COMPARISON OF PROVENANCES OF *GLIRICIDIA*

AFRC, Institute of Grassland and Environmental Research

Hurley, Maidenhead, Berkshire, SL6 5LR

EMC Contract Number X0169

Final Report

Covering period 1 October 1989 to 31 March 1992

CONTENTS

OBJECTIVES

The original contract of 6 months was followed by two 3 month contracts and one 6 month contract. The overall objectives were as follows:

1. To identify *Gliricidia* provenances with differences in chemical composition being independent of climate or location using samples taken from trees grown on one site.
2. To evaluate *in-vitro* gas production measurements conducted on the *Gliricidia* samples.
3. To develop Near Infra-red (NIR) reflectance calibrations to predict gas production parameters.
4. To obtain and interpret NIR spectra on samples of tree leaves from Nepal.
5. To develop NIR calibrations to predict NRI tannin analysis of the Nepal samples.

This report is divided into two sections, the first dealing with the *Gliricidia* samples and the second the Nepal samples.

1

CLASSIFICATION OF *GLIRICIDIA* PROVENANCES

1.1 WORK CARRIED OUT

Seventy-five samples of *Gliricidia* were received comprising bulk samples from 25 different provenances, and a further 10 samples from 5 of the provenances. Two samples arrived bearing the same sample numbers (viz: 25/84/4 and 25/84/7) and were therefore omitted from the sample sets analysed. The remaining 73 samples were scanned in duplicate.

Evaluation and classification of *Gliricidia* provenances took place using only Near Infra-red (NIR) reflectance spectral data and employing multivariate statistical techniques, including Biplot, Principal Component Analysis (PCA), Discriminant, Cluster and Canonical Variate Analysis (CVA) (1). For each sample 700 data points were obtained. However, this would have been an enormous amount of data to handle, and so for the majority of the analyses the data for each sample was condensed to 140 data points. This was achieved by averaging every set of 5 data points. Codes for the 25 different provenances are given in Table 1.1. On the basis of the chemical data from samples from the previous year, the Oxford Forestry Institute (OFI) selected five of these provenances as being the most different. The 5 provenances chosen were:

P2 Top overall in performance, productivity etc. but susceptible to very high aphid attack.

P5 Good performance with very low aphid attack.

P7 Lowest N content.

P10 Highest N content.

P17 Highest digestibility.

For these provenances a further 10 samples were supplied (taken from 10 different trees of the same provenance). For this report these samples were coded a, b, c, d, e, f, g, h, i and j.

Initially, the 10 samples from these provenances were treated as replicates. After discussions with the OFI, we were informed that they should in fact be treated as individual samples. Both types of statistical analyses are presented in this report, as they show some interesting features.

Gas production analysis (2) was used to rank samples using gas production as an indicator of *in-vitro* degradability. Initially the technique was performed using the 25 bulk samples. Subsequently, a second experiment using 5 of the bulk samples used in the initial run and a further 15 samples (ie. a, e and i samples) was undertaken.

Calibration equations were developed from the NIR spectral data for the prediction of gas

Table 1.1 Codes for the 25 *Gliricidia* provenances.

Code	Provenance
P1	13/84
P2	14/84
P3	15/84
P4	16/84
P5	17/84
P6	24/84
P7	25/84
P8	30/84
P9	31/84
P10	33/85
P11	34/85
P12	35/85
P13	38/85
P14	39/85
P15	40/85
P16	41/85
P17	1/86
P18	10/86
P19	11/86
P20	12/86
P21	13/86
P22	24/86
P23	43/87
P24	72/87
P25	75/87

production parameters and the other chemical parameters measured.

1.2 RESULTS

1.2.1 Classification of provenances

A. Principal component, biplot, canonical variate and discriminant analyses were used to classify the different *Gliricidia* provenances. Samples from the same provenances were treated as replicates during canonical variate and discriminant analyses.

Principal Component Analysis (PCA)

GENSTAT (3) was used to calculate the principal components using 140 data points for each sample. Principal components are linear combinations of the NIR data. The first stage of this analysis involved calculation of latent roots. It can be seen from Table 1.2.1, that virtually 100% of the variation was accounted for by the first 13 roots. The first root informs us how much variation can be accounted for by the first component, the second root, the amount of variation accounted for by the second root and so on for the other roots. The principal components are then calculated from the latent roots. Graphical presentation of the pair-wise components allows the natural grouping of the samples to be observed. It is known that Component 1 accounts for a large amount of physical effects eg. particle size and path length (4). Figure 1.2.1a shows Component 1 versus Component 2 for the 25 different provenances. This clearly separates out P23 on the Component 2 axis and P5 on the Component 1 axis. As Component 1 is dominated by physical effects it is useful to look at further components to determine whether the difference may be attributed to a physical difference. Figure 1.2.1b displays Component 2 versus Component 3. These components represent the chemical parameters. P23 was set apart from the other provenances on the Component 2 axis, but P5 was found to be in-line with the other samples. This suggested that physical differences may have accounted for P5 showing differences in Figure 1.2.1a but that the difference in P23 could not be accounted for by a physical difference, and warrants further investigation. The other provenances were located in similar positions to those in the first graph (Figure 1.2.1a).

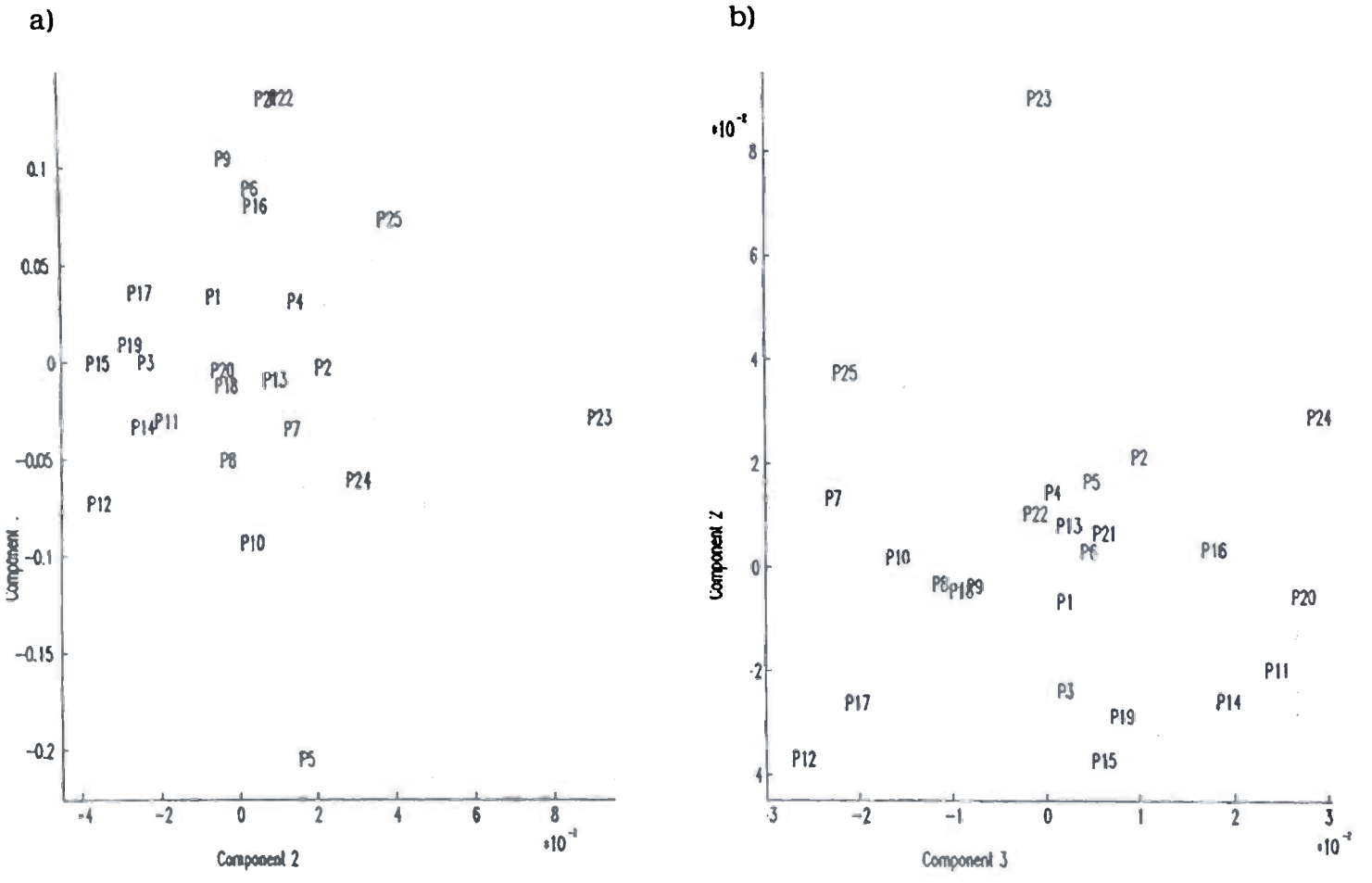
Biplot Analysis

Biplot analysis (5) produces a graphical representation of the relationships between provenances, giving an overall view of the data. Again, 140 data points were used in the calculations, to investigate the distribution of the 5 provenances (P2, P5, P7, P10 and P17)

Table 1.2.1 Percentage of variation accounted for each of the latent roots.

Latent Root	1	2	3	4	5	6	7
% Variation Accounted For	83.16	10.78	3.59	0.80	0.77	0.32	0.30
Latent Root	8	9	10	11	12	13	14
% Variation Accounted For	0.11	0.06	0.04	0.03	0.01	0.01	0.00
Latent Root	15	16	17	18	19	20	
% Variation Accounted For	0.00	0.00	0.00	0.00	0.00	0.00	

Figure 1.2.1 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the 25 different provenances.



and the 25 different provenances. Figure 1.2.2 showed the distribution of the 5 provenances, and indicated P5 to be the most different provenance. Figure 1.2.3 showed the distribution of the 25 different provenances and how the 20 other provenances were situated with respect to P2, P5, P7, P10 and P17.

Canonical Variate Analysis (CVA)

CVA calculates the inter-group distances which are Mahalanobis distances (6). CVA of the 5 provenances P2, P5, P7, P10 and P17 used the replicated samples for each provenance (n=10 with the exception of P7 when n=8). This analysis used 16 principal components in the calculations (16 principal components were chosen as they accounted for virtually 100% of the variation). Table 1.2.2 shows the inter-group distances between the five chosen provenances. The maximum distance was found between provenances P5 and P17. This value is taken to be D_{max}, and in this case was equal to 10.431. These values were also displayed as a similarity matrix, all values being scaled by D_{max}. The similarity index (SI) was calculated as follows:

$$SI = 1 - (D^2 / D^2_{max})$$

where D is the distance between the groups, or in this case provenances.

By definition, the provenances exhibiting the maximum inter-group distance had a similarity of 0.0. Table 1.2.3 shows the similarity values for the 5 provenances, with P5 and P17 having a similarity value of 0.0. Alternatively, the inter-provenance relationships can be represented in the form of a minimum spanning tree as shown in Figure 1.2.4. This is a single linkage cluster analysis and the similarities were calculated as follows:

$$(1 - D^2 / D^2_{max} \quad x\ 100$$

The minimum spanning tree joins the nearest neighbours ie. the most similar samples. The above data was also presented as hierarchical clusters ie. a dendrogram, as shown in Figure 1.2.5. The scale is from 0 to 100% similarity. This diagram showed that P5 and P10 had 80% similarity, P2 and P7 75% similarity and that P17 did not cluster until 40% similarity. This indicated that P17 was a very different provenance from the remaining 4. The minimum spanning tree and dendrogram use the same information but display it in a slightly different manner.

Figure 1.2.2 Distribution of provenaces P2, P5, P7, P10 and P17.

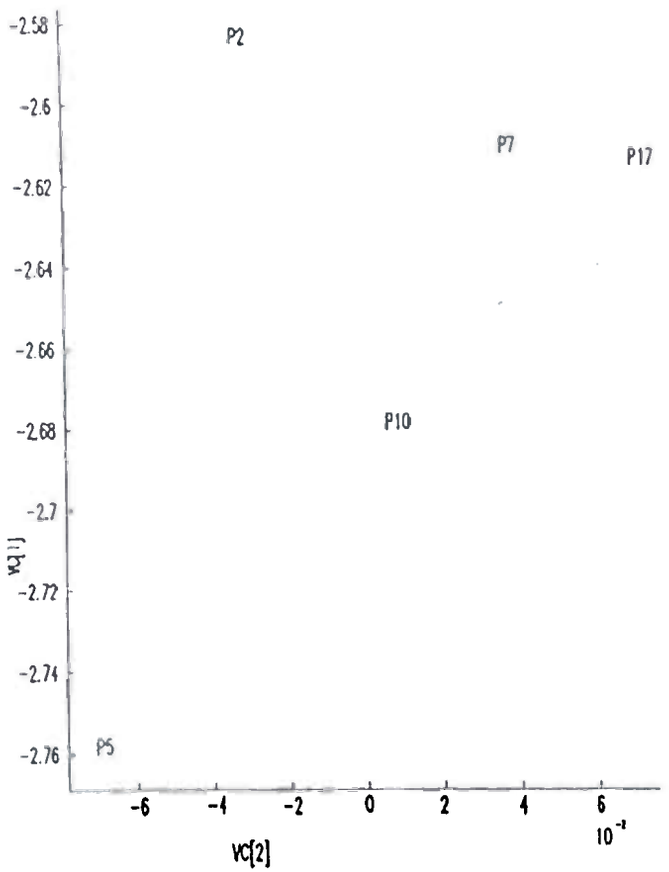


Figure 1.2.3 Distribution for the 25 different provenances.

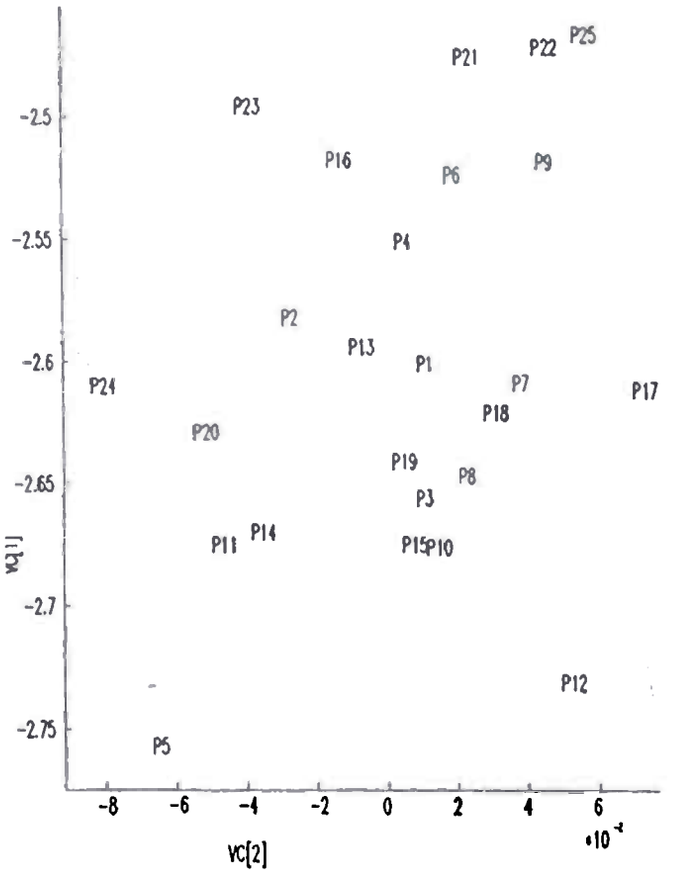


Table ter grou istances for rovenances of *G ricidia* (P2 P7 1(7).

DISTANCES

P2	000				
	85	000			
P7	933	.350	000		
	1.		.23	000	
.7	.712	431	608	7 844	0.000
	P2	P	P7	P:.0	P1'

Tab' 2. Si lari ies for provenan of *Gliricidia* (P2, P5 P7 P P17).

DISTANCES

P2	000				
P5	648	000			
P7	7'	97	000		
P1	756	80	520	1.000	
	.303	000	.319	435	000
			P7	P	P .7

Figure 1.2.4 Minimum Spanning Tree for the 5 different provenances of *Gliricidia* (P2, P5, P7, P10 and P17)

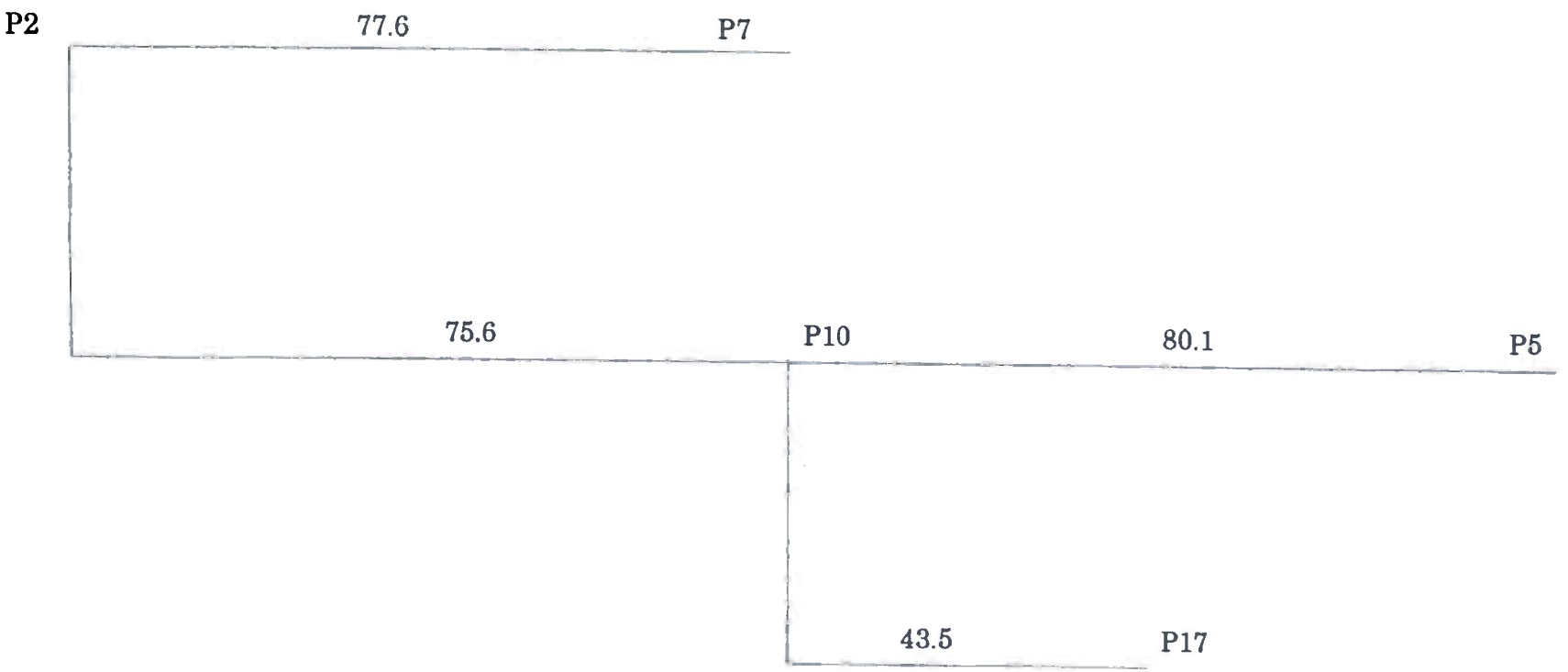
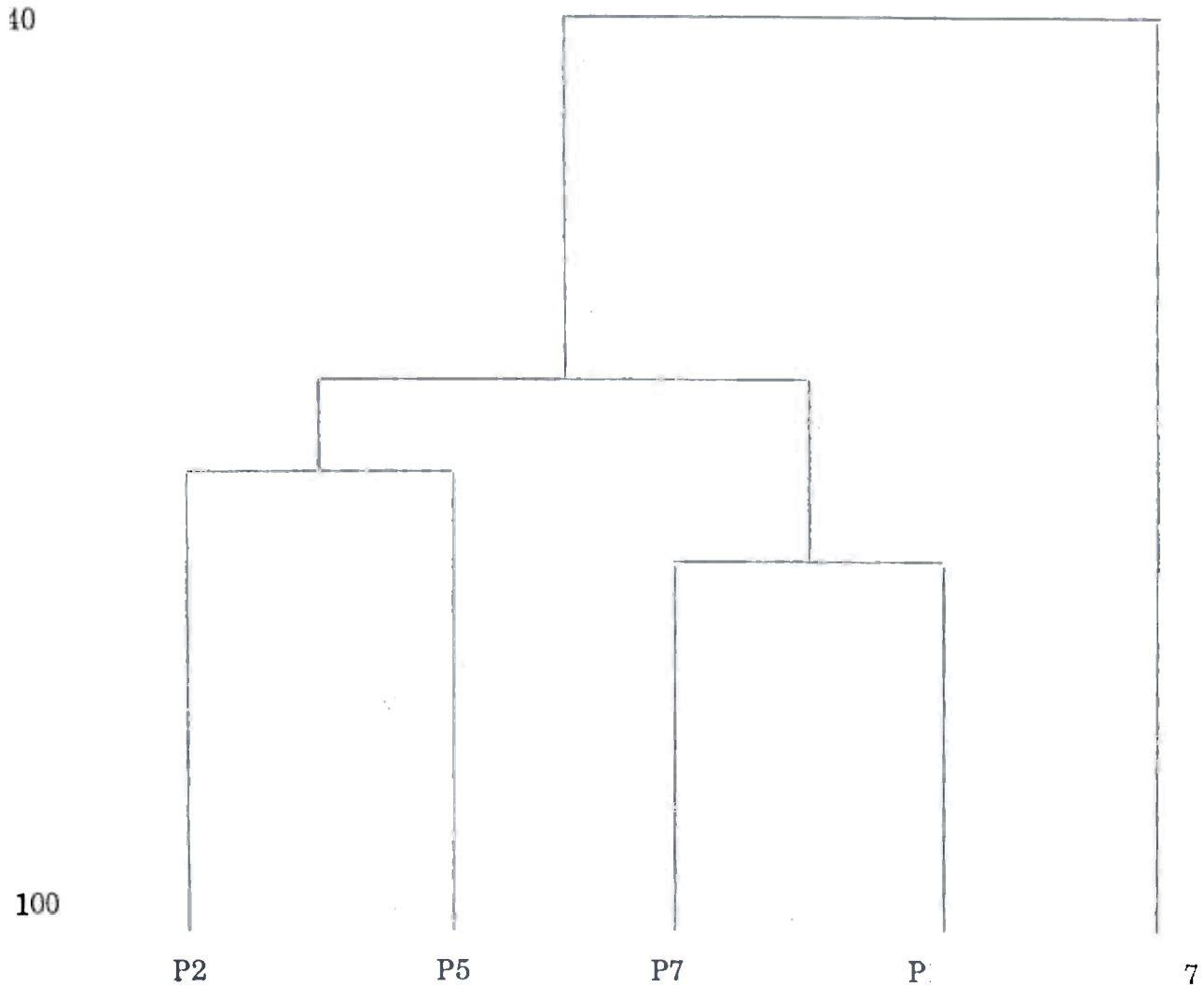


Figure .2.5 Dendrogram for the ifferent provenances of *Gliricidia* (P2, P5 P7 P1 and 7)



It is useful to look at the neighbourhood of each sample. The nearest neighbours table (Table 1.2.4) shows the levels of similarity for each of the 4 nearest neighbours. It is apparent the provenances P5 and P17 are different from each other.

Discriminant Analysis

The replicate samples for the 5 provenances were used as a library/learning set to predict the 20 unknown provenances. This gave an indication of which of the 5 provenances each of the remaining 20 were most similar or most related to. Table 1.2.5 indicates that the bulk samples from the 5 provenances which were treated as unknowns, were all identified correctly. Fifty per cent of the samples were found to be most like P17, which is an artifact population. Only P23, which is in fact a different species, was classified as being most like P5, which is a unique provenance and none were classified as being most like P2, which is also a unique provenance.

B. Statistical analyses were performed to investigate intra-provenance differences. All samples from the same provenance were treated as individuals ie. the a-i samples and the bulk sample. Only Biplot and PCA could be performed as before, but cluster analysis was performed in addition.

PCA

The principal components were calculated as before. The pair-wise plots, Component 1 versus Component 2 and Component 2 versus Component 3 are presented in Figures 1.2.6a -1.2.10a and 1.2.6b - 1.2.10b, for the individual provenances P2, P5, P7, P10 and P17, respectively. These figures give an indication of intra-provenance variation. The graphs for P2 (Figures 1.2.6a and 1.2.6b), show a split between P2b, P2d and P2g and the other samples in Component 1 and a split between P2a and the remaining samples in Component 3. Figures 1.2.7a and b display P5i set apart in Component 1, and P5a and P5d set apart in Component 2. For the P7 provenance, P7f was set well apart from the other samples in Component 1 (Figure 1.2.8a) and P7i at the extreme of Component 3 (Figure 1.2.8b). In Figure 1.2.10a, P17j can be seen split from the other samples in Component 1 and P17f in Component 2, the latter can also be seen clearly in Figure 1.2.10b. When the data from the 5 provenances was graphed together (Figure 1.2.11a and b), samples from the same provenance do not appear to cluster together. A certain

Table 1.2.4 Nearest neighbours table for P2, P5, P7, P10 and P17.

	PROVENANCES				
	P2	P5	P7	P10	P17
Nearest set	P7	P10	P2	P5	P10
Percentage similarity	77.6	80.1	77.6	80.1	43.5
Next nearest	P10	P2	P10	P2	P7
Percentage similarity	75.6	64.8	52.0	75.6	31.9
Next nearest	P5	P7	P17	P7	P2
Percentage similarity	64.8	19.7	31.9	52.0	30.2
Next nearest	P17	*	P5	P17	*
Percentage similarity	30.2	0	19.7	43.5	0

Table 1.2.5 Original groupings, and new allocation of provenances from Discriminant Analysis.

ORIGINAL	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
ALLOCATED	P7	P2	P7	P17	P5	P17	P7	P10	P17	P10	P10

ORIGINAL	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22
ALLOCATED	P17	P17	P10	P7	P10	P17	P17	P10	P17	P10	P17

ORIGINAL	P23	P24	P25
ALLOCATED	P5	P17	P17

Figure 1.2.6 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for provenance P2.

