The Determination of the Homogeneity of a *Vitis vinifera* L. cv. Cape Riesling Vineyard

A. C. DE LA HARPE[1] AND J. H. VISSER[2]

[1] Viticultural and Oenological Research Institute, Private Bag X5026, 7600 Stellenbosch, Republic of South Africa.

[2] Department of Botany, University of Stellenbosch, 7600 Stellenbosch, Republic of South Africa.

Submitted for publication: December 1982
Accepted for publication: August 1983

The value of Principal Component and Stepwise Discriminant analyses in selecting uniform plants for experimental purposes is discussed. Twenty seven variables were taken into account to establish the homogeniety (uniform plants) of 297 *Vitis vinifera* L. cv. Cape Riesling vines. A detailed study of the relationship and interrelationship of these variables resulted in 208 vines being selected as a uniform population. This selection provides the researcher with the possibility of using single vines as experimental units. However, it must be pointed out that Principal Component and Stepwise Discriminant analyses can only be used as an aid to normal statistical evaluation of experimental results and not as substitute for statistical experimental design.

The main statistical tools in compensating for variability are replication, randomization and blocking (Hammer, 1981). Replication normally involves multiple experimental units, and together with randomization it results in valid estimates of the experimental error (variance). Biological variation can be minimized by selecting more uniform plants at the pretreatment stage and then using replication and randomization for treatment applications (Hammer, 1981). According to Hammer (1981) this will allow the scientist to detect differences between treatments with fewer replications. The complexity of biological material, with intercorrelating variables, has as result that single variables cannot be treated as independant components of a factor (Broschat, 1979).

The problem of identification of uniform plants at the pretreatment stage could, therefore, be solved by measuring the appropriate variables and subsequently performing a Principal Component analysis (PCA) decreasing the dimensionality of the data.

PCA has been successfully used in psychology (Hotelling, 1936) and in the biological and horticultural sciences for a number of years (Orlocki, 1967; Sneath & Sokal, 1973; Gladon & Stadby, 1976; Oliver, Siddiqi & Goward, 1978; Leegwater & Leegwater, 1981 and Van Rooyen & Tromp, 1982).

The purpose of this study was to select relatively uniform vines in a *Vitis vinifera* L. cv. Cape Riesling vineyard by means of different growth and quality parameters with the aid of Stepwise Discriminant analysis (SDA) and PCA in order to decrease the large number of vines per treatment needed for physiological studies on this specific vineyard. The relatively small number of plants in the vineyard made the normal randomized block design, which necessitates a large number of experimental units per replication, impossible.

MATERIALS AND METHODS

Experimental vineyard: A 17 year old vineyard on the experimental farm Nietvoorbij, Stellenbosch, South Africa was used in this study. It consists of 297 vines of *V. vinifera* cv. Cape Riesling grafted onto 99 Richter, planted in a vineyard consisting of 4 soil types, namely a Southwold, Avalon, Glencoe and Kanonkop series (soil series as described by MacVicar, C. W. & Soil Survey Staff, 1977). The vines are trellised on a Perold system (Zeeman, 1981) and spur pruned to 16 buds. Kg^{-1}shoots. $Vine^{-1}$. Rainfall was supplemented by two 200mm irrigations by means of overhead sprinklers on 19/11/81 and 5/1/82.

Variables: The investigation was done in two phases. In phase I, the 22 growth variables depicted in Table 1 were measured on 2 shoots per cordon, and the respective mean values of these measurements were used as data points. The leaf area of a vine was determined by measuring the area of individual leaves with a model LT-3000 Li-Cor portable area meter and summated. Leaf dry mass was determined after drying to constant mass at 80°C. The vines were visually evaluated by 5 judges and grouped into 3 categories: sick and poorly developed vines taken as 100; normally developed vines as 500, and well developed vines as 900. All measurements were done at harvesting time.

In phase II, five quality variables were measured — total soluble must solids in °Balling; pH; total titratable must acidity ($g.1^{-1}$); the total number of bunches per vine, and yield per vine.

Data processing: The data were processed using a BMD-07M SDA programme (Health Sciences Computing Facility, UCLA) and a PCA programme forming part of the pattern recognition system "Arthur" (Harper, Duewer & Kowalski, 1977). The subroutines used in the

TABLE 1
Variables measured in a *Vitis vinifera* L. cv. Cape Riesling vineyard.

Variable Number	Variables	Unit
	Phase I	
1.	Shoot length cordon 1	cm shoot. cordon 1^{-1}
2.	Shoot length cordon 2	cm shoot. cordon 2^{-1}
3.	Total shoot length of both cordons	cm shoot. cordons^{-1}
4.	Spurs cordon 1	—
5.	Spurs cordon 2	—
6.	Spurs per vine	—
7.	Number of leaves per shoot of cordon 1	* —
8.	Number of leaves per shoot of cordon 2	** —
9.	Total number of leaves of the shoots of variables 7 and 8	—
10.	Total leaf area per shoot of cordon 1	*cm^2. shoot^{-1}
11.	Total leaf area per shoot of cordon 2	**cm^2. shoot^{-1}
12.	Total leaf area of both shoots	cm^2. shoot^{-1}
13.	Mean area per leaf of the shoots of cordon 1	*cm^2. leaf^{-1}
14.	Mean area per leaf of the shoots of cordon 2	**cm^2. leaf^{-1}
15.	Total mean area per leaf of both shoots	cm^2. leaf^{-1}
16.	Total dry leaf mass per shoot of cordon 1	*g. total leaf number^{-1}
17.	Total dry leaf mass per shoot of cordon 2	**g. total leaf number^{-1}
18.	Total dry leaf mass of both shoots	g. total leaf number^{-1}
19.	Mean dry mass per leaf of the shoots of cordon 1	*g. leaf^{-1}
20.	Mean dry mass per leaf of the shoots of cordon 2	g. leaf^{-1}
21.	Total mean dry mass per leaf of the shoots of both cordons	g. leaf^{-1}
22.	Evaluation of the vines	
	Phase II	
	Phase I plus the following 5 variables	
23.	Total soluble solids	°B
24.	Total titratable acids	g.l^{-1}
25.	pH	—
26.	Yield per vine	Kg
27.	Number of bunches per vine	—

* Mean of the two shoots of the 2nd spur on cordon 1
** Of the two shoots of the 2nd spur on cordon 2.

"Arthur" programme are listed in Table 2.

The BMD-07M programme was executed on a Burroughs 7800 computer of the Department Agriculture and the "Arthur" programme on a Univac 1100 computer of the University of Stellenbosch.

TABLE 2
Programmes of Arthur as performed on the data set.

Programme	Programme function
Phase I	
Input	Creates a data matrix as output to a binary file that is compatible with all other routines in Arthur.
Utilit	Provides a line printer listing of the data matrix and/or the distance matrix.
Scale	Scales the data to same proportions. The scaling factors are derived from the n data vectors of the training set and applied to all the data.
Correl	Calculates all feature — feature and feature — property covariances and correlations.
Kaprin	The extraction of the eigenvalues and eigenvectors of the data dispersion matrix as performed.
Katran	Creates a new data matrix from the first K factors of the data.
Varvar	Produces line printer plots of a data matrix.
Kaprin-Kavari-Katran-Varvar	Perform a principal component analysis plus rotation of eigenvalues with plotting.
Phase II	
Input	Same as phase I
Scale	Same as phase I
Kaprin	Same as phase I
Katran	Same as phase I
Varvar	Same as phase I
Kavari	Executes a Varimax rotation on the eigenvectors.
Katran	As phase I but with Kavari results.
Varvar	Same as phase I.

