<i>Hibiscus microcarpus</i>				+	
<i>Verbena tenuisecta</i>					+
<i>Athrixia elata</i>		+			



TABLE 4.22.- Synoptic version of Table 4.19 using  
PHYTOTAB-PC. Total separation units=4.  
Classification efficiency=75%

Community:	1	2	3	4	5
<i>Chaetacanthus</i> sp.	4	+	+		
<i>Hermannia</i> cf. <i>grandifolia</i>	4				
<i>Hibiscus microcarpus</i>	3				
<i>Helichrysum nudifolium</i>			3		
<i>Phyllanthus maderaspatensis</i>			3		
<i>Aristida congesta</i> subsp. <i>barbicollis</i>				3	
<i>Indigofera zeyheri</i>				2	
<i>Sutera</i> sp.				1	
<i>Themeda triandra</i>				3	4
<i>Heteropogon contortus</i>			3	5	3 4
<i>Eragrostis chloromelas</i>	5	5	5	5	5
<i>Elionurus muticus</i>	5	5	5	5	5
<i>Hyparrhenia anamesa</i>	5	5	5	5	5
<i>Conyza podocephala</i>	4	5	5	5	5
<i>Trachypogon spicatus</i>	3	5	5	4	5
<i>Vernonia oligocephala</i>	5	3	5	4	5
<i>Scabiosa columbaria</i>	4	3	5	5	4
<i>Crabbea angustifolia</i>	5	3		+	3
<i>Cymbopogon excavatus</i>	4	3	+	4	2
<i>Brachiaria serrata</i>	2	5		3	2
<i>Anthospermum pumilum</i>	2	+	+		2
<i>Tephrosia capensis</i>	2		+	+	1
<i>Melinis repens</i>		3	+	3	
<i>Eragrostis capensis</i>			5	+	+
<i>Tagetes minuta</i>		+	+	3	
<i>Helichrysum rugulosum</i>		+			
<i>Aristida congesta</i> subsp. <i>congesta</i>					+
<i>Gazania krebsiana</i>	+				
<i>Dicoma zeyheri</i>					+
<i>Eragrostis racemosa</i>		+			
<i>Sonchus wilmsii</i>					+

TABLE 4.23.- Re-classification of 4.21 & 4.22 and emphasizing community diagnostic species. Total separation units=32. Classification efficiency=76%. Relevè numbers for Table 4.22 have been renumbered from 6 to 10



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

Community number: 0 0 0 0 1 0 0 0 0 0  
5 4 6 8 0 9 3 7 2 1

*Eragrostis capensis* 5 + +  
*Indigofera zeyheri* 2  
*Salvia runcinata* 2  
*Tragus berteronianus* 2  
*Eragrostis pseudosclerantha* 2 +  
*Hibiscus microcarpus* + 3  
*Helichrysum rugulosum* + 3  
*Lippia scaberrima* + 2  
*Convolvulus sagittatus* 2

<i>Eragrostis chloromelas</i>	5	5	5	5	5	5	5	5	5	5
<i>Hyparrhenia anamesa</i>	5	5	5	5	5	5	5	5	5	5
<i>Elionurus muticus</i>	5	5	5	5	5	5	4	5	5	5
<i>Conyza podocephala</i>	5	+	4	5	5	5	5	5	5	5
<i>Vernonia oligocephala</i>	5	4	5	5	5	4	2	3	4	2
<i>Cymbopogon excavatus</i>	5	5	4	+	2	4	3	3		+
<i>Hermannia cf. grandifolia</i>	5	4					2	4	+	2
<i>Anthospermum pumilum</i>	5	2	+	2			4	+	+	2
<i>Heteropogon contortus</i>	4		5	4	3	2	3	4	3	
<i>Crabbea angustifolia</i>		5		3	+	2	3		2	
<i>Scabiosa columbaria</i>		4	5	4	5		3	5		
<i>Tephrosia capensis</i>		2	+	1	+	3			+	
<i>Trachypogon spicatus</i>		3	5	5	4		5			
<i>Brachiaria serrata</i>		2		2	3		5			
<i>Sutera sp.</i>	4			1		2		5		
<i>Tagetes minuta</i>				+		3	3	+		2
<i>Melinis repens</i>				+		3		3	+	
<i>Aristida congesta subsp. congesta</i>	4					+	2			
<i>Verbena brasiliensis</i>							1			2
<i>Chaetacanthus sp.</i>	5	4	+							
<i>Themeda triandra</i>						4	3			
<i>Aristida congesta subsp. barbicollis</i>						3		2		
<i>Helichrysum nudifolium</i>							3	1		
<i>Phyllanthus maderaspatensis</i>							3	2		+
<i>Verbena tenuisecta</i>						+				
<i>Athrixia elata</i>										+
<i>Protasparagus suaveolens</i>										+
<i>Gazania krebsiana</i>						+				
<i>Dicoma zeyheri</i>							+			
<i>Eragrostis racemosa</i>										+
<i>Sonchus wilmsii</i>							+			+



when describing vegetation unit variation for a single study area.

The classification results of the two 10 x 10 m contiguously sampled quadrats (Tables 4.18 & 4.19) show that discontinuities can be found in contiguously sampled data and that such data can be classified. Furthermore, high classification efficiencies are obtained with detailed sampling, indicating the effectiveness of the PHYTOTAB-PC programs for large scale work.

Table 4.21 is a synoptic version of Table 4.18; Table 4.22 is a synoptic version of Table 4.19; and Table 4.23 shows the result of combining Tables 4.21 and 4.22 and re-classifying objectively to produce a single classification, using the PHYTOTAB-PC program package. No subjective decisions were made in the re-classification process. In Table 4.23 only community diagnostic species non-diagnostic species have been shown to facilitate comparisons. The community diagnostic species-groups are reduced with the combined classification from a total of eight to five. The only species-group which remains unchanged is the first species-group in Table 4.21 which is the last diagnostic species-group in Table 4.23. The integrity of the original relevè-groups, which formed the synoptic matrices, remains unchanged but the species-groups and the relationships between the relevè-groups has altered considerably. Therefore, re-classification of the separate synoptic tables has not improved the synthesis which confirms that vegetation unit variation sampling was inadequate.



Inadequate vegetation unit variation sampling, as illustrated above, can also explain why division of a data set into subsets, as suggested by Coetzee (1982), to improve classifiability, appears to improve vegetation classifications. It is suggested that division of a data set has the effect of increasing the number of vegetation units by dividing vegetation units at a particular scale into variations of the vegetation units, so that a variation is regarded as a vegetation unit. This can only increase the number of vegetation units by:

$$I = J$$

where I = increase in vegetation units; and

J = number of subsets in which vegetation units occur more than once.

Clearly this can not improve the results because the increase can only be obtained by division of common vegetation units. Although the resulting vegetation units could be regarded as vegetation units at a larger scale they could also be highly arbitrary. This is dependent on the manner in which the original data set is subdivided. However, the reason for subdivision is usually to improve a classification so that it can be assumed that such subdivisions are not based on a verified classification i.e. confirmed floristic relationships, but possibly some or other habitat factor. Such floristic and habitat correlations can only be reliably shown after a classification, so that subdividing a data set to improve a classification can only be considered a doubtful practice. This is confirmed by Van Staden (1992) where the study area could be divided into Arid Bushveld and Mixed Bushveld (Acocks 1975, 1988). However, his koppies which are



TABLE 4.24. - PHYTOTAB-PC random classification (1) of the first data set in the second study area. Total separation units=260. Classification efficiency=40%

Relevè number: 00111001 12 20 2201 210 200111  
82043417 81 37 4251 563 069952

<i>Eragrostis pseudosclerantha</i>	11	1		1			
<i>Aristida congesta</i> subsp. <i>barbicollis</i>	1	1	1				
<i>Chaetacanthus</i> sp.		1	1				1
<i>Convolvulus sagittatus</i>				11			
<i>Lippia scaberrima</i>					11		1
<i>Scabiosa columbaria</i>					1	1	1
<i>Helichrysum rugulosum</i>			1				1 1
<i>Sutera</i> sp.	1			1	1 1	1	1 11
<i>Eragrostis chloromelas</i>	34323343	23	33	3333	323	233333	
<i>Hyparrhenia anamesa</i>	34333333	33	33	3333	333	33333	
<i>Conyza podocephala</i>	111	11	1	11	111	111	11
<i>Elionurus muticus</i>	2222221	2	23	2222	2	2	222222
<i>Anthospermum pumilum</i>	11	11	11	11	1	1	1
<i>Vernonia oligocephala</i>	1	11	1	1		1	111
<i>Heteropogon contortus</i>		22	32	2	2	2	1
<i>Cymbopogon excavatus</i>	2	3322		3		3	2
<i>Hermannia</i> cf. <i>grandifolia</i>		11	11	1	1	1	11
<i>Tagetes minuta</i>	1		1	1	1	1	11
<i>Aristida congesta</i> subsp. <i>congesta</i>	21		2		2		1
<i>Tephrosia capensis</i>	1+			1	1	1	
<i>Crabbea angustifolia</i>	1	1	1				1
<i>Phyllanthus maderaspatensis</i>		1	1		1		1
<i>Salvia runcinata</i>					1	1	1
<i>Tragus berteronianus</i>		1	1				1
<i>Verbena brasiliensis</i>			1	1		1	1
<i>Helichrysum nudifolium</i>	1					1	
<i>Protasparagus suaveolens</i>							1
<i>Melinis repens</i>							1
<i>Hibiscus microcarpus</i>			1				
<i>Verbena tenuisecta</i>			+				
<i>Athrixia elata</i>					1		



TABLE 4.25. - PHYTOTAB-PC random classification (2) of the first data set in the second study area. Total separation units=272. Classification efficiency=37%

Relevè number: 0211112 1010 210 2 21 021 01000  
5404190 3859 362 5 17 622 48173

<i>Lippia scaberrima</i>	1	1				1	
<i>Scabiosa columbaria</i>	11		1				
<i>Helichrysum rugulosum</i>		11					1

<i>Crabbea angustifolia</i>	1	1	11	1
-----------------------------	---	---	----	---

<i>Chaetacanthus</i> sp.	1					1	1
<i>Aristida congesta</i> subsp. <i>congesta</i>		2	1		1		221

<i>Eragrostis chloromelas</i>	3332332	3333	324	3	33	333	32433
<i>Hyparrhenia anamesa</i>	333333	3333	334	3	33	333	33333
<i>Elionurus muticus</i>	2222222	2222	2	2	2	222	2 132
<i>Conyza podocephala</i>	111 1 1	1111	111	1	11	111	11 11
<i>Anthospermum pumilum</i>	1 1	11	1	11	1 1	1111	
<i>Heteropogon contortus</i>	2 2122	2 2	2		2 1	23	
<i>Cymbopogon excavatus</i>		3 3 3 3	2 3		3 3	2 232	
<i>Vernonia oligocephala</i>	1 11 1	111	1 1 1			1 1	
<i>Hermannia</i> cf. <i>grandifolia</i>	1 11		1 1		1	1111	
<i>Sutera</i> sp.	11 1	11				1 11	
<i>Tagetes minuta</i>		1 1	1	11	11	1	
<i>Tephrosia capensis</i>	1	1	1+		1	1	
<i>Phyllanthus maderaspatensis</i>	1	1	1	1	1		
<i>Tragus berteronianus</i>		1 1		1	1		
<i>Verbena brasiliensis</i>	1			11	1		
<i>Salvia runcinata</i>	1		1		1 1		
<i>Eragrostis pseudosclerantha</i>		11	1			1	
<i>Aristida congesta</i> subsp. <i>barbicollis</i>		1	1	1			
<i>Helichrysum nudifolium</i>	1		1				
<i>Convolvulus sagittatus</i>					1	1	
<i>Protasparagus suaveolens</i>		1					
<i>Melinis repens</i>		1					
<i>Hibiscus microcarpus</i>					1		
<i>Verbena tenuisecta</i>							+
<i>Athrixia elata</i>	1						



TABLE 4.26. - PHYTOTAB-PC random classification (3) of the first data set in the second study area. Total separation units=263. Classification efficiency=39%

Relevè number: 200201 101 112 2101 0 02012101  
563155 244 804 2619 9 23830771

<i>Scabiosa columbaria</i>	11	1				
<i>Sutera</i> sp.	1 11 1 1	1	1	1	1	
<i>Lippia scaberrima</i>	1	1 1				
<i>Hermannia</i> cf. <i>grandifolia</i>	11 1 11 11 11					1
<i>Eragrostis pseudosclerantha</i>		1			1 11	
<i>Aristida congesta</i> subsp. <i>barbicollis</i>					1 1 1	
<i>Helichrysum nudifolium</i>					1	1
<i>Crabbea angustifolia</i>		1		1	111	
<i>Eragrostis chloromelas</i>	333333	332	233	3243	3	43332333
<i>Hyparrhenia anamesa</i>	333333	333	333	3333	3	4333 333
<i>Elionurus muticus</i>	222222	222	22	2 12	2	22222 32
<i>Conyza podocephala</i>	111111	11	111	11	1	11111111
<i>Anthospermum pumilum</i>	111	1 1 1 1	1	1	1	11 11
<i>Vernonia oligocephala</i>	1 1 11	11	1	1	1	1 11
<i>Cymbopogon excavatus</i>	332	323		2	3 2	33 3
<i>Heteropogon contortus</i>	1	2	2 2 2	32	2	222 1
<i>Tagetes minuta</i>	11	1		11		1 1 1
<i>Tephrosia capensis</i>		1 1		1	+	1 1
<i>Aristida congesta</i> subsp. <i>congesta</i>	11			2	1 2	2
<i>Phyllanthus maderaspatensis</i>		1	1 1			1 1
<i>Tragus berteronianus</i>		1			1	1 1
<i>Verbena brasiliensis</i>	1			1		1 1
<i>Salvia runcinata</i>	1	1		1		1
<i>Chaetacanthus</i> sp.	1		1	1		
<i>Helichrysum rugulosum</i>			1	1		1
<i>Convolvulus sagittatus</i>	1		1			
<i>Protasparagus suaveolens</i>		1				
<i>Melinis repens</i>		1				
<i>Hibiscus microcarpus</i>						1
<i>Verbena tenuisecta</i>				+		
<i>Athrixia elata</i>			1			





TABLE 4.27. - PHYTOTAB-PC random classification (4) of the first data set in the second study area. Total separation units=296. Classification efficiency=32%

Relevè number:	11110202	12120120	00	21	01	100
	24396430	75837019	28	21	46	551
<i>Chaetacanthus</i> sp.	1	1				1
<i>Helichrysum rugulosum</i>	1	1	1			
<i>Convolvulus sagittatus</i>			1	1		
<i>Cymbopogon excavatus</i>	333	3	23	3	3	3
<i>Crabbea angustifolia</i>	1	1		1	1	1
<i>Tragus berteronianus</i>	1	1	1		1	
<i>Phyllanthus maderaspatensis</i>	1	1	1	1		1
<i>Anthospermum pumilum</i>	111	111	1	1	1	1
<i>Scabiosa columbaria</i>		1				11
<i>Eragrostis chloromelas</i>	32333332	33233333	43	33	32	334
<i>Hyparrhenia anamesa</i>	33333333	33333333	43	33	33	333
<i>Elionurus muticus</i>	22222222	2	23222	22	22	2
<i>Conyza podocephala</i>	1	1	1111	11111111	11	11
<i>Heteropogon contortus</i>	22212	2	2	22		1
<i>Vernonia oligocephala</i>	1	11	11	1	1	1
<i>Hermannia cf. grandifolia</i>	1	1	11	11		11
<i>Tagetes minuta</i>	1	1	1	1	1	11
<i>Sutera</i> sp.	11	11		1	1	
<i>Aristida congesta</i> subsp. <i>congesta</i>		1	1	2	12	
<i>Tephrosia capensis</i>	1			1	+1	1
<i>Verbena brasiliensis</i>			1	1	11	
<i>Eragrostis pseudosclerantha</i>	1		1		11	
<i>Salvia runcinata</i>	1	1			1	1
<i>Lippia scaberrima</i>		1			1	1
<i>Aristida congesta</i> subsp. <i>barbicollis</i>	1		1		1	
<i>Helichrysum nudifolium</i>					1	1
<i>Protasparagus suaveolens</i>						1
<i>Melinis repens</i>						1
<i>Hibiscus microcarpus</i>			1			
<i>Verbena tenuisecta</i>						+
<i>Athrixia elata</i>		1				



geographically and floristically part of the Mixed Bushveld would be separated from the footslopes which floristically are part of the Arid Bushveld.