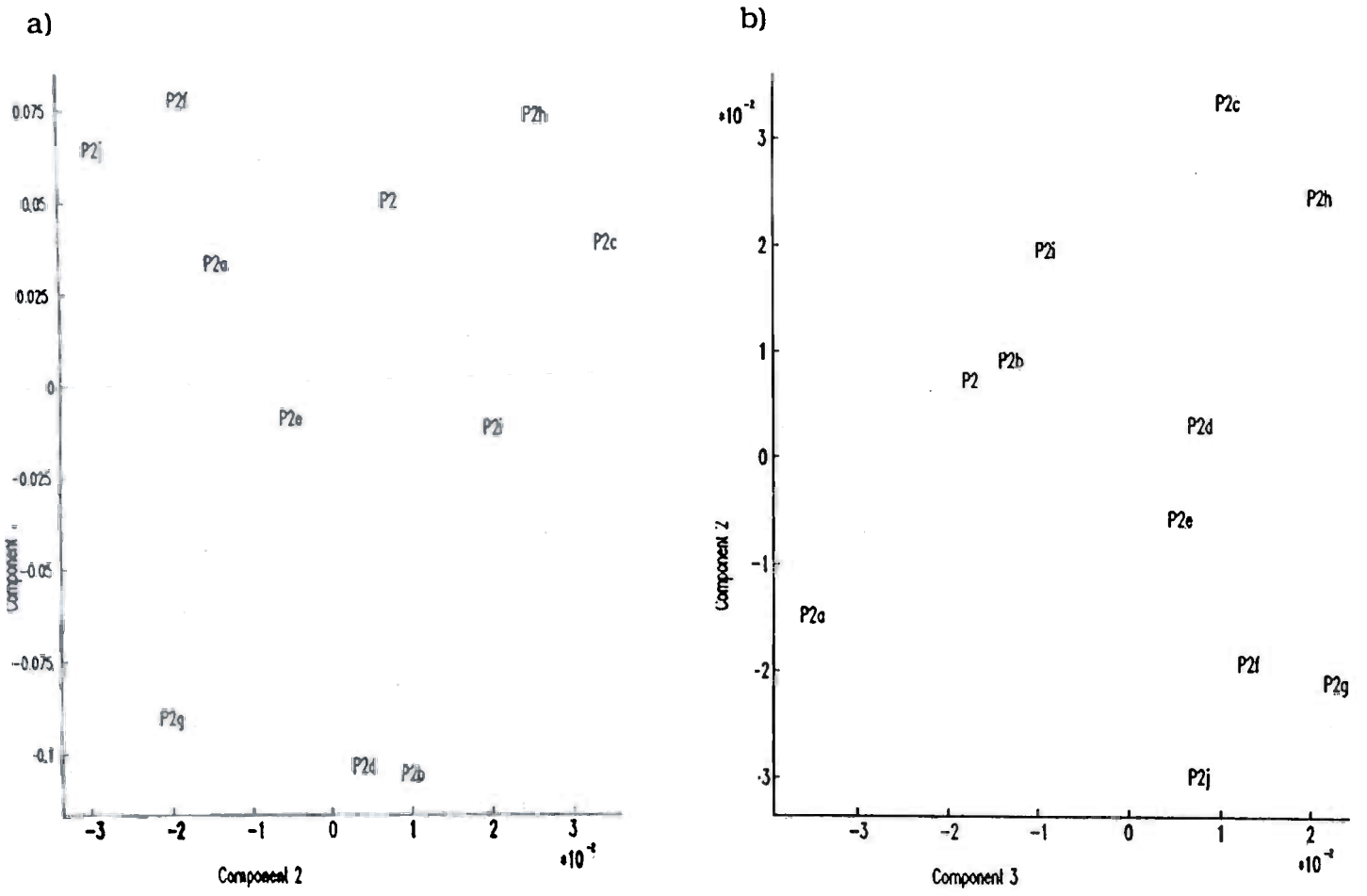


Figure 1.2.7 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for provenance P5.

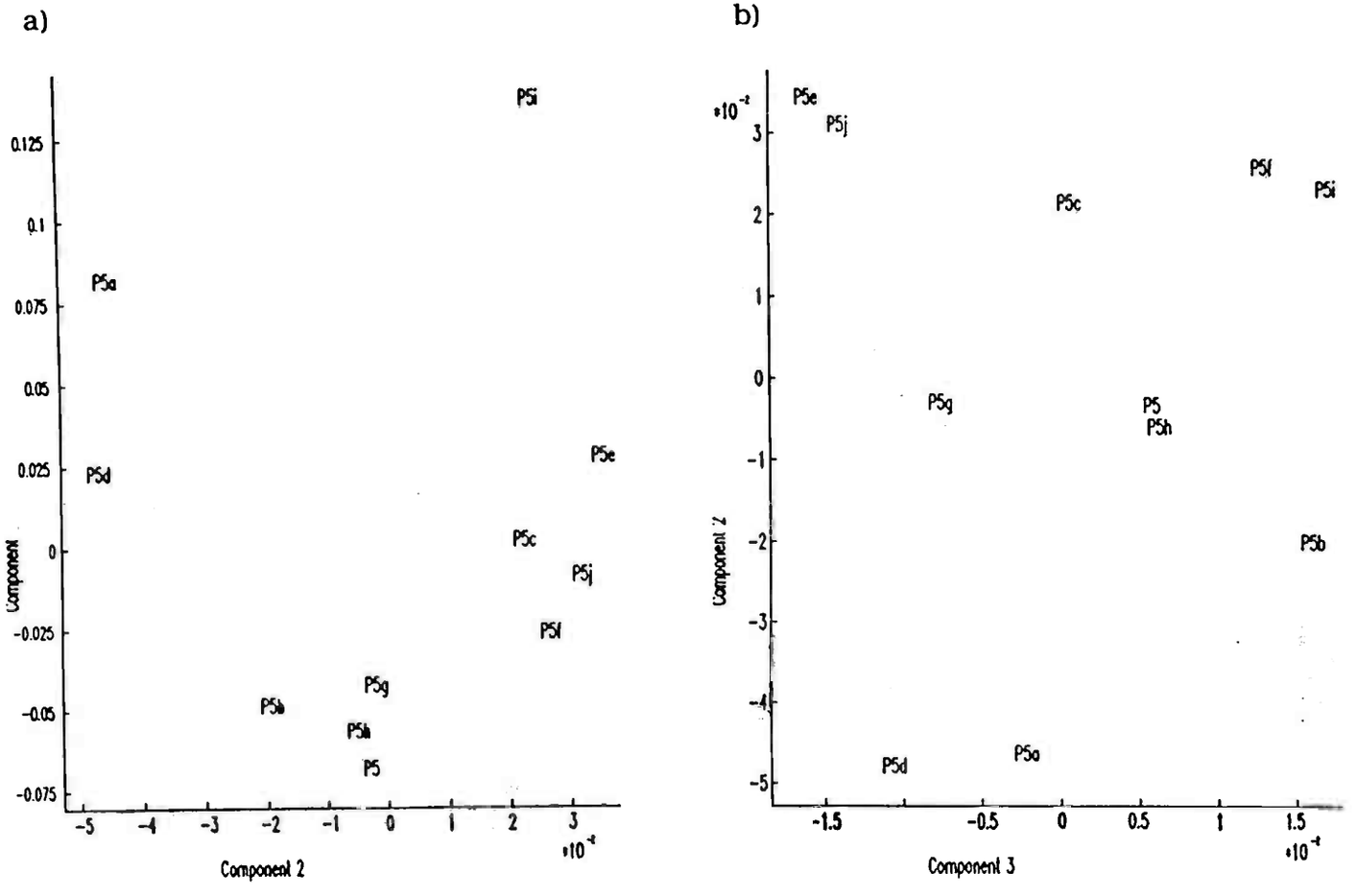


Figure 1.2.8 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for provenance P7.

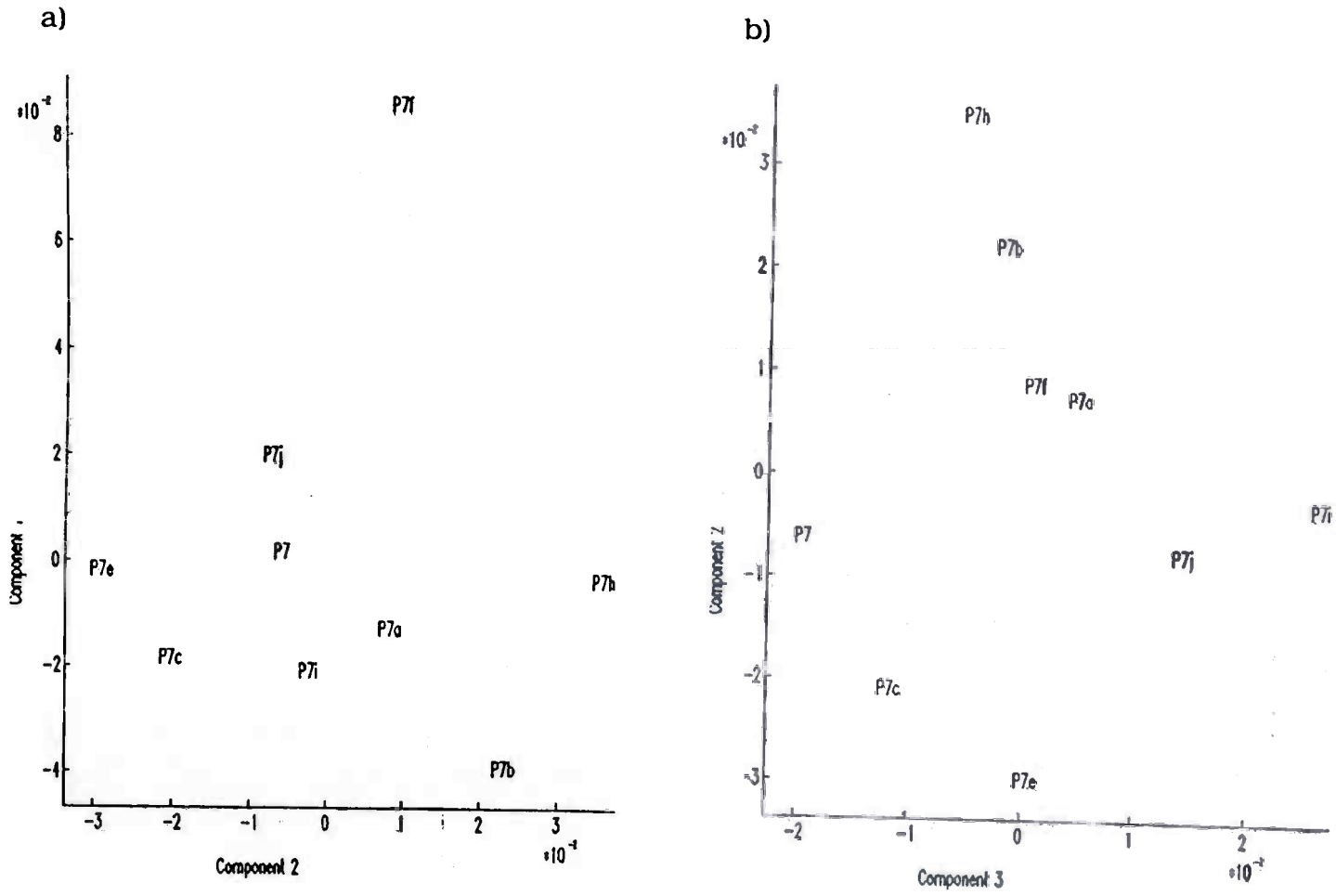


Figure 1.2.9 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for provenance P10..

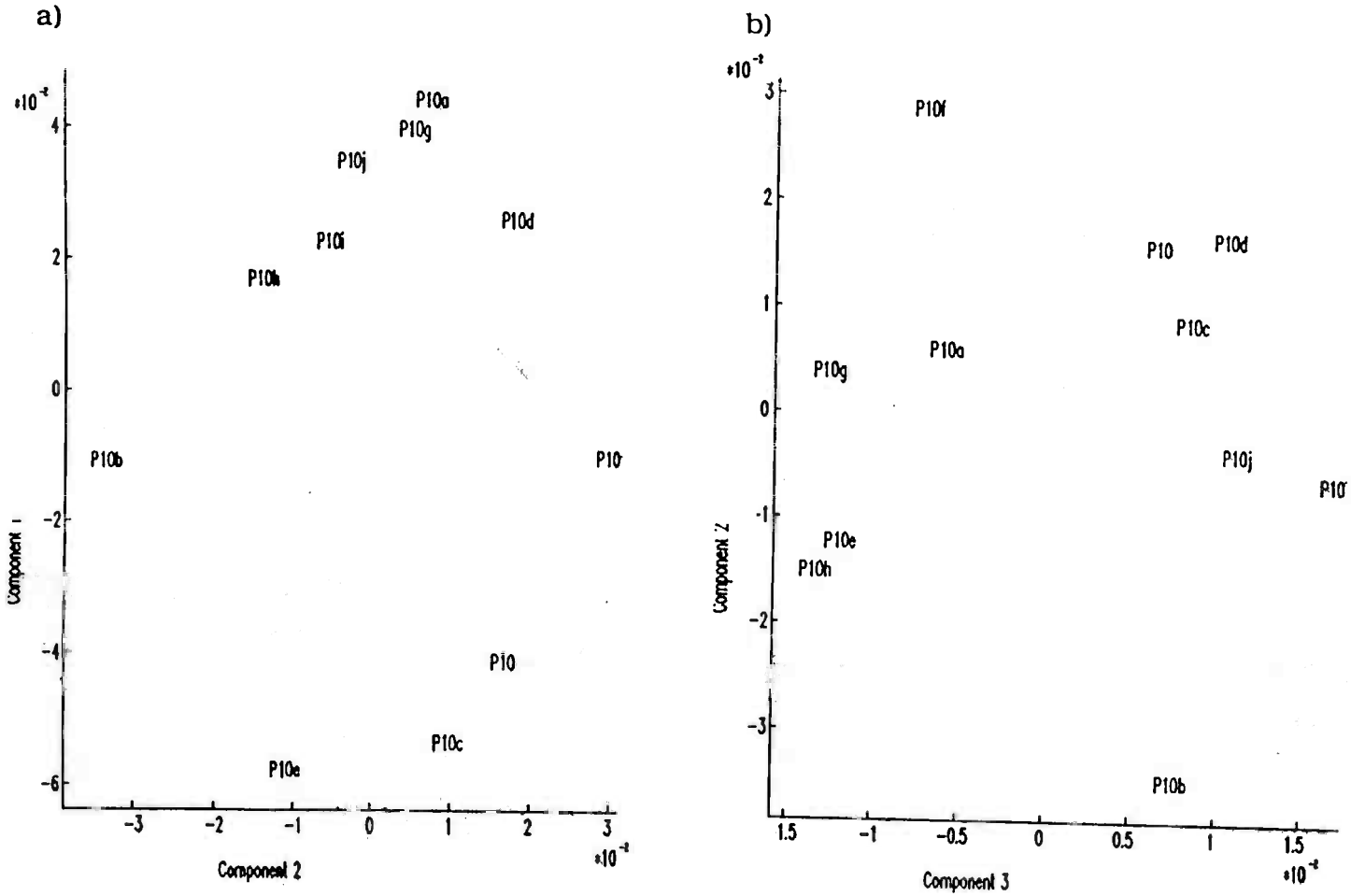


Figure 1.2.10 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for provenance P17.

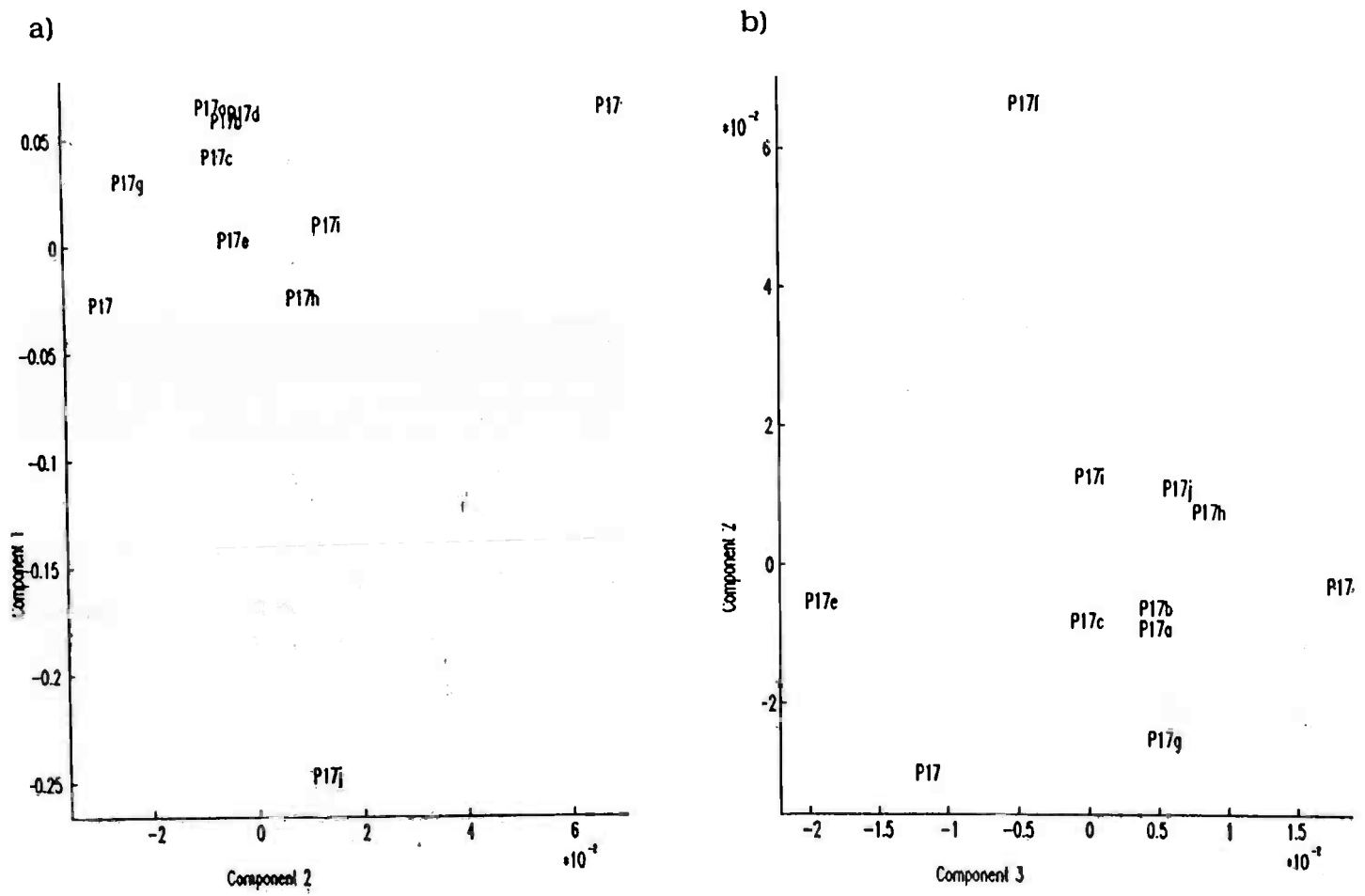
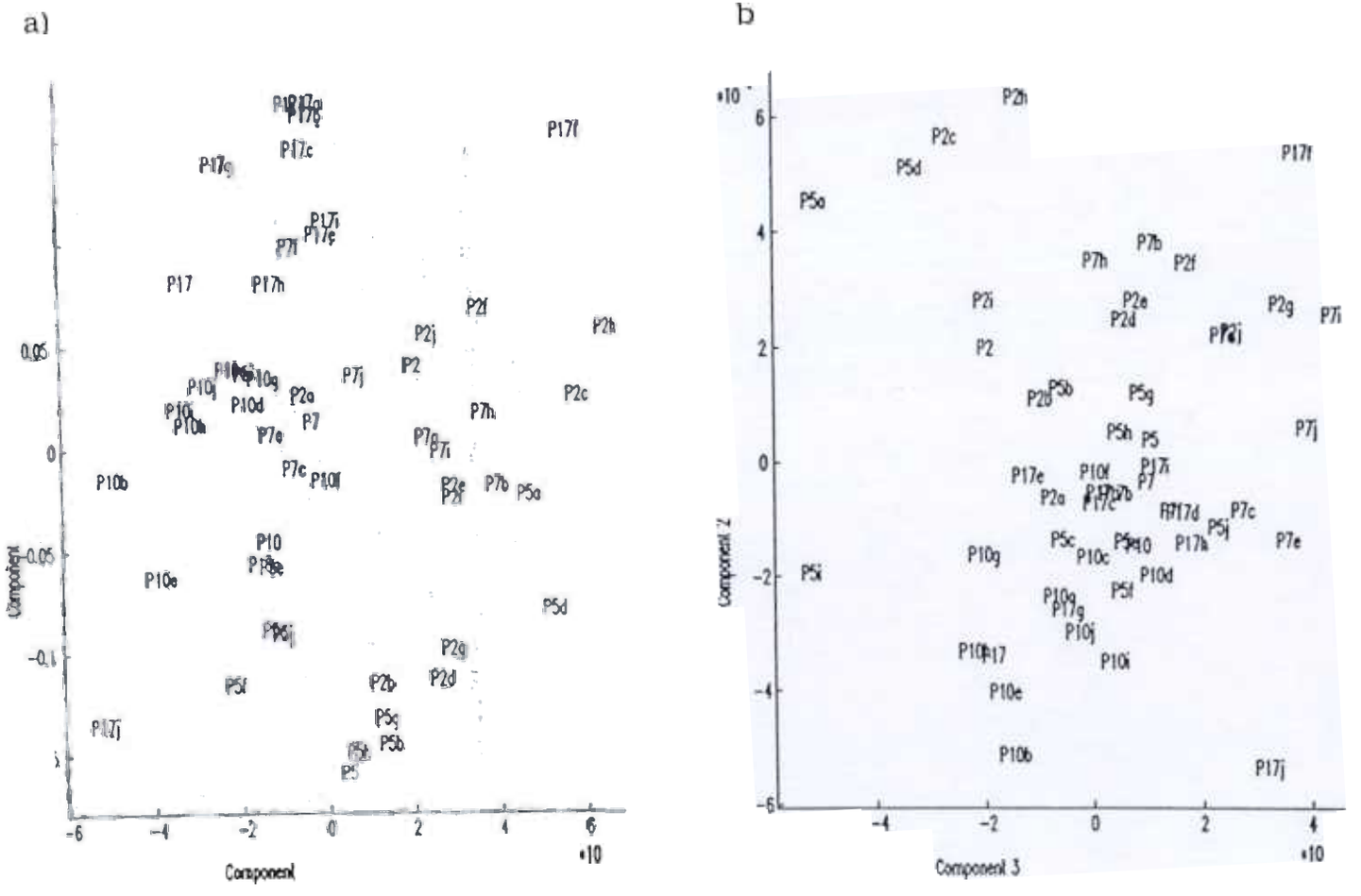


Figure Principal Componen plots howin a) C mponen versi C mponei 2 and b) Componen 2 versus Componen 3. for all samples from provenances P2 P5 P7 P 0 and P. 7



amount of overlap between the provenances was evident. The intra-provenance variation can be seen more clearly in Figure 1.2.12 which shows the 3-dimensional representation of Component 1 versus Component 2 versus Component 3.

Biplot Analysis

The distribution of samples from each of the provenances (P2, P5, P7, P10 and P17) are presented in Figures 1.2.13, 1.2.14, 1.2.15, 1.2.16 and 1.2.17, respectively. P5a, P5d and P17j were found to be set apart from the other samples from the same provenance.

Cluster Analysis

Samples were treated in the same sets as described previously. When the 25 provenances were analysed P5 and P23 were found to be the most different from the remainder (Figure 1.2.18). Individual cluster analyses were performed for the 5 provenances, P2, P5, P7, P10 and P17. The resulting dendrograms are given in Figures 1.2.19, 1.2.20, 1.2.21, 1.2.22 and 1.2.23, respectively. The dendrograms suggested that P2a, P5a, P5i, P7f, P10f, P17f and P17j were somewhat different from the other samples in their respective provenances. Tables 1.2.6, 1.2.7, 1.2.8, 1.2.9 and 1.2.10 show the nearest neighbours tables for provenances P2, P5, P7, P10 and P17, respectively, indicating the levels of similarity between samples and their 5 nearest neighbours. Cluster analysis was also performed using all of the above samples together. Overlapping between the different provenances was apparent.