Prior to PCA all data were scaled to a standard deviation (S_i) of 1 and zero mean. The normalized standard deviation is defined as

$$S_i = \frac{\sigma_i}{x_i}$$

where σ = standard deviation
x_i = weighted mean

$$\text{and } \bar{x} = \frac{\sum_{j=1}^{n}\left[\frac{x_{i \cdot j}}{U_{i \cdot j}^2}\right]}{\sum_{j=1}^{n}\left[\frac{1}{U_{i \cdot j}^2}\right]}$$

where $U_{i \cdot j}$ is the uncertainty associated with the feature $X_{i \cdot j}$

$$x_{i \cdot j} = \frac{\left[x_{i \cdot j} - \bar{x}_{i \cdot j}\right]}{\left[\sum_{j=1}^{n}\left[x_{i \cdot j} - \bar{x}_{i \cdot j}\right]^2\right]^{\frac{1}{2}}}$$

and where n = total number of data vectors in the training data set, and x is the i th feature associated with the j th data vector.

RESULTS AND DISCUSSION

Phase I: Table 3 represents the scaled data with the mean, standard deviation, normalized standard deviation as previously defined as well as minimum and maximum values. Three of the PCA factors have eigenvalues (the sum of the variances) greater than 1 and are retained for discussion (Table 4). They account for 65% of the variance in the original variables with the remaining 35% caused by random variation.

The first PCA factor with an eigenvalue of 8,7 accounts for 39,6% of the variance of the original variables. This factor has relatively high factor loadings on the total number of leaves, leaf area and leaf dry mass of all measured shoots, indicating that leaf canopy variables dominate this factor. Factor 2 has an eigenvalue of 3,5 explaining 15,8% of the total variance. The variables with the highest factor loadings are the number of leaves per shoot of cordon 1, the total leaf area and the total leaf dry mass of cordon 1. Figure 1 represents a plot of factor 1 (representing mainly total leaf canopy) against factor 2 (representing mainly total leaf area). From this plot it can be deduced that the vineyard consists of two groups of vines, separated mainly by factor 2.

Factor 3 has an eigenvalue of 2,1 representing 9,5% of the variance in the original variables. The highest loadings in this factor are the spur variables (Table 4). This may be interpreted that factor 3 is a general growth factor or component.

In Fig. 2 factor 1 (X-axis) and factor 3 (Y-axis) are plotted. It is evident that the leaf canopy factor (factor 1) correlates with the growth factor (factor 3) and that the grouping of the vines is well defined. In Fig. 3 the leaf canopy of cordon 1 (factor 2, X-axis) is plotted against the growth factor (factor 3, Y-axis). Once again the vines seem to be well grouped into clusters indicating uniform vines as far as the leaf covering and other growth parameters are concerned. A further indication of the grouping is given in the totals on the Y-axis showing the total of plotted vines on the 2-dimensional plane.

TABLE 3
The scaled data of phase I with the mean, standard deviation, normalized standard deviation and minimum and maximum values.

Variable number	Variables	Mean	Standard deviation	Normalized std. deviation	Minimum	Maximum
	Phase I					
1.	Shoot length cordon 1	*140,80	96,31	0,68	0,00	514,00
2.	Shoot length cordon 2	*128,50	92,07	0,72	0,00	608,00
3.	Total shoot length of both cordons	*266,40	146,30	0,55	0,00	768,00
4.	Spurs per cordon 1	2,70	1,07	0,39	0,00	6,00
5.	Spurs per cordon 2	2,73	1,03	0,39	1,00	5,00
6.	Spurs per vine	5,43	1,59	0,29	0,00	10,00
7.	Number of leaves per shoot of cordon 1	*24,76	15,16	0,61	0,00	64,00
8.	Number of leaves per shoot of cordon 2	*19,51	14,51	0,74	0,00	59,00
9.	Total number of leaves of the shoots of variables 7 and 8	*44,42	21,54	0,48	0,00	102,00
10.	Total leaf area per shoot of cordon 1	*1392,00	952,60	0,68	0,00	6136,00
11.	Total leaf area per shoot of cordon 2	*1134,00	867,00	0,76	0,00	3933,00
12.	Total leaf area of both shoots	*2508,00	1309,00	0,52	0,00	7966,00
13.	Mean area per leaf of the shoots of cordon 1	*48,99	23,95	0,49	0,00	204,50
14.	Mean area per leaf of the shoots of cordon 2	*46,13	27,35	0,59	0,00	171,50
15.	Total mean area per leaf of both shoots	*53,79	18,16	0,34	0,00	146,80
16.	Total leaf dry mass per shoot of cordon 1	*7,05	4,89	0,69	0,00	32,41
17.	Total leaf dry mass per shoot of cordon 2	*5,82	4,52	0,78	0,00	20,67
18.	Total leaf dry mass of both shoots	*12,85	6,88	0,53	0,00	37,81
19.	Mean dry mass per leaf of the shoots of cordon 1	*0,38	0,50	0,22	0,00	2000,00
20.	Mean dry mass per leaf of the shoots of cordon 2	*0,24	0,13	0,57	0,00	0,49
21.	Total mean dry mass per leaf of the shoots of both cordons	*0,27	0,09	0,32	0,00	0,53
22.	Evaluation of the vines	2537,00	655,20	0,26	1000,00	3000,00

* Discrepancies in the data set are attributable to computer rounding off.

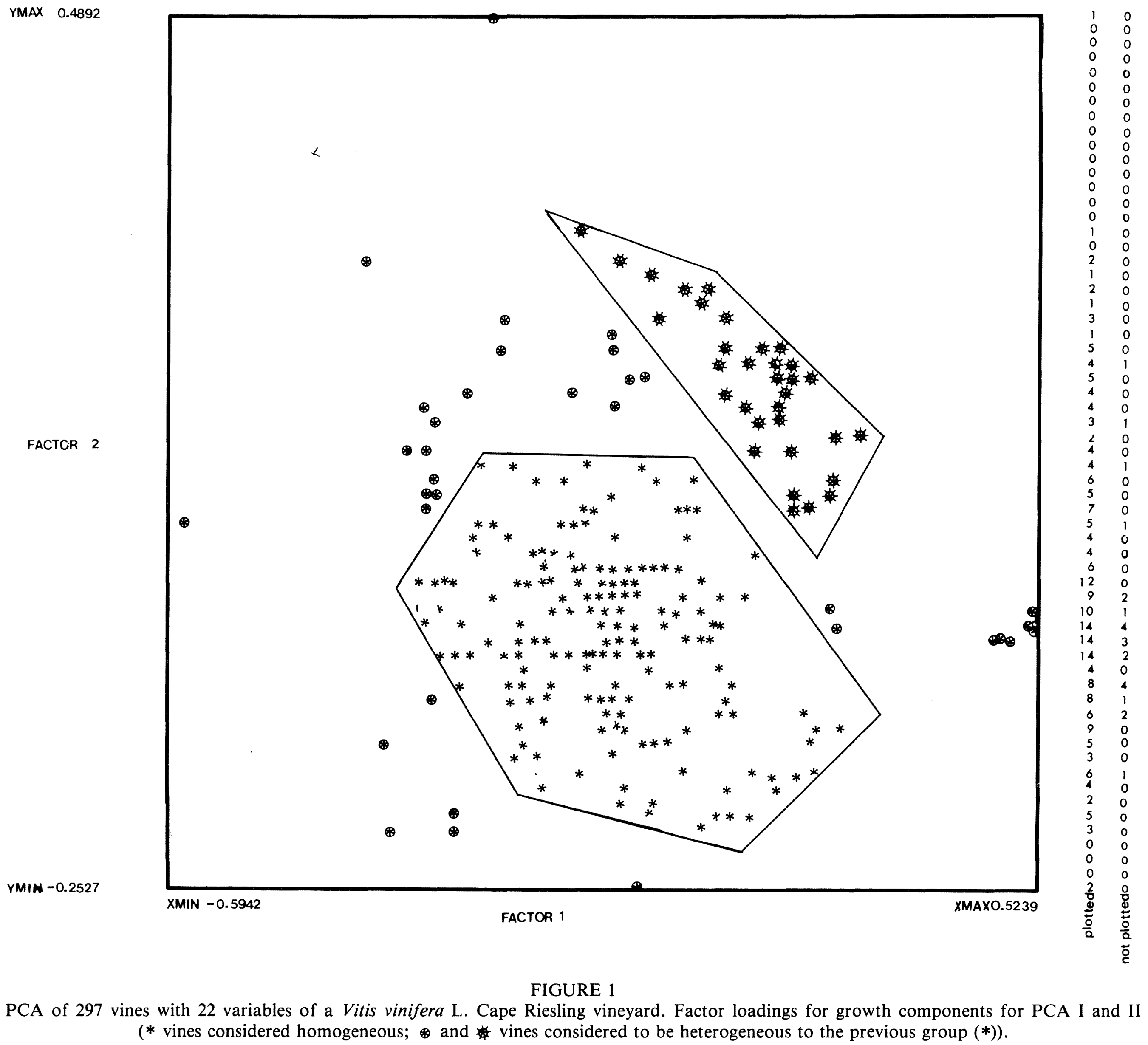

FIGURE 1
PCA of 297 vines with 22 variables of a *Vitis vinifera* L. Cape Riesling vineyard. Factor loadings for growth components for PCA I and II (* vines considered homogeneous; ⊛ and ✵ vines considered to be heterogeneous to the previous group (*)).