In Tables 4.24 to 4.27 the relevè sequences have been generated with a random number generator available on the PHYTOTAB-PC package. The data is the same as that for Table 4.18. Relevè-groups have been formed and species sequenced programmatically. Pattern is evident in all four Tables and classification efficiencies are 40% or less. The number of diagnostic species is not necessarily correlated with the classification efficiency, because Table 4.27 has 32% classification efficiency and nine diagnostic species, whereas, Table 4.25 has 37% classification efficiency and only six diagnostic species. These Tables indicate that a classification efficiency of 40% or less is comparable to that which can be obtained with a random relevè sequence and that not much credence should be given to such classifications.

These results are only four of many tested to show that a classification is not unique and that many arbitrary solutions are possible, so that virtually any relevè sequence could produce some sort of pattern. The presence of noise in the form of outliers or gaps in a species distribution over the relevè sequence preclude the attainment of 100% classification efficiency.

What then is an adequate classification? It appears from the results that classification efficiencies of 60% or higher can be considered adequate, but the user will not know if, for example,



65% is the best obtainable or if through further sequencing a better classification can be obtained. In all the tests conducted, no higher classification efficiency was obtained than that obtained with the PHYTOTAB-PC programs. It must be remembered that the classification efficiency relates to relevè sequence and not relevè-group delimitation or species sequence, both of which can change without affecting the classification efficiency value. These changes could, however, effect pattern in terms of species sequencing.

The relevè sequence determines which relevès can be grouped to form relevè-groups and also the relationships between relevè-groups. The delimitation of relevè-groups according to a given relevè sequence is influenced by the scale at which relevè-groups are to be recognized. If the species are not sequenced, no pattern will be evident at this stage, but neither the relevè-group delimitation nor the relationships between relevè-groups will be affected and the classification efficiency value will remain the same. The advantage of this approach to classification, is that with correct species sequencing the adequacy of a classification can be inferred, to an extent, by pattern strength through relevè-group delimitation and species sequencing. The PHYTOTAB-PC programs, sequence species according to minimum noise. The user can increase pattern strength through selective re-sequencing of species. This is, however, based on subjective decisions as to what is an adequate balance between noise and pattern.

Gauch (1982) regards noise (unco-ordinated occurrences) and



pattern (co-ordinated occurrences) as opposites, in that if noise is decreased then pattern is increased. He further regards noise as unquantifiable. As has been shown, pattern is unquantifiable. Noise, however, can be attributed to three sources, namely, i) that which is related to a relevè sequence and can be quantified in terms of included blanks; ii) that which is related to a species sequence and relates mainly to outliers; and iii) that which is related to relevè-group delimitation and also can relate to outliers. The first-mentioned can be quantified in terms of separation units which determine the relevè sequence that produces least noise overall, for a given data set. This should also be the sequence in which pattern, after relevè-group delimitation and species sequencing, is strongest. This supports hypothesis (iv) in Section 1.3. However, species sequence-related noise, is not quantifiable because of differences in the amount that could be permitted by different users for a particular data set. The flexibility, in species sequencing is an advantage because it permits matrix simplification, in terms of grouping species groups to form gradients, without loss of information, as suggested in Section 3.5.2.1. The last-mentioned source of noise is relevè-group delimitation which is essentially related to scale and can affect species sequencing, as has been shown. The use of noise to obtain a relevè sequence and hence classify the vegetation can be described as a minimum entropy method.

The relevè-group delimitation programs allow for some flexibility in relevè-group delimitation so that the classification can match the scale of the stratification, but all the relevè-groups are de-



TABLE 4.28. - Commonality index of relevès with species occurrences in relevès and relevè-groups represented, from PHYTOTAB-PC, for the first data set, in the second study area

Relevè number	Commonality Index	Species occurrences	Relevè-group
23	125	5	1
19	175	7	1
10	175	7	1
22	200	8	1
4	200	8	5
25	200	8	5
20	200	8	1
11	225	9	3
18	225	9	1
9	225	9	3
21	225	9	1
16	250	10	3
1	250	10	4
6	275	11	3
7	275	11	3
14	275	11	4
15	275	11	2
24	275	11	2
8	275	11	3
3	275	11	4
12	300	12	3
17	300	12	3
13	300	12	3
2	300	12	3
5	300	12	2



limited at the same scale. Where mixed scales are present, the user can adjust relevè-group delimitation by inserting or removing relevè-group delimiters.

Objectivity in a classification is inversely proportional to the number of decisions required to complete the classification. It is for this reason that stratification and sampling should be according to scale so that fewer decisions are required in the classification process, thereby increasing objectivity.

Although field proficiency can increase with repetition the complexity and processing time increase exponentially with data set size. Furthermore, redundancy also appears to increase with data set size. It is, therefore, suggested that for maximum efficiency, data sets should not exceed 150 relevès, where possible. Automatic relevè sequencing with the PHYTOTAB-PC program package is also limited to a product of species and relevès of 186 000 which limits species to 1240 for 150 relevès. The other programs in the package do not have such limitations and are dependent on hard disk space.

Table 4.28 gives the commonality sequence output for relevès, using the PHYTOTAB-PC program package, together with relevè-group numbers in which each relevè occurs. The first seven relevès occur in either of the extreme relevè-groups, whereas the last 14 occur in the central relevè-groups. It is the identification of the extreme relevès that facilitate the heuristic approach to relevè sequencing. Outliers can confuse this pattern in that an extreme



TABLE 4.29. - Commonality index of species with species occurrences and position in the classified matrix, from PHYTOTAB-PC for the first data set in the second study area

Species	Commonality Index	Species occurrences	Position in matrix
<i>Verbena tenuisecta</i>	10	1	
<i>Protasparagus suaveolens</i>	11	1	single
<i>Melinis repens</i>	11	1	occurrences
<i>Athrixia elata</i>	11	1	
<i>Hibiscus microcarpus</i>	12	1	
<i>Convolvulus sagittatus</i>	18	2	
<i>Helichrysum nudifolium</i>	21	2	
<i>Helichrysum rugulosum</i>	24	3	
<i>Lippia scaberrima</i>	27	3	
<i>Chaetacanthus</i> sp.	32	3	mainly
<i>Scabiosa columbaria</i>	34	3	diagnostic
<i>Aristida congesta</i> subsp. <i>barbicollis</i>	36	3	species
<i>Verbena brasiliensis</i>	38	4	
<i>Salvia runcinata</i>	42	4	
<i>Eragrostis pseudosclerantha</i>	44	4	
<i>Tragus berteronianus</i>	45	4	
<i>Crabbea angustifolia</i>	47	5	
<i>Phyllanthus maderaspatensis</i>	57	5	
<i>Aristida congesta</i> subsp. <i>congesta</i>	66	6	
<i>Tephrosia capensis</i>	68	6	
<i>Tagetes minuta</i>	82	8	
<i>Sutera</i> sp.	88	8	
<i>Hermannia</i> cf. <i>grandifolia</i>	97	10	
<i>Heteropogon contortus</i>	116	12	
<i>Vernonia oligocephala</i>	118	12	
<i>Cymbopogon excavatus</i>	123	12	
<i>Anthospermum pumilum</i>	142	13	
<i>Elionurus muticus</i>	213	22	
<i>Conyza podocephala</i>	216	22	general
<i>Hyparrhenia anamesa</i>	236	24	occurrences
<i>Eragrostis chloromelas</i>	244	25	



TABLE 4.30. - Similarity co-efficients, using the PHYTOTAB-PC program package, for the relevès in the first quadrat, in the second study area

Relevè number	Similarity co-efficient
23	0,000
11	77,778
22	77,273
21	85,000
6	71,429
7	84,615
2	82,143
8	76,667
3	78,571
14	78,571
1	80,769
25	69,231
4	100,000
10	83,333
9	80,000
20	77,273
13	71,429
17	75,000
12	70,588
16	73,333
5	73,333
15	76,667
24	78,571
19	69,231





relevè, with a species outlier from a central relevè-group, could cause that relevè to occupy a central position. The exact relevè, however, is not required as only an approximation of the starting relevè is needed to save processing time.

Table 4.29 gives the commonality sequence output for species, using the PHYTOTAB-PC program package, together with the position in the classified matrix that the species occupy. This sequence is not required for species classification but is included for completeness of processing and could be of benefit to users.

No pattern is formed in the classified matrix with single and general occurrence species and these are, therefore, not regarded as diagnostic. Not all the species in the middle group (Table 4.29), are necessarily diagnostic, as their distribution included in the relevant study area, could be inadequate.

Table 4.30 gives the initial similarity co-efficients for each successive pair of relevès. The starting relevè (23) is that obtained from the commonality sequence (Table 4.28). Only two relevès are identical, namely, numbers 25 & 4 (100% similarity). Similarity sequencing, unlike the commonality sequencing, is repeated for a data set and is included to save processing time in relevè-group construction.

Classification efficiencies for some published classifications are as follows: Van Staden (1992) 60%; Scheepers (1975)



Kroonstad 48%; Bethlehem 44%; Leistner (1967) 62% and Westfall et al. (1985) 64%.

The classification efficiency of 60% obtained by Van Staden (1992) is considered good because of the integrity of the vegetation units; correspondence between stratification and vegetation units; and correspondence between environment and vegetation units. These results were achieved in an area with weak environmental gradients and no visibly distinct vegetation unit borders. The classification efficiency of 62% obtained by Leistner (1969) is attributed to a low species richness and clearly visible differences in the floristic composition of the vegetation units. This facilitates visual sequencing of relevés and species. Scheepers (1975) worked in generally overgrazed grassland with a high species richness and generally weak environmental gradients. The difficulty experienced in classifying such data sets is shown in the classification efficiencies of 48% (Kroonstad) and 44% (Bethlehem). These results support classification efficiency values as a reliable method of assessing the efficiency of a classification.

The last-mentioned data set (Westfall et al. 1985) was reclassified using the PHYTOTAB-PC program package. Relevé-group delimiters were inserted in three relevé-groups in which obvious subdivisions were possible and removed between relevé-groups which had no community diagnostic species. Removal and insertion of delimiters was necessary because the original stratification, and

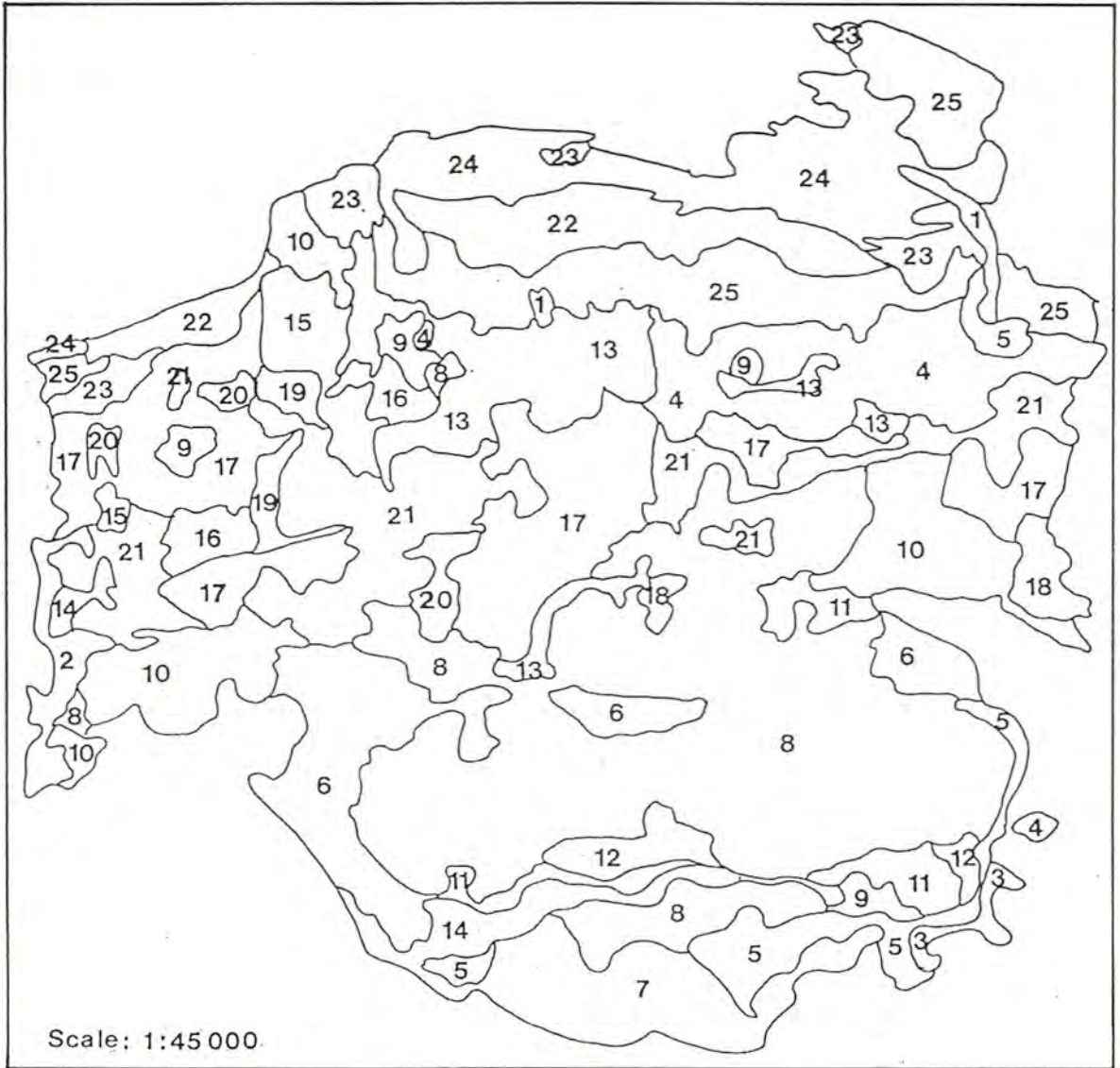


FIGURE 4.7. - The spatial relationships of relevé-groups formed by the re-classification of relevés in Table 4.31. Figures refer to the relevé-group numbers used in the re-classification.



hence sampling, was not according to scale. The results of the re-classification, without altering the programmatic relevè and species sequences, are shown in Table 4.31 (back pocket). Classification efficiency is increased from 64% to 68% and 25 communities can be identified in contrast to the 22 of the original classification (Table 4.32 back pocket). The number of species represented in community diagnostic species-groups is 65 in both classifications, but the re-classification is based on minimum outliers, so that community diagnostic pattern could be enhanced by inclusion of more species with outliers.