1.2.2 Gas Production

A preliminary investigation was undertaken to examine whether different provenances of *Gliricidia* had the same or different degrees of predicted *in-vivo* degradability. Values for gas production rates and gas pool sizes were determined under EMC X0162. Initially the 25 provenances were fermented and ranked according to the gas pool sizes, as shown in Table 1.2.11. Gas production rates were in the range 0.0219 - 0.0421 h^{-1}, and it was found that samples with high gas production rates did not have necessarily correspondingly high gas pool sizes. It was noted that P23 gave the highest gas production rate. In a subsequent experiment, due to sample availability, different samples had to be chosen which were representative of the range of gas production. P9, P12, P14, P20 and P23 were selected as well as samples a, e and j for provenances P2, P5,

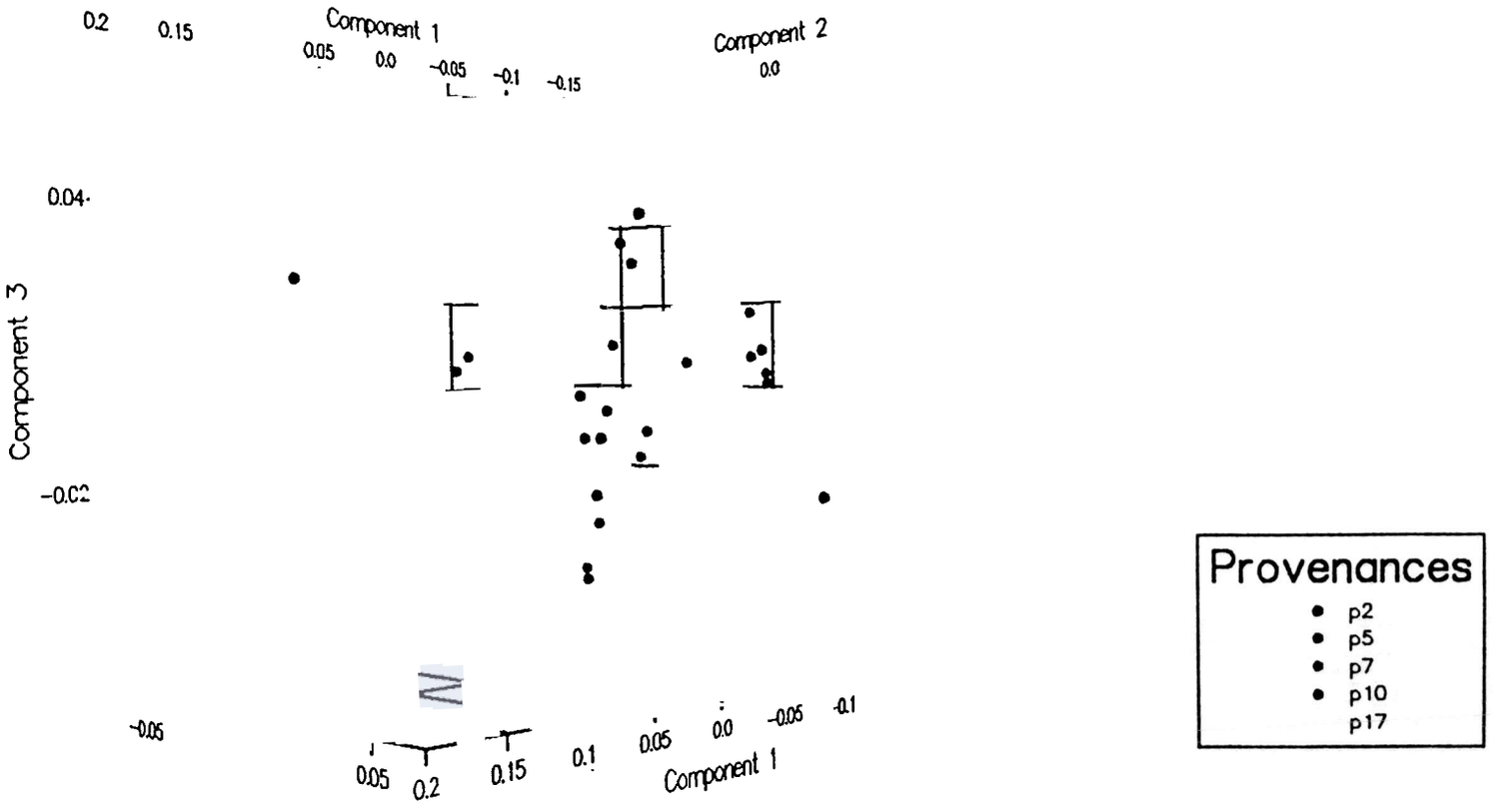

Compor 1 vs Co ıponent 2 vs Component 3
Component 1
Component 2
Component 3
Provenances
p2
p5
p7
p10
p17

Figure .2. 3 Distribution of provenace P2

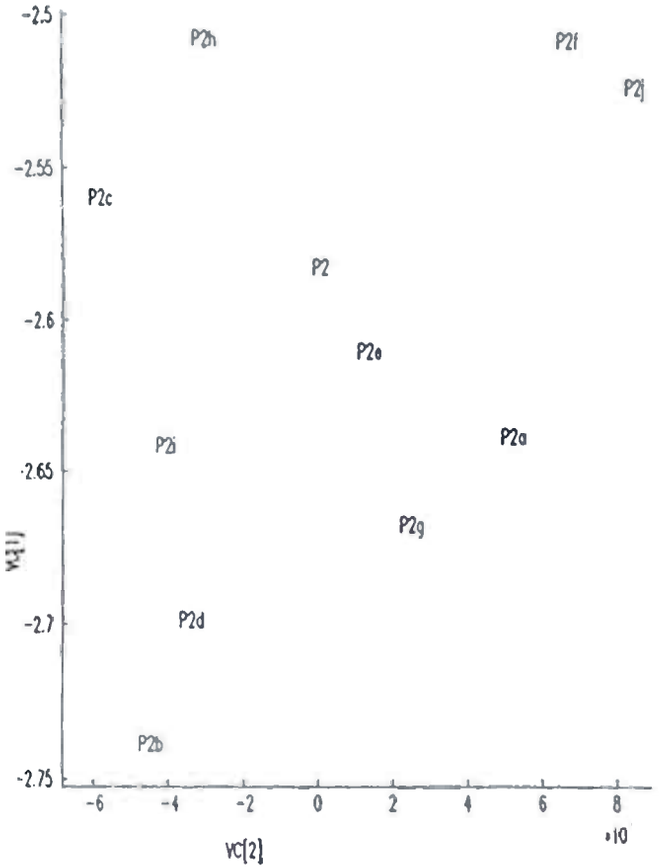
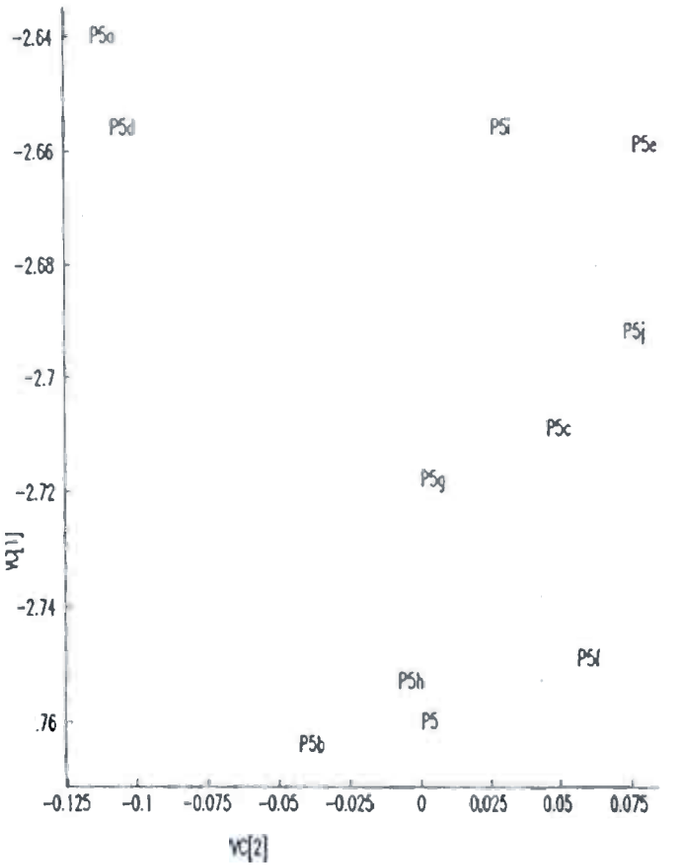
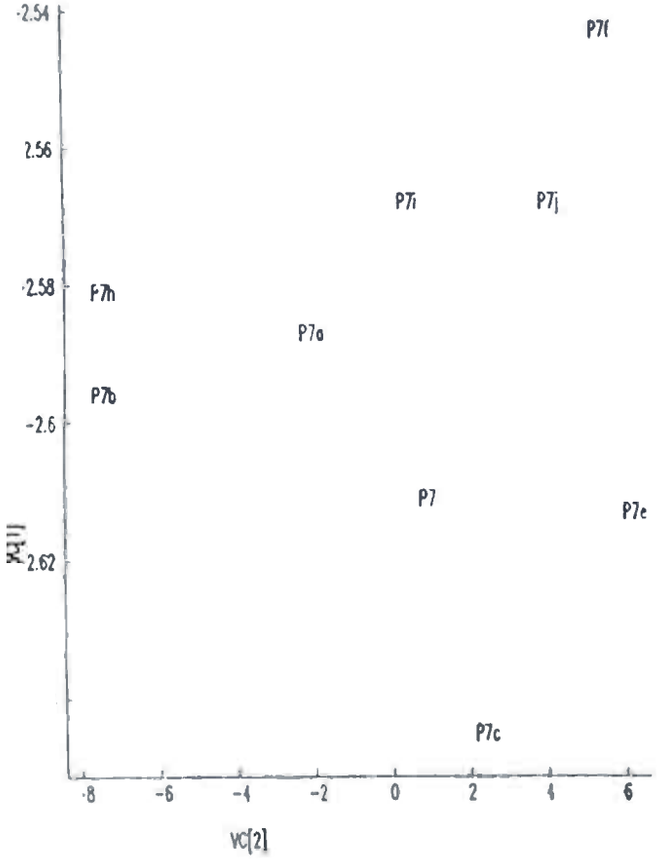


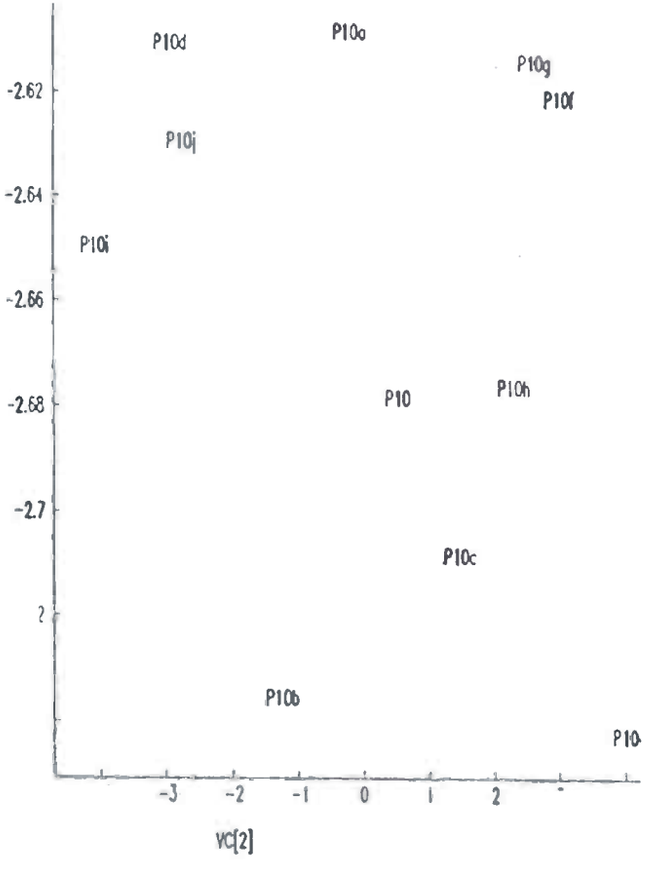
Figure .2 Distribution of provenance P5



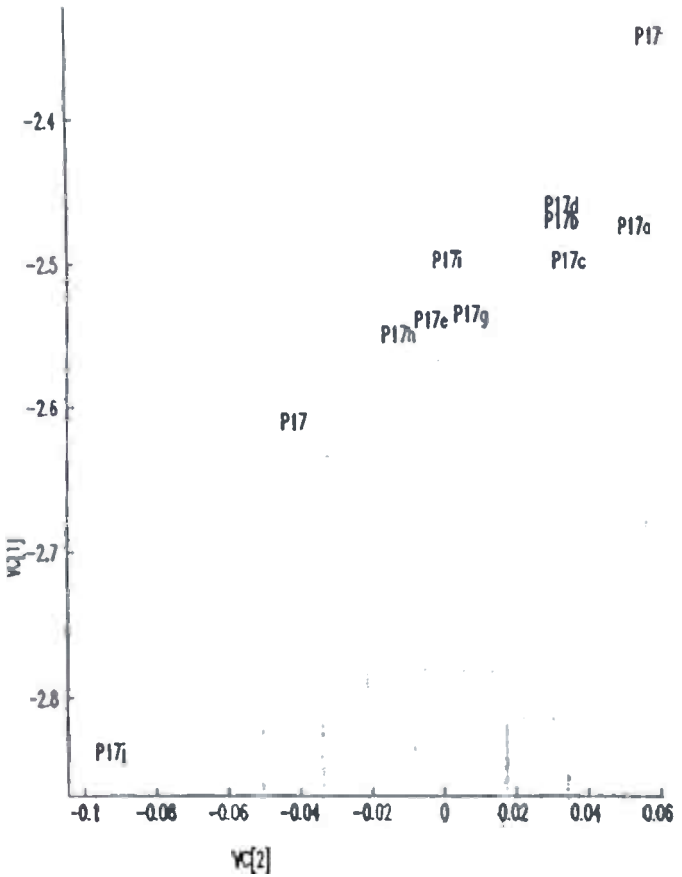
Figu Distribution of prove ace P7



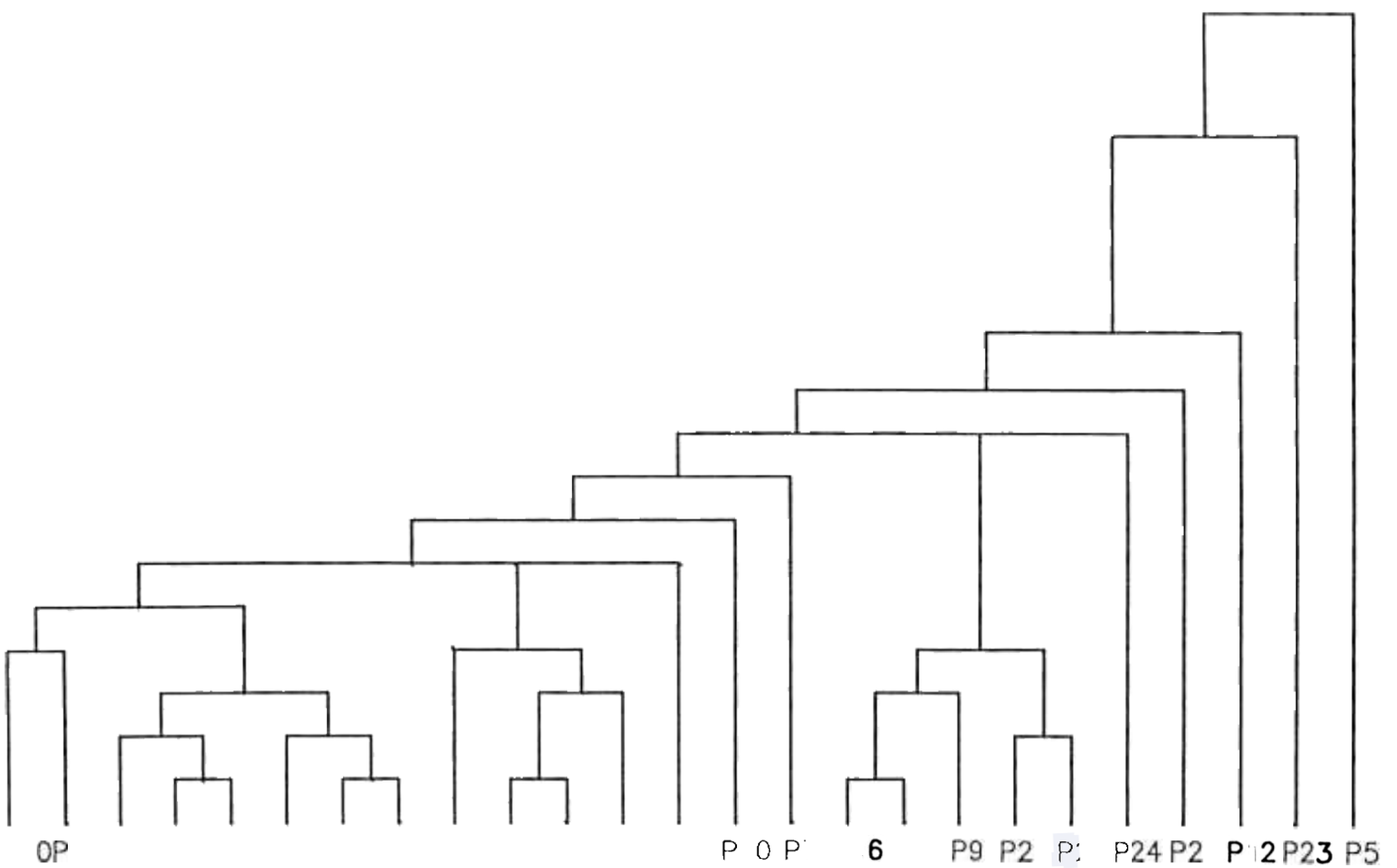
Figu 6 Distributio provenance P 0.



Igu Distribution of provena P



Figu Dendrogram f th 25 P provenances of *Gliricidia*



Figu Dendrogram for prov ance P2

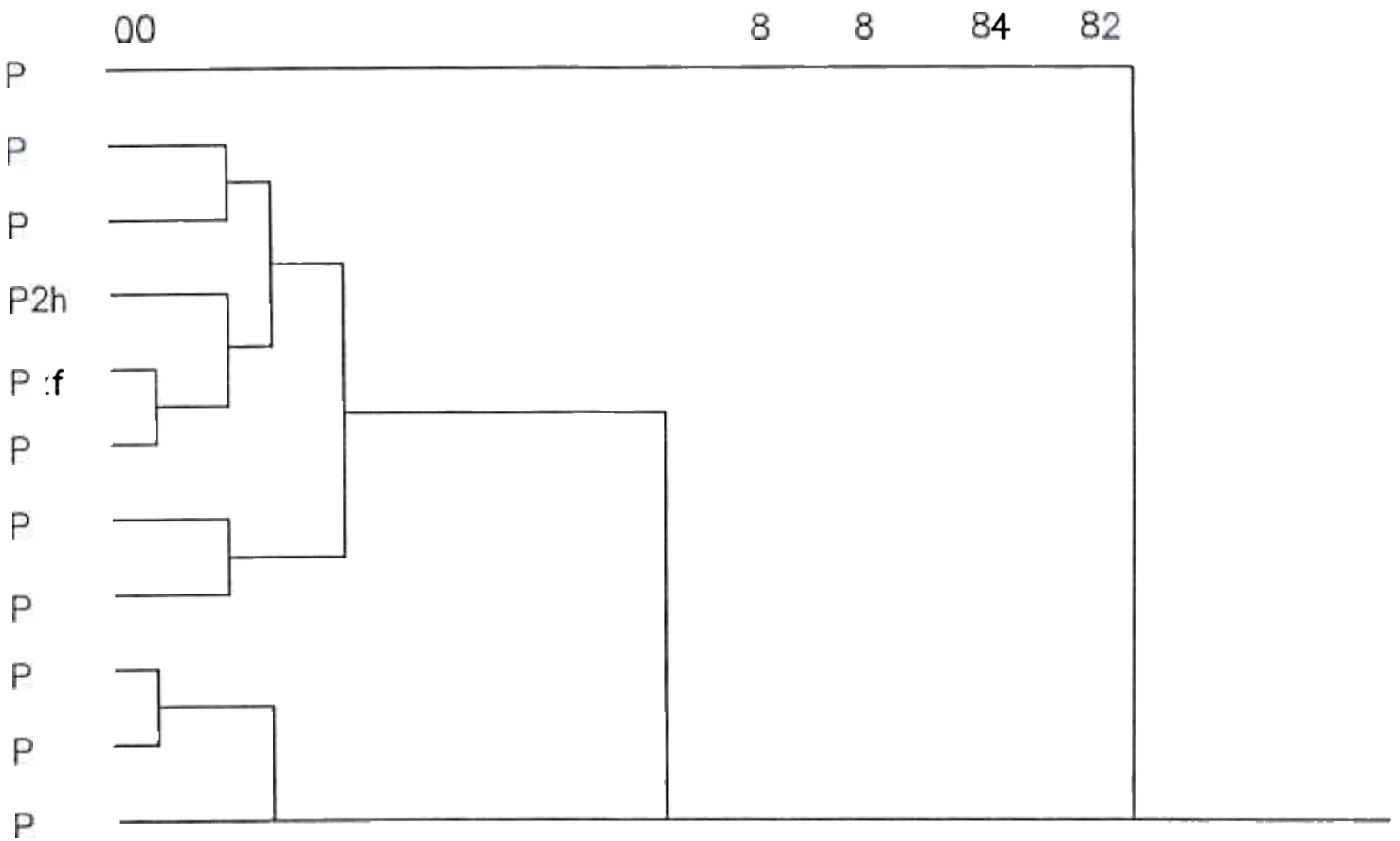


Figure 1.2.20 Dendrogram for provenance P5.

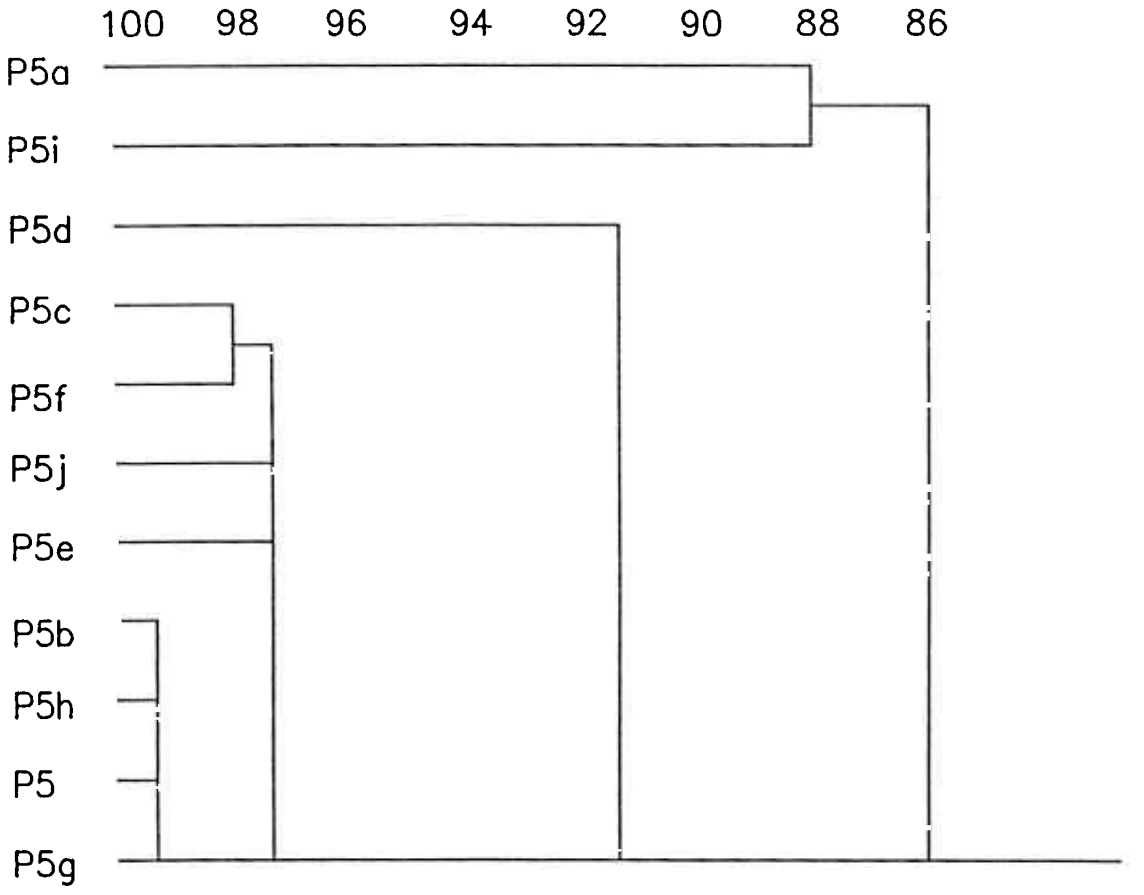
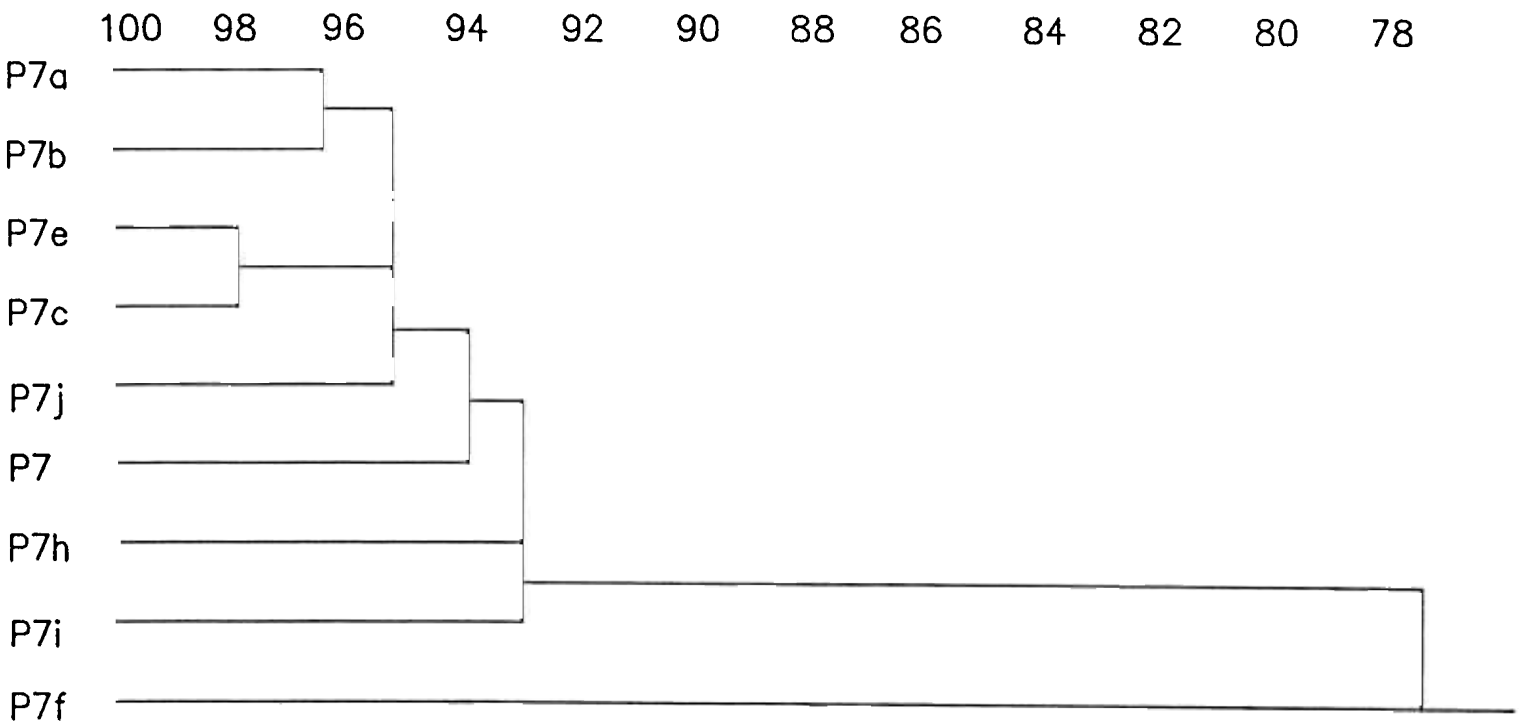
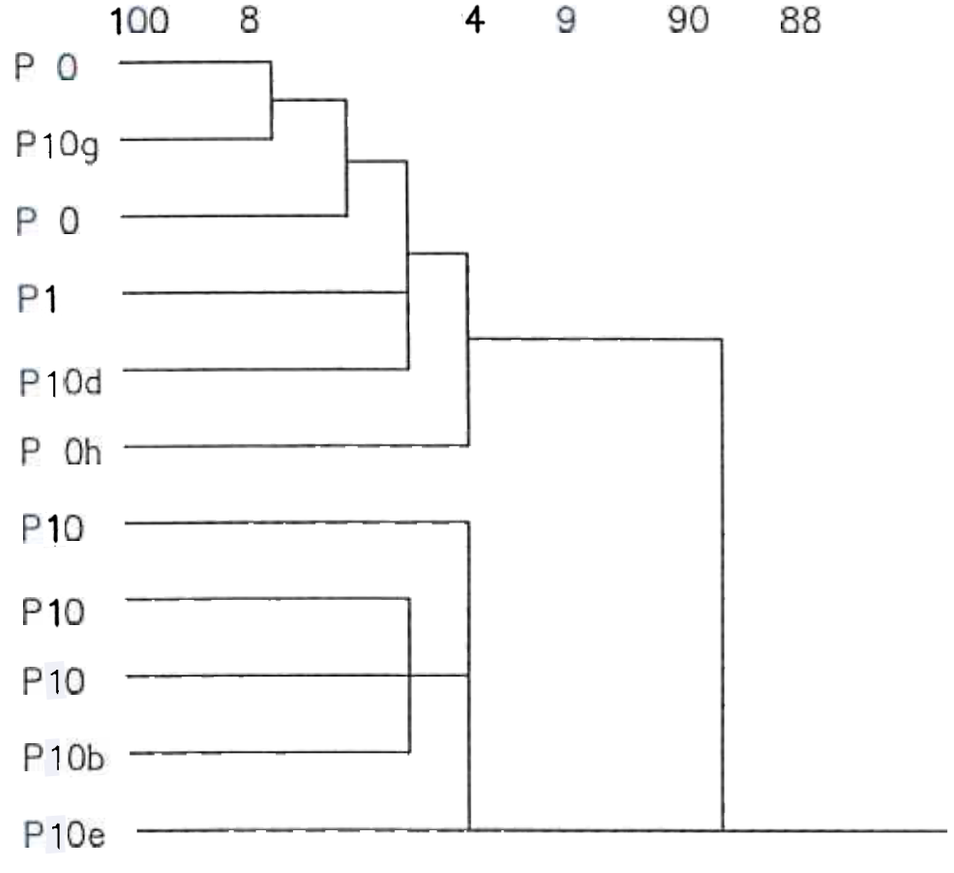


Figure 1.2.21 Dendrogram for provenance P7.