TABLE 4
Factor loadings for the first 3 Eigenvalues for 22 variables (Programmes used: Input, Utilit, Scale, Correl, Kaprin, Katran, Varvar)

Variable number	Variables	Factor 1	Factor 2	Factor 3
	Phase I			
1.	Shoot length cordon 1	−0,1819	−0,0187	−0,0198
2.	Shoot length cordon 2	−0,1658	−0,0462	−0,0915
3.	Total shoot length of both cordons	−0,2229	−0,0428	−0,0689
4.	Spurs per cordon 1	−0,1167	+0,0043	−0,3900
5.	Spurs per cordon 2	−0,1364	−0,0837	−0,3501
6.	Spurs per vine	−0,1672	−0,0516	−0,4901
7.	Number of leaves per shoot of cordon 1	−0,1972	+0,3836	+0,1089
8.	Number of leaves per shoot of cordon 2	−0,2367	−0,2877	+0,2254
9.	Total number of leaves of the shoots of variables 7 and 8	−0,2976	+0,0677	+0,2146
10.	Total leaf area per shoot of cordon 1	−0,2061	+0,4049	+0,0395
11.	Total leaf area per shoot of cordon 2	−0,2527	+0,2991	+0,2031
12.	Total leaf area of both shoots	−0,3138	+0,0975	+0,1621
13.	Mean area per leaf of the shoots of cordon 1	−0,2010	+0,2236	−0,1987
14.	Mean area per leaf of the shoots of cordon 2	−0,2151	−0,2868	−0,0802
15.	Total mean area per leaf of both shoots	−0,2190	+0,0714	−0,2584
16.	Total leaf dry mass per shoot of cordon 1	−0,2050	+0,3928	+0,1050
17.	Total leaf dry mass per shoot of cordon 2	−0,2480	+0,3004	+0,2098
18.	Total leaf dry mass of both shoots	−0,3089	+0,0819	+0,2119
19.	Mean dry mass per leaf of the shoots of cordon 1	−0,0046	−0,0614	+0,1666
20.	Mean dry mass per leaf of the shoots of cordon 2	−0,2204	−0,3061	−0,0625
21.	Total mean dry mass per leaf of the shoots of both cordons	−0,2349	+0,0494	−0,1882
22.	Evaluation of the vines	−0,0647	−0,0104	−0,1187
	Eigenvalues	8,7	3,5	2,1
	Factor percentage responsible for variance	39,6	15,8	9,5
	Cumulative percentage of variance	39,6	55,4	64,9

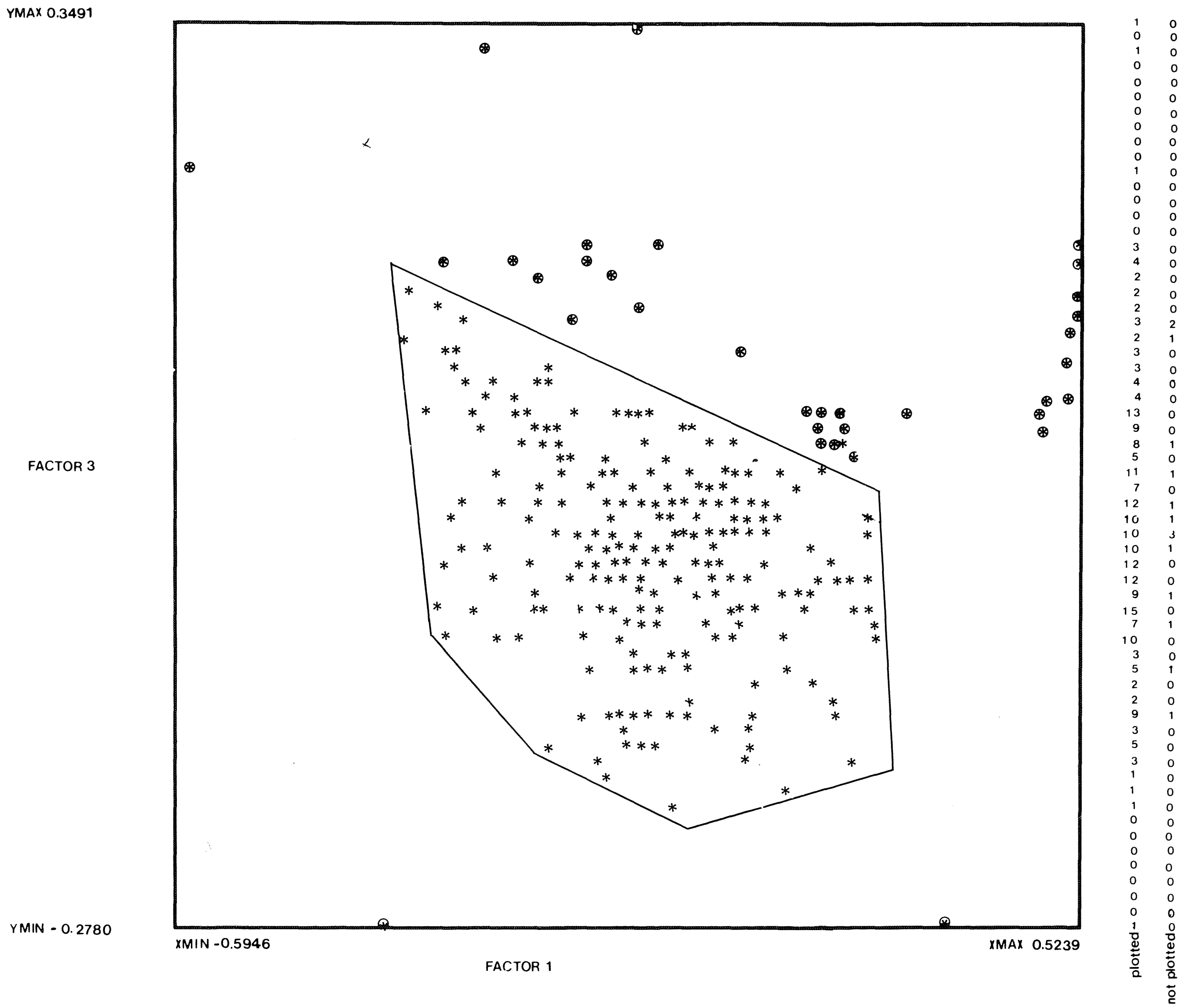

FIGURE 2
PCA of 297 vines with 22 variables of a *Vitis vinifera* L. cv. Cape Riesling vineyard. Factor loadings for growth and leaf canopy components for PCA I and III (* homogeneous and ⊛ heterogeneous.)