Figure 4.7 shows the spatial relationships of the re-classified relevè-groups (Table 4.31). Relevè-groups 1 and 3 are the same for both classifications (Tables 4.31 & 4.32) but in the re-classification their positions are switched thereby making forest associated with North-Eastern Mountain Sourveld the most extreme group on the left of Table 4.31. Relevè-group 3 is then better associated geographically and floristically with Sour Bushveld than in Table 4.32. The extreme relevè-group on the right of Table 4.31 represents North-Eastern Mountain Sourveld as was the case in the original classification. The differences in the classifications are illustrated by a grouped number sequence comparison where the mean correspondence is 36%. This shows a poor comparison between relevè-groups of the two classifications. Generally, the re-classified relevè-groups appear to correspond better with the environment than the original classification. Furthermore, the spatial integrity of the re-classified relevè-groups was such that a much better correspondence was obtained with vegetation pattern



on the aerial photograph, used to show spatial relationships between the re-classified relevè-groups (Figure 4.7), than was obtained with the original vegetation map.

As with Tables 4.19 and 4.20 relevè-groups without community diagnostic species, are regarded as floristically intermediate to the adjacent relevè-groups. Relevè-groups, such as 5 and 8 with a low constancy of community diagnostic species, could probably be improved by including more species in their respective species-groups, although more outliers would be included. Such refinement of the matrix was not the purpose of the classification, however, but rather to show the basic pattern obtainable from the program.

A classification is not necessarily invalid if community diagnostic species have a low constancy. This can be as a result of widely spaced species for which the dimensions of the sampling units (quadrats) are inadequate. To ensure adequate sampling unit dimensions, for community diagnostic species, would require a classification, prior to sampling, to determine which species are diagnostic for communities and suitable sampling unit dimensions for their sampling.

The absence of a community diagnostic species in a particular relevè can be regarded as relevè-related noise and its presence can be inferred by the other species present in the relevè. This is fundamental to the Braun-Blanquet approach (Mueller-Dombois & Ellenberg 1974). The sequencing of species and relevès based on visual pattern formation, however, can be greatly facilitated by



recognition of relevè-groups, in which community diagnostic species have a high constancy. However, this should not preclude the inclusion of relevès in such groups, where the community diagnostic species are absent, if the total floristic composition of the relevès indicates inclusion. It is, therefore, suggested that "total floristics" refers not to sampling, which is clearly not the case (section 3.2.4) but rather to the classification in which all species present in the matrix must be considered.

It is for these reasons that the PHYTOTAB-PC classification (Table 4.31) is considered a better classification than the original visual classification (Table 4.32). The re-classification also supports the hypothesis (iii) that more than one solution is possible in a vegetation data set. However, the problem of what is the "best" classification, still occurs, hence the need for verification.

#### 4.4 VERIFICATION

The results for this section are shown in Van Staden (1992) as fieldwork is ongoing in the main study area.

Apart from the classification efficiency values and pattern strength, the relationships between the classified vegetation units and stratification units and differentiating environmental factors, are the main criteria in assessing the adequacy of a classification. Such relationships can also have the practical value of improving vegetation map quality as well as the under-



standing of vegetation and hence management implications, in terms of limiting environmental factors.

It is probable that each species distribution, in a study area, could be linked to one or other environmental influence. Classification has the effect of averaging such influences, on a vegetation unit basis, so that the main environmental influences are apparent. It is, therefore, necessary to group the correct species, on a relevè-group basis to show such influences, hence, environmental correlation supporting a classification. The following guidelines are suggested for assessing a classification:

- i. scale should be appropriate, for example, seasonality can not be expected to be differentiating at 1:50 000 scale as it is usually applicable at biome scale;
- ii. gradients should be present. It is unlikely that a different environmental factor will differentiate each vegetation unit in a study area;
- iii. environmental relationships should be relatively simple, because of the averaging effect. Greater complexity could be expected with individual species distributions; and,
- iv. environmental relationships should be logical in context. In other words the relationships should make sense for the particular study area.

Furthermore, the use of the PHYTOTAB-PC program package for determining the classification and environmental factor correspondence emphasized the following:

- i. environmental relationships are often hierarchical;
- ii. class intervals for grouping environmental continua are not



TABLE 4.33. - Alphabetical listing of species selected from the PHYTOBAS data bank from undisturbed dune crests with low rainfall (less than 250 mm) from Leistner (1967)

---

<i>Acacia erioloba</i>	<i>Jatropha erythropoda</i>
<i>Acacia haematoxylon</i>	<i>Lapeirousia littoralis</i>
<i>Acrotome inflata</i>	<i>Limeum arenicolum</i>
<i>Aristida meridionalis</i>	<i>Limeum fenestratum</i>
<i>Boscia albitrunca</i>	<i>Limeum sulcatum</i>
<i>Brachiaria glomerata</i>	<i>Oxygonum delagoense</i>
<i>Bulbostylis hispidula</i>	<i>Phyllanthus omahekensis</i>
<i>Centropodia glauca</i>	<i>Plexipus pumilus</i>
<i>Chamaesyce inaequilatera</i>	<i>Plinthus sericeus</i>
<i>Citrullus lanatus</i>	<i>Pollichia campestris</i>
<i>Crotalaria spartioides</i>	<i>Requienia sphaerosperma</i>
<i>Cynanchum orangeanum</i>	<i>Sesamum sp.</i>
<i>Eragrostis lehmanniana</i>	<i>Stipagrostis amabilis</i>
<i>Heliotropium ciliatum</i>	<i>Stipagrostis uniplumis</i>
<i>Hermannia tomentosa</i>	

---





necessarily equal; and,

- iii. vegetation unit limits need not necessarily correspond with changes in environmental factor values at the same points.

Classification adequacy can also be assessed by the integrity of the vegetation units to be mapped. These should form mappable units, except where outliers occur. The occurrence of outliers can often be attributed to mixed scales such as in Figure 4.7. Ground truth after a classification is not to assess the classification but to assess the reliability of mapped borders as well as the relevancy of community diagnostic species as indicators for the entire communities they represent.

An adequate classification is of little value, however, if its relevancy is low. The relevancy of a classification is directly proportional to the uses that can be derived from it.

#### 4.5 DERIVATIVES

Community structure, community composition analyses, stand analyses, growth form analyses, community cover assessments, and species cover relationships are programmatically derived using the PHYTOTAB-PC program package and are illustrated by Van Staden (1992). These derivatives increase the understanding of vegetation component interactions and the uses for a classification, considerably. No similar programs are available.

An example of the data bank derivatives is given in Table 4.33 in



which species occurring on dune crests with less than 250 mm mean annual rainfall, are listed. In this case relevè numbers with the required habitat are input. It is necessary to know which relevès are required as the data bank only contains floristic information. A GIS could also be used to select relevès where relevant environmental data sets (coverages) are available.

Where species presence over more than one vegetation unit is required, then species representing all the relevant vegetation units can be selected to retrieve associated species from the data bank using Boolean "and" logic. This could, for example, be applied to determine common species for a stand so that only those not listed, need be recorded in the field. The PHYTOBAS data bank can also be used for a single data set.

A static perspective of vegetation units, which are inherently dynamic, can limit the application potential of vegetation ecology. The dynamic succession approach suggested by Clements (1916) and Bews (1916) appear to confuse scale both spatially and temporally so that the concepts used have little practical relevance. Inferring dynamics of vegetation units from sampling units, representing a moment in time requires that:

- i the vegetation units be comparable i.e. sampled at the same scale;
- ii. the condition of the vegetation units relative to some reference be known; and,
- iii. the trend or direction of change can be inferred, for



TABLE 4.33. - PHYTOTAB-PC applications excluding data input,  
transfer, corrections and file listing

---

**A Sequencing**

Automatic relevè sequence	Random sequence
Automatic species sequence	Reverse sequence
Alphanumeric sequence	User sequence
Ascending sequence	

**B Classification processing**

Checklist compilation	Species cover relationships
Community composition analysis	Stand phase analysis
Community cover assessment	Synoptic matrix
Community & habitat correlation	

**C Internal utilities**

Co-ordinate processing	Number set comparisons
Digital mapping	Plant identification key
File merging	Sample & stand dimensions
Format conversion	Statistics

**D External utilities**

3-D ordination plotting  
Species name search, checking and author additions

**E Data bank**

Data set splitting	Information retrieval
Data sets synthesis	Species spelling checker

---



practical relevance.

The bench mark concept or reference is essential if trend is to be inferred. However, fixed benchmarks, as suggested by Foran *et al.* (1978) can result in such scale and succession differences to the vegetation units for which they serve as reference, that their validity could be questioned. The required number of such benchmarks for relevant-scale work would also be prohibitive. It appears far more feasible for benchmarks to be constructed from relevès representing a particular vegetation unit. Such a benchmark could be: a synoptic relevè; relevès representing the central part of a vegetation unit; or a synthetic relevè based on species composition from the relevès representing the vegetation unit, such as is used with the community composition analysis. Trend can be inferred from a species composition gradient, of relevès representing the vegetation unit, relative to the benchmark. Thus succession appropriate to scale and practical time-span can be determined as is shown with the stand phase analysis. It is, therefore, suggested that veld condition assessment techniques, including vegetation monitoring, for whatever purposes, can only be effective for large areas, if based on an adequate classification.

#### 4.6 PHYTOTAB-PC

The applications possible with the PHYTOTAB-PC program package are summarized in Table 4.33. The package includes online manuals as well as online fault-finding. Processing speed is dependent on

matrix size. Automatic relevè sequencing is the most time-consuming of the programs and is dependent on matrix size and processor speed. For example, the programs have been tested on an 80486 processor which can halve processing time.

The PHYTOTAB-PC program package is essentially a research tool which can facilitate objectivity in vegetation analyses. The addition of programs to test various results such as comparison of grouped number sequences and the statistical utility program are for research purposes. If a purely production package were required a far simpler package could be developed. Several of the options are unique, such as automatic relevè sequencing, automatic species sequencing, environmental factor correspondence, vegetation component analysis, stand phase analysis and plant identification key generation. These coupled with data bank facilities ensure a powerful and comprehensive tool for vegetation analyses.

The amount of data collected during fieldwork in a typical project is vast. For example, 150 relevès with a total of 400 species represents a matrix of 60 000 cells. To this must be added cover, growth form, and environmental information. Reduction of this information to meaningful pattern requires considerable processing. This is achieved with the PHYTOTAB-PC program package and flexibility in application is still maintained. Data integrity and security is assured by writing all relevant files to separate diskettes. However, a package of this nature with the flexibility offered, presupposes a fundamental knowledge of vegetation classification theory. Without this background a researcher is unlikely



to be able to apply the programs effectively even though they are menu driven and a comprehensive online manual is available.

#### 4.7 REFERENCES

- ACOCKS, J.P.H. 1975. Veld types of South Africa. 2nd edition.  
*Memoirs of the Botanical Survey of South Africa* 40: 1-128.
- ACOCKS, J.P.H. 1988. Veld types of South Africa. 3rd edition.  
*Memoirs of the Botanical Survey of South Africa* 57: 1-146.
- BEWS, J.W. 1916. Account of the chief types of vegetation in South Africa with notes on plant succession. *Journal of Ecology* 4: 129-159.
- CLEMENTS, F.C. 1916. *Plant succession: An analysis of the development of vegetation*. Carnegie Institute, Washington.
- COETZEE, B.J. 1982. *Phytosociology, vegetation structure and landscapes of the Central District, Kruger National Park*. D.Sc. thesis, University of Pretoria, Pretoria.
- DENT, M.C., LYNCH, S.D. & SCHULZE, R.E. 1987. Mapping mean annual and other rainfall statistics over southern Africa. *Agricultural Catchments Research Unit Report No. 27*. Department of Agricultural Engineering, University of Natal, Pietermaritzburg.
- EDWARDS, D. 1967. A plant ecology survey of the Tugela Basin. *Memoirs of the Botanical Survey of South Africa* 36: 1-285.
- FORAN, B.D., TAINTON, N.M. & BOOYSEN, P. de V. 1978. The development of a method for assessing veld condition in three grassveld types in Natal. *Proceedings of the Grassland Society of southern Africa* 13: 27-33.



- FREUND, J.E. & WILLIAMS, F.J. 1958. *Modern business statistics*. Pitman, London.
- GAUCH, H.G. 1982. *Multivariate analysis in community ecology*. Cambridge, New York.
- HILL, M.O. 1979a. *TWINSPAN - A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes*. Unpublished report. Cornell University, Ithaca, New York.
- HILL, M.O. 1979b. *DECORANA - A FORTRAN program for detrended correspondence analysis and reciprocal averaging*. Unpublished report. Cornell University, Ithaca, New York.
- LEISTNER, O.A. 1967. The plant ecology of the southern Kalahari. *Memoirs of the Botanical Survey of South Africa* 38: 1-172.
- LONDO, G. 1976. The decimal scale for relevès of permanent quadrats. *Vegetatio* 33: 61-64.
- MUELLER-DOMBOIS, D. & ELLENBERG, H. 1974. *Aims and methods of vegetation ecology*. Wiley, New York.
- MUNSELL SOIL COLOR CHARTS, 1954. Munsell Color co., Baltimore.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1986. Biomes of southern Africa - An objective categorization. *Memoirs of the Botanical Survey of South Africa* 54: 1-98.
- SCHEEPERS, J.C. 1975. *The plant ecology of the Kroonstad and Bethlehem areas of the Highveld Agricultural Region*. D. Sc. thesis, University of Pretoria, Pretoria.
- TER BRAAK, C.J.F. 1986. Canonical correspondence analysis: A new Eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.