Figu 22 D ogram for provenance P 0.



Figu :.23 Dendrogram for provenance

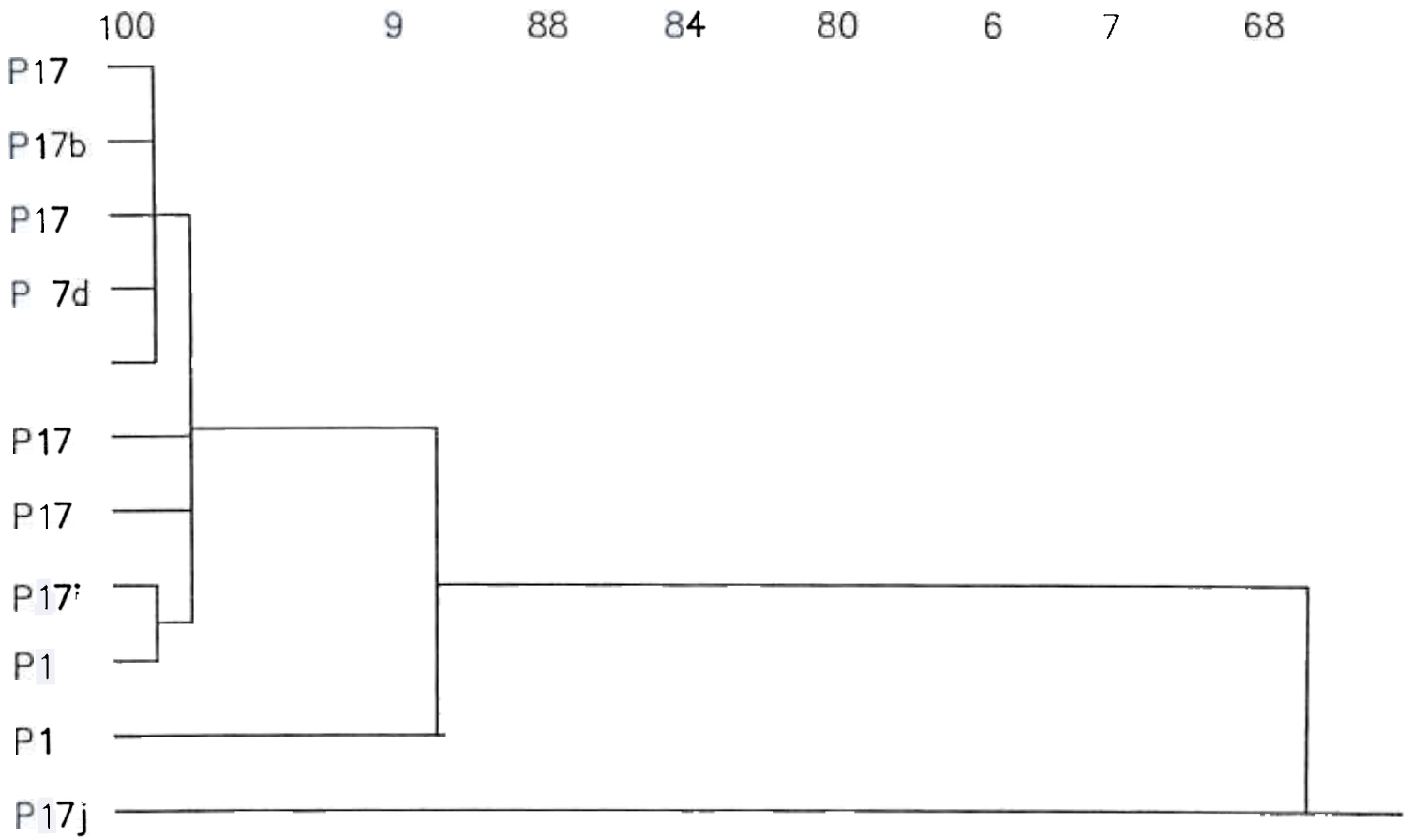


Table 1.2.6 Nearest neighbours table for provenance P2.

Sample	Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity
P2a	P2	81.4	P2h	79.9	P2f	78.7	P2j	76.3	P2c	74.9
P2b	P2d	99.2	P2g	94.8	P2i	90.4	P2e	90.1	P2	77.3
P2c	P2	98.7	P2h	97.9	P2f	96.7	P2i	96.0	P2e	96.0
P2d	P2b	99.2	P2g	97.8	P2e	90.0	P2i	88.8	P2c	76.3
P2e	P2i	98.7	P2	96.0	P2c	96.0	P2j	94.3	P2f	92.4
P2f	P2j	99.3	P2h	98.2	P2	97.0	P2c	96.7	P2e	92.4
P2g	P2d	97.8	P2b	94.8	P2e	85.8	P2i	81.7	P2c	69.0
P2h	P2f	98.2	P2c	97.9	P2	97.2	P2j	97.0	P2i	91.2
P2i	P2e	98.7	P2	96.7	P2c	96.0	P2j	92.2	P2h	91.2
P2j	P2f	99.3	P2	97.1	P2h	97.0	P2c	95.9	P2e	94.3
P2	P2c	98.7	P2h	97.2	P2j	97.1	P2f	97.0	P2i	96.7

'able 1.2.7 Nearest neighbours table for provenance P5.

Sample	Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity
P5a	P5i	88.3	P5d	86.4	P5e	79.7	P5c	75.1	P5j	73.0
P5b	P5h	99.3	P5	98.2	P5g	97.9	P5f	97.2	P5c	95.0
P5c	P5f	98.6	P5j	97.5	P5e	96.9	P5b	95.0	P5h	94.8
P5d	P5c	91.9	P5e	91.7	P5b	88.6	P5j	88.5	P5g	87.8
P5e	P5j	97.9	P5c	96.9	P5f	93.9	P5d	91.7	P5g	89.8
P5f	P5c	98.6	P5b	97.2	P5h	97.2	P5j	96.4	P5g	96.3
P5g	P5h	99.1	P5	98.6	P5b	97.9	P5f	96.3	P5c	94.6
P5h	P5	99.4	P5b	99.3	P5g	99.1	P5f	97.2	P5c	94.8
P5i	P5a	88.3	P5e	59.7	P5d	57.7	P5c	51.9	P5j	49.4
P5j	P5e	97.9	P5c	97.5	P5f	96.4	P5g	93.6	P5h	92.9
P5	P5h	99.4	P5g	98.6	P5b	98.2	P5f	95.3	P5c	91.6

Table 1.2.8 Nearest neighbours table for provenance P7.

Sample	Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity
P7a	P7b	96.7	P7e	95.4	P7j	95.0	P7	94.4	P7c	93.7
P7b	P7a	96.7	P7i	92.2	P7e	91.6	P7c	90.8	P7	90.3
P7c	P7e	97.7	P7a	93.7	P7	91.6	P7b	90.8	P7j	86.2
P7e	P7c	97.7	P7a	95.4	P7	92.3	P7b	91.6	P7 i	90.5
P7f	P7h	77.8	P7j	74.6	P7	60.9	P7a	56.7	P7e	43.3
P7h	P7a	93.4	P7j	92.3	P7	91.3	P7b	86.9	P7e	82.0
P7i	P7a	93.0	P7b	92.2	P7e	90.5	P7j	87.8	P7c	85.5
P7j	P7a	95.0	P7h	92.3	P7	90.9	P7e	89.9	P7i	87.8
P7	P7a	94.4	P7e	92.3	P7c	91.6	P7h	91.3	P7j	90.9

Table 1.2.9 Nearest neighbours table for provenance P10.

Sample	Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity
P10a	P10g	97.6	P10j	95.7	P10d	92.6	P10i	92.1	P10h	91.4
P10b	P10c	95.0	P10e	93.9	P10	93.7	P10f	80.3	P10i	73.9
P10c	P10	95.6	P10b	95.0	P10e	89.7	P10f	87.2	P10i	68.9
P10d	P10i	95.8	P10a	92.6	P10j	89.9	P10g	89.6	P10f	88.9
P10e	P10b	93.9	P10c	89.7	P10	86.8	P10f	76.2	P10i	69.6
P10f	P10	94.1	P10d	88.9	P10c	87.2	P10i	83.7	P10b	80.3
P10g	P10a	97.6	P10j	96.1	P10h	94.8	P10i	91.4	P10d	89.6
P10h	P10g	94.8	P10j	92.6	P10a	91.4	P10i	89.6	P10d	82.4
P10i	P10d	95.8	P10j	95.3	P10a	92.1	P10g	91.4	P10h	89.6
P10j	P10g	96.1	P10a	95.7	P10i	95.3	P10h	92.6	P10d	89.9
P10	P10c	95.6	P10f	94.1	P10b	93.7	P10e	86.8	P10d	82.9

Table 1.2.10 Nearest nighbours table for provenance P17.

Sample	Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity	Next Nearest Neighbour	Level of Similarity
P17a	P17b	99.6	P17d	99.3	P17c	99.2	P17g	98.3	P17e	95.7
P17b	P17c	99.7	P17a	99.6	P17d	99.5	P17g	98.5	P17i	97.2
P17c	P17b	99.7	P17a	99.2	P17g	99.0	P17d	98.7	P17e	98.4
P17d	P17b	99.5	P17a	99.3	P17c	98.7	P17g	97.5	P17i	96.1
P17e	P17c	98.4	P17	98.3	P17i	98.0	P17h	97.7	P17g	97.6
P17f	P17i	91.5	P17h	87.3	P17b	87.3	P17d	87.2	P17c	86.2
P17g	P17c	99.0	P17b	98.5	P17a	98.3	P17e	97.6	P17d	97.5
P17h	P17i	99.0	P17e	97.7	P17	96.0	P17c	96.0	P17b	94.7
P17i	P17h	99.0	P17e	98.0	P17c	97.8	P17b	97.2	P17d	96.1
P17j	P17h	66.3	P17	64.0	P17e	56.7	P17h	56.6	P17c	42.1
P17	P17e	98.3	P17g	96.1	P17h	96.0	P17c	94.8	P17i	94.6

Table 1.2.11 Gas production run 1.

MOST GAS PRODUCED	**P1**
	P2
	P18
	P4
	P8
	P6
	P11
	P23
	P22
	P5
	P13
	P20
	P15
	P17
	P21
	P3
	P25
	P9
	P1
	P24
	P19
	P7
	P16
LEAST GAS PRODUCED	P14

P7, P10 and P17. The ranking of the samples according to gas pool size can be seen in Table 1.2.12. The absolute values obtained differed from those found in the first experiment but the ranking orders for gas production rate and gas pool size were the same for the 5 provenances repeated. The results of this work suggested that samples within a provenance could have different degrees of *in-vivo* degradability.

1.2.3 Calibration development

Chemical composition and nutritional value (assessed by the *in-vitro* technique (7) were assessed for the 25 provenances. Data for dry matter (DM), ash, crude protein (CP) and DM Digestibility (DMD), organic matter Digestibility (OMD) and digestible OM in the DM (DOMD) are given in Table 1.2.13. Calibrations were developed for the above chemical parameters and two gas production parameters, Total Gas Produced (TGP) and Total Gas Produced minus the Control (TGP-C).

Calibration equations were developed using the following sets of samples:-

i) SET A - 25 bulk samples

ii) SET B - 53, a - i samples for P2, P5, P7, P10 and P17

iii) SET C - 25 *Gliricidia* samples from another experiment plus SET A

A Partial Least Squares (PLS) regression technique (8) was used to develop the calibration equations.

SET A

A mathematical procedure was applied to the spectral data to rank individual samples on the basis of their Mahalanobis ('H') distance from the average spectrum in the set (9). The calculated distances are given in Table 1.2.14. This ranking technique is useful for identifying extreme samples. In this case P23 would be classified as being extreme with an 'H' distance greater than 3. However, this sample was retained in the calibration set at this stage as it was considered that it could be a valid extension of the population. The calibration procedure employs a cross validation process (10) which is a technique based only on the calibration data. It is similar to prediction, as it only tests predictors on data which were not involved in the calibration process. The cross validation statistics give an indication of how well the equation will perform. Table 1.2.15 shows the calibration statistics ie. number of samples used in the calibration (n), standard error of calibration

Table 1.2.12 Gas production run 2.

MOST GAS PRODUCED	P7e
	P10j
	P17e
	P17a
	P2e
	P12
	P2j
	P2a
	P5e
	P5a
	P5j
	P7a
	P7j
	P23
	P17j
	P20
	P10e
	P10a
	P9
LEAST GAS PRODUCED	**P14**

Table 1.2.13 Laboratory analyses for the 25 *Gliricidia* provenances

Sample No.	Code	%DM	%ASH	%CP	%DMD	%DOMD	%DMD
P1	13/84	94.68	10.49	21.63	66.93	57.04	63.73
P2	14/84	95.30	10.46	19.63	65.66	56.55	63.15
P3	15/84	95.05	9.53	21.69	62.82	53.93	59.61
P4	16/84	94.92	9.90	20.25	64.72	55.60	61.71
P5	17/84	94.75	9.69	19.25	61.41	52.51	58.14
P6	24/84	94.85	9.12	19.00	64.18	55.63	61.21
P7	25/84	95.16	10.41	21.38	66.20	56.67	63.26
P8	30/84	95.36	9.11	18.50	66.24	57.43	63.18
P9	31/84	94.87	9.20	19.63	65.24	56.48	62.20
P10	33/85	94.87	9.17	20.06	64.41	55.91	61.55
P11	34/85	94.70	9.17	20.13	60.96	52.83	58.16
P12	35/85	95.21	9.23	21.94	63.77	54.89	60.47
P13	38/85	95.04	9.75	21.38	62.16	53.39	59.16
P14	39/85	94.90	9.48	22.31	58.96	49.66	54.85
P15	40/85	95.28	9.76	22.12	63.84	54.53	60.42
P16	41/85	95.33	10.39	20.75	64.19	54.16	60.44
P17	1/86	95.97	9.61	21.94	62.60	63.08	69.79
P18	10/86	94.99	8.64	20.44	66.08	57.34	62.76
P19	11/86	94.95	9.35	20.19	61.36	51.77	57.11
P20	12/86	94.90	9.80	20.69	61.35	51.59	57.19
P21	13/86	95.83	9.62	20.44	65.00	56.01	61.97
P22	24/96	95.36	9.01	20.06	70.72	61.88	68.01
P23	43/87	94.19	11.22	16.63	53.92	44.99	50.68
P24	72/87	94.84	10.70	19.44	62.88	53.23	59.61
P25	75/87	95.71	10.34	21.00	69.33	59.50	66.37

Table 1.2.14 Mahalanobis distances for the 25 provenances

Position Number	Sample Number	H Distance
1	P4	0.115
2	P13	0.167
3	P18	0.385
4	P19	0.415
5	P2	0.422
6	P6	0.478
7	P3	0.647
8	P17	0.667
9	P9	0.678
10	P15	0.694
11	P10	0.802
12	P14	0.833
13	P22	0.841
14	P11	0.853
15	P21	0.876
16	P7	0.971
17	P16	1.049
18	P1	1.110
19	P8	1.199
20	P20	1.215
21	P25	1.333
22	P24	1.674
23	P12	1.855
24	P5	1.948
25	P23	3.772

Table 1.2.15 Calibration and cross validation statistics for SET A (25 bulk *Gliricidia* samples)

Variable	n	SEC	R^2	SECV	CVR^2
DM	24	0.28	0.34	0.32	0.36
ASH	25	0.30	0.77	0.41	0.55
CP	20	0.16	0.97	0.38	0.84
DMD	19	0.28	0.99	1.01	0.83
DOMD	21	0.53	0.96	1.30	0.80
OMD	21	0.37	0.99	1.34	0.83
TGP-C	24	4.69	0.59	7.22	0.10
TGP	24	4.44	0.66	6.55	0.33

Table 1.2.16 Prediction statistics for SET B (predicted using the equations in Table 1.2.15)

Variable	BIAS	SEP(C)	R^2
DM	-2.03	1.06	0.12
CP	-1.73	0.76	0.77

Table 1.2.17 Prediction statistics for SET C (predicted using the equations in Table 1.2.15)

Variable	BIAS	SEP(C)	R^2
DM	-2.45	1.76	0.06
ASH	-0.20	0.39	0.87
CP	-3.56	0.83	0.46
DMD	8.75	2.58	0.36
DOMD	11.45	2.05	0.59
OMD	11.75	2.19	0.70

(SEC) and fraction of explained variance (R^2) for the calibration. Cross validation statistics ie. standard error for cross validation (SECV) and fraction of explained variance (CVR^2) for SET A are also presented. Equations for predicting DM are very difficult to develop when the range in DM of the samples is small. The range differed by less than 1.5%, so the poor calibration statistics were not unexpected. Low SEC values were achieved for these equations, but the R^2 values were also very low. In addition, the number of samples used to develop the equations was small. However, the equations for the other constituents gave good calibration and cross validation statistics despite the small number of samples.

Equations developed from SET A were used to predict values for samples in Sets B and C. Prediction statistics (bias, standard error of prediction (corrected for bias) (SEP(C)) and R^2) obtained when predicting parameters for SET B and for SET C are presented in Table 1.2.16 and 1.2.17, respectively. The very large biases found between the sample sets, possibly were a result of the samples being analysed in different laboratories, different ranges in constituents covered in the different sets and also SET C samples were collected in a different season.

SET B

Equations relating DM and CP to spectral data were developed for SET B samples. These variables were the only laboratory analyses available for these samples. The calibration and cross validation statistics are given in Table 1.2.18. As would be expected, given a wider range in DM, calibration statistics were better than those obtained for SET A. The equations derived were used to predict values for samples in SET A. The prediction statistics are given in Table 1.2.19, and again a very large bias is apparent between SET A and B.

SET C

This set consisted of a total of 50 *Gliricidia* samples from two seasons. The calibration and cross validation statistics are given in Table 1.2.20. The equations derived were used to predict DM etc. for samples in SET B. Table 1.2.21 shows the prediction errors obtained for SET B.

Table 1.2.18 Calibration and cross validation statistics for SET B

Variable	n	SEC	R^2	SECV	CVR^2
DM	50	0.21	0.96	0.27	0.94
CP	51	0.41	0.93	0.63	0.84

Table 1.2.19 Prediction statistics for SET A (predicted using the equations in Table 1.2.18)

Variable	BIAS	SEP(C)	R^2
DM	1.93	0.54	0.27
CP	1.47	0.68	0.73

Variable	n	SEC	R^2	SECV	CVR^2
DM	49	1.11	0.62	1.27	0.52
ASH	48	0.31	0.85	0.35	0.80
CP	46	0.33	0.91	0.47	0.81
DMD	46	0.67	0.95	1.09	0.87
DOMD	47	0.90	0.90	1.38	0.77
OMD	48	0.92	0.93	1.69	0.77
TGP-C	24	4.72	0.58	7.24	0.09
TGP	24	4.50	0.66	6.58	0.33

Table 1.2.21 Prediction statistics for SET B (predicted using the equations in Table 1.2.20)

Variable	BIAS	SEP(C)	R^2
DM	-2.03	1.06	0.12
CP	-1.73	0.76	0.77

1.2.4 NIR spectra

Quantitative statistical analysis of spectral data revealed differences between *Gliricidia* provenances. Qualitative assessment of spectra was undertaken to identify regions of the spectrum which may be associated with these differences. The 5 provenances, P2, P5, P7, P10 and P17, were studied together. Further, P23 was found to be different to the other provenances so P2, P5 and P23 were examined together. Log 1/R spectra for these provenances are shown in Figures 1.2.24 and 1.2.25. Log 1/R spectra are difficult to interpret, so mathematical transformations were used to extract any useful information. In this particular case, a 2nd derivative and Standard Normal Variate (SNV) and De-Trend (DT) (collectively named SDT) (11) transformations were used. The mathematical transformation SDT removes a large amount of particle size and multi-collinearity effects. The 2nd derivative spectra (Figure 1.2.26) revealed differences in the 1654 - 1700 nm and 2070 - 2220 nm regions of the spectra. P7 appeared to differ from the other provenances in the 2202 nm region. In Figure 1.2.27, P23 shows distinct differences in several regions of the spectrum, particularly in the regions of 1686 and 2280 nm. SDT transformed spectra are given in Figures 1.2.28 and 1.2.29 for the 5 and 3 provenances, respectively. Figure 1.2.29 shows distinct differences between the spectrum for P23 and the other 2 provenances.