TABLE 5
The mean and standard deviation of the two categories as classified by stepwise discriminant analysis

Variable number	Variables	Means Category A	Means Category B	Grand Means over Categories	Standard Deviation A	Standard Deviation B
	Phase I					
1.	Shoot length cordon 1	*140,84	165,56	145,09	96,30	134,38
2.	Shoot length cordon 2	*128,50	169,07	135,47	92,06	142,28
3.	Total shoot length of both cordons	*266,41	338,56	278,80	146,31	219,99
4.	Spurs per cordon 1	2,70	2,47	2,66	1,06	1,27
5.	Spurs per cordon 2	2,73	2,49	2,69	1,03	1,36
6.	Spurs per vine	5,43	4,96	5,35	1,58	2,11
7.	Number of leaves per shoot of cordon 1	*24,75	28,45	25,39	15,16	23,79
8.	Number of leaves per shoot of cordon 2	*19,50	25,49	20,53	14,51	24,00
9.	Total number of leaves of the shoots of variables 7 and 8	*44,41	53,94	46,05	21,53	31,57
10.	Total leaf area per shoot of cordon 1	*1 392,20	1 851,02	1 470,99	952,56	1 711,11
11.	Total leaf area per shoot of cordon 2	*1 134,31	1 777,97	1 244,84	867,59	1 665,91
12.	Total leaf area of both shoots	*2 507,85	3 585,67	2 692,93	1 309,05	2 267,65
13.	Mean area per leaf of the shoots of cordon 1	*48,98	47,79	48,78	23,94	30,87
14.	Mean area per leaf of the shoots of cordon 2	*46,13	56,28	47,87	27,34	71,16
15.	Total mean area per leaf of both shoots	*53,78	62,11	55,21	18,15	30,77
16.	Total leaf dry mass per shoot of cordon 1	*7,04	9,30	7,43	4,88	8,25
17.	Total leaf dry mass per shoot of cordon 2	*5,82	8,61	6,30	4,51	8,67
18.	Total leaf dry mass of both shoots	*12,84	17,92	13,72	6,88	11,24
19.	Mean dry mass per leaf of the shoots of cordon 1	*8,37	0,24	6,98	127,52	0,14
20.	Mean dry mass per leaf of the shoots of cordon 2	*0,23	0,22	0,23	0,13	0,17
21.	Total mean dry mass per leaf of the shoots of both cordons	*0,27	0,30	0,27	0,08	0,10
22.	Evaluation of the vines	2 536,58	2 392,15	2 511,78	665,18	723,28

* Discrepancies in the data set are attributable to computer rounding off.

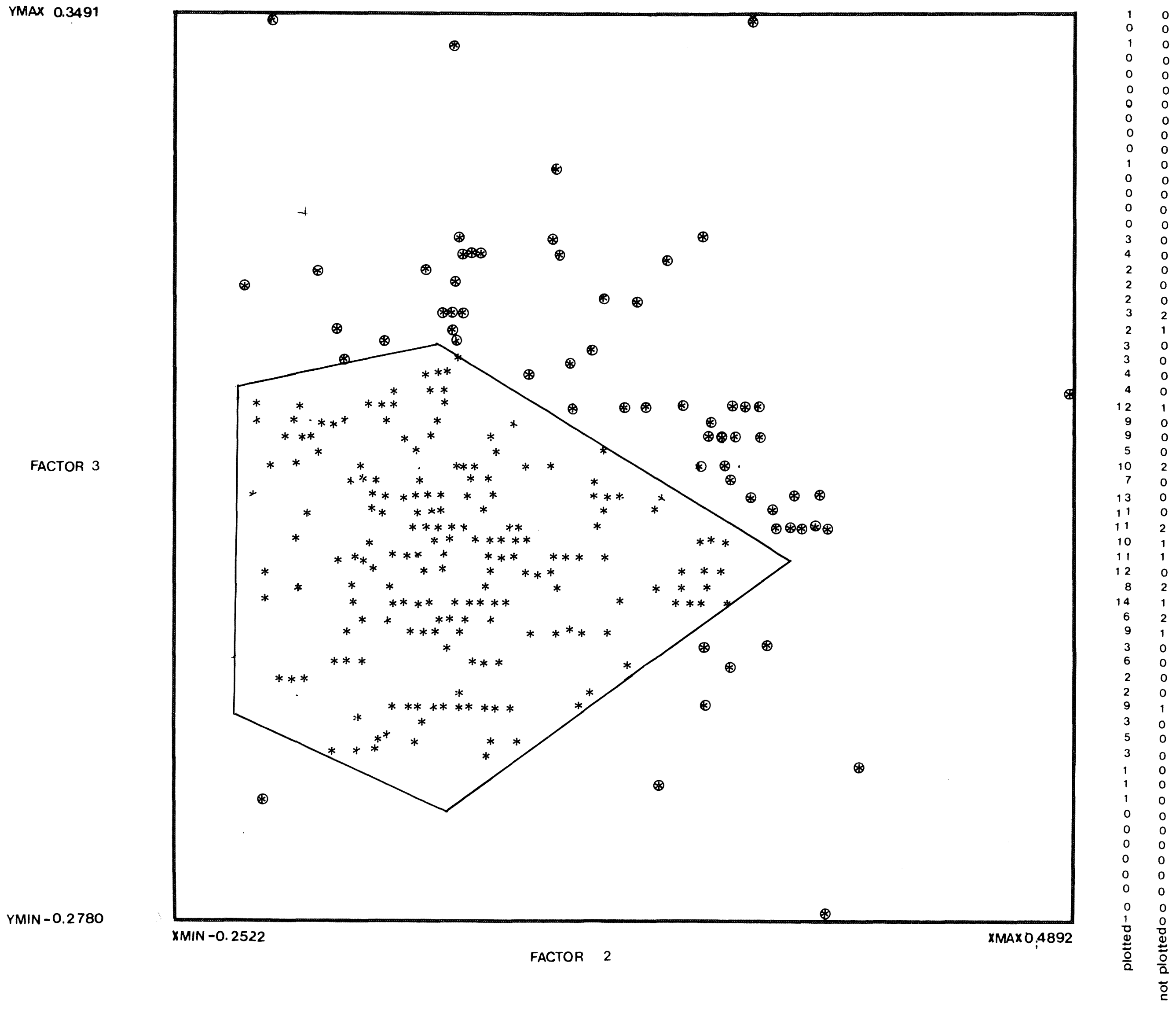

FIGURE 3
PCA of 297 vines with 22 variables of a *Vitis vinifera* L. cv. Cape Riesling vineyard. Factor loadings for leaf canopy and growth components for PCA II and III (* homogeneous and ⊛ heterogeneous).