- VAN STADEN, J.M. 1992. *Die fitososiologie van die Steenbokpan omgewing in die noordwes-Transvaal*. M. Sc. thesis, University of Pretoria, Pretoria.
- WESTFALL, R.H. 1981. *The plant ecology of the farm Groothoek, Thabazimbi District*. M.Sc. thesis, University of Pretoria, Pretoria.
- WESTFALL, R.H. & PANAGOS, M.D. 1988. The plant number scale - an improved method of cover estimation using variable-sized belt transects. *Bothalia* 18: 289-291.
- WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. (1985). The plant ecology of the farm Groothoek, Thabazimbi District. II. Classification. *Bothalia* 15: 655-688.





## CHAPTER 5. CRITICAL EVALUATION AND CONCLUSIONS

### Scale

The tendency for vegetation heterogeneity to increase with decreasing scale necessitates increasing cognizance of scale, as scale decreases. This is particularly important for: vegetation unit recognition and hence mapping; determining the influence of environment on vegetation units; and vegetation sampling requirements at small-scale. At large-scale, however, where sampling unit dimensions are such that the sampling unit is implicitly representative of a stand, then cognizance of scale can often be implicit. Implicit scale recognition is from about 1:8 000 to about 1:50 000 based on a 200 m<sup>2</sup> sampling unit or quadrat. For larger scales a smaller sampling unit would be required. For uniformity, however, and to facilitate communication, it is recommended that scales in vegetation ecology work be explicit.

### Stand area

Definition of stand area is the means by which scale is made explicit for vegetation sampling in this work. The stand is integral to the Braun-Blanquet method but relies on researcher decisions in its implementation. By linking scale to stand area these decisions are reduced and the stand becomes easier to delimit in the field. It is, therefore, recommended that stand area be linked to scale, as a simple means of expressing scale, in vegetation ecology work.

Although not tested it is hypothesized that the scale-defined



stand can also relate to the minimum sustainable area for the conservation of a vegetation unit.

### **Reconnaissance**

This work confirms the necessity of a reconnaissance, prior to sampling, as recommended in the Braun-Blanquet method. However, actions during a reconnaissance are made more explicit for vegetation stratification, by means of both satellite imagery and aerial photography, to determine vegetation variation and produce a preliminary vegetation map. Other actions recommended during reconnaissance are less critical and will depend to a large extent on particular methods adopted.

Although it is possible to obtain the basic data associated with a reconnaissance from an analysis of aerial photographs a comparison of field reconnaissance with aerial photograph analysis was not made.

### **Stratification**

This work suggests that, apart from stratifying vegetation for stratified random sampling purposes, a vegetation stratification can be used for a preliminary vegetation map which can also serve as a hypothesis, to be tested by the classification.

Suggested procedures for the use of satellite imagery for small-scale stratification and, although not tested, the potential of larger-scale satellite imagery, for large-scale stratification, are detailed. Application of these computerized procedures for



vegetation stratification has lead to the following suggestions for simplifying the visual use of aerial photography for small scale stratification:

- i. that vegetation be primarily stratified according to vegetation structure and cover as perceived by textural and contextual vegetation pattern;
- ii. that secondary and further divisions of the primary units be on the basis of topography and other physical factors; and,
- iii. that scale can be taken into account by minimum area comparisons.

The use of aerial photography in this way is a visual pattern analysis of aerial photographs rather than the aerial photograph interpretation method whereby topography forms the primary stratification. Thus a stratification of vegetation will be vegetation-based rather than topography-based. Vegetation pattern analysis can also make stereoscopy redundant for small-scale work. These methods have not been adequately tested but preliminary results of a single farm stratified in this way, appear satisfactory.

#### **Stand location**

Computerized random number generation for a 4 mm grid map overlay, together with grid co-ordinate conversion to degrees, minutes and seconds, reduces decision-making and facilitates objective stand location considerably, for random sampling. Subjective stand location, as can also be applied in the Braun-Blanquet method, is not considered here because of the experience required, which could increase considerably with decreasing scale, in its application.



Precision in stand location requirements will increase in direct proportion to the reliance placed, in the future, on geographic information systems for environmental data. Precision in stand location will also be required when vegetation data is input to geographic information systems. Visual stand location, in this study, proved adequate for classification purposes. However, indirect methods of obtaining environmental data, necessitated greater precision in stand location. Altimeters and optical rangefinders proved inadequate for the precision required. Trigonometrical techniques were not tried because of the cost of apparatus, time required and the inability to detect three beacons at many stands.

Geographic positioning by satellite is recommended for precise stand location and, although not tested, shows potential for saving time where random stand location is applied. Precise location of stands has a further advantage of enhancing the value of sampling data as such data can be used for monitoring purposes.

#### **Sampling unit area**

The method of sub-sampling a vegetation stand, applied in this study, is not recommended because: a) sub-sampling is time-consuming; b) the methods applied are based on minimum area which can lead to a low degree of constancy in a classified matrix; and c) sub-quadrat sampling has to change from the criterion of rooted plants, in the case of the smaller growth forms, to overhanging plants, for the larger growth forms. It is not certain that, even with the change in criteria, an adequate sample of the larger



growth forms is obtained.

However, the use of a defined stand, as recommended, introduces flexibility into sampling options. Stands can be sampled with a representative quadrat of suitable size, as is often the case, but the limits of representation will be known; large stands can be sub-sampled using quadrats, as applied in this study, but the area of such quadrats may have to be enlarged; point methods can be used within the stand area; and informal sampling using, for example, the plant number scale with pre-determined lower limits, can be applied. The stand defines the limits for the application of these methods, which should enhance repeatability, and hence scientific validity.

The use of the species-area regression can be used for areas up to one hectare, for comparative purposes. Although the regression seems adequate for these purposes, it could probably be improved with more data.

#### **Sampling unit location**

Sampling units, within the defined stand, which is circular, can be located with vectors, the components of which are direction and distance, from the stand centre, unless the sampling unit area approximates that of the stand, in which case the sampling unit location will be that of the stand. Where sampling is by a plot-less method such as a point method or informal sampling, then the criteria for sampling location should be made known to ensure repeatability. In the case of very large stands i.e. sampling at



very small scales such as 1:500 000 or less then geographic positioning by satellite can be employed. The importance of recording sampling unit location is firstly, to ensure, repeatability for scientific validity, and secondly, to enhance the value of the data so that the data can be used for benchmarks, monitoring and other purposes.

#### **Plant identification and verification**

The methods applied in this work have improved the species knowledge of the researchers concerned and permitted detection of infra-specific differences in plants, which might otherwise have gone undetected. Apart from the advantage of plant knowledge gained on a systematic basis, it is postulated that making known the criteria by which plants are identified can only improve the scientific validity of vegetation ecology. In the International Metric System (SI) standards have been determined for physical observations. No similar standards exist for botanical observations. Voucher specimens serve as a reference and not as a standard. Furthermore, character and character states used for plant identification in southern Africa vary depending on the systems used, hence, the criteria by which plants are identified should be made known. Without this information, the degree to which vegetation ecology work could be repeated, is very uncertain. The recommendation regarding plant identification is, therefore, aimed at the criteria for identification and not necessarily the methods applied in this study.



### **Species cover**

Species canopy cover, which is used in the Braun-Blanquet method, is preferred to basal cover estimations because more information can be derived from the former. Selection of cover class scale is dependent on the aim of a study. A simple, minimum class scale, such as the Braun-Blanquet scale, is likely to produce better cover pattern in a classified matrix, than a scale with many classes. The plant number scale, on the other hand, is preferred where cover values are treated arithmetically, and greater precision is required of cover values, than can be obtained with visual estimation techniques. No advantage can be found for scales with an intermediate number of cover classes, such as the Domin-Krajina cover-abundance scale.

An alternative method of cover determination is that which can be obtained with a point method, such as the wheel-point method. However, care should be exercised because such methods often rely on linear proportionate cover which should be converted to area cover.

### **Floristic data recording**

The recommended minimum floristic data to be recorded at each sampling unit are: species presence; canopy cover for each species; and growth form for each species. The last-mentioned is far easier to determine on-site, than by means of literature or herbarium specimens. Where precision is used in determining cover, such as with a point method or the plant number scale, then the addition of mean canopy diameter for each species, permits species



densities to be calculated.

No time-saving advantage was found with computerized field data capture, which also entailed a greater burden in the field.

#### **Habitat data**

The recommended minimum stand description data to be recorded for each stand are: stand/relevè number; date; and stand centre coordinates in degrees, minutes and seconds. In addition, the data required for herbarium labels, such as, major and minor localities and farm name can be included.

It is further recommended that increasing use be made of indirect means of environmental data capture, such as geographic information systems, as these become available, to decrease time spent in the field. Where field observations are made, careful consideration needs to be given to techniques, to ensure adequate samples, because the criteria for floristic sampling is not necessarily the same as that required for environmental sampling.

#### **Classification**

The aims of a classification are: floristic field data reduction to a comprehensible form, through the grouping of relevès, based on floristic similarity, to form relevè-groups and the grouping of relevè-groups, also based on floristic similarity, so that these can correspond with environmental gradients; and species grouping, based on occurrence in relevès, to emphasize the relationships between relevè-groups.





The classification of a floristic data set can result in more than one solution. A classification should, therefore, be verified to determine classification adequacy.

It is recommended that species-groups, common to two or more relevè-groups, be sequenced to correspond with environmental gradients, where possible. This simplifies the matrix, by reducing the number of species-groups, without loss of species-relationships information and provides more information on gradient relationships.

Care should be exercised in making inferences about vegetation units which are not entirely included in a study area because their total floristic variation is often unknown. This is particularly relevant to small-scale work where large vegetation units are often only partially included in a study area. When synthesizing two or more data sets cognizance should, therefore, be taken of partially included vegetation units.

Justification for splitting data sets to improve species constancy in relevè-groups or improving the classifiability of data sets could not be shown. It is suspected that improvements are obtained by grouping variation within larger units to form separate vegetation units. However, more work is required in this regard.

#### **Verification**

It is recommended that classification verification include: classification efficiency values; degree of integrity of mapping



units; degree of correspondence between mapping units and stratification; degree of correspondence between classification and differentiating environmental factors; and pattern strength in the constancy of community diagnostic species.

### **Derivatives**

The application potential of a classification is directly proportional to the information which can be derived therefrom.

Plant communities are derivatives of a classification and have immense value in land-use practices and planning, at various scales, because of their integrating effect on environmental influences and suitability as mapping units.

Plant community definition, in terms of species constancy in community diagnostic species, in a classified matrix, can be directly proportional to the degree of inter- and intra-community environmental change. However, poor community definition can also be caused by inadequate sampling unit area.

Environmental gradients can be derived directly from classified matrices and more importantly, natural discontinuities in these gradients can be ascertained from vegetation unit borders, within the gradients.

Vegetation structure can be derived from the recommended floristic data making separate structural analyses, as required in the Braun-Blanquet method, redundant.



Community composition analysis is a new approach to analysing community composition, in terms of species cover, frequency, growth forms and structure. The initial results indicate improvements in assessing vegetation condition and suitability for detecting both cover and composition change. However, more analyses are required to confirm these results.

Stand phase analysis combines the components of successional theory with floristically determined data to infer trend and status of stands in relation to the community. As with the community composition analysis, this is a new approach which appears promising but requires more analyses for confirmation.

Community cover assessment for determination of the adequacy of cover within a community, relates to minimum cover for soil conservation. However, these lower limits are based on experience in limited vegetation types and more input is required from other vegetation types before any reliance can be placed on the results.

#### **PHYTOTAB-PC**

This program package is a comprehensive package, suitable for many aspects of vegetation analysis, with new features, not available in other programs. The package can reduce decision-making, and hence reduce observer bias, with a corresponding increase in objectivity, in vegetation analysis data processing. Time spent on classification can be reduced considerably. The derivative programs have the potential for enhancing the uses to which classifications can be put, thereby increasing the application



potential of the Braun-Blanquet approach. Although the programs are fully documented on-line, and are menu driven, the number of programs and permutations possible, preclude use without an adequate background in Braun-Blanquet methodology and training in program use.

Caution is advised, as has been mentioned, when applying the community cover assessment program as modifications will probably be necessary. Furthermore, the automatic relevè sequencing programs can not be compared with the permutation approach, in all but the smallest data sets, so that the maximum classification efficiency for such larger matrices, can not be known.

Application of the recommended methods can reduce decision-making in vegetation ecology, decrease time spent on certain processes and increase the relevancy of classifications, especially for small-scale work, without conflicting with the basic principles of the Braun-Blanquet method. The derivatives should justify expanded use and application of classifications, especially in agriculture and conservation. In the case of conservation, vegetation should receive the highest priority in South Africa because adequate conservation of vegetation will ensure conservation of other primary natural resources such as soil and soil water.



OPSOMMING

Objektiwiteit in stratifisering, monsterneming

en klassifisering van plantegroei

deur

ROBERT HOWARD WESTFALL

Promotor: Prof. Dr. G.K. Theron

Mede-promotor: Dr. N. van Rooyen

in die

Departement Plantkunde

vir die graad

PHILOSOPHIAE DOCTOR (PLANTKUNDE)

Die doelwitte van hierdie studie was om waarnemers bevooroordeelings in die stratifiserings-, monsternemings- en klassifiseringsprosesse van plantegroei-wetenskap te verminder, om sodoende herhaalbaarheid en voorspelbaarheid, in hierdie prosesse te verbeter, asook om die geldigheid van plantegroei-klassifikasies te verbeter.

Die doelwitte is bereik deur: stratifiserings-, monsternemings- en klassifiseringsprosesse met skaal in verband te bring, deur middel van skaal-gedefiniëerde plantegroei-stande; die verbetering van klein-skaal stratifisering en dus plantegroei kartering deur middel van skaalverwante, plantegroei-versterkte, satelliet-beelde; die ontwikkeling van verbeterde metodes vir plant bedekkingsbepalings; plantegroei-klassifikasie deur minimum entropie; en die bepaling van plantegroei status deur spesie samestelling volgens bedekking-tot-frekwensie-verhoudings, binne groeivorms. 'n



Rekenaarpakket om objektiwiteit by plantegroei-analise te verbeter en om tyd van analyses te verminder, is ontwikkel.