Figure 1.2.24 Log 1/R Sectra for P2, P5, P7 P10 and P17

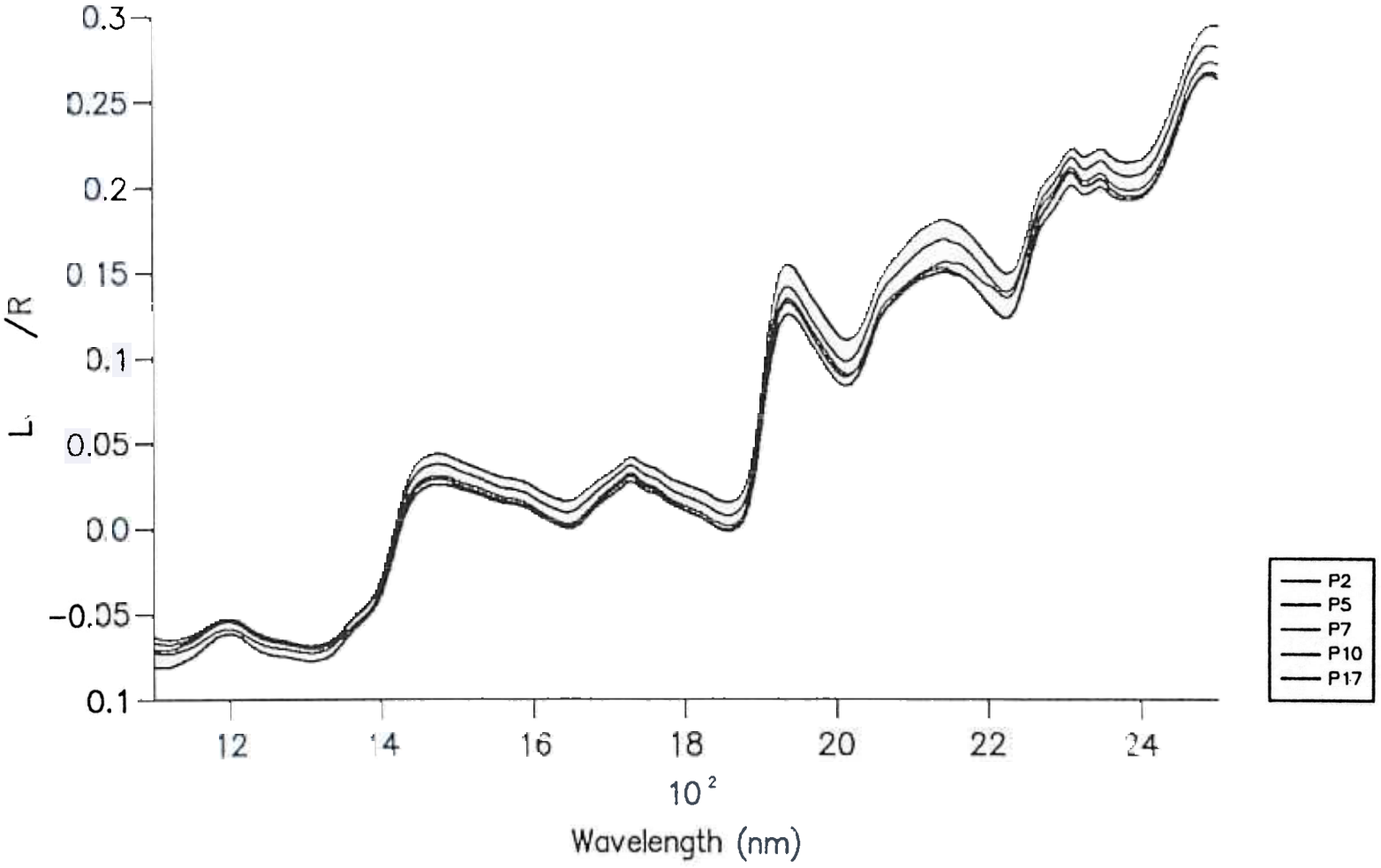


Figure 1.2.25 Log 1/R) Spectra for P2, P5 and P23

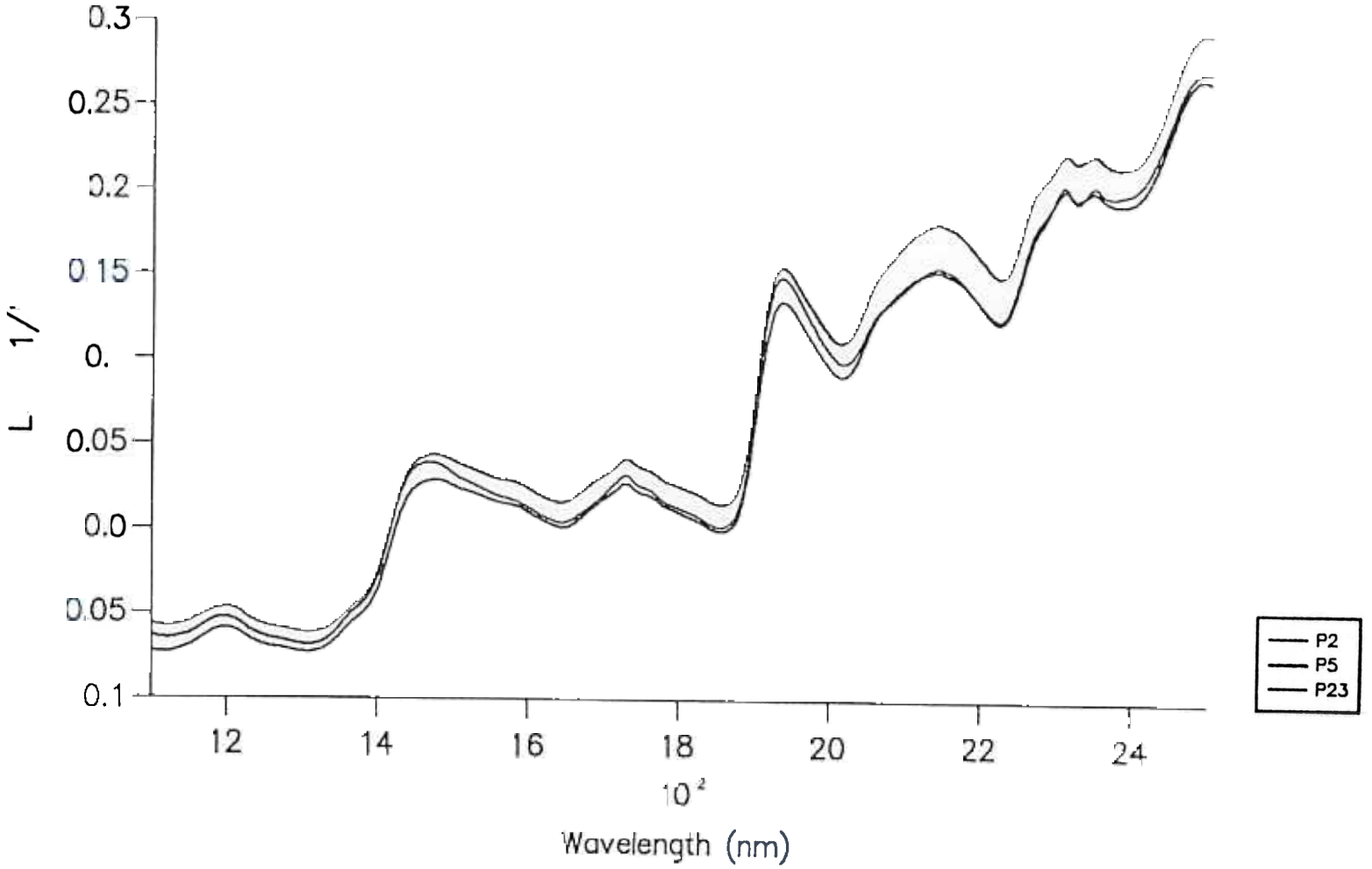


Figure 1.2.26 Second Derivative Spectra for P2, P5, P7, P10 and P17

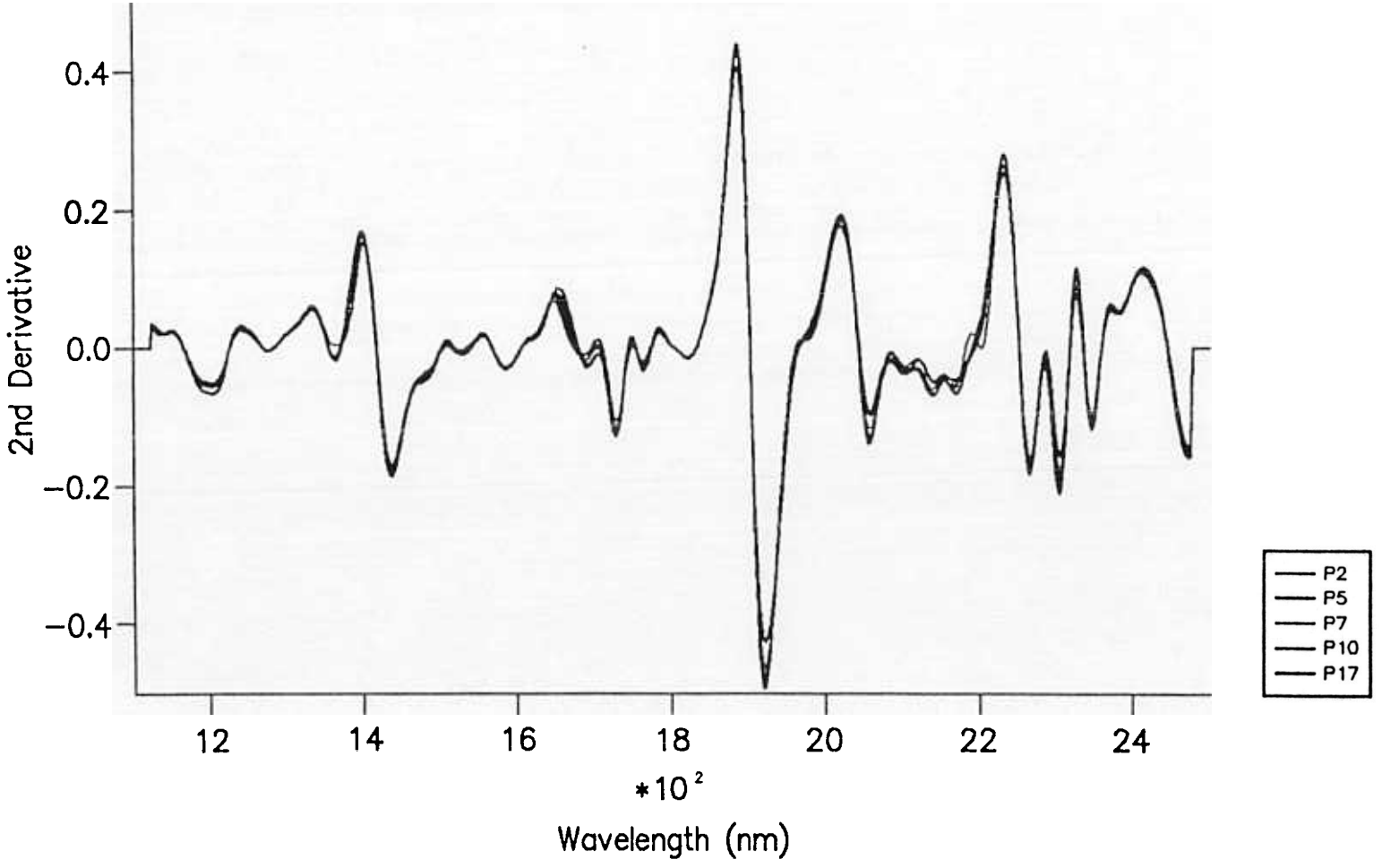
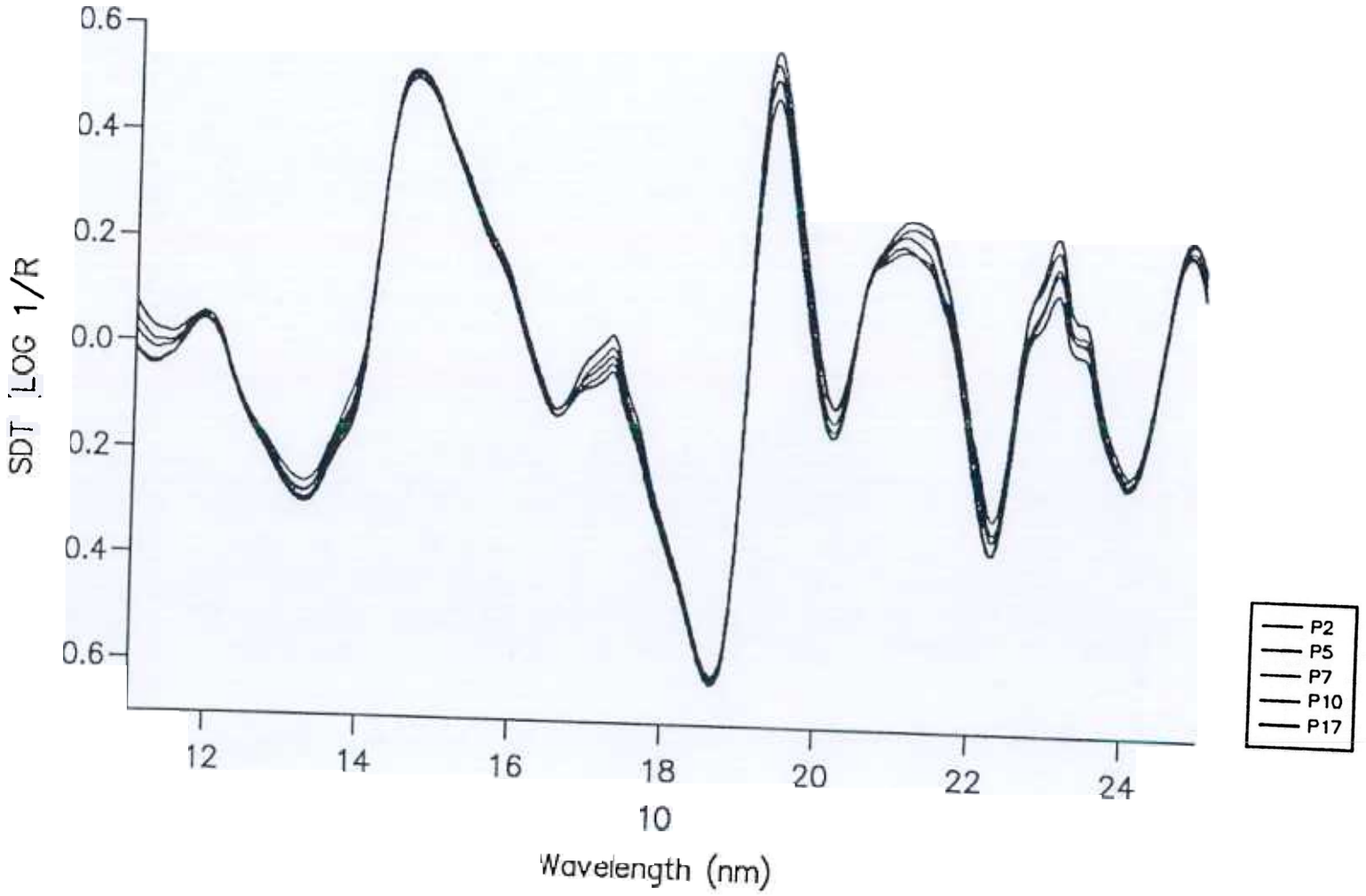
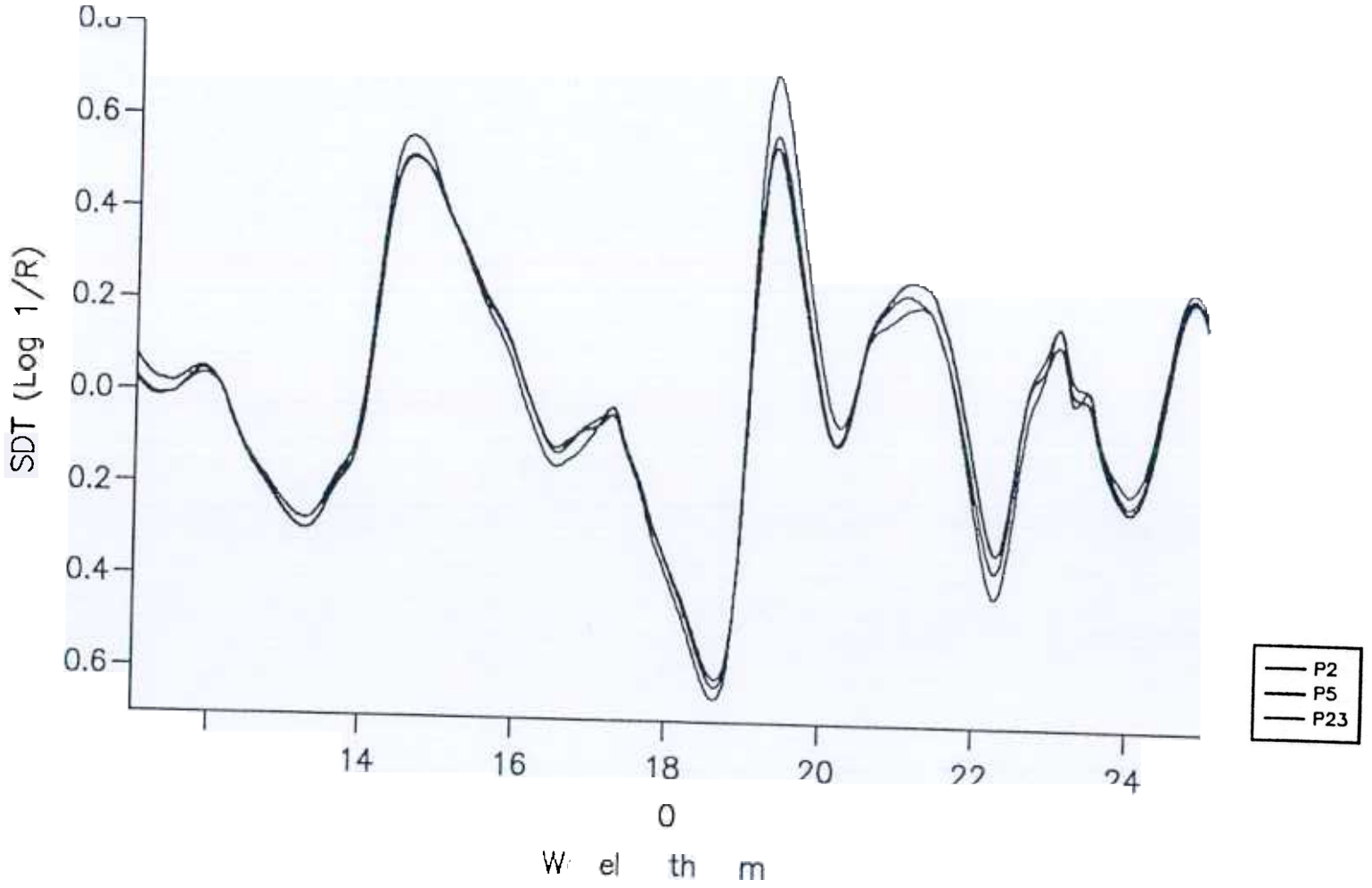


Figure 1.2.28 SDT Spectra for P2, P5, P7, P10 and P17



F ıul 2.29 SDT Sp c ra or P2 P5 and P23



2

NEPAL SAMPLES

2.1 WORK CARRIED OUT

Samples from various species of trees (ie. replicate a and b samples for species Al, Bh, Ci, Ct, Ds, Fg, Fn, Fr, Fs, Lp, Pc, Ql and Qs) and samples from different parts of trees (ie. with codes, b=bottom, i=intermediate, o=old, t=top and y=young) from Nepal were scanned in duplicate. The month of sampling was denoted by the last letter of the code, for example, Fgan and Fgad are samples from the Fg species for November and December, respectively. The samples were classified as follows for the statistical analyses:-

A. Monthly samples for November, December, January,February, March, April and May.

B. Samples from different parts of the trees for the Al and Qs species (ie. 1,2,3 samples) for the months of November, January and March.

C. March samples dried at 80°C.

All statistical analyses were performed as for the *Gliricidia* samples.

2.2 RESULTS

2.2.1 Statistical Analysis

A. Monthly Samples

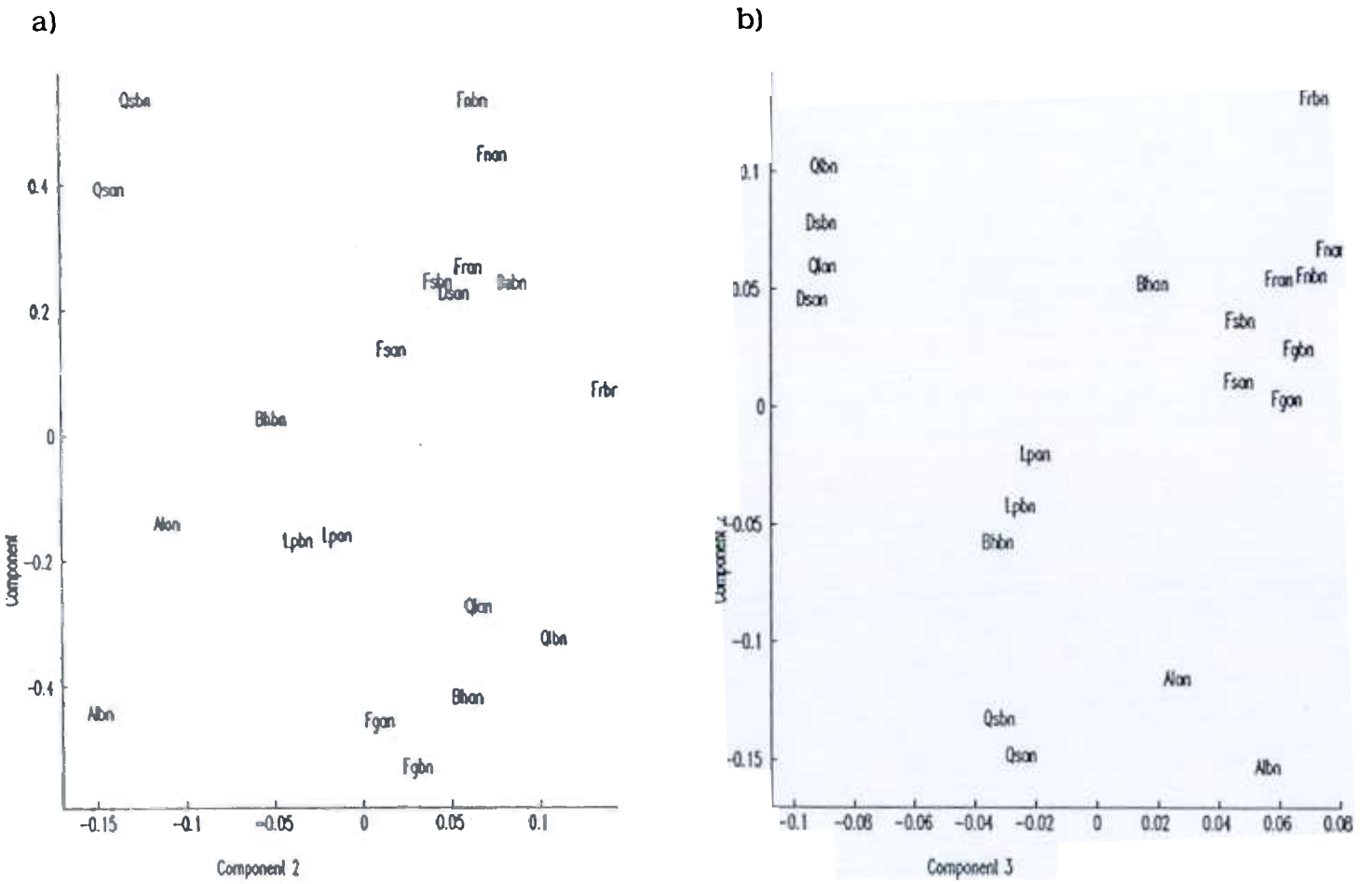
Discriminant analyses

The spectra from the Nepal samples were analysed to ascertain if they were similar to any of the *Gliricidia* samples. The *Gliricidia* samples were used as the library sample set and the Nepal samples were then treated as unknowns. All of the Nepal samples were most similar to provenance P5.

PCA

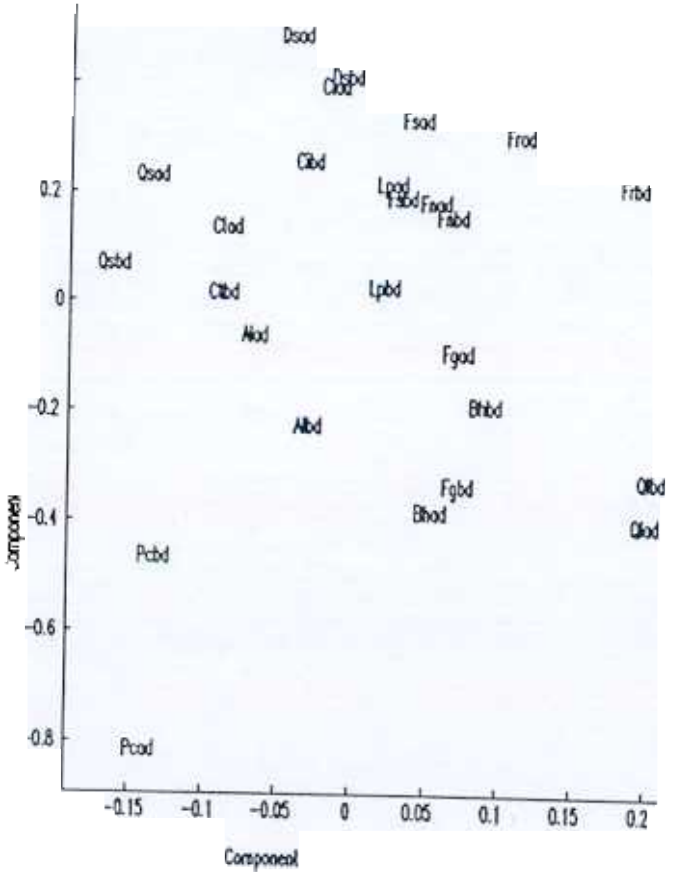
Graphing of pair-wise components from PCA allows us to look at the natural grouping of the samples. Figures 2.2.1, 2.2.2, 2.2.3, 2.2.4, 2.2.5, 2.2.6 and 2.2.7 a and b show Component 1 versus Component 2 and Component 2 versus Component 3 for the November, December, January, February, March, April and May samples, respectively. With the exception of Fgan and Fgbn on the Component 1 axis, the other F species from the November samples appear to cluster together (Figure 2.2.1a). The graph of Component 2 versus Component 3 (Figure 2.2.1b) shows all of the F species grouped together. This suggests that there was a physical difference between Fgan and Fgbn and the other species. This graph also shows the Ql and Ds species set apart from the other species on the Component 3 axis. The December samples (Figure 2.2.2a and b) included

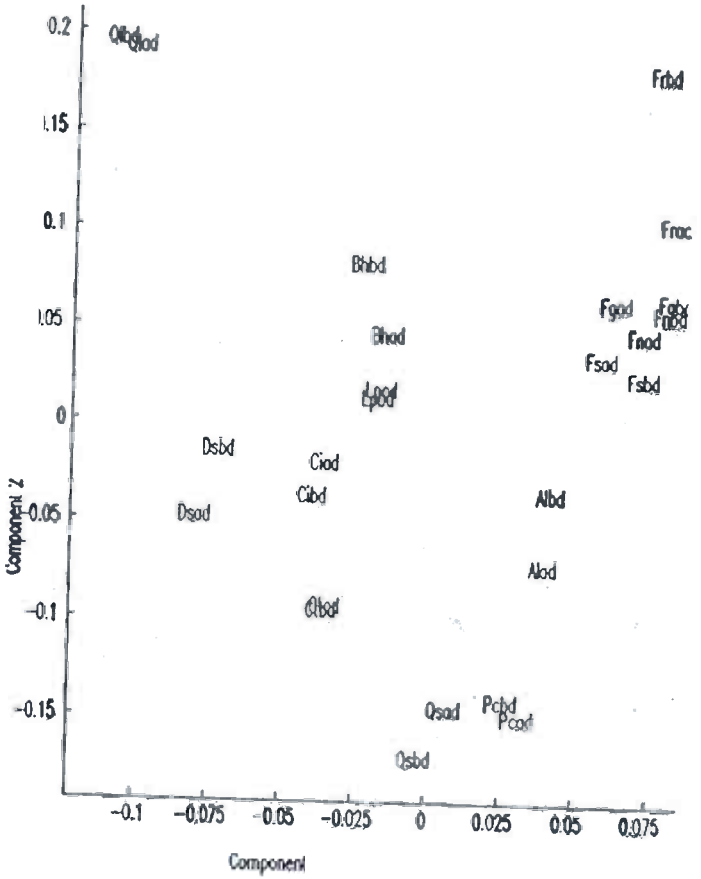
Figure 2.2.1 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the November Nepal samples.



gu Pr al Componen pl ts sh win Componen Compone Comp vers Compon 3,
fc th Dec mbe Nepal pl

b)





Figu Principal Componen plots showing Componen versu Compo en 2 an b) Componen versus Componen 3, f th an ary Nepal am es

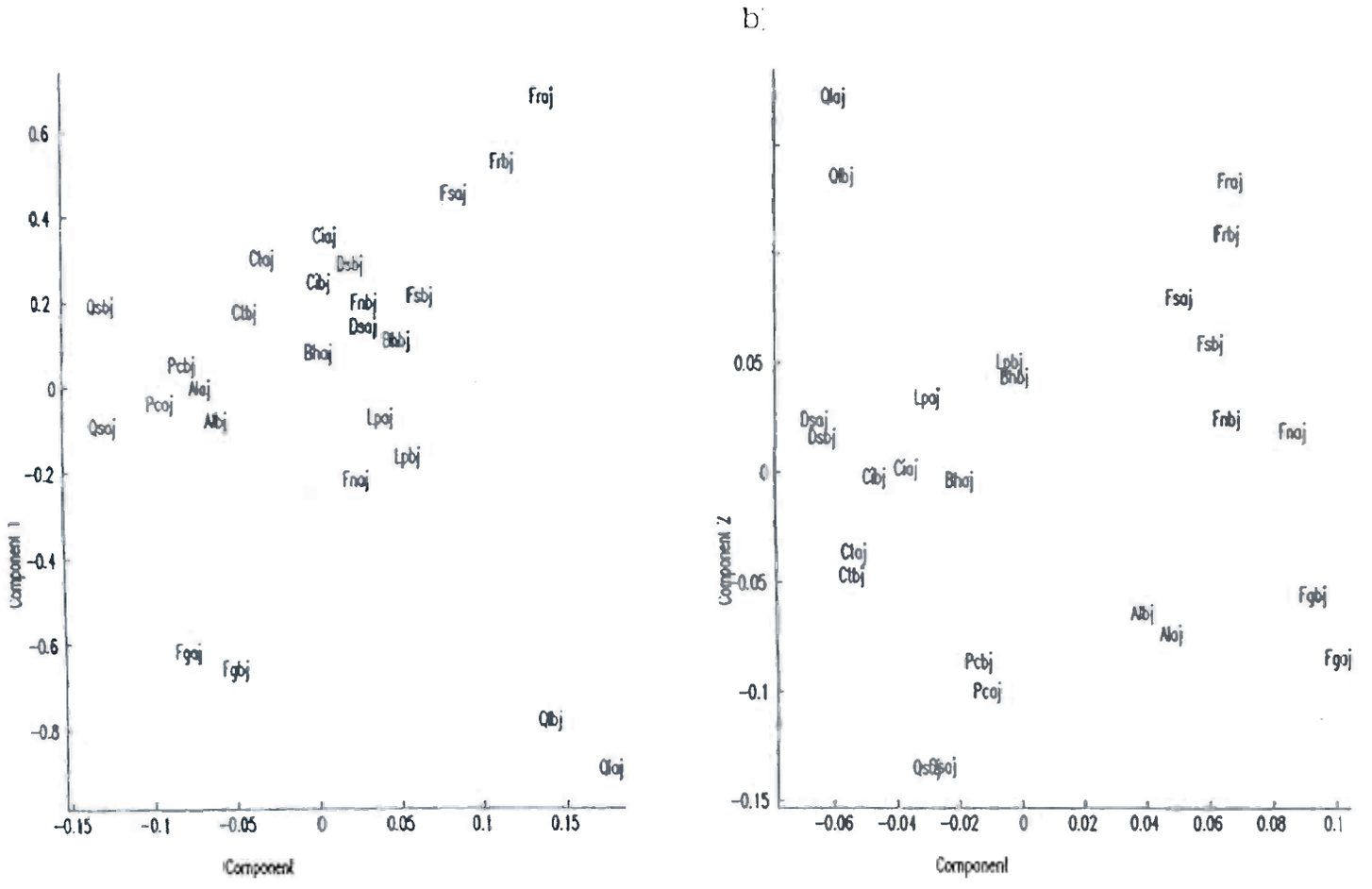


Figure 2.2.4 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the Febraury Nepal samples.