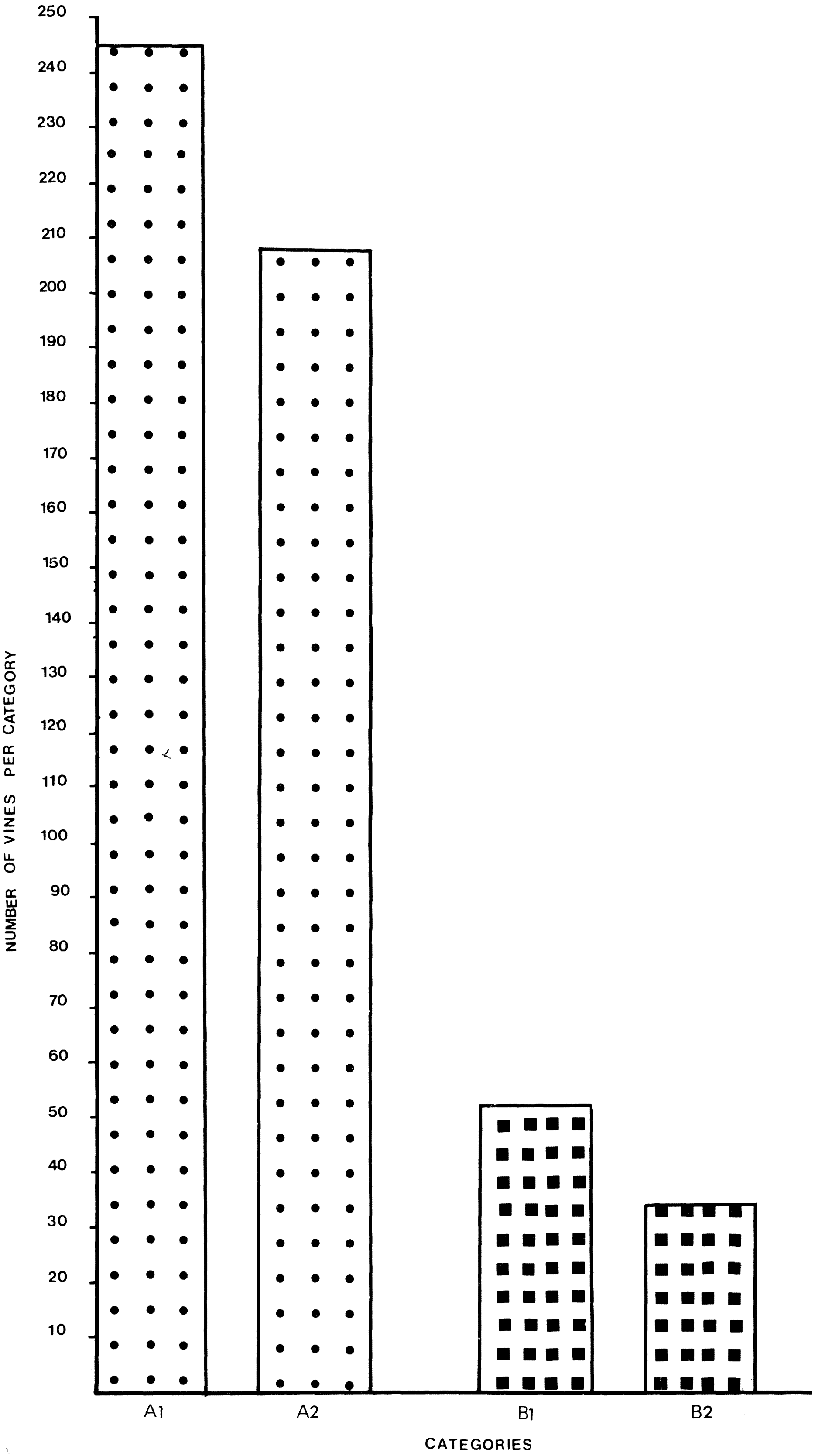

FIGURE 4
Vines grouped into homogeneous and heterogeneous categories by PCA and SDA techniques. (A = vines selected by PCA as homogeneous; A1 = vines selected by SDA as homogeneous; B = heterogeneous vines; B1 = heterogeneous vines grouped by SDA).

TABLE 6

The means and standard deviations of the five categories of vines as defined by the four soil types and heterogeneous group as pointed out by PCA in which they are growing as classified by stepwise discriminant analysis.

Variable number	Variables	Means Category A	Means Category B	Means Category C	Means Category D	Means Category E	Grand Means of Categories	Standard Deviation A	Standard Deviation B	Standard Deviation C	Standard Deviation D	Standard Deviation E
	Phase I											
1.	Shoot length cordon 1	*211,04	133,80	137,29	87,18	165,56	145,09	107,27	89,68	75,44	76,78	134,38
2.	Shoot length cordon 2	*185,86	116,48	140,89	77,92	169,07	135,47	129,31	61,34	64,63	87,81	142,28
3.	Total shoot length of both cordons	*391,30	250,20	279,94	154,38	338,56	278,80	167,57	107,72	101,91	129,87	219,99
4.	Spurs per cordon 1	*2,98	2,75	2,91	2,10	2,47	2,66	1,16	0,97	0,94	1,95	1,27
5.	Spurs per cordon 2	*2,62	3,32	2,63	1,90	2,49	2,69	0,92	0,88	0,76	1,03	1,36
6.	Spurs per vine	*5,60	6,07	5,55	4,00	4,96	5,35	1,48	1,38	1,33	1,42	2,11
7.	Number of leaves per shoot of cordon 1	*25,08	26,59	21,55	28,46	24,92	25,39	15,71	13,35	12,38	19,73	23,79
8.	Number of leaves per shoot of cordon 2	*20,88	20,45	26,87	7,92	25,49	20,53	14,26	12,01	14,28	12,28	24,00
9.	Total number of leaves of the shoots of variables 7 and 8	*46,92	47,04	48,43	32,64	53,94	46,05	21,47	19,37	19,20	24,15	31,57
10.	Total leaf area per shoot of cordon 1	*1513,41	1406,67	1172,51	1500,37	1851,02	1470,99	1151,98	754,25	715,96	232,49	1711,11
11.	Total leaf area per shoot of cordon 2	*1263,37	1130,38	1625,14	422,80	1777,97	1244,84	875,72	728,73	829,73	680,35	1665,91
12.	Total leaf area of both shoots	*2776,78	2484,87	2797,66	1943,17	3585,67	2692,93	1429,42	169,88	993,07	485,08	2267,65
13.	Mean area per leaf of the shoots of cordon 1	*49,98	51,11	51,49	41,33	47,79	48,78	33,99	17,46	15,73	28,39	30,87
14.	Mean area per leaf of the shoots of cordon 2	*51,32	51,10	56,27	20,42	56,28	47,87	25,63	22,77	20,88	28,02	71,16
15.	Total mean area per leaf of both shoots	*57,99	53,59	56,99	46,20	62,11	55,21	20,77	12,66	12,16	25,91	30,77
16.	Total leaf dry mass per shoot of cordon 1	*6,96	7,11	6,29	7,87	9,30	7,43	4,89	4,01	3,94	6,87	8,25
17.	Total leaf dry mass per shoot of cordon 2	*6,16	5,72	8,71	2,28	8,61	6,30	4,25	3,65	4,61	3,55	8,67
18.	Total leaf dry mass of both shoots	*13,04	12,84	15,00	10,15	17,92	13,71	6,72	6,17	6,17	8,14	11,24
19.	Mean dry mass per leaf of the shoots of cordon 1	*0,23	0,25	34,76	0,22	0,24	6,98	0,12	0,17	262,61	0,15	0,14
20.	Mean dry mass per leaf of the shoots of cordon 2	*0,25	0,25	0,30	0,10	0,22	0,23	0,12	0,19	0,10	0,14	0,17
21.	Total mean dry mass per leaf of the shoots of both cordons	*0,27	0,26	0,30	0,23	0,30	0,27	0,18	0,14	0,06	0,13	0,10
22.	Evaluation of the vines	2460,00	2454,54	2689,65	2580,00	2392,15	2511,78	761,57	641,64	568,37	641,74	723,28

Category A: Southwold series
Category B: Avalon series
Category C: Glencoe series
Category D: Kanonkop series.

* Discrepancies in the data set are attributable to computer rounding off.