Die resultate van hierdie studie het tot 'n vermindering in visuele volgorde-bepaling, van plantegroei-matrikse, veral van opnames teen 'n klein-skaal, gelei. Daar is ook 'n ooreenstemmende afname in tyd wat vir klassifikasie benodig word. Die plantegroei-stand, volgens skaal gedefiniëer, verseker buigbaarheid ten opsigte van metodes wat toegepas kan word. Aanbevelings sluit die noodsaaklikheid vir bevestiging van 'n klassifikasie in, en verskeie bevestigingstegnieke word beskryf. Afgeleide toepassings, voortspruitende uit klassifikasies sluit spesies-samestellings-analise ter bepaling van spesiesverhoudings op 'n plantgemeenskaps basis, bepaling van sleutelspesies op spesies kompetisievermoë gebaseer asook stand fase-analise vir die monitering van veranderings in plantegroei-stande, in. Behalwe plantegroei klassifikasie en afgeleide toepassings sluit die rekenaarpakket fasiliteite vir veld indentifikasie van plantspesies, statistiese dataverwerking, plantgemeenskap en omgewing korrelasie en 'n floristiese databank in.

Hierdie werk beklemtoon die belangrikheid van plantegroei-klassifikasies en beveel dat die hoogste nasionale prioriteit aan die hulpbron plantegroei toegeken word omdat korrekte plantegroei-bestuur die bewaring van grond en grondwater ook kan verseker.



SUMMARY

Objectivity in stratification, sampling  
and classification of vegetation

by

ROBERT HOWARD WESTFALL

Promoter: Prof. Dr G.K. Theron

Co-promoter: Dr N. van Rooyen

in the

Department of Botany

for the degree

PHILOSOPHIAE DOCTOR (BOTANY)

The aims of this study were to reduce observer bias in the stratification, sampling and classification processes in vegetation science so as to improve repeatability and predictability, as well as to increase the relevancy of vegetation classifications.

The aims were achieved by: relating the stratification, sampling and classification processes to scale by means of scale-defined vegetation stands; improving small-scale stratification and hence vegetation mapping by the use of scale-related vegetation-enhanced satellite imagery; developing a method for improved plant cover estimations; vegetation classification by minimum entropy; and assessing vegetation state through species composition according to cover-to-frequency ratios, within growth forms. A computer program package was developed to facilitate objectivity in the analysis of vegetation data and reduce time spent on analyses.



The results of this study have lead to a reduction in visual sequencing of vegetation matrices, especially those involving small-scale work, with a corresponding decrease in time required for classification. The vegetation stand, defined according to scale, introduces greater flexibility in the sampling methods that can be applied. Recommendations include the necessity for verifying a classification and several methods of verification techniques have been described. Classification derivatives include species composition analysis for determining species relationships on a community basis, as well as determining key species based on species competitive ability, and stand phase analysis for monitoring changes in vegetation stands. Apart from vegetation classification and derivatives the computer program package includes facilities for field identification of plant species, statistical treatment of data, community and habitat correlations and a floristic data bank.

This work emphasizes the importance of vegetation classifications and recommends that the vegetation resource be given the highest national priority because correct vegetation management can also ensure conservation of soil and soil water.





#### ACKNOWLEDGEMENTS

The assistance and co-operation of the following persons and organizations are gratefully acknowledged:

- (1) Prof. G.K. Theron and Dr N. van Rooyen for their leadership, guidance and encouragement throughout this work.
- (2) Dr J.C. Scheepers for guidance and encouragement.
- (3) The President, Agricultural Research Council and the Director, Roodeplaat Grassland Institute, for the use of their excellent facilities and for permission to use the results of an official facet for thesis purposes.
- (4) The staff of the National Herbarium, Pretoria for identification of the herbarium voucher specimens.
- (5) Dr O.A. Leistner and Prof. A.E. van Wyk for comments and suggestions on the characters and character states used in the plant identification aid.
- (6) The Librarians at the Mary Gunn Library, National Botanical Institute, particularly Mrs E. Potgieter for assistance with literature searches.
- (7) Mr J.M. van Staden for photographic work and testing of many hypothesis.
- (8) Mr M.D. Panagos for capable technical assistance, data processing and illustrating the characters used in plant identification.
- (9) Mrs J. Schaap for the drawings.
- (10) Miss A.P. Backer for accurate typing of the first draft.



## CURRICULUM VITAE

Robert Howard Westfall was born on the 17th December, 1944 in Kokstad, Cape Province. He completed high school at Fish Hoek, Cape Province in 1962. In 1976 he obtained a B.Sc. degree at the University of Pretoria with Botany and Zoology as major subjects.

In 1977 he was appointed as a Professional Officer at the Botanical Research Institute (Department of Agricultural Technical Services) in Pretoria. He was seconded, for completion of the B.Sc. (Hons.) degree, to the University of Pretoria in 1978 with ecology and taxonomy as major subjects. An M.Sc. degree was awarded to him in 1981 by the University of Pretoria on submission of a thesis entitled "The plant ecology of the farm Groothoek, Thabazimbi district.

In 1989 he was transferred to the Grassland Research Centre, which became the Roodeplaat Grassland Institute within the Agricultural Research Council in 1992, where he occupies the post of Principal Agricultural Researcher. During his period of service he has been engaged in ecological research and has concentrated on South African savanna vegetation and methodology in vegetation research.

Mr Westfall is a member of several scientific societies such as the South African Association of Botanists, the Grassland



Society of southern Africa and the South African Institute of Ecologists. He has been an active member of these societies, presenting papers at congresses and serving on the Northern Transvaal Branch of the South African Association of Botanists. He has served as treasurer for several years on this committee. He also serves on Departmental committees and was the convenor of the National Task Group for monitoring and evaluating vegetation change. He is registered as a natural scientist with the South African Council for Natural Scientists. The following is a list of published papers:

- (1) VAN DER MEULEN, F., MORRIS, J.W. & WESTFALL, R.H. 1978. A computer aid for the preparation of Braun-Blanquet tables. *Vegetatio* 38: 129-134.
- (2) VAN DER MEULEN, F. & WESTFALL, R.H. 1979. A vegetation map of the western Transvaal Bushveld. *Bothalia* 12: 731-735.
- (3) VAN DER MEULEN, F. & WESTFALL, R.H. 1980. Structural analysis of Bushveld vegetation in Transvaal, South Africa. *Journal of Biogeography* 7: 337-348.
- (4) WESTFALL, R.H., DEDNAM, G., VAN ROOYEN, N. & THERON, G.K. 1982. PHYTOTAB - A program package for Braun-Blanquet tables. *Vegetatio* 49: 35-37.
- (5) WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. 1983. A homogeneity index based on species diversity, in Sour Bushveld. *Bothalia* 14: 299-301.



- (6) WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. 1983. The plant ecology of the farm Groothoek, Thabazimbi District. I. Ordination. *Bothalia* 14: 785-790.
- (7) WESTFALL, R.H., EVERSON, C.S. & EVERSON, T.M. 1983. The vegetation of the protected plots at Thabamhlope Research Station. *South African Journal of Botany* 2: 15-25.
- (8) WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. 1983. Veld condition assessments in Sour Bushveld. *Proceedings of the Grassland Society of southern Africa* 18: 73-76.
- (9) WESTFALL, R.H. & PANAGOS, M.D. 1984. A cover meter for canopy and basal cover estimations. *Bothalia* 15: 241-244.
- (10) WESTFALL, R.H. & DREWES, R. 1984. Grass root pattern in an Orange Free State Floodplain. *Bothalia* 15: 293-294.
- (11) RUTHERFORD, M.C. & WESTFALL, R.H. 1984. Sectors of the Transvaal province of South Africa. *Bothalia* 15: 294-295.
- (12) WESTFALL, R.H. 1984. Review of the vegetation of the Subantarctic Islands Marion and Prince Edward by N.J.M. Gremmen. Junk, The Hague. 1982. *Bothalia* 15: 340.
- (13) WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. 1985. The plant ecology of the farm Groothoek, Thabazimbi District. II. Classification. *Bothalia* 15: 655-688.
- (14) WESTFALL, R.H. 1985. PHYTOCAP. A field data capture program for the PHYTOTAB program package. *Bothalia* 15: 749-750.



- (15) WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. 1986. The plant ecology of the farm Groothoek, Thabazimbi District. III. An annotated checklist. *Bothalia* 16: 77-82.
- (16) PANAGOS, M.D., WESTFALL, R.H. & SCHEEPERS, J.C. 1986. Baseline data for the vegetation of two protected plots at the Matimba Power Station, Ellisras, NW Transvaal. *Bothalia* 16: 89-91.
- (17) WESTFALL, R.H. 1986. Review of Sampling methods and taxon analysis in vegetation science (Handbook of Vegetation Science. Part 4) edited by R. Knapp. Junk, The Hague. 1984. *South African Journal of Botany* 52: 192-193.
- (18) WESTFALL, R.H., GLEN, H.F. & PANAGOS, M.D. 1986. A new identification aid combining features of a polyclave and an analytical key. *Botanical Journal of the Linnean Society* 92: 65-73.
- (19) RUTHERFORD, M.C. & WESTFALL, R.H. 1986. The biomes of southern Africa - an objective categorization. *Memoirs of the botanical Survey of South Africa* 54: 1-98.
- (20) WESTFALL, R.H. & MALAN, O.G. 1986. A method for vegetation stratification using scale-related, vegetation-enhanced satellite imagery. *Bothalia* 16: 263-268.
- (21) WESTFALL, R.H. 1986. PHYTOLOC - a random number generator and sample set location program for stratified random vegetation sampling. *Bothalia* 16: 270-271.



- (22) WESTFALL, R.H., VAN STADEN, J.M. & PANAGOS, M.D. 1987.  
Predictive species area relations and determination of subsample size for vegetation sampling in the Transvaal Waterberg. *South African Journal of Botany* 53: 44-48.
- (23) WESTFALL, R.H. & DE WET, B.C. 1988. New programs for preliminary sequencing of species and relevés in phytosociological data sets *Bothalia* 18: 122-123.
- (24) MALAN, O.G. & WESTFALL, R.H. 1987. A new strategy for vegetation mapping with the aid of Landsat MSS data. *Advances in Space Research* 7,11: 97-103.
- (25) WESTFALL, R.H. & PANAGOS, M.D. 1988. An improved method of cover estimation using variable sized belt transects. *Bothalia* 18: 289-291.
- (26) DEALL, G.B. & WESTFALL, R.H. 1989. Improving the resolution of floristic/habitat pattern correlations on phytosociological tables. *Bothalia* 19: 263-266.
- (27) PANAGOS, M.D. & WESTFALL, R.H. 1989. Plant collecting apparatus for taxonomic and ecological studies 1. A lightweight plastic plant press for on-site specimen pressing. *Bothalia* 19: 266-267.
- (28) WESTFALL, R.H. 1989. Plant collecting apparatus for taxonomic and ecological studies 2. COLDAT: A field-data capture program for collector's data and herbarium labels. *Bothalia* 19: 267-268.
- (29) WESTFALL, R.H., BRITZ, P.J. & PANAGOS, M.D. 1989. Plant collecting apparatus for taxonomic and ecological studies 3. A new top-loading plant press for off-site specimen pressing. *Bothalia* 19: 268-269.



- (30) PANAGOS, M.D. & WESTFALL, R.H. 1989. Plant collecting apparatus for taxonomic and ecological studies 4. Drier-transporters for plant presses. *Bothalia* 19: 269-270.
- (31) PANAGOS, M.D., BRITZ, P.J. & WESTFALL, R.H. 1989. Plant collecting apparatus for taxonomic and ecological studies 5. A gas drier for field drying of plant specimens. *Bothalia* 19: 270-272.
- (32) WESTFALL, R.H. & PANAGOS, M.D. 1989. Plant collecting apparatus for taxonomic and ecological studies 6. A transportable map cabinet for vehicle and office use. *Bothalia* 19: 272-273.
- (33) WESTFALL, R.H., PANAGOS, M.D. & VAN STADEN, J.M. 1989. Plant collecting apparatus for taxonomic and ecological studies 7. A transportable camping kitchen for vehicle use. *Bothalia* 19: 273-274.
- (34) DEALL, G.B., THERON, G.K. & WESTFALL, R.H. 1989. The vegetation ecology of the Eastern Transvaal Escarpment in the Sabie area. 2. Floristic classification. *Bothalia* 19: 69-89.



REFERENCES

- ACOCKS, J.P.H. 1953. Veld types of South Africa. *Memoirs of the Botanical Survey of South Africa* 28: 1-192.
- ACOCKS, J.P.H. 1975. Veld types of South Africa. 2nd edition. *Memoirs of the Botanical Survey of South Africa* 40: 1-128.
- ACOCKS, J.P.H. 1988. Veld types of South Africa. 3rd edition. *Memoirs of the Botanical Survey of South Africa* 57: 1-146.
- ADAMSON, R.S. 1938. *The vegetation of South Africa*. Whitefriars, London.
- BAKER, H.A. 1970. A key for the genus *Erica* using edge-punched cards. *Journal of South African Botany* 36: 151-156.
- BEWS, J.W. 1916. Account of the chief types of vegetation in South Africa with notes on plant succession. *Journal of Ecology* 4: 129-159.
- BEWS, J.W. 1918. *The grasses and grasslands of South Africa*. Davis, Pietermaritzburg.
- BIANCHI, A.T.J. 1931. Een nieuwe determinatie-methode. *Tectona* 24: 884-893.
- BOLUS, H. 1886. Sketch of the flora of South Africa. *Official handbook of the Cape of Good Hope* pp. 286-317. Richards, Cape Town.
- BOLUS, H.M.L. 1958. Key to the genera of Mesembryanthemaceae. *Notes on Mesembryanthemum and allied Genera* 2: 392-403.
- BOUGHEY, A.S., BRIDGES, K.W. & IKEDA, A.G. 1968. An automated biological identification key. *Museum of Systematic Biology, University of California, Irvine, Research Series*, 2: i-xix, 1-35.