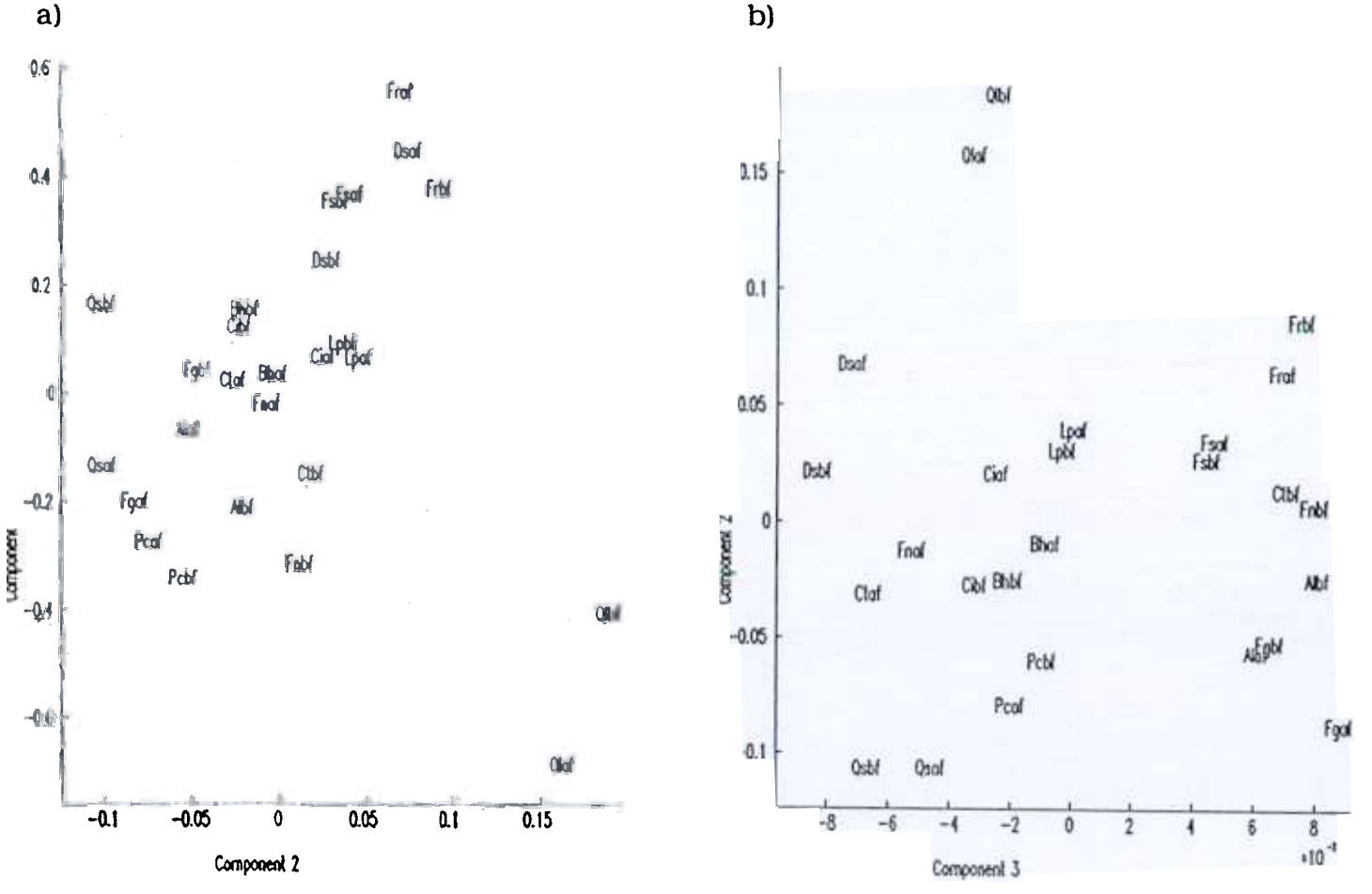
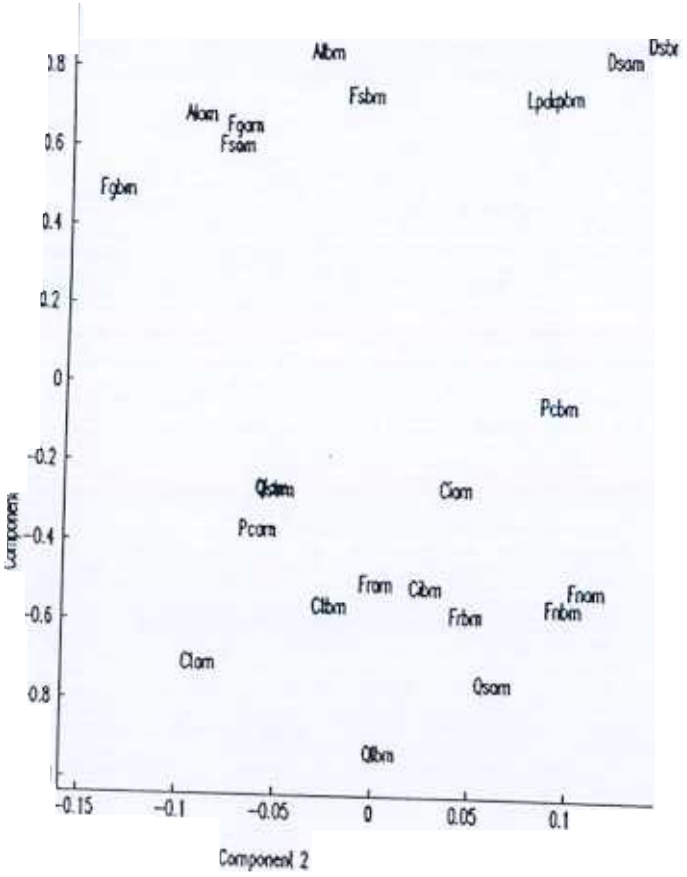
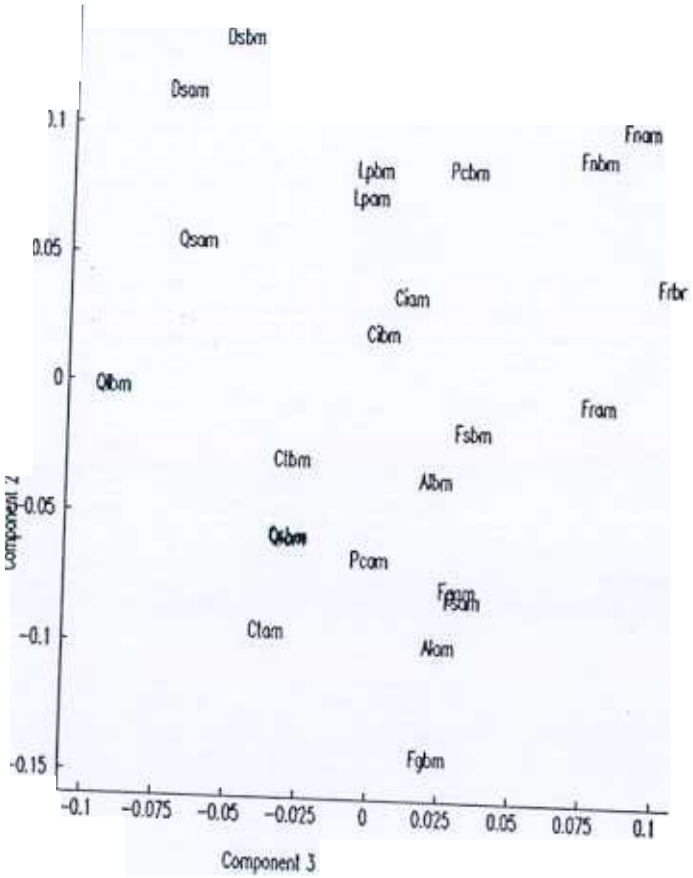


Figure 2.2.5 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the March Nepal samples.

a)

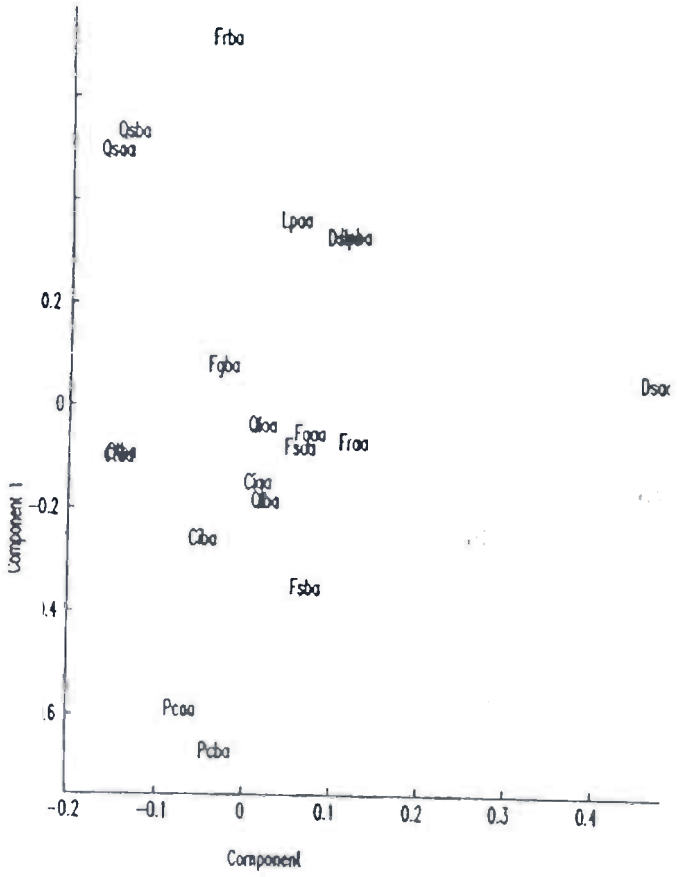


b)



F 'ri pal Compon pl ts showing Com Compon C mpo vers C po
th April Nepal ples.

b)



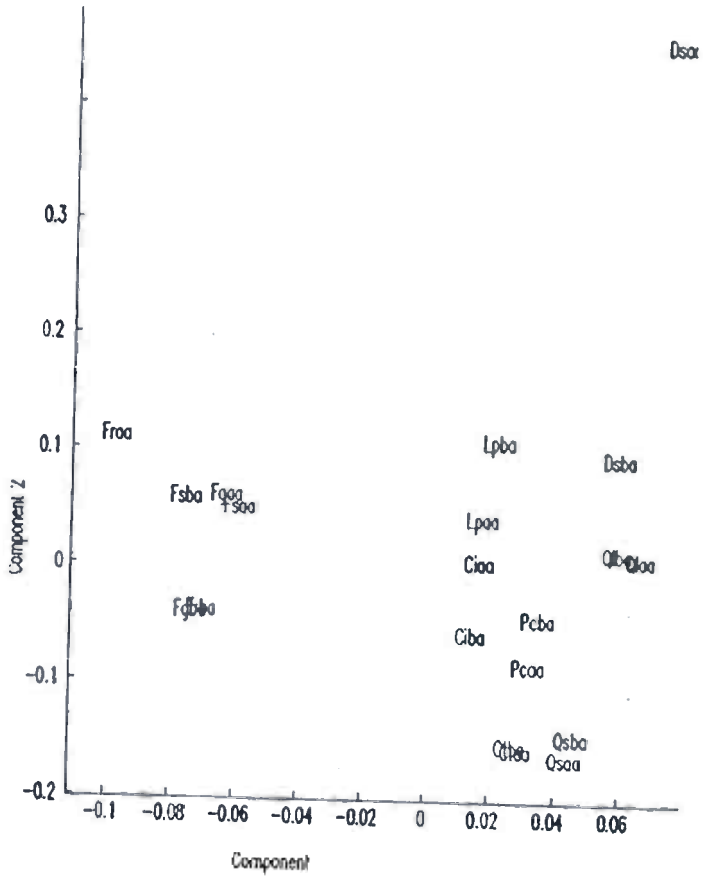
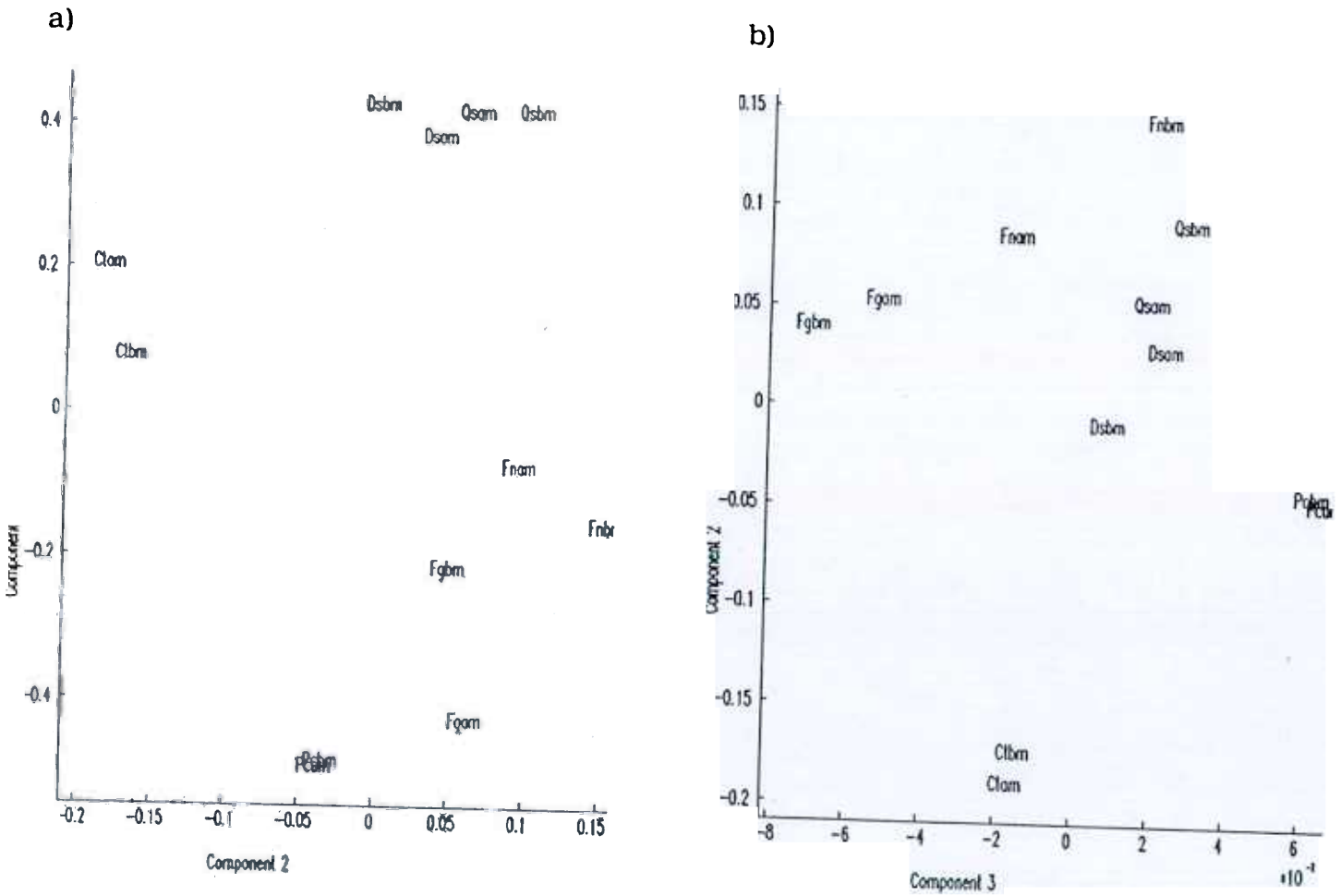


Figure 2.2.7 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the May Nepal samples.



more species than the previous month. As a result the distribution was slightly different, but the same general trends were found as with the November samples. Plotting Component 2 against Component 3, showed the F species clustering at the right hand side of the plot, Al, Bh, Qs, Al clustering in the middle and Ds and Ql clustering at the left hand side. Similar patterns were found for the January (Figure 2.2.3a and b) and February (Figure 2.2.4a and b) samples. However, in these months, Ql appears to be set apart from the other species. The F species in February and March (Figure 2.2.5a and b) did not appear in such a distinct group as in earlier months. The plot of Component 1 versus Component 2 for the March samples showed the Fg, Fs, Al, Lp and Ds species set apart from the other samples on the Component 1 axis. Figures 2.2.6a and b display the plots for the April samples, and show the F species grouped together at the right hand side of the plot and the Ds and Ql species at the left hand side. Only samples from 6 species were available for May, as shown in Figure 2.2.7a and b.

Biplot

Graphs indicating the distribution of the different species are given in Figure 2.2.8, 2.2.9, 2.2.10, 2.2.11, 2.2.12, 2.2.13 and 2.2.14 for November, December, January, February, March, April and May samples, respectively. The F and Al species are situated at the right hand side of Figure 2.2.8, and these species contain predominantly condensed tannins (12). Moving towards the right hand side species which contain a mixture of hydrolysable and condensed tannins can be found and at the left hand side, species which contain predominantly hydrolysable tannins and those which are relatively free from tannins. Plots for the other months showed similar distributions, with the exception of February and April which showed a mirror image of this distribution, with species containing condensed tannins at the left side and those being relatively free of tannins at the right side.

B. 1,2,3 SAMPLES

These are samples from different parts of trees of the same species. Samples were from two different species (Al and Qs) for three different months (November, January and March).

Figure 2.2.8 Distribution of Nepal samples from November.

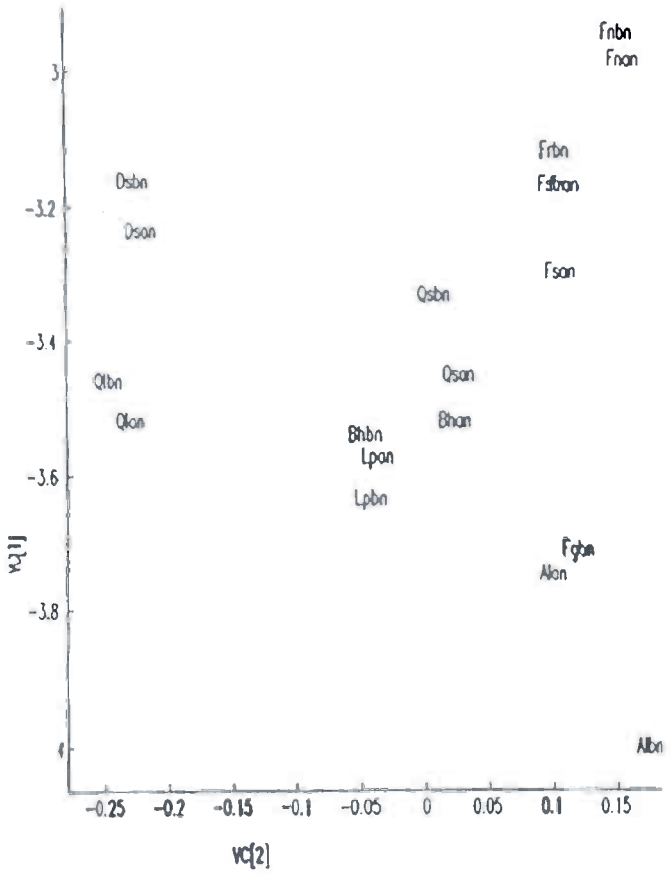


Figure 2.2.9 Distribution for the Nepal samples from December.

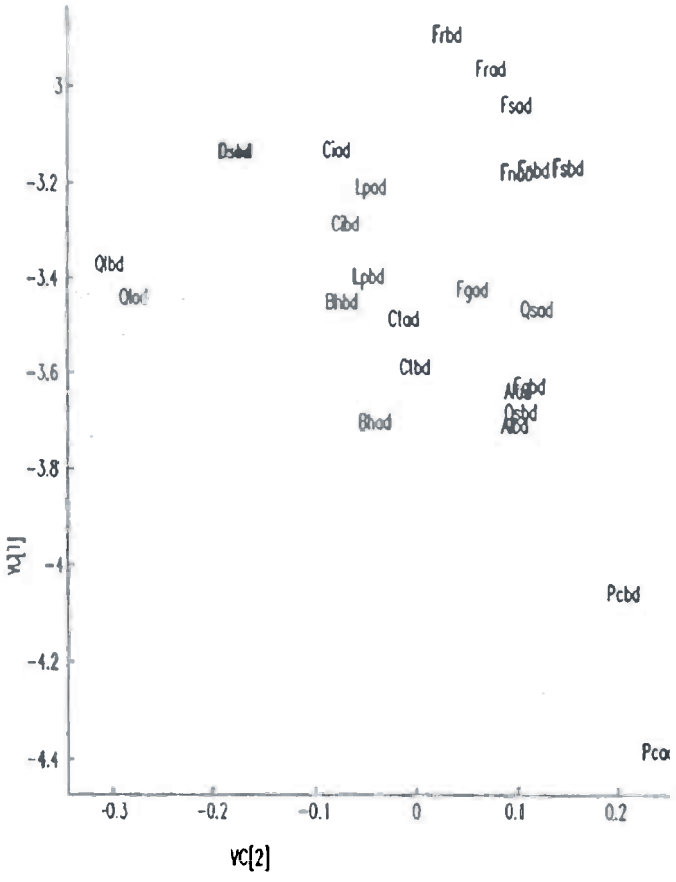


Figure 2.2.10 Distribution of Nepal samples from January.

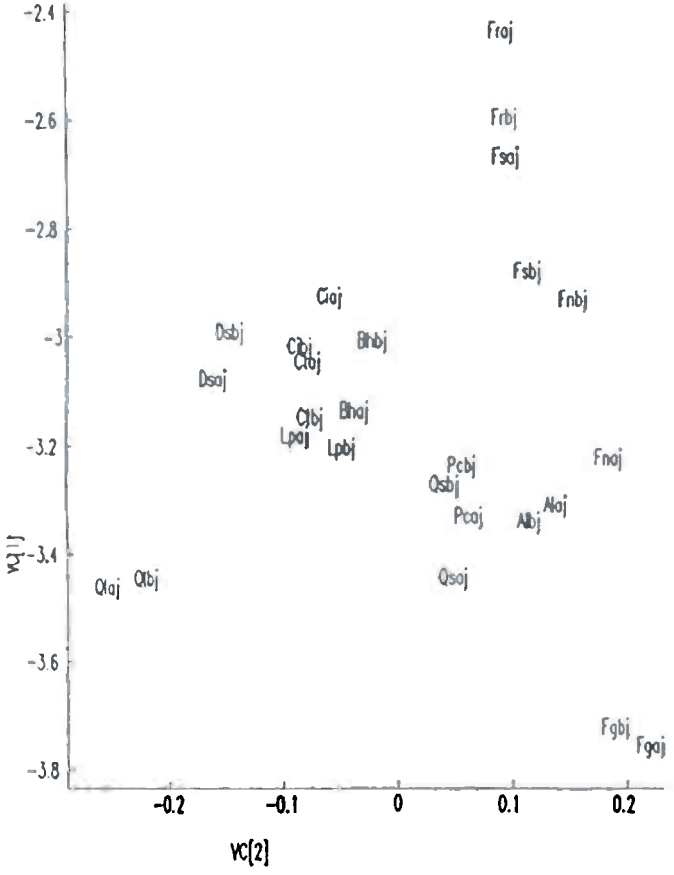
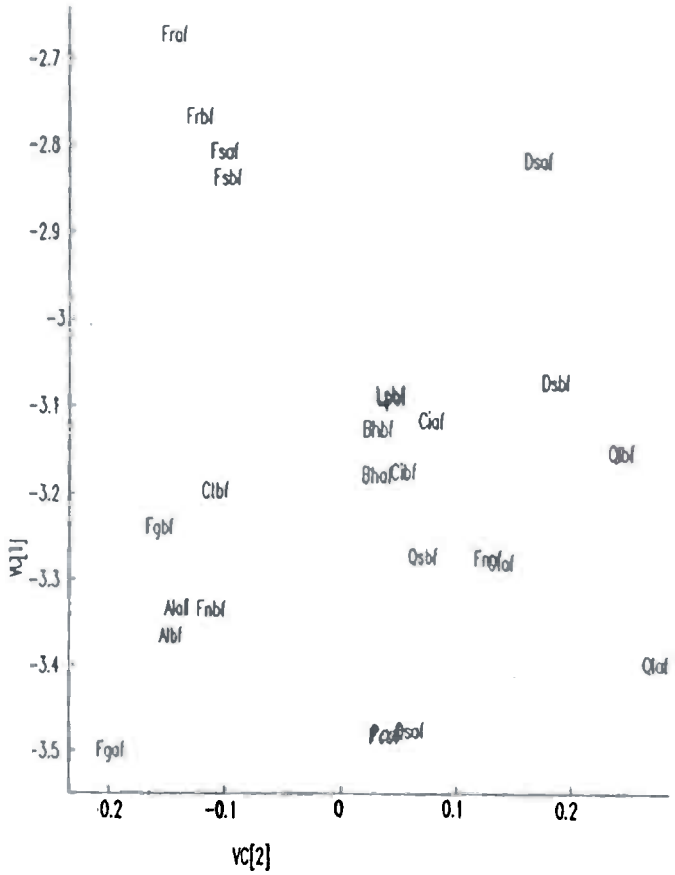
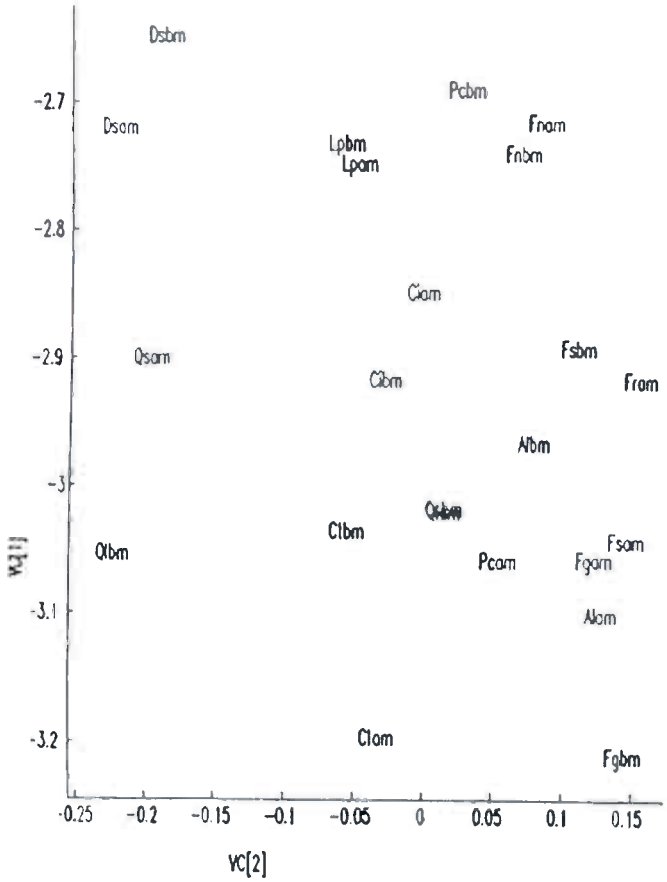


Figure 2.2.11 Distribution for the Nepal samples from February.