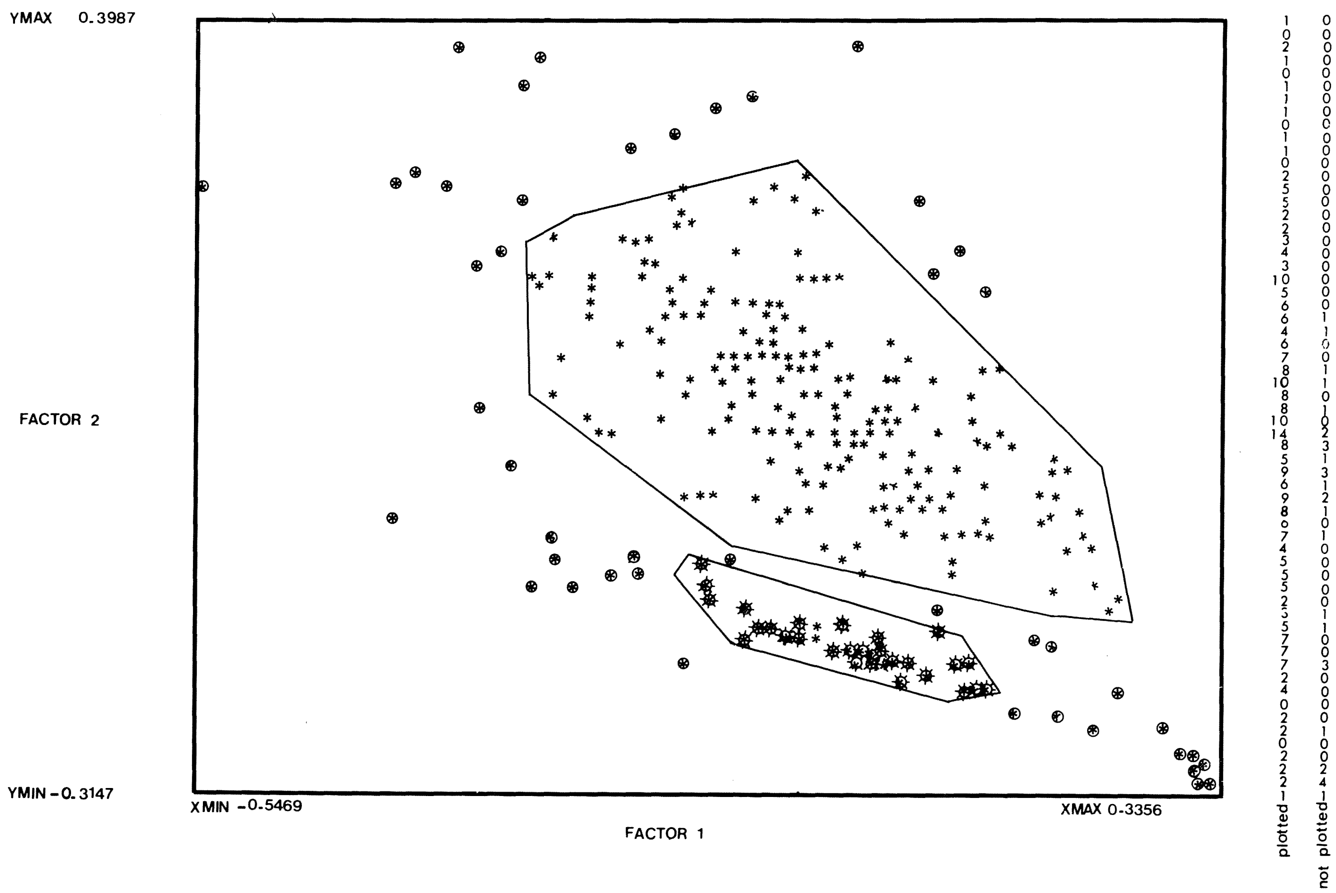

FIGURE 5

PCA of 245 vines (homogeneous group A) with 27 variables of a *Vitis vinifera* L. cv. Cape Riesling vineyard. Factor loadings for total leaf cover and total leaf cover of cordon 2 for PCA I and II (* vines considered homogeneous and ⊛ and ✵ heterogeneous).

TABLE 7
The scale data of phase I & II with the mean, standard deviation, normalized standard deviation, and minimum and maximum values.

Variable number	Variables	Mean	Standard deviation	Normalized std. deviation	Minimum	Maximum
	Phase I					
1.	Shoot length cordon 1	*146,00	104,10	0,71	0,00	514,00
2.	Shoot length cordon 2	*134,70	101,20	0,75	0,00	664,00
3.	Total shoot length of both cordons	*278,90	162,10	0,58	0,00	769,00
4.	Spurs per cordon 1	2,66	1,11	0,42	0,00	6,00
5.	Spurs per cordon 2	2,70	1,09	0,40	0,00	5,00
6.	Spurs per vine	5,36	1,70	0,32	0,00	10,00
7.	Number of leaves per shoot of cordon 1	*25,51	16,99	0,67	0,00	93,00
8.	Number of leaves per shoot of cordon 2	*20,39	16,22	0,79	0,00	72,00
9.	Total number of leaves of the shoots of variables 7 and 8	*46,03	23,43	0,51	0,00	117,00
10.	Total leaf area per shoot of cordon 1	*1480,00	1130,00	0,76	0,00	6340,00
11.	Total leaf area per shoot of cordon 2	*1237,00	1040,00	0,84	0,00	5956,00
12.	Total leaf area of both shoots	*2693,00	1539,00	0,57	0,00	8419,00
13.	Mean area per leaf of the shoots of cordon 1	*48,94	25,25	0,52	0,00	204,50
14.	Mean area per leaf of the shoots of cordon 2	*48,00	38,70	0,81	0,00	458,10
15.	Total mean area per leaf of both shoots	*55,51	20,98	0,38	0,00	200,60
16.	Total leaf dry mass per shoot of cordon 1	*7,48	5,66	0,76	0,00	32,41
17.	Total leaf dry mass per shoot of cordon 2	*6,25	5,33	0,85	0,00	26,70
18.	Total leaf dry mass of both shoots	*13,71	7,84	0,57	0,00	38,75
19.	Mean dry mass per leaf of the shoots of cordon 1 .	7,05	116,60	16,55	0,00	2000,00
20.	Mean dry mass per leaf of the shoots of cordon 2	0,23	0,14	0,60	0,00	0,61
21.	Total mean dry mass per leaf of the shoots of both cordons	0,28	0,09	0,32	0,00	0,53
22.	Evaluation of the vines	2520,00	659,30	0,26	1000,00	3000,00
	Phase II					
	Phase I plus the following 5 variables					
23.	Total soluble solids	19,77	2,16	0,11	0,00	23,10
24.	Total titratable acids	7,90	1,16	0,15	0,00	11,80
25.	pH	3,45	0,30	0,09	0,00	3,69
26.	Yield per vine	4,99	2,25	0,45	0,40	11,10
27.	Number of bunches per vine	26,27	10,86	0,41	0,00	74,00

* Discrepancies in the data set are attributable to computer rounding off.