- BRAUN-BLANQUET, J. 1928. *Pflanzensoziologie*. 1 Aufl. Springer, Wien.
- BRAUN-BLANQUET, J. 1951. *Pflanzensoziologie*. 2 Aufl. Springer, Wien.
- BREDENKAMP, G.J. & THERON, G.K. 1978. A synecological account of the Suikerbosrand Nature Reserve: The phytosociology of the Witwatersrand geological system. *Bothalia* 12: 513-529.
- BURCHELL, W.J. 1822. *Travels in the interior of southern Africa*. Vol. 1. Paternoster, London.
- BURCHELL, W.J. 1824. *Travels in the interior of southern Africa*. Vol. 2. Paternoster, London.
- CLEMENTS, F.C. 1916. *Plant succession: An analysis of the development of vegetation*. Carnegie Institute, Washington.
- COETZEE, B.J. 1975. A phytosociological classification of the Rustenburg Nature Reserve. *Bothalia* 11: 561-580.
- COETZEE, B.J. 1982. *Phytosociology, vegetation structure and landscapes of the Central District, Kruger National Park*. D.Sc. thesis, University of Pretoria, Pretoria.
- COETZEE, B.J., VAN DER MEULEN, F., ZWANZIGER, S., GONSALVES, P. & WEISSER, P.J. 1976. A phytosociological classification of the Nylsvley Nature Reserve. *Bothalia* 12: 137-160.
- COETZEE, B.J., VAN WYK, P., GERTENBACH, W.P.D., HALL-MARTIN, A. & JOUBERT, S.C.J. 1981. 'n Plantekologiese verkenning van die Waterberggebied in die Noord-Transvaalse bosveld. *Koedoe* 24: 1-23.
- DEALL, G.B. & WESTFALL, R.H. 1989. Improving the resolution of floristic/habitat pattern correlations on phytosociological tables. *Bothalia* 19: 263-266.



- DENT, M.C., LYNCH, S.D. & SCHULZE, R.E. 1987. Mapping mean annual and other rainfall statistics over southern Africa. *Agricultural Catchments Research Unit Report No. 27*. Department of Agricultural Engineering, University of Natal, Pietermaritzburg.
- DUKE, J.A. 1969. On tropical tree seedlings. I. Seeds, seedling systems and systematics. *Annals of the Missouri Botanical Gardens* 56: 125-161.
- EDWARDS, D. 1967. A plant ecology survey of the Tugela Basin. *Memoirs of the Botanical Survey of South Africa* 36: 1-285.
- EDWARDS, D. 1983. A broad-scale structural classification of vegetation for practical purposes. *Bothalia* 14: 705-712.
- ELLIOTT, J.M. 1983. *Some methods for the statistical analysis of samples of benthic invertebrates*. Kendall, Wilson.
- ENGLER, A. 1882. *Versuch einer Entwicklungsgeschichte der Pflazenwelt insbesondere der Florengebiete seit der Tertiärperiode*. Engelmann, Leipzig.
- EYRE, S.R. 1963. *Vegetation and soils. A world picture*. Arnold, London.
- FORAN, B.D., TAINTON, N.M. & BOOYSEN, P. de V. 1978. The development of a method for assessing veld condition in three grassveld types in Natal. *Proceedings of the Grassland Society of southern Africa* 13: 27-33.
- FREUND, J.E. & WILLIAMS, F.J. 1958. *Modern business statistics*. Pitman, London.
- F.S.S.A., 1974. Manual of soil analysis methods. *The Fertilizer Society of South Africa* 37: 1-65.



- GABRIEL, H.W. & TALBOT, S.S. 1984. Glossary of landscape & vegetation ecology for Alaska. *BLM-Alaska Technical Report* 10: 1-137.
- GAUCH, H.G. 1982. *Multivariate analysis in community ecology*. Cambridge, New York.
- GIBBS RUSSELL, G.E., REID, C., VAN ROOY, J. & SMOOK, L. 1985. List of species of southern African plants. Edition 2. Recent literature and synonyms. Part 1. Cryptogams, Gymnosperms, Monocotyledons. *Memoirs of the Botanical Survey of South Africa* 51: 1-152.
- GIBBS RUSSELL, G.E., WELMAN, W.G., RETIEF, E., IMMELMAN, K.L., GERMISHUIZEN, G. PIENAAR, B.J., VAN WYK, M. & NICHOLAS, A. 1987. List of species of southern African plants. Edition 2. Recent literature and synonyms. Part 2. Dicotyledons. *Memoirs of the Botanical Survey of South Africa* 56: 1-270
- GLEN, H.F. 1974. A taxonomic revision of the subtribe *Gibbaeinae* (*Mesembryanthemaceae*). M.Sc. thesis, University of Cape Town.
- GOODALL, D.W. 1953. Objective methods for the classification of vegetation. i. The use of positive interspecific correlation. *Australian Journal of Botany* 1: 39-63.
- GOODALL, D.W. 1961. Objective methods for the classification of vegetation. iv. Pattern and minimal area. *Australian Journal of Botany* 9: 162-193.
- GOODALL, D.W. 1968. Identification by computer. *BioScience* 18: 485-488.
- GRISEBACH, A. 1872. *Die vegetation der Erde nach ihrer klimatischen Anordnung*. Engelmann, Leipzig.



- GYLLENBERG, H.G. 1965. A model for the computer identification of micro-organisms. *Journal of General Microbiology* 39: 401-405.
- HANSEN, B. & RAHN, K. 1969. Determination of Angiosperm families by means of a punched-card system. *Dansk botanisk Archiv* 26: 1-46 & 172 punched cards.
- HENKEL, J.S. 1934. *A field book of the woody plants of Natal and Zululand*. Robinson, Durban.
- HERRE, A.G.J. 1971. *Genera of Mesembryanthemaceae*. Tafelberg, Cape Town.
- HILL, M.O. 1979a. *TWINSpan - A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes*. Unpublished report. Cornell University, Ithaca, New York.
- HILL, M.O. 1979b. *DECORANA - A FORTRAN program for detrended correspondence analysis and reciprocal averaging*. Unpublished report. Cornell University, Ithaca, New York.
- IRVINE, L.O.F. 1941. *The major veld types of the northern Transvaal and their grazing management*. D.Sc. (Agric.) thesis, University of Pretoria, Pretoria.
- ITO, K. 1979. A tentative study of stratification diagrams. In Miyawaki, A. & Okuda, S. (eds). *Vegetation und Landschaft Japans*. pp.357-368. The Yokohama Phytosociological Society, Yokohama.
- JANSEN, H. 1982. The geology of the Waterberg Basins in the Transvaal, Republic of South Africa. *Memoirs of the Geological Survey of South Africa* 71: 1-98.
- LAMARCK, J.B.A.P.M. de. 1778. *Flore Francaise* 3 vols. Imprimerie Royale, Paris.



LAPAGE, S.P., BASCOMB, S. WILCOX, W.R. & CURTIS, M.A. 1970.

Computer identification of bacteria. In A. Baillie & R.J. Gilbert (eds), *Automation, Mechanisation and Data Handling in Microbiology* 1-22. Academic Press, London.

LAPAGE, S.P., BASCOMB, S. WILCOX, W.R. & CURTIS, M.A. 1973.

Identification of bacteria by computer: general aspects and perspectives. *Journal of General Microbiology* 77: 273-290.

LASSERRE, M., MALAN, O.G. & TURNER, B. 1983. *The application of principal component analysis to Landsat MSS data*. Proceedings of seminar on Principal component analysis in the atmosphere and earth sciences, Pretoria 7-8 February 1983.

LEENHOUTS, P.W. 1966. Keys in biology: a survey and a proposal of a new kind. *Proceedings of the Koninklijke Nederlandsche Akademie van Wetenschappen Amsterdam, Series C*. 69: 571-596.

LEISTNER, O.A. 1967. The plant ecology of the southern Kalahari. *Memoirs of the Botanical Survey of South Africa* 38: 1-172.

LE ROUX, C.J.G., GRUNOW, J.O., MORRIS, J.W., BREDEKAMP, G.J. & SCHEEPERS, J.C. 1988. A classification of the vegetation of the Etosha National Park. *South African Journal of Botany* 54: 1-10.

LICHTENSTEIN, H. 1811. *Reisen im südlichen Africa in der Jahren 1803, 1804, 1805, 1806*. Vol. 1. Salfeld, Berlin.

LICHTENSTEIN, H. 1812. *Reisen im südlichen Africa in der Jahren 1803, 1804, 1805, 1806*. Vol. 2. Salfeld, Berlin.

LONDO, G. 1976. The decimal scale for relevès of permanent quadrats. *Vegetatio* 33: 61-64.



- MACVICAR, C.N. 1973. *Soil map, Republic of South Africa.*  
1: 2 500 000. Soil and Irrigation Research Institute,  
Pretoria.
- MACVICAR, C.N., DE VILLIERS, J.M., LOXTON, R.F., VERSTER, E.,  
LAMBRECHTS, J.J.N., MERRYWEATHER, F.R., LE ROUX, J., VAN  
ROOYEN, T.H. & HARMSE, H.J.VON M. 1977. *Soil classifica-  
tion: A binomial system for South Africa.* Department of  
Agricultural Technical Services, Pretoria.
- MACVICAR, C.N., LAMBRECHTS, J.J.N., BENNIE, A.T.P., BRUCE, R.W.,  
DE VILLIERS, J.M., DOHSE, T.E., ELLIS, F., ELOFF, J.F., FEY,  
M.V., GREY, D.C., HARMSE, H.J.VON M., HARTMANN, M.O., HENSLEY,  
M., IDEMA, S.W.J., LAKER, M.C., LOXTON, R.F., MERRYWEATHER,  
F.R., MEYER, J.H., MICHAEL, D., PATERSON, D.G., SCHLOMS,  
B.H.A., SCHOEMAN, J.L., SCHONAU, A.P.G., SCOTNEY, D.M.,  
SNYMAN, K., TURNER, D.P., VAN NIEKERK, B.J., VAN ROOYEN, T.H.,  
VERSTER, E. & YAGER, T.U. 1991. *Soil classification: A  
taxonomic system for South Africa. Memoirs on the Agricultural  
Natural Resources of South Africa* 15: 1-257.
- MALAN, O.G. & LAMB, A.D. 1985. *Display of digital image data;  
quantitative and optimal.* Proceedings of the third South  
African symposium on Digital image processing, Durban 22-23  
July 1985.
- MARLOTH, R. 1908. *Das Kapland, insonderheit der Reich der  
Kapflora, das Waldgebiet und die Karoo pflanzengeografisch  
dargestellt.* Fischer, Jena.
- MEESTER, J. 1965. The origins of the southern African mammal  
fauna. *Zoologica Africana* 1: 87-93.



- MERXMUELLER, H. 1972. *Prodromus einer Flora von Suedwest Afrika*.  
Cramer, Lehre.
- METCALFE, C.R. & CHALK, L. 1979. *Anatomy of the Dicotyledons*. Vol.  
1, 2nd edition. Oxford University Press, Oxford.
- MEYER, P.G. 1969. *Eintuehrung in die Pflanzenwelt Suedwestafrikas*.  
S.W.A. Wissenschaftliches Gesellschaft, Windhoek.
- MORSE, L.E. 1971. Specimen identification and key construction  
with time-sharing computers. *Taxon* 20: 269-282.
- MORSE, L.E. 1974. Computer programs for specimen identification,  
key construction and description printing using taxonomic data  
matrices. *Publications of the Museum, Michigan State  
University,, Biological Series* 5: 1-128.
- MORSE, L.E. 1975. Recent advances in the theory and practice of  
specimen identification. In R.J. Pankhurst (Ed.), *Biological  
identification with Computers* pp. 11-52. Academic Press, London.
- MUELLER-DOMBOIS, D. & ELLENBERG, H. 1974. *Aims and methods of  
vegetation ecology*. Wiley, New York.
- MUNSELL SOIL COLOR CHARTS, 1954. Munsell Color co., Baltimore.
- NEWELL, I.M. 1970. Construction and use of tabular keys. *Pacific  
Insects* 12: 25-37.
- OSBORNE, D.V. 1963. Some aspects of the theory of dichotomous  
keys. *New Phytologist* 62: 144-160.
- PANKHURST, R.J. 1971. Botanical keys generated by computer.  
*Watsonia* 8: 357-368.
- PANKHURST, R.J. 1978. *Biological identification*. Edward Arnold,  
London.
- PARTON, W.J. & INNIS, G.S. 1972. *Some graphs and their functional  
forms*. U.S. International Biological Program, Grassland Biome



Project Technical Report 153: 1-41.

- POLE EVANS, I.B. 1922. The main botanical regions of South Africa. *Memoirs of the Botanical Survey of South Africa* 4: 49-53.
- POLE EVANS, I.B. 1936. A vegetation map of South Africa. *Memoirs of the Botanical Survey of South Africa* 15: 1-23.
- POLUNIN, O. 1969. *Flowers of Europe: A field guide*. Oxford University Press, London.
- POORE, M.E.D. 1956. The use of phytosociological methods in ecological investigations. iv. General discussion of phytosociological problems. *Journal of Ecology* 44: 28-50.
- REHMANN, A. 1880. Geo-botanische Verhältnisse von Süd-Afrika. *Botanische Centralblad* 1: 1119-1128.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1984. Sectors of the Transvaal Province of South Africa. *Bothalia* 15: 294-295.
- RUTHERFORD, M.C. & WESTFALL, R.H. 1986. Biomes of southern Africa - An objective categorization. *Memoirs of the Botanical Survey of South Africa* 54: 1-98.
- RYPKA, B.W., CLAPPER, W.E., BOWEN, I.G. & BABB, R. 1967. A model for the identification of bacteria. *Journal of General Microbiology* 46: 407-424.
- SCHEEPERS, J.C. 1975. *The plant ecology of the Kroonstad and Bethlehem areas of the Highveld Agricultural Region*. D.Sc. thesis, University of Pretoria, Pretoria.
- SCHEEPERS, J.C. (Unpubl.) *Geomorphology classes*. Grassland Research Centre, Pretoria.
- SCHEINER, S.M., & ISTOCK, C.A. 1987. Affinity analysis: methodologies and statistical inference. *Vegetatio* 72: 89-93.





- SCHÖNLAND, S. 1922. South African botanical literature. *Memoirs of the Botanical Survey of South Africa* 4: 69-85.
- SCHOEUW, J.F. 1823. *Grundzüge einer allgemeinen Pflanzen-geographie und pflanzengeographischer Atlas*. Reimer, Berlin.
- SCHULZE, B.R. 1947. The climates of South Africa according to the classifications of Köppen and Thornthwaite. *South African Geographic Journal* 29: 32-42.
- SCHULZE, B.R. 1965. *Climate of South Africa. Part 8. General Survey*. Government Printer and Weather Bureau, Pretoria.
- SNEATH, P.H.A. & SOKAL, R.R. 1973. *Numerical Taxonomy*. Freeman, San Francisco.
- TAINTON, N.M., BRANSBY, D.I. & BOOYSEN, P. de V. 1976. *Common Veld and Pasture Grasses of Natal*. Shuter & Shooter, Pietermaritzburg.
- TER BRAAK, C.J.F. 1986. Canonical correspondence analysis: A new Eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.
- TER BRAAK, C.J.F. 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio* 69: 69-77.
- TINLEY, K.L. 1977. *Framework of the Gorongosa ecosystem*. Unpublished D.Sc. thesis, University of Pretoria, Pretoria.
- TROLLOPE, W.S.W. & POTGIETER, A.L.F. 1983. *Characteristics of fire behaviour in the Kruger National Park*. Report. Department of Agronomy, University of Fort Hare, Alice.