Figu D tribu io of Nep ples
Mar



Figur 3 Dis tbu io fo :h Nepal ples
fr April

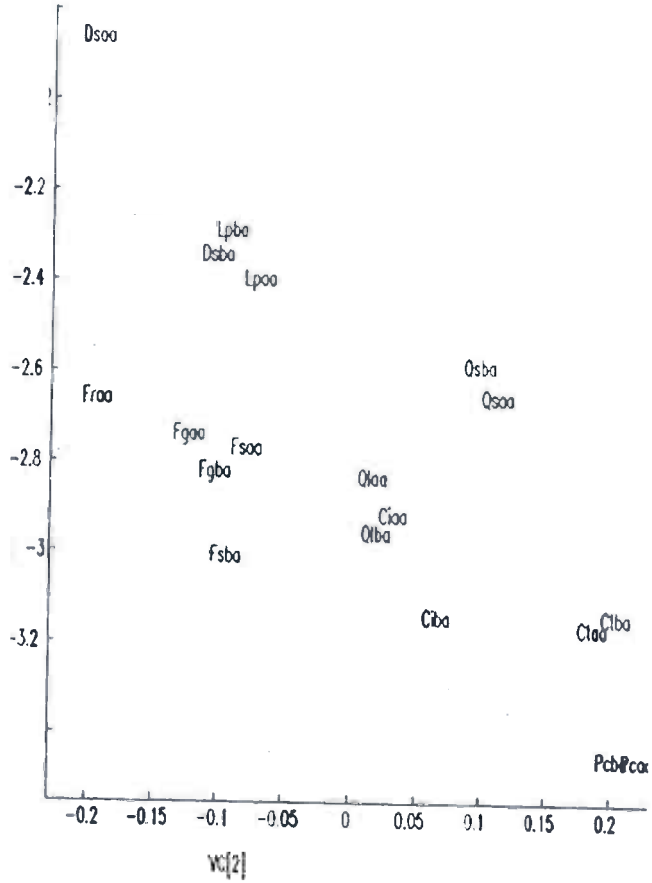
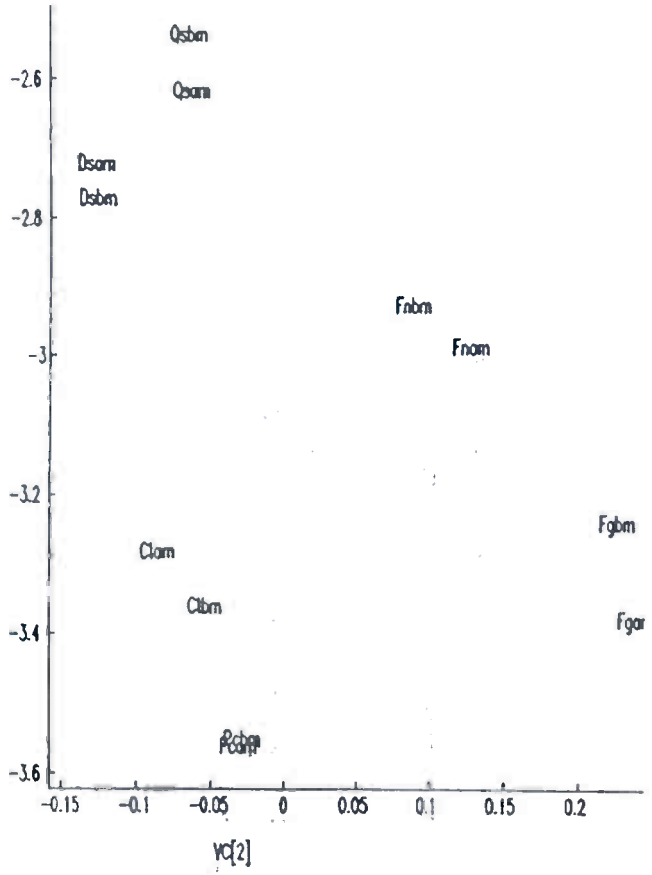


Figure 2.2.14 Distribution of Nepal samples from May.



PCA

PCA plots for Al species are given in Figures 2.2.15, 2.2.16 and 2.2.17 for November, January and March, respectively. There appeared to be some clustering of the 1's, 2's and 3's. Similar patterns were found for several, but not all of the component plots for the Qs samples (Figures 2.2.18, 2.2.19 and 2.2.20). In some of the these graphs, in particular Figures 2.2.15, 2.2.16 and 2.2.19, a trend was apparent showing differences according to the age of the tree. Samples from young trees were located at one side of the graph, moving across to samples from intermediate trees and then samples from older trees at the other side.

Biplot Analysis

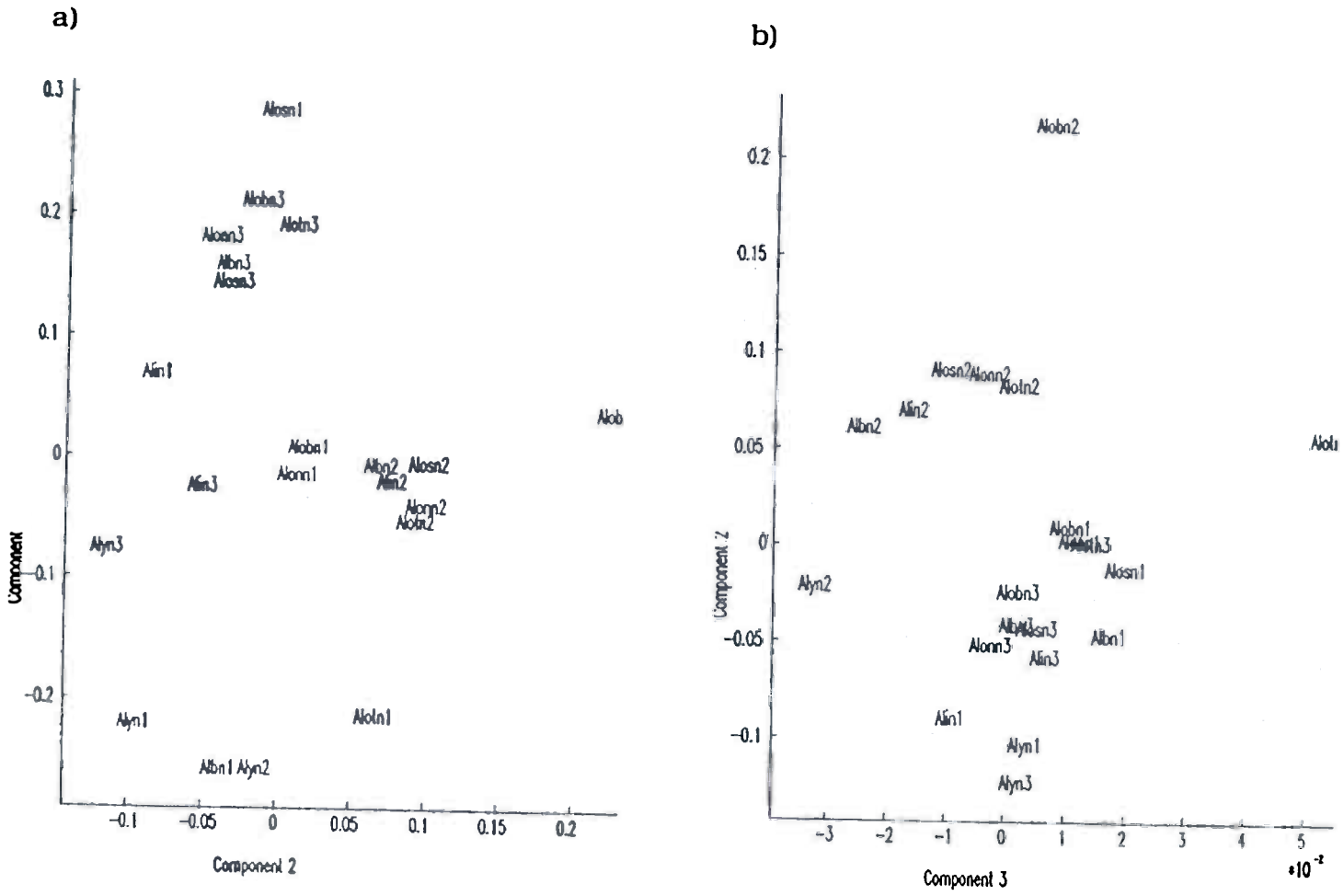
Figure 2.2.21 shows the distributions of the November Al samples. The 1, 2 and 3 samples did not cluster separately as seen with PCA, but there was some indication of the trend found relating to the age of the trees. Samples from young trees can be seen in the bottom right corner, moving diagonally across to the intermediate samples and then the old samples in the top left corner. The January samples for the same species are shown in Figure 2.2.22 and follow a similar trend to the November samples, however the March samples (Figure 2.2.23) do not, they in fact show a clustering of the 1's, 2's and 3's. The November Qs samples (Figure 2.2.24) showed a trend where the young tree samples were clustered together at the top of the diagram, with intermediate samples in the middle and samples from older trees at the bottom. Some clustering is apparent for the January and March samples in Figures 2.2.25 and 2.2.26, respectively.

C. Samples Dried at 80°C

Some of the Nepal samples for the month of March were dried at 80°C, but samples of the same material dried at 60°C were also received. This allowed an investigation of differences between samples dried at the two temperatures. Other work (12) has suggested that there is an effect of drying temperature on protein precipitation assay measurements. Samples dried at 80°C gave lower precipitation activity values than the corresponding samples dried at 60°C. NIR spectra from the two sets of samples, namely Ds, Fg, Fr, Fs and Lp species and samples from different parts of Qs trees, were obtained, and analysed using PCA and Biplot analysis.

Figure 2.2.27 shows distinct differences between the samples dried at 60°C and the

Figure 2.2.15 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the Al species (November).



Figu Principal Componen pl ts howin Compo Co po and b) Co po Compo en for th Al speci (January).

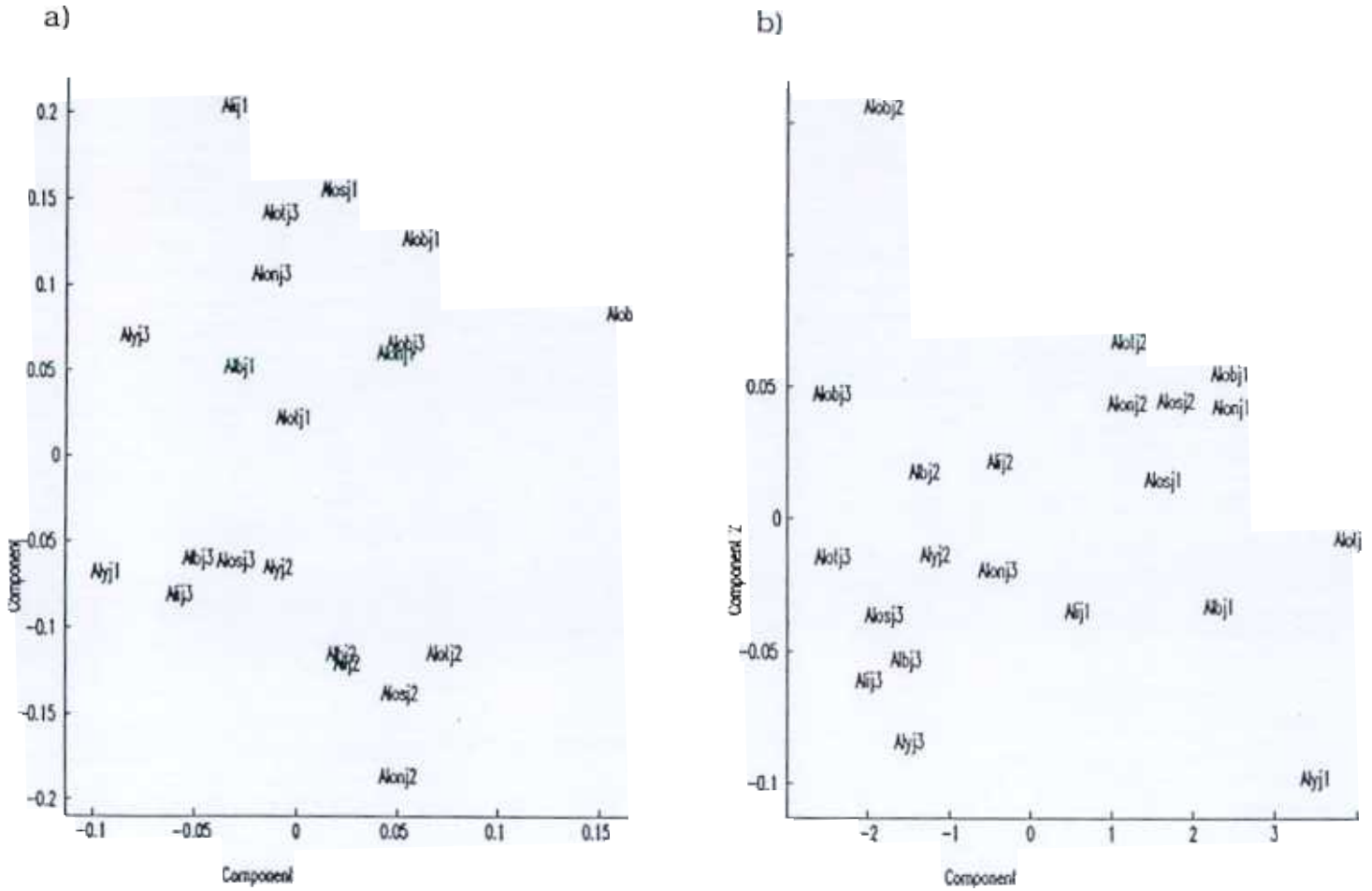


Figure 2.2.17 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the Al species (March).

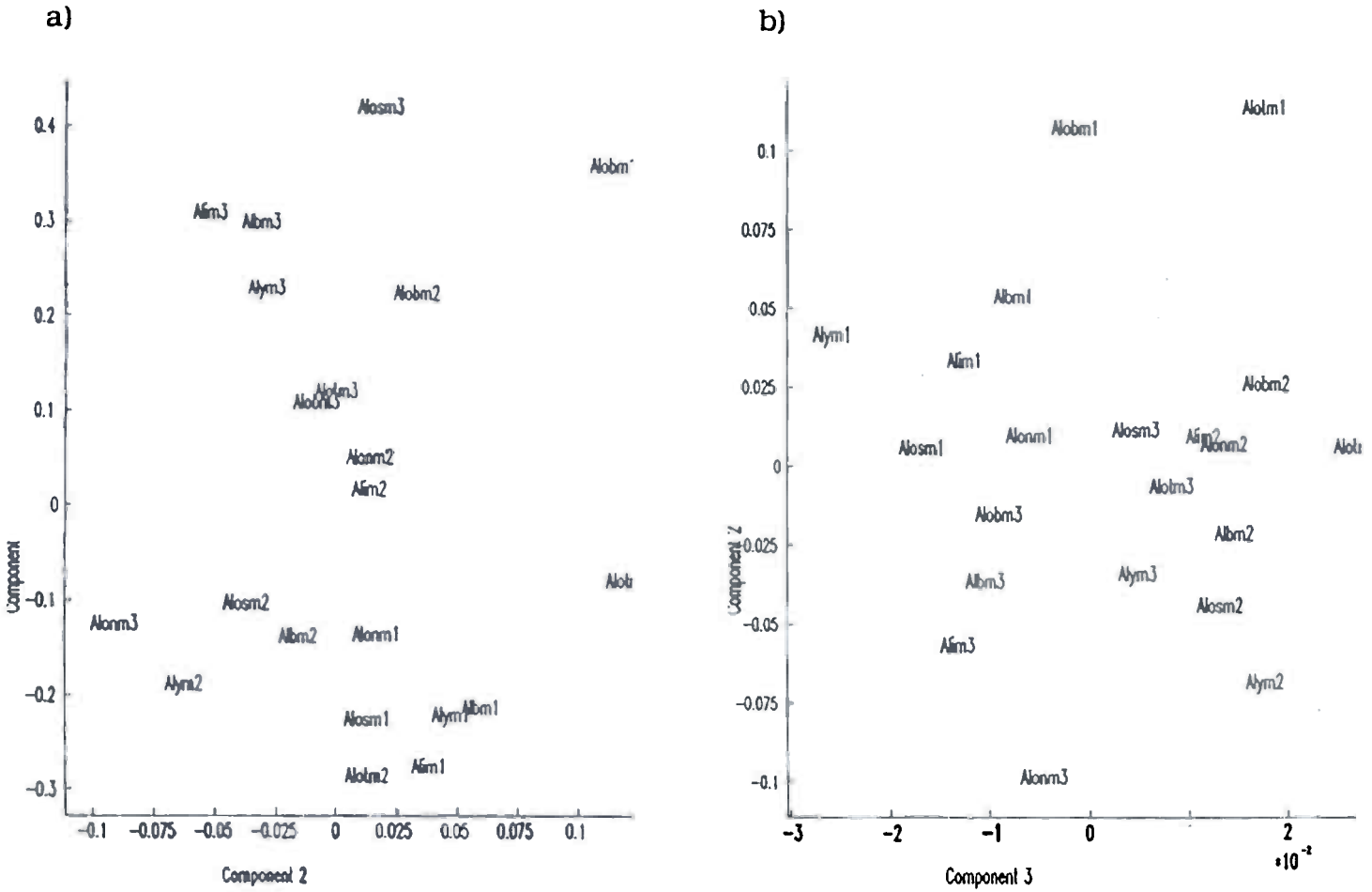
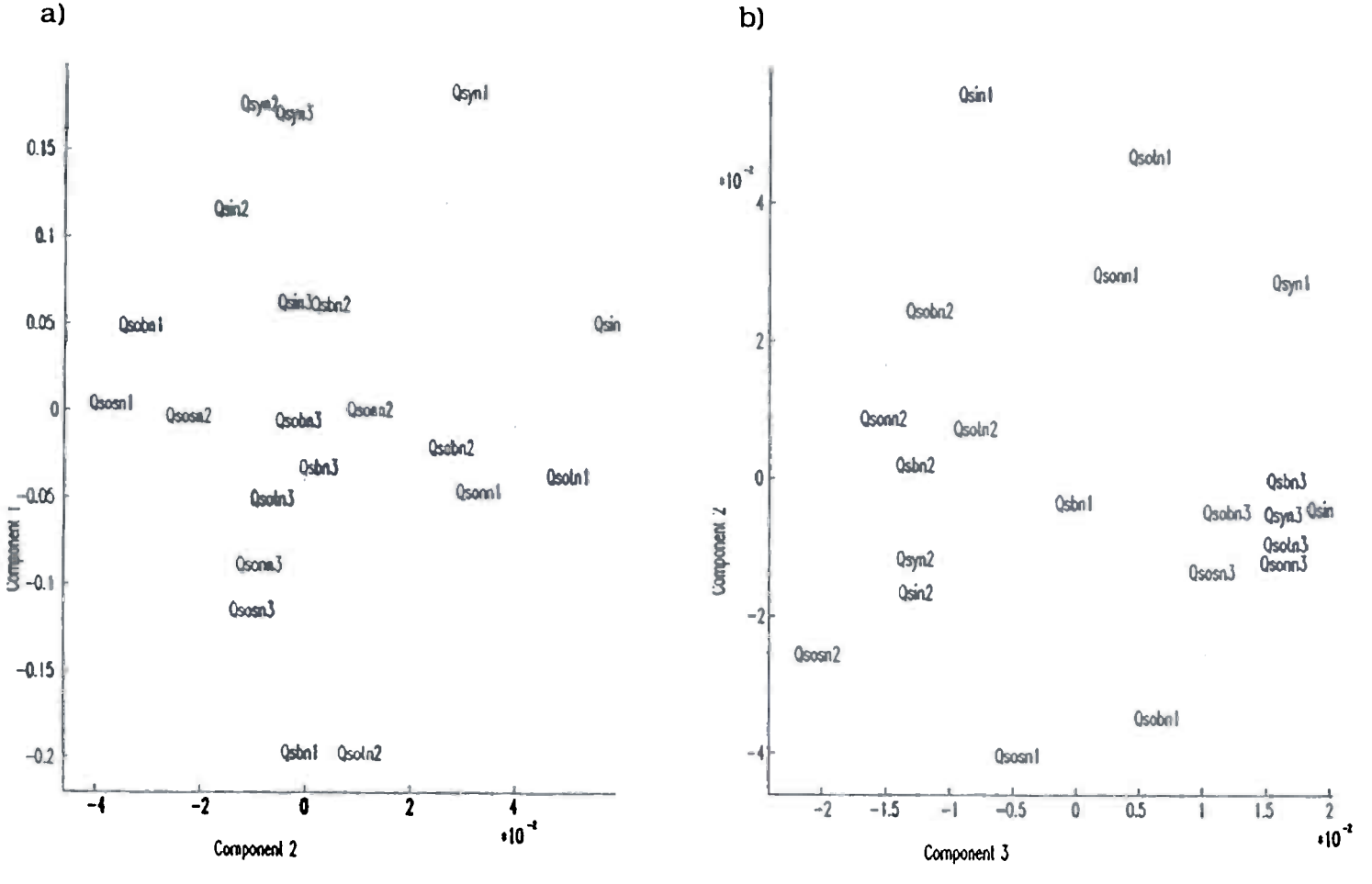


Figure 2.2.18 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the Qs species (November).



Figur rincipal C pc nen plots showin Co C po p
Compo en fo th Qs speci ary

b)

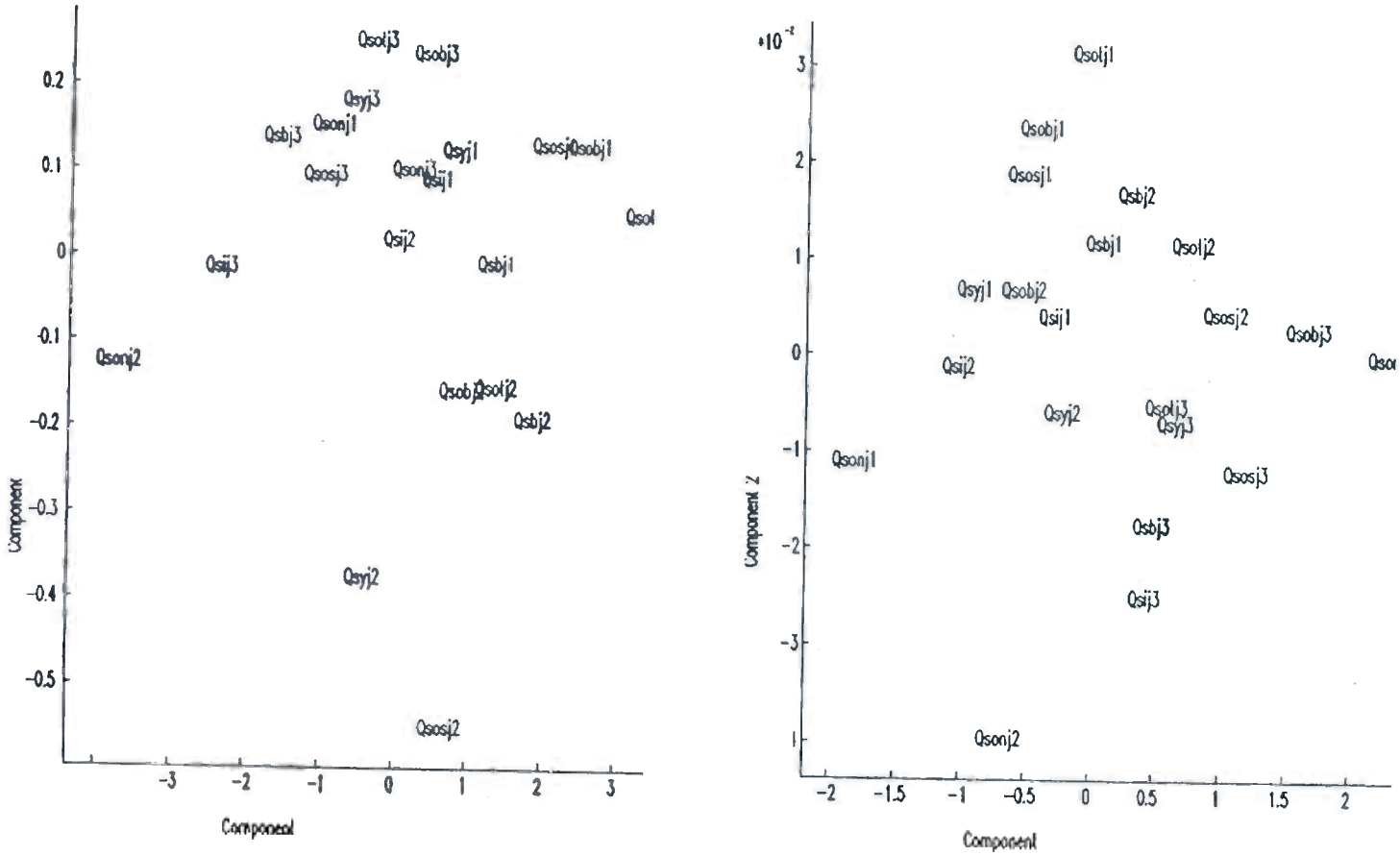


Figure 2.2.20 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the Qs species (March).

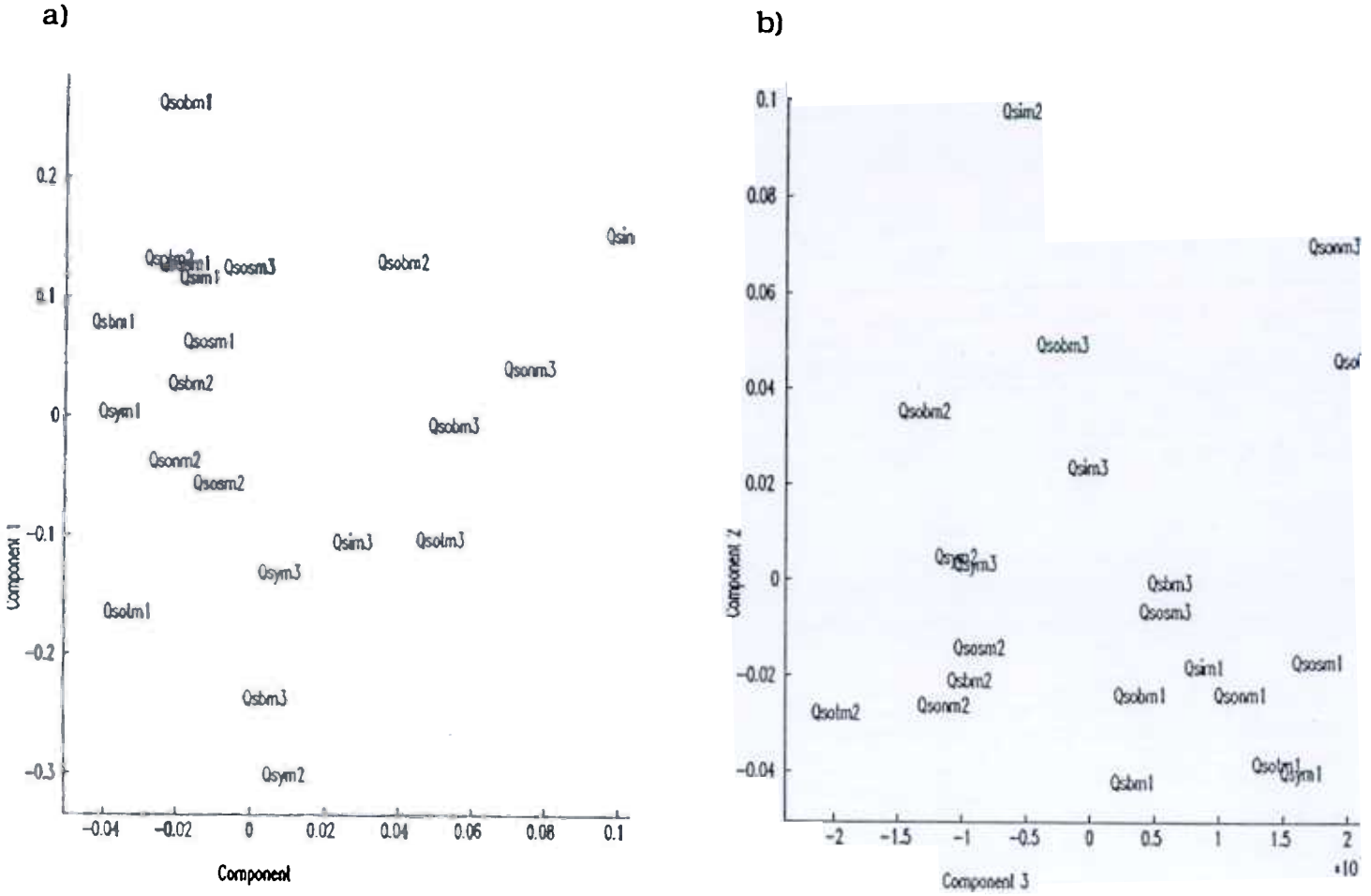


Figure 2.2.2 Distribution of Al species from November.