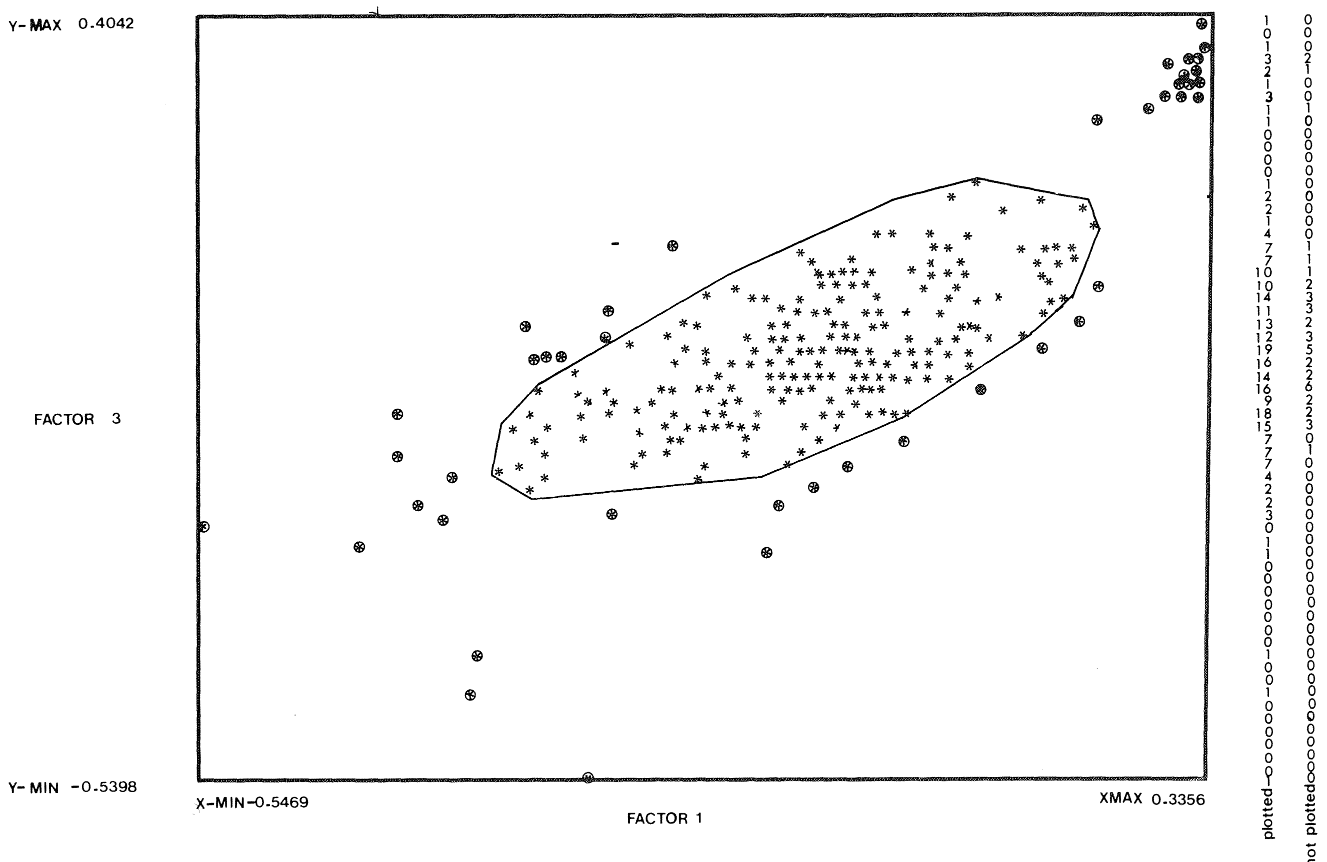

FIGURE 6
PCA of 245 vines (homogeneous group A) with 27 variables of a *Vitis vinifera* L. cv. Cape Riesling vineyard. Factor loadings for for total leaf cover and average leaf cover for PCA I and III (* vines considered homogeneous; and ⊛ heterogeneous).

TABLE 8

Factor loadings for the first 4 Eigenvalues after rotation for 27 variables. (Programmes used: Input, Utilit, Scale, Correl, Kaprin, Katran, Varvar)

Variable number	Variables	Factor 1	Factor 2	Factor 3	Factor 4
	Phase I				
1.	Shoot length cordon 1	−0,1205	+0,1501	−0,0400	−0,0644
2.	Shoot length cordon 2	−0,1172	+0,0469	−0,0699	−0,0348
3.	Total shoot length of both cordons	−0,1516	+0,1236	−0,0642	−0,0606
4.	Spurs per cordon 1	−0,0323	+0,0456	−0,1120	−0,5037
5.	Spurs per cordon 2	−0,0227	+0,0500	−0,0171	−0,4559
6.	Spurs per vine	−0,0356	+0,0618	−0,0839	−0,6206
7.	Number of leaves per shoot of cordon 1	−0,4146	−0,0085	+0,0058	−0.0337
8.	Number of leaves per shoot of cordon 2	−0,0176	−0,4315	+0,0098	−0,0596
9.	Total number of leaves of the shoots of variables 7 and 8	−0,3083	+0,2921	+0,0202	−0,0788
10.	Total leaf area per shoot of cordon 1	−0,4119	+0,0279	−0,1210	−0,0224
11.	Total leaf area per shoot of cordon 2	−0,0087	+0,4189	−0,1554	−0,0189
12.	Total leaf area of both shoots	−0,3048	+0,2552	−0,1976	−0,0293
13.	Mean area per leaf of the shoots of cordon 1	−0,2646	−0,0291	−0,3309	−0,1828
14.	Mean area per leaf of the shoots of cordon 2	+0,0395	+0,2208	−0,4160	−0,0828
15.	Total mean area per leaf of both shoots	−0,1313	+0,0713	−0,5655	−0,0962
16.	Total leaf dry mass per shoot of cordon 1	−0,4183	−0,6172	−0,0500	+0,0042
17.	Total leaf dry mass per shoot of cordon 2	−0,0213	+0,4314	−0,0398	−0,0409
18.	Total leaf dry mass of both shoots	−0,3164	+0,2810	−0,0643	−0,0240
19.	Mean dry mass per leaf of the shoots of cordon 1	−0,0748	+0,0272	+0,0215	−0,0649
20.	Mean dry mass per leaf of the shoots of cordon 2	+0,0273	+0,3121	−0,2346	−0,1963
21.	Total mean dry mass per leaf of the shoots of both cordons	−0,1809	+0,1130	−0,3729	−0,1336
22.	Evaluation of the vines	+0,0129	+0,0062	−0,2258	+0,0049
	Phase II				
	Phase I plus the following 5 variables				
23.	Total soluble solids	−0,0124	+0,0446	−0,1170	+0,0279
24.	Total titratable acids	+0,0718	−0,0303	+0,1380	−0,0811
25.	pH	+0,0165	+0,0408	−0,0427	+0,0061
26.	Yield per vine	−0,0655	+0,0416	+0,0217	−0,0744
27.	Number of bunches per vine	−0,0540	+0,0358	−0,0009	−0,0777
	Eigenvalues	5,4	4,9	2,4	2,3
	Factor percentage responsible for variance	24,8	23,6	11,2	11,1
	Cumulative percentage of variance	24,8	48,4	59,6	70,7

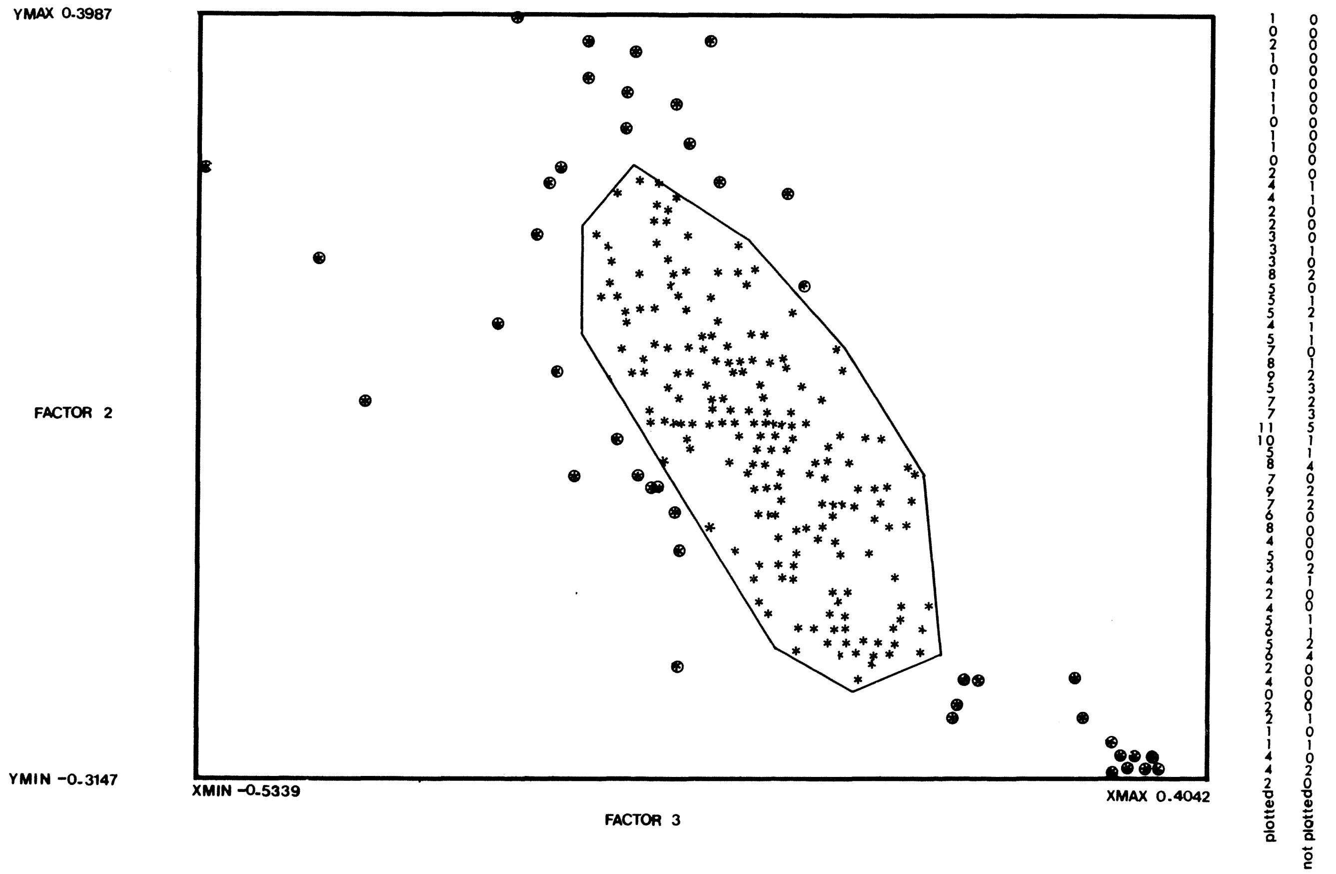

FIGURE 7

PCA of 245 vines (homogeneous group A) with 27 variables of a *Vitis vinifera* L. cv. Cape Riesling vineyard. Factor loadings for total leaf cover of cordon 2 and average leaf cover for PCA II and III (* vines considered homogeneous and ⊛ heterogeneous).