- VAN DER MEULEN, F. 1979. *Plant sociology of the western Transvaal Bushveld, South Africa. A syntaxonomical and synecological study.* Cramer, Vaduz.
- VAN ROOYEN, N., THERON, G.K. & GROBBELAAR, N. 1981. A floristic description and structural analysis of the plant communities of the Punda-Milia-Pafuri-Wambiya area in the Kruger National Park, Republic of South Africa. I. The hygrophilous communities. *Journal of South African Botany* 47: 213-246.
- VAN STADEN, J.M. 1992. *Die fitososiologie van die Steenbokpan omgewing in die noordwes-Transvaal.* M. Sc. thesis, University of Pretoria, Pretoria.
- VOSS, L.G. 1952. The history of keys and phylogenetic trees in biology. *Journal of the Science Laboratory, Denison University* 43: 1-25.
- WEBER, W.A. & NELSON, P.P. 1972. *Random access key to the genera of Colorado mosses.* University of Colorado Museum, Boulder.
- WERGER, M.J.A. 1972. Species-area relationships and plot size: with some examples from South African vegetation. *Bothalia* 10: 583-594.
- WERGER, M.J.A. 1973. *Phytosociology of the Upper Orange River Valley, South Africa.* D.Sc. thesis, University of Nijmegen, Nijmegen.
- WERGER, M.J.A. 1974. On concepts and techniques applied in the Zürich-Montpellier method of vegetation survey. *Bothalia* 11: 309-323.



- WERGER, M.J.A. 1978. Biogeographical division of southern Africa. In Werger, M.J.A. (ed.). *Biogeography and ecology of southern Africa*. pp.145-170. Junk, The Hague.
- WESTFALL, R.H. 1981. *The plant ecology of the farm Groothoek, Thabazimbi District*. M.Sc. thesis, University of Pretoria, Pretoria.
- WESTFALL, R.H. (in prep.). SPECOM: A program for interrogating and comparing the PRECIS synonym list with species name data sets.
- WESTFALL, R.H., DEDNAM, G. VAN ROOYEN, N. & THERON, G.K. 1982. PHYTOTAB - A program package for Braun-Blanquet tables. *Vegetatio* 49: 35-37.
- WESTFALL, R.H. & DE WET, B.C. 1988. New programs for preliminary sequencing of relevés and species in phytosociological data sets. *Bothalia* 18: 122-123.
- WESTFALL, R.H. & MALAN, O.G. 1984. A comparison of vegetation units derived from vegetation-enhanced satellite imagery with the vegetation units derived from the floristic classification of the farm Groothoek, Thabazimbi District. Proceedings of the symposium on Pattern recognition in remote sensing and geophysics, Johannesburg, 2-3 July 1984.
- WESTFALL, R.H. & MALAN, O.G. 1986. A method for vegetation stratification using scale-related, vegetation-enhanced satellite imagery. *Bothalia* 16: 263-268.
- WESTFALL, R.H. & PANAGOS, M.D. 1984. A cover meter for canopy and basal cover estimations. *Bothalia* 15: 241-244.
- WESTFALL, R.H. & PANAGOS, M.D. 1988. The plant number scale - an improved method of cover estimation using variable-sized belt transects. *Bothalia* 18: 289-291.



- WESTFALL, R.H., VAN ROOYEN, N. & THERON, G.K. 1985. The plant ecology of the farm Groothoek, Thabazimbi District. II. Classification. *Bothalia* 15: 655-688.
- WESTFALL, R.H., VAN STADEN, J.M. & PANAGOS, M.D. 1987. Predictive species-area relations and determination of subsample size for vegetation sampling in the Transvaal Waterberg. *South African Journal of Botany* 53: 241-244.
- WHITE, F. 1981. *Vegetation map of Africa*. Scale 1:5 000 000. UNESCO, Paris.



APPENDIX I

*ILLUSTRATED MORPHOLOGY FOR PLANT IDENTIFICATION IN THE FIELD*

The character and character states used for field identification are sequenced in three main groups, namely, stemless or pseudo-stem, leafless and stems and leaves, from general to particular. This sequence is used for both encoding and identifying plants, as it was found that a set sequence facilitates encoding and identification. The program PHYTO 00 therefore, only sequences species to form dichotomies. Poaceae characters include those that are applicable to other plants as well as a group (numbers 187-224) which are exclusive to the Poaceae. Certain character states such as the presence of tendrils are redundant in the concept of climbing plants, in which they are implicitly included.

In field identification it was found that the efficiency of allocating a character state to a character of a plant is inversely proportional to the number of character states used. Fewer character states generally imply fewer problems with transitions between character states. The formula  $x$  gives the number of plants that can be differentiated where  $x = \text{number of character states per character}$  and  $y = \text{the number of characters}$ . For example, 4 character states per character for 8 characters provides 65 536 combinations. It is unlikely that so few characters and character states could be selected to differentiate



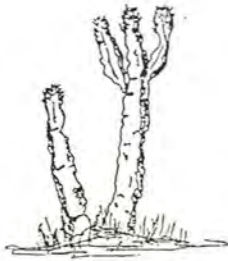
all the species in southern Africa. Nevertheless, this illustrates that judicious selection of characters and character states could be far more efficient, with further research, than those used at present.

Tables 4.4 to 4.9 show the results of the application of the character and character states, listed in this appendix, for field identification of plants.



\* NOT MUTUALLY EXCLUSIVE  
PHOTOSYNTHESIZING ORGANS PRESENT ( REDUNDANT )

MAIN AXIS SHAPE



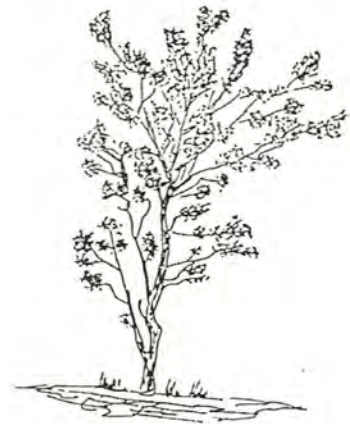
1. LEAFLESS ( main axis photosynthetic )



2. STEMLESS OR PSEUDOSTEM  
( main axis formed by separate  
photosynthetic organs )



3. STEMS AND LEAVES (photosynthetic organs separate from main axis )





MAIN AXIS ORIENTATION



4. ERECT



4. ERECT



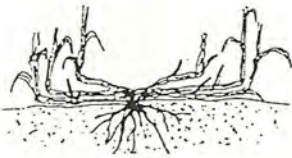
5. GENICULATE ( abruptly bent )



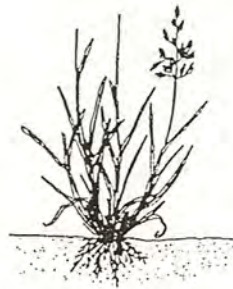
6. TWINING



7. CLIMBING ( tendrils/hooks redundant )



8. DECUMBENT

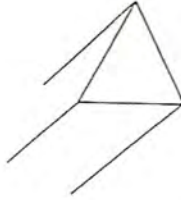


9. TUFTED/CLUMPED

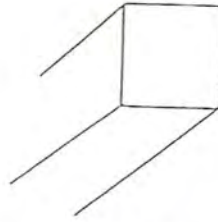




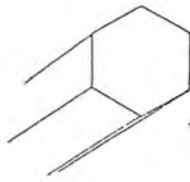
MAIN AXIS OUTLINE IN TRANSVERSE SECTION



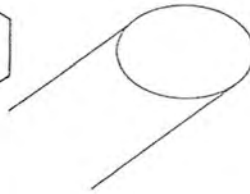
10. TRIANGULAR



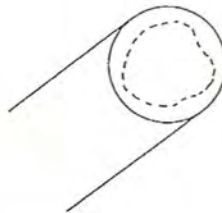
11. QUADRANGULAR



12. HEXAGONAL



13. ELLIPTIC/FLATTENED

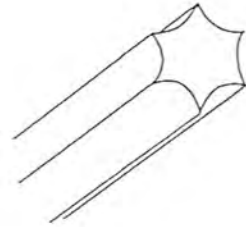


14. ROUND/IRREGULAR

MAIN AXIS SURFACE FEATURES \*



15. SMOOTH



16. RIDGED



17. STICKY



18. FISSURED ( longitudinal )



19. FISSURED ( rectangular )



20. STRIPPED/PEELED/FLAKY ( bark )

MAIN AXIS APPENDAGES ( trunk ) \*



21. KNOBS



22. THORNS SINGLE  
CURVED



23. THORNS SINGLE  
RETROSE



24. SPINES SINGLE  
STRAIGHT



25. THORNS AND SPINES PAIRED OR SINGLE, CURVED AND STRAIGHT



26. THORNS PAIRED  
CURVED



27. SPINES PAIRED  
STRAIGHT



28. THORNS IN THREES  
CURVED



29. PRICKLES ( short straight  
or curved non-woody,  
protuberances )



30. KNOTS  
( Poaceae )

THORN/KNOB ARRANGEMENT  
( trunk )



31. SCATTERED



32. IN ROWS

BRANCHING FROM MAIN AXIS



33. UNBRANCHED



34. OPPOSITE

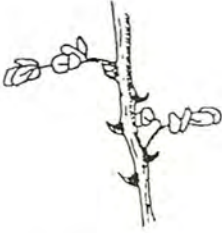


35. RECTANGULAR



36. SCATTERED

BRANCH/BRANCHLET APPENDAGES



37. THORNS SINGLE CURVED



38. THORNS SINGLE RETRORSE



39. SPINES SINGLE STRAIGHT



40. THORNS AND SPINES PAIRED OR SINGLE, CURVED AND STRAIGHT



41. THORNS PAIRED CURVED



42. SPINES PAIRED STRAIGHT



43. THORNS IN THREES CURVED



44. PRICKLES ( short straight or curved,  
non-woody protuberances )



45. SPINESCENT BRANCHLET

SAP ( at stem, petiole or if succulent, leaf )



46. CLEAR



47. YELLOW/  
RED



48. MILKY

LEAF ARRANGEMENT ( separate photosynthetic organs )



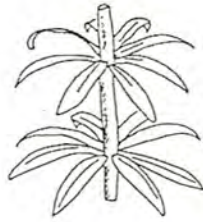
49. OPPOSITE



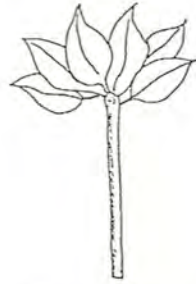
50. DECUSSATE



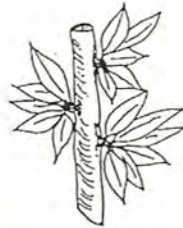
51. SPIRAL/ALTERNATE



52. WHORLED



53. CLUSTERED TERMINALLY ON  
BRANCHES



54. CLUSTERED ON ABBREVIATED  
BRANCHLETS/STEMS



55. DISTICHOUS ( two rows )

LEAF ATTACHMENT



56. SESSILE ( petiole less than 0,5 mm )



57. AMPLEXICAUL (enlarged base, petiole, stipule embracing stem)



58. PERFOLIATE (stem appears to pass through leaf base)



59. DECURRENT (leaf base runs down stem)



60. SHEATHING (lower part of leaf ± encircles stem)

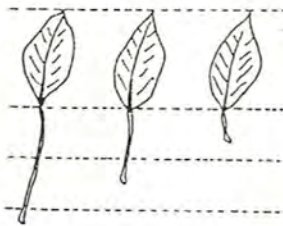


61. PELTATE (leaf attached by lower surface not margin)

62. PETIOLE GREATER THAN LAMINA LENGTH

63. PETIOLE HALF TO LAMINA LENGTH

64. PETIOLE LESS THAN HALF LAMINA LENGTH



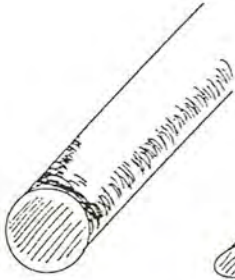
62.

63.

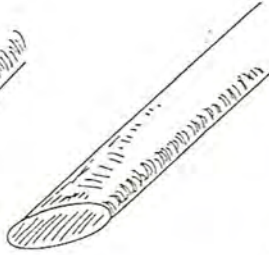
64.



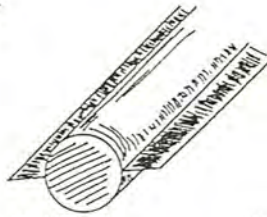
PETIOLE SHAPE



65. CYLINDRICAL



66. FLATTENED ( one or both sides )

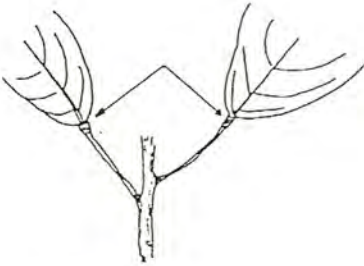


67. WINGED

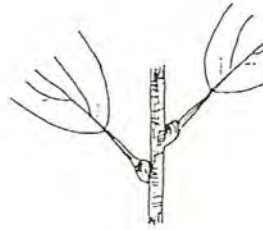


68. CHANNELED ABOVE ( grooved )

PETIOLE APPENDAGES \*



69. APEX SWOLLEN



70. BASE SWOLLEN



71. SINGLE BASAL GLAND



72. PAIRED BASAL GLANDS





LIGULES



73. ABSENT/INCONSPICUOUS



74. FRINGE OF HAIRS



75. MEMBRANE STRAIGHT



76. MEMBRANE ROUNDED



77. MEMBRANE POINTED



78. MEMBRANE NOTCHED



STIPULES ( unmodified - young growth or stipular scar )



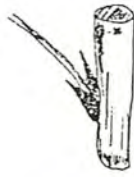
79. EXSTIPULATE



80. SINGLE



81. INTERPETIOLATE



82. PAIRED - FREE



83. PAIRED - JOINED



84. PAIRED - UNEQUAL



85. LOBED

LEAF FORM

petiolules greater than or equal to 0,5 mm  
for compound leaves



86. SIMPLE



87. TWO



88. THREE



89. ONCE DIGITATE/PALMATE



90. TWICE DIGITATE/PALMATE



91. ONCE PINNATE



92. TWICE PINNATE



93. PARIPINNATE



94. IMPARIPINNATE

LEAFLET ATTACHMENT\*



95. SESSILE



96. PETIOLATE



97. SESSILE/TERMINAL  
PETIOLATE

LATERAL LEAFLET ARRANGEMENT ( excluding terminal leaflet )



98. ALTERNATE LEAFLETS



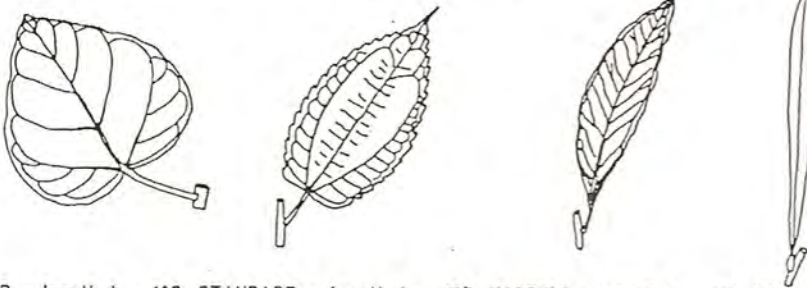
99. OPPOSITE LEAFLETS



100. ALL INSERTED AT ONE  
POINT



LEAF SHAPE RATIO ( compound leaves - entire lamina )

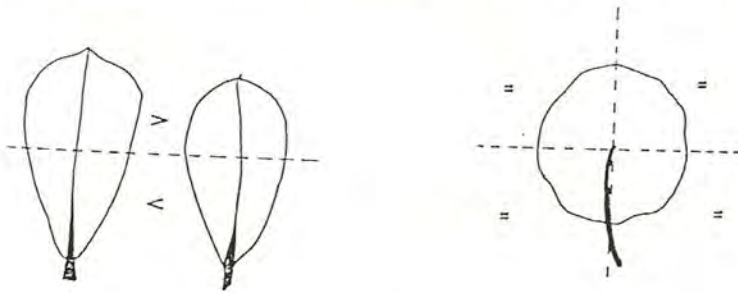


101. BROAD - length to width equal to or less than 1  
 102. STANDARD - length to width greater than 1 to 3  
 103. NARROW - length to width greater than 3 to 9  
 104. LINEAR - length to width greater than 9

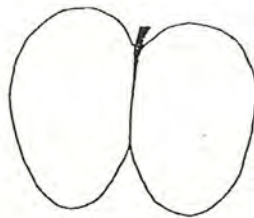
LEAF SHAPE ( folded transversely with apex on base - lamina )



105. ELLIPTIC  
 106. LINEAR ( parallel sided - grasslike )  
 107. OVATE ( broadly/narrowly )



108. OBOVATE ( broadly/narrowly )  
 109. ORBICULAR ( round )



110. BUTTERFLY



111. ASYMMETRIC

LEAF OR LEAFLET MARGINS \*



112. ENTIRE



113. ENTIRE/UNDULATE



114. CILIATE



115. CONVOLUTE ( rolled upon itself )



116. INVOLUTE ( rolled inwards )



117. REVOLUTE ( rolled down and back )



118. CRENATE ( scalloped /rounded teeth )



119. SERRATE ( sharp forward pointing teeth )



120. DENTATE ( sharp outward pointing teeth - middle of lamina )



121. TOOTHED ENTIRE MARGIN



122. TOOTHED APEX ONLY



123. TOOTHED BASE ONLY



124. LOBED, UP TO HALFWAY TO MIDRIB



125. CUT, GREATER THAN HALFWAY TO MIDRIB



NERVATION, LAMINA



126. OPAQUE ( nerves not visible )



127. PARALLEL



128. PINNINERVED ( feather-like - major side veins parallel towards margin )



129. PINNATE/NET ( side veins diminish to form net pattern )



130. LOOPED ( major side veins joined by loops near margin )

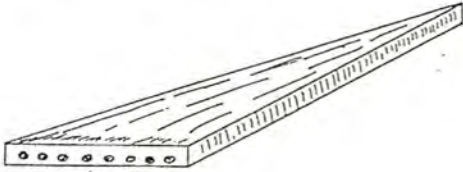


131. TRIPLINERVED ( midrib with two major side veins originating from base of blades )

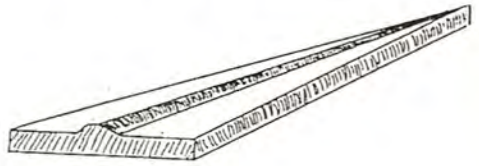


132. DIGITATELY NERVED ( midrib with more than two major side veins originating from base of blade )

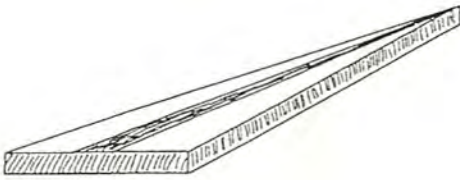
NERVATION RELIEF OF MIDRIB - PRINCIPAL RIB, DISTINCTLY THICKER THAN SIDE VEINS \*



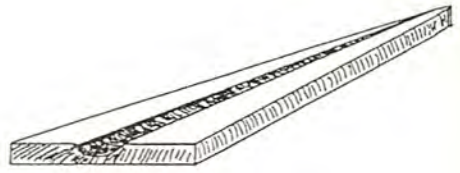
133. NO DISTINCT MIDRIB



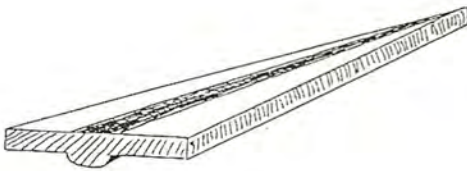
134. UPPER RAISED



135. UPPER LEVEL



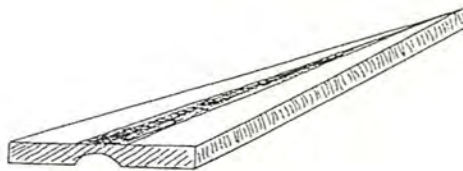
136. UPPER DEPRESSED



137. LOWER RAISED ( keeled )



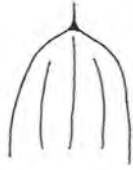
138. LOWER LEVEL



139. LOWER DEPRESSED



LEAF APEX ( ENTIRE LAMINA )



140. MUCRONATE



141. EMARGINATE



142. OBTUSE/ROUNDED



143. ACUTE/POINTED

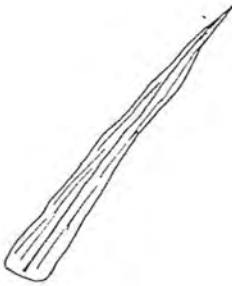


144. JUTTING/TWISTED



145. HOODED ( emarginate when flattened )

LEAF BASE



146. TRUNCATE



147. OBTUSE/ROUNDED



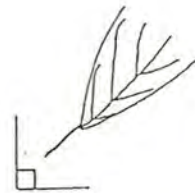
148. AURICULATE



149. SAGITTATE



150. CORDATE ( heartshaped )

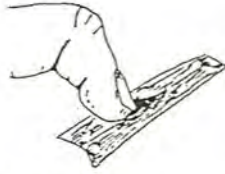


151. ACUTE/TAPERING

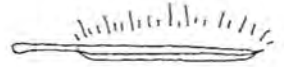
LEAF BLADE ( lamina ) COLOUR/TEXTURE \*



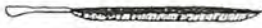
152. FLESHY



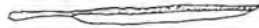
153. GLAUCOUS ( wax/powder bloom on bluish green )



154. GLOSSY



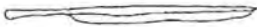
155. DISTINCTLY BICOLOROUS



156. WHITE SILVERY BELOW



157. RUSSET BROWN BELOW



158. GLABROUS ( both surfaces )



159. STELLATE/TUFTED HAIRS



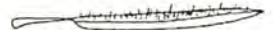
160. BRISTLY/SCABRID



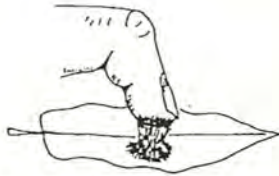
161. WOOLLY/FELTED



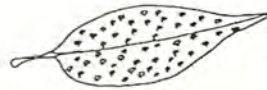
162. SPARSELY HAIRY



163. HAIRS ON SINGLE SURFACE



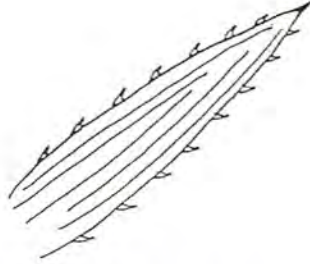
164. STICKY ( viscid )



165. SCALY



LEAF APPENDAGES \*



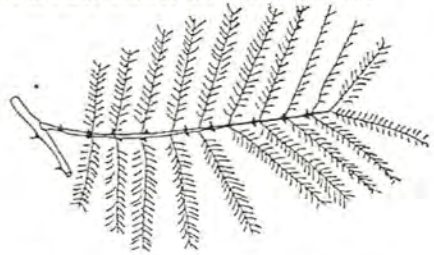
166. SPINES/THORNS ON MARGINS



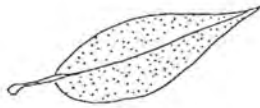
167. SPINES/THORNS ON LOWER SURFACE



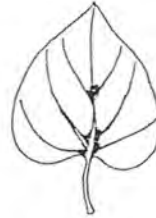
168. SPINES/THORNS ON UPPER SURFACE



169. SPINES/THORNS ON RACHIS



170. SPOTS/GLANDS



171. DOMATIA

LEAF SMELL



172. NEUTRAL



173. AROMATIC



174. FOETID



SHOOT APEX



175. SMALL INCONSPICUOUS



176. COVERED BY SCALES

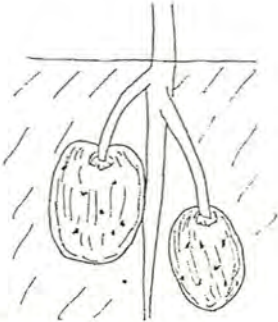


177. COVERED BY VELVETY/RUSSET  
BROWN HAIRS



178. COVERED BY LEAF PRIMORDIA

ROOTS/UNDERGROUND STRUCTURES



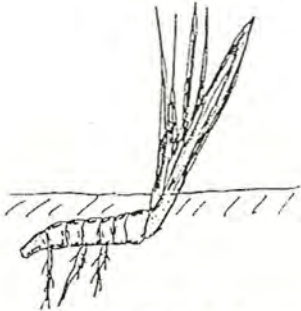
179. TUBER ( abrupt swelling  
- vertical, starch )



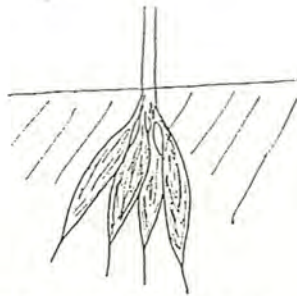
180. CORM ( fibrous leaf  
bases )



181. BULB ( fleshy leaf  
bases )



182. ROOTSTOCK ( horizontal )

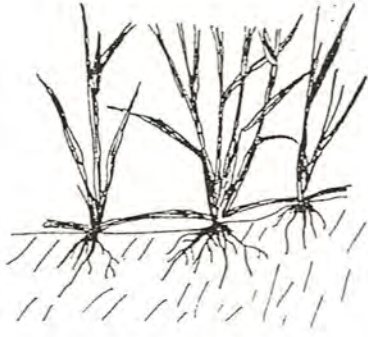


183. SWOLLEN ROOT ( gradual  
swelling )



184. RHIZOMES

ABOVEGROUND STRUCTURES



185. STOLONS



186. ROOTING FROM NODES/KNOTS

OTHER CHARACTERS FOR GRASS IDENTIFICATION  
CULM \*



187. CULMS WITH HAIRS  
ABOVE NODES



188. CULMS WITH HAIRS  
ON NODES



189. CULMS TUBEROUS AT  
BASE



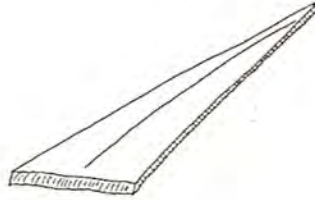
190. LEAVES MAINLY ON CULMS ( leafy culms )



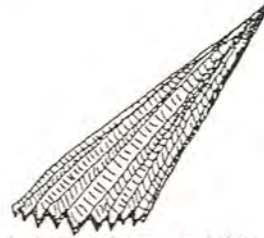
191. LEAVES MAINLY BASALLY AGGREGATED



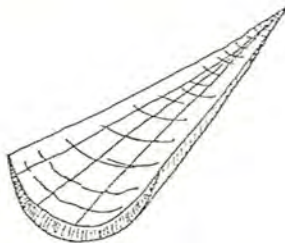
SHEATH/LAMINA \*



192. LAMINA FLAT



193. LAMINA FOLDED - Plicate



194. LAMINA CURVED



195. LAMINA ROLLED ± TUBULAR



196. LAMINA BASE BEARDED



197. LAMINA BASE WITH TWO TUFTS OF HAIRS



198. HAIRS BELOW LAMINA BASE



199. BASE GREATER THAN SHEATH



200. BASE EQUALS SHEATH



201. BASE LESS THAN SHEATH



202. SHEATH MARGINS JOINED (at least one quarter of length)



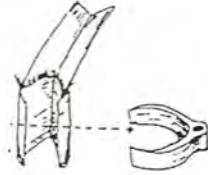
203. FREE SHEATH MARGINS



FRESH SHEATH \*



204. FLABELLATE AT BASE  
( fan shaped )



205. KEELED



206. GLABROUS



207. ENTIRE EXTERIOR  
HAIRY



208. BASE HAIRY



209. APEX HAIRY



210. ENTIRE EXTERIOR  
TINGED PURPLE/RED



211. BASE PURPLE/RED



212. APEX PURPLE/RED



OLD LEAF SHEATH



213. GLABROUS



214. ENTIRE EXTERIOR  
HAIRY



215. BASE HAIRY



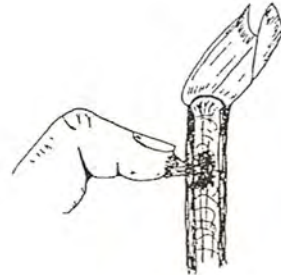
216. APEX HAIRY



217. INTERIOR TINGED  
SHINY ORANGE BROWN



218. INTERIOR TINGED  
PURPLE



219. INTERIOR VISCID ( sticky )





OLD DEAD LEAVES



220. ABSENT



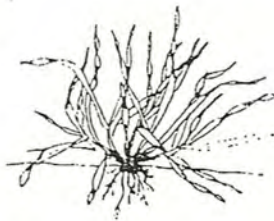
221. LOOSLEY CURLED WAVY



222. A FEW DISTINCT CURLS



223. TIGHTLY CURLED



224. OLD BLADES TWISTED  
INTO CORKSCREWS