Outlying vines, not considered part of the clusters, were eliminated from further experimentation. These are the vines where the relative distance between any 2 vines is too large in relation to the average distance of the other vines to one another. The assessment of the distances is a subjective choice of the authors and may lead to criticism as far as objectivity is concerned. However, it must be kept in mind that the purpose of this grouping was to obtain an indication of the homogeneity of the data set and to provide the researcher with sufficient scope when selecting experimental units. The fact must be emphasized that this is not a statistical analysis for each variable alone but an analysis for the complete set of variables.

After the vines were classed into a homogeneous group A (the 245 vines considered in the cluster) and a heterogeneous group B (the 52 vines not considered part of the cluster), SDE was performed on the data set of groups A and B. The mean values, as well as the standard deviation of the variable for the 2 groups are given in Table 5.

The results of the SDA indicated that 208 of the original 245 vines considered to be homogeneous (85%) could be retained as category A vines, whereas 34 of the original 52 vines considered to be heterogeneous vines (65%) were retained in category B (Fig. 4). Although the percentage grouping for category B is low, the vines excluded from this group had not been taken into consideration for category A because of the relatively large distances between them and those of category A. This low percentage may be the result of some unexplained variance in the data set. After establishing the homogeneous group of vines (A), another SDA was performed on the data, this time classing the vines according to the 4 soil types. Table 6 gives the mean values and standard deviation of the 22 growth parameters.

From the Southwold series 58 vines (73%), the Avalon series 102 vines (79%), the Glencoe series 44 vines (80%) and the Kanonkop series 41 vines (75%) were selected to be part of the homogeneous group, indicating that in this specific vineyard the 4 soil types had little or no effect on the growth parameters of the vines growing on that particular soil type during this season.

Phase II: As a supplement to the existing data, 5 additional parameters, including some grape quality parameters, were determined. The Arthur programme was used on the data set including the 5 additional parameters, and the results are listed in Tables 7 and 8.

Seven of the PCA factors have eigenvalues greater than 1 and were retained in the analysis. They account for 100% of the variance in the original variables. After the data was rotated by the Varimax rotation algorithm KAVARI, the first PCA factor explains 24,8% of the variance of the original variables (Table 8). This factor has relatively high factor loadings on leaf canopy (surface) and growth variables such as total number of leaves and total leaf area per leaf of shoots on both cordons, which is similar to factor 1 in phase I where leaf cover and growth variables played an important role in the clustering of the vines. Factor 2 has an eigenvalue of 4,9 and explains 23,6% of the total variance. The variables with the highest factor loadings are total leaf area of the shoots on cordon 2, the total dry leaf mass of the shoots on cordon 2, and the average leaf mass per leaf of the shoots on cordon 2. This factor may, therefore, be interpreted to be relating to leaf cover in general and to growth parameters of the vines.

Fig. 5 represents the plot of the total leaf cover (factor 1, X-axis) to the total leaf cover of cordon 2 shoots (factor 2, Y-axis). In this plot the two groups of vines which were present in cluster 1 (Fig. 1) of phase I are still evident, although the cluster seems to be more compact with much smaller relative distances between groups (Fig. 5). This is because of the additional clustering effect of the extra parameters measured.

Factor 3 has an eigenvalue of 2,4 explaining 11,2% of the total variance in the original data set. The highest factor loadings in this factor are the area per leaf of the shoots on cordon 1, cordon 2, and both cordons as well as the average leaf dry mass of the vine. This may once again be interpreted as being a growth factor.

In Fig. 6 factor 1 was plotted (X-axis) against factor 3 (Y-axis). Compared to cluster 1 (Fig. 1) the additional clustering effect of the 5 extra parameters is evident. In Fig. 7 factor 2 (Y-axis) was plotted against factor 3 (X-axis).

Factors 4 and 7 represent 11,1% and 9,4% (not shown) respectively of the variance in the original variables and have eigenvalues of 2,3 and 1,9. The highest factor loadings are on the growth parameters namely spurs and shoot length, and may be interpreted as growth factors.

Factors 5 and 6 have eigenvalues of 2,2 and 2,0 (not shown) respectively, with relatively high loadings on the parameters, such as pH, yield per vine and number of bunches per vine.

CONCLUSION

In most PCA factors the leaf area was important in the clustering process, although a number of factors affect the final selection. The more uniform vines were those with approximately the same leaf surface and growth variables, whereas those rejected for experimental purposes deviated from the above. In the selection of homogeneous vines, it appears that instead of measuring 27 factors, one could concentrate on variables for determining leaf canopy.

When all measured variables were taken into account, it is evident that the 4 soil types had little or no effect on the homogeneity of the different vines in the vineyard during this growth season. Quality variables, such as °B, TTA and pH, resulted in better defined clusters and should, therefore, be used in future studies of this nature. The programmes used in this study are powerful and handy tools in the hands of the viticulturist, enabling him simultaneously to take into account groups of variables as well as their relationships with other groups. Combined with the normal statistical tools, such as randomization and replication, they may lead to a better understanding of the data.

LITERATURE CITED

BROSCHAT, T. K., 1979. Principal component analysis in horticultural research. *Hortscience* **14**(2), 114-117.

GLADON, R. J. & STABY, G. L., 1976. Opening of immature Chrysanthemums with sucrose and 8-hydroxyquinoline citrate. *Hortscience* **11**, 206-208.

HARPER, A. M., DUEWER, D. L. & KOWALSKI, B. R., 1977. Arthur and experimental data analysis: The heuristic use of a Polyalgorithm. Documentation for Arthur, Version 1-9-77. Siberatory for Chemometrics, Dept. of Chemistry, Univ. Georgia, Athens, Georgia.

HAMMER, P. A., 1981. Controlling variability. *Hortscience* **16**(5), 628-630.

HOTELLING, H., 1936. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417-441, 498-520.

LEEGWATER, D. C. & LEEGWATER, J. A., 1981. The use of a microcomputer in the classification of grape brandies by pattern recognition. *J. Sci. Food Agric.* **32**, 1115-1118.

MACVICAR, C. N. & SOIL SURVEY STAFF, 1977. Soil classification — A binomial system for South Africa. *Sci. Pamphlet* **390**, 152 pp, Government Printer, Pretoria.

OLIVER, J. E., SIDDIQI, A. H. & GOWARD, S. N., 1978. Spatial patterns of climate and irrigation in Pakistan. A multivirate statistical approach. *Arch. Met. Geoph. Biol. Ser. B.* **25**, 345-357.

ORLOCKI, L., 1967. An agglomerative method for classification of plant communities *J. Ecol.* **44**, 193-206.

SNEATH, P. H. A. & SOKAL, R. R., 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco.

VAN ROOYEN, P. C. & TROMP, A., 1982. Chenin blanc wine volatiles and the intensity of a guava-like flavour. *S. Afr. J. Enol. Vitic.* **3**(1): 1-7.

ZEEMAN, A. S., 1981. Oplei. *In:* Wingerdbou in Suid-Afrika 185-198. Ed. J. D. Burger & J. Deist., V.O.R.I.