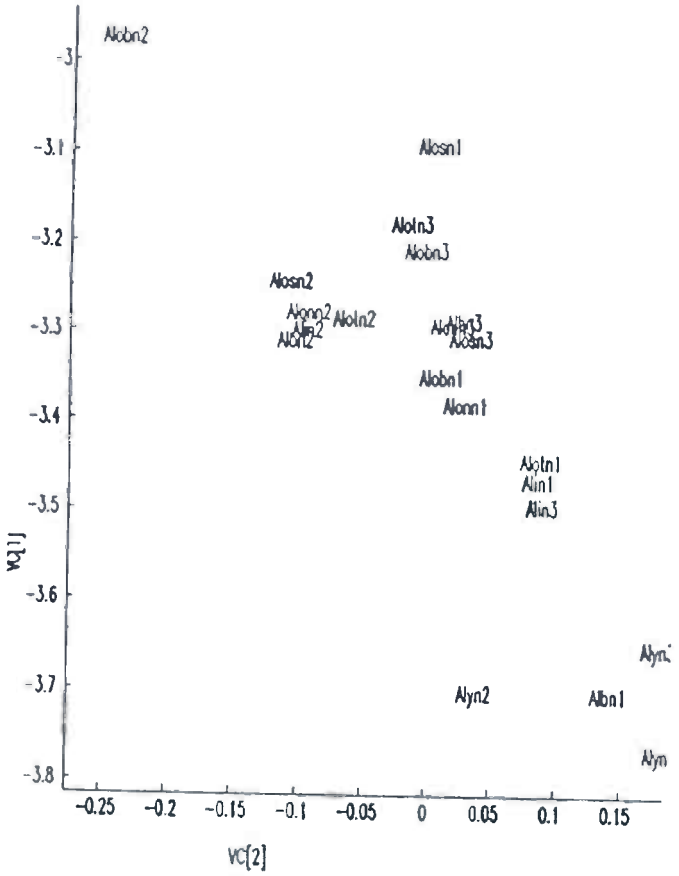
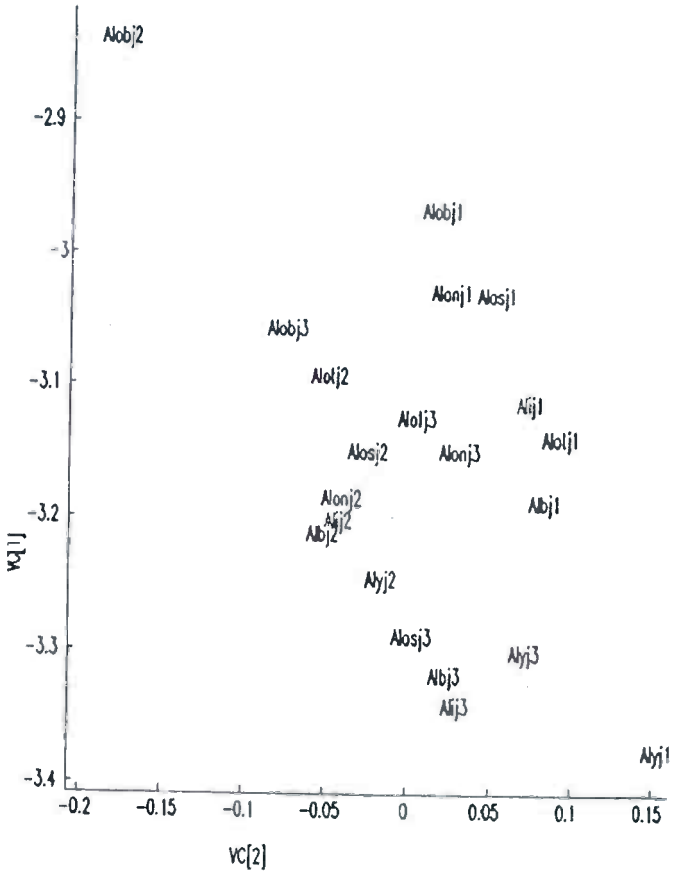
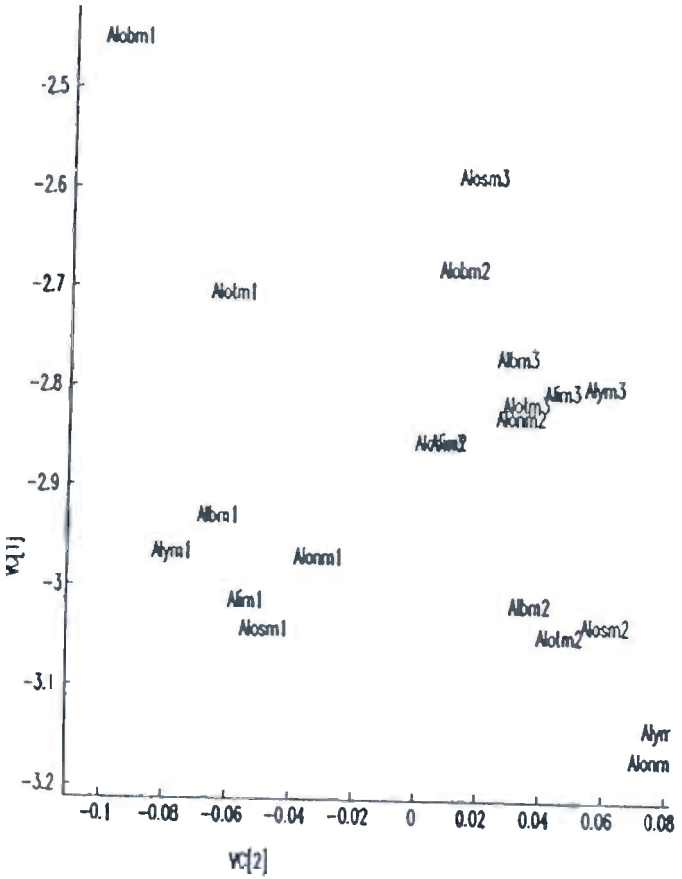


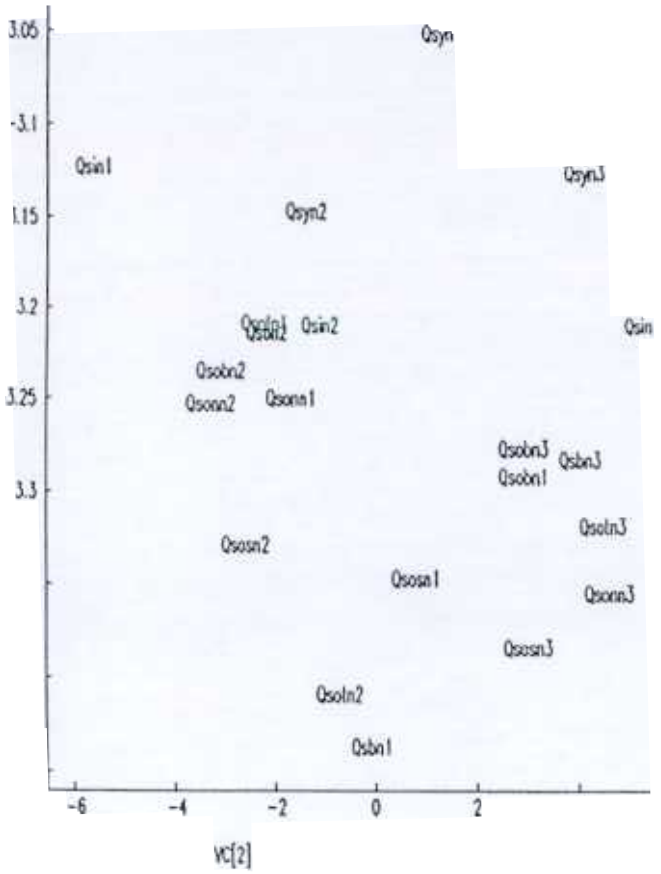
Figure 2.2.22 Distribution Al species from January.



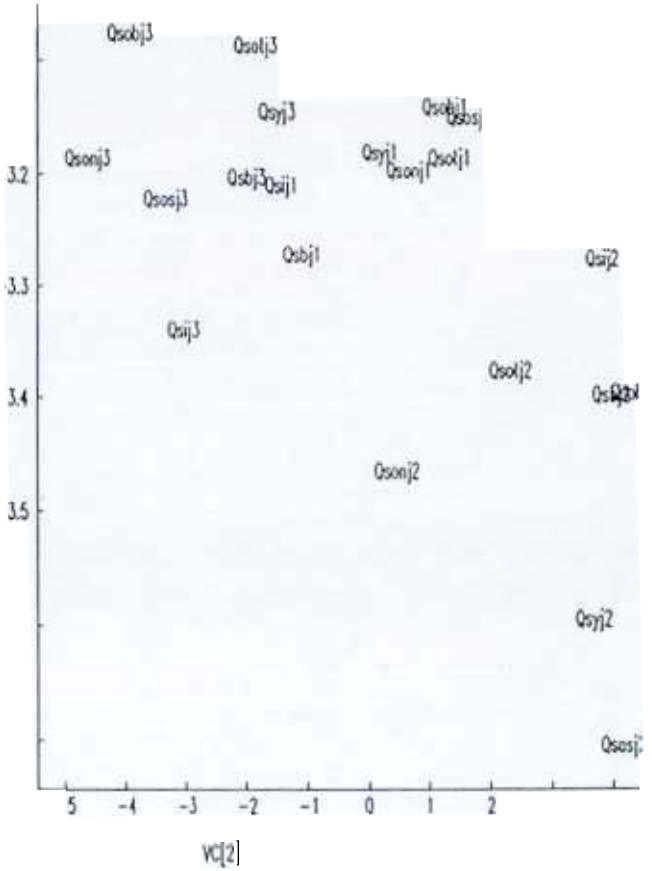
Figur .2 23 Distribution of Al specI
fr March



Figu 2 .24 Distribution of Q species from November



Figu 2.2. 5 Distrib tl Qs pec es from J uary



Figu .26 Distribution of Qs species from March

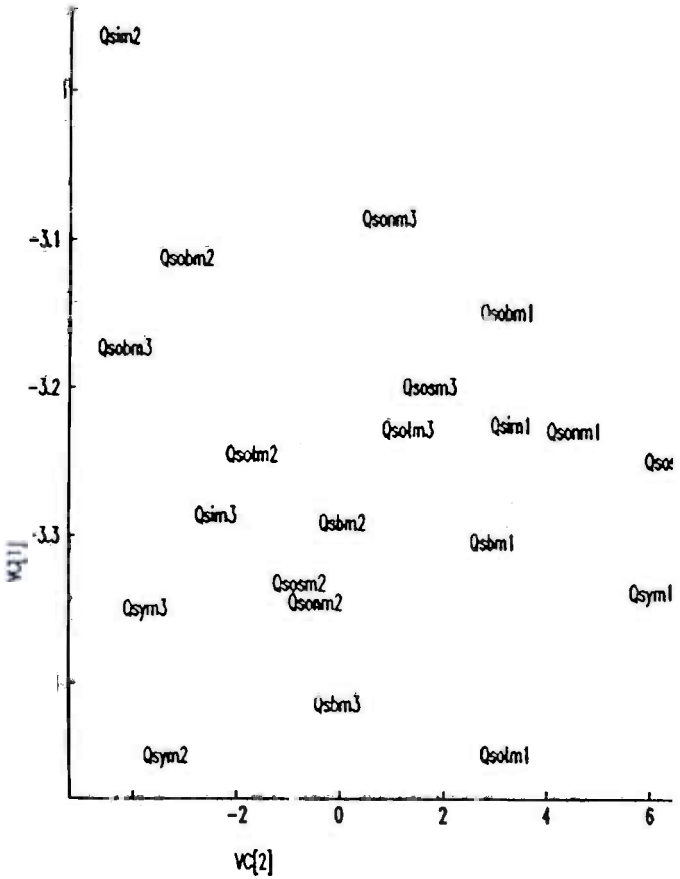
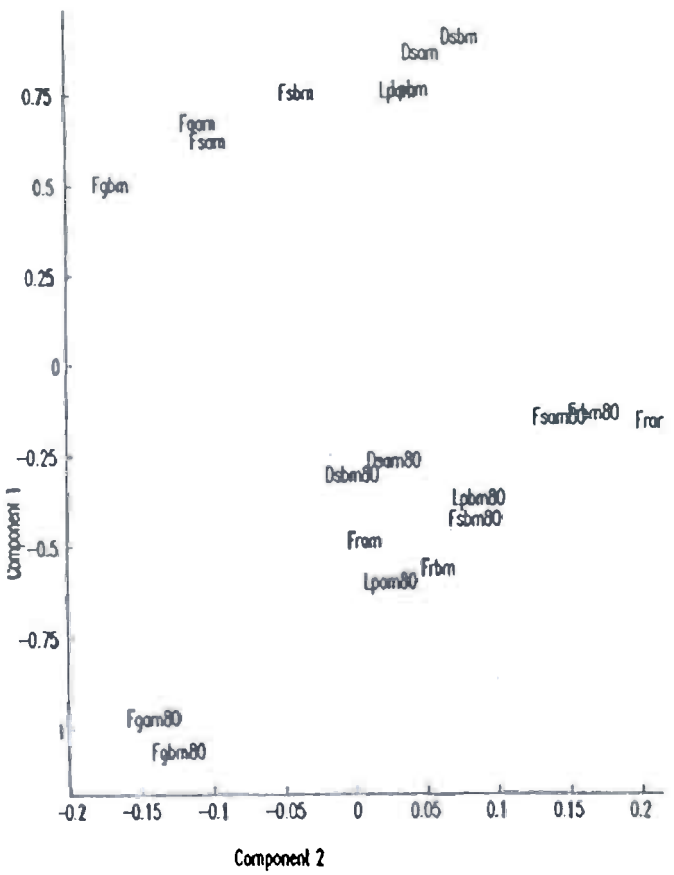
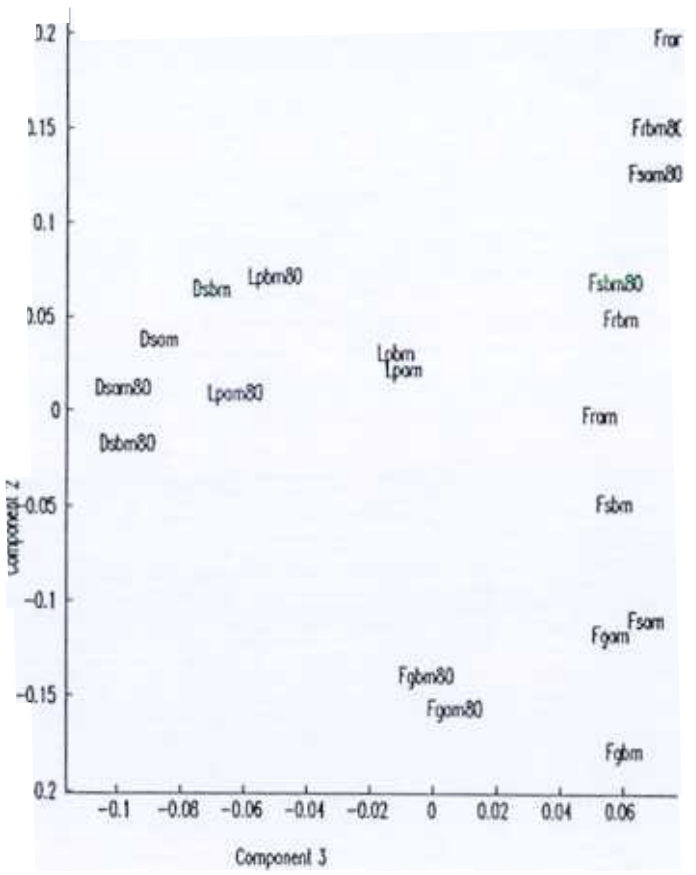


Figure 2.2.27 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the March Nepal samples dried at 60°C and 80°C.

a)



b)



corresponding ones dried at 80°C in Component 1, with the exception of the Fr species. This may be due to a difference in colour between the samples dried at the different temperatures. When Component 2 was plotted against Component 3, the samples dried at the different temperatures were found to sit closer together.
Biplot analysis (Figure 2.2.28) shows a very similar pattern to the latter PCA plot.

PCA plots for the different Qs samples are given in Figure 2.2.29. These plots show the majority of the samples dried at 80°C clustering together, but no distinct split between these and the samples dried at 60°C.
A similar pattern was obtained from the Biplot analysis of the Qs species (Figure 2.2.30).

2.2.2 Calibration Development

Calibrations were developed using a Partial Least Squares (PLS) regression technique for each individual month and then the samples were grouped together and an overall calibration was developed for the whole set. Calibration and cross validation statistics are shown in Tables 2.2.1, 2.2.2, 2.2.3, 2.2.4 and 2.2.5 for moisture, ash, crude protein (CP), protein precipitate assay (PrA) and acid butanol assay (Abut) (12), respectively.
The equations developed for each month were based on between 15 and 26 samples. This is not an ideal number of samples on which to base a calibration but the equations do give some indication of what can be achieved. Equations developed for moisture content are given in Table 2.2.1. The best calibration and cross validation statistics were achieved with the January samples and the poorest with the February samples. The statistics for the ash equations (Table 2.2.2) were similar for all months with the exception of the February equation which gave a somewhat larger SEC than the other months. Table 2.2.3 shows the statistics for the CP equations, all of which were acceptable, but again, the equation for the February equation tended to show poorer statistics. The statistics for the PrA equations (Table 2.2.4) were best for the November samples, but this equation was developed using only 15 samples and therefore this analysis should be interpreted with caution. The Abut analysis was only available on the January samples, the statistics are given in Table 2.2.5.

All of the equations developed should not be just accepted as being accurate but have to be validated properly with samples not used in the calibration set. Cross validation

Figure 2.2.28 Principal Component plots showing a) Component 1 versus Component 2 and b) Component 2 versus Component 3, for the Qs species (March) dried at 60°C and 80°C.

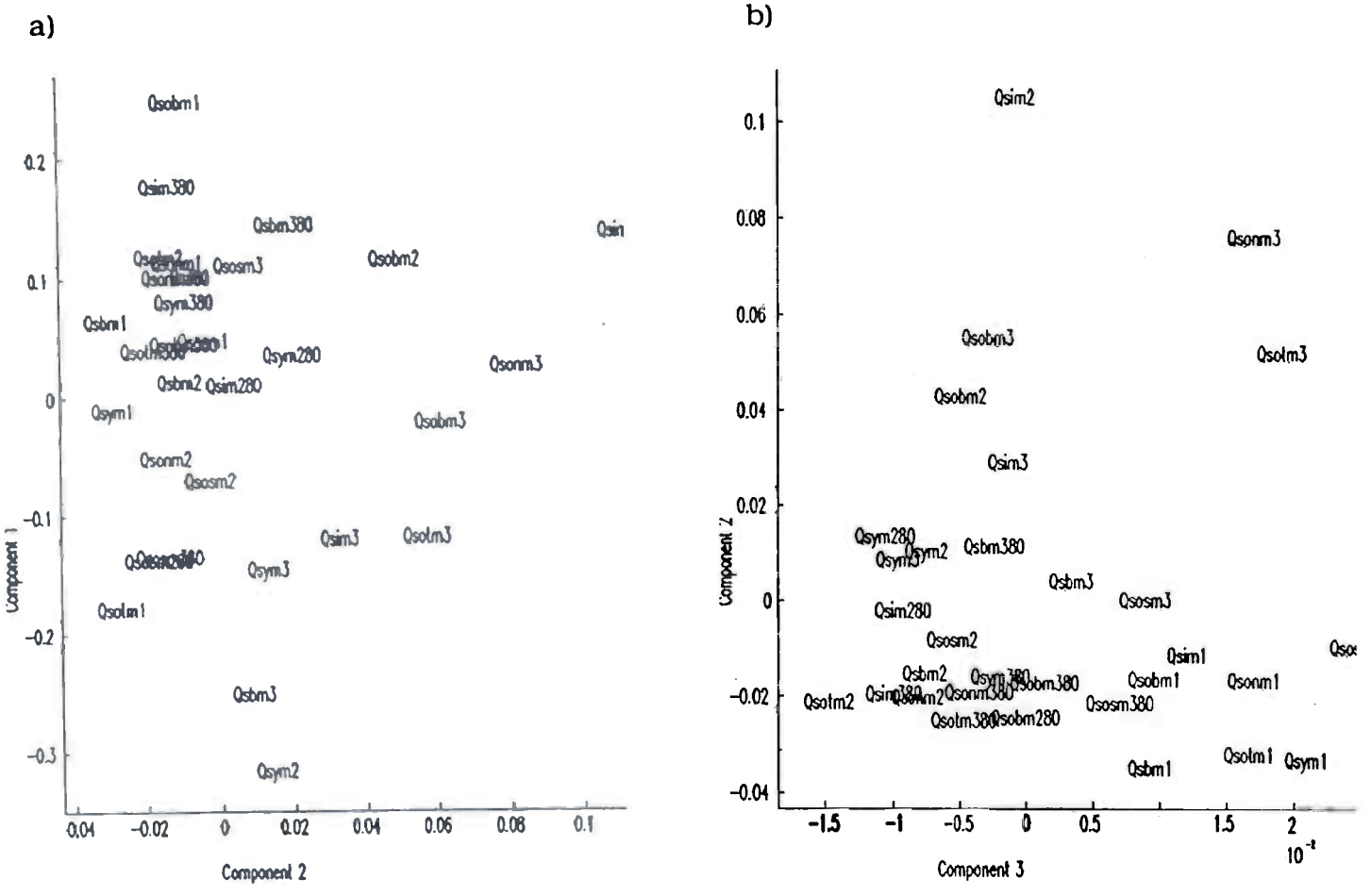


Figure 2.2.29 Distribution of Nepal samples dried at 60°C and 80°C.

Figure 2.2.30 Distribution of Qs species dried at 60°C and 80°C.

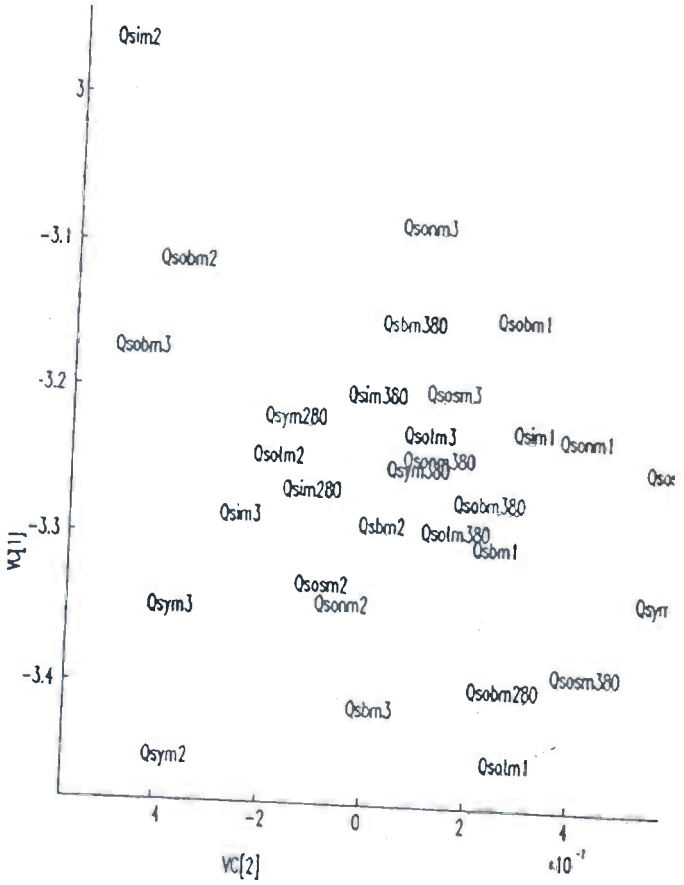


Table 2.2.1 Calibration and Cross Validation Statistics for Moisture

Equation	n	SEC	R^2	SECV	1-VR
November	20	0.25	0.92	0.44	0.75
December	25	0.32	0.75	0.37	0.68
January	26	0.16	0.94	0.24	0.86
February	25	0.64	0.88	0.95	0.72
March	23	0.45	0.84	0.60	0.70
April	17	0.21	0.94	0.43	0.75
All	145	0.73	0.60	0.80	0.54

Table 2.2.2 Calibration and Cross Validation Statistics for Ash

Equation	n	SEC	R^2	SECV	1-VR
November	19	0.51	0.99	1.18	0.94
December	26	0.73	0.98	1.29	0.94
January	25	0.69	0.98	1.02	0.96
February	26	1.31	0.94	1.63	0.92
March	24	1.09	0.96	1.49	0.93
April	18	0.30	1.00	0.49	0.99
All	142	0.77	0.98	0.87	0.97

Table 2.2.3 Calibration and Cross Validation Statistics for Crude Protein

Equation	n	SEC	R^2	SECV	1-VR
November	19	0.74	0.89	1.06	0.70
December	24	0.42	0.97	0.64	0.93
January	26	0.51	0.94	0.98	0.76
February	25	0.84	0.82	1.11	0.69
March	22	0.89	0.86	1.12	0.78
April	20	0.54	0.98	0.94	0.94
All	145	0.64	0.94	0.76	0.92

Table 2.2.4 Calibration and Cross Validation Statistics for PrA

Equation	n	SEC	R^2	SECV	1-VR
November	15	11.95	0.97	19.90	0.92
December	24	39.03	0.84	65.15	0.56
January	24	31.24	0.82	40.98	0.69
February	24	60.21	0.57	79.80	0.33
March	20	27.79	0.90	43.31	0.76
April	18	26.82	0.92	47.64	0.75
All	137	49.20	0.74	58.23	0.63

Table 2.2.5 Calibration and Cross Validation Statistics for Abut

Equation	n	SEC	R^2	SECV	1-VR
January	24	6.49	0.89	12.05	0.66

statistics give an indication of how each of these equations will perform given that future samples are similar to those used the calibration set. Each of the equations for the individual months were used to predict the samples from the other months. The prediction statistics (Bias, standard error of prediction(corrected for bias) SEP(C) and R^2) are given in Tables 2.2.6, 2.2.7, 2.2.8 and 2.2.9. Prediction statistics were poor and large biases were found between months for certain constituents, in particular PrA, but also moisture and CP in some cases. These results indicate that the calibrations for some months do not cover the necessary range in constituents for other months. The January equation appeared to predict parameters for the other months most successfully. The February samples appear to be most different from the other months.

The protein precipitate assay is not a well defined chemical technique and we do not know exactly what it is measuring. The extraction procedure is unlikely to remove all of the phenolic compounds/tannins. Thus, it is possible that the precipitate analysis and the NIR are measuring different entities. This has to be taken into account when developing equations for these constituents.

Table 2.2.6 Prediction Statistics for December, January and February using the November Equation

Predicting		BIAS	SEP(C)	R^2
December	DM	-0.41	0.67	0.45
	ASH	-0.35	1.33	0.94
	CP	0.37	1.86	0.40
	PrA	2.92	86.55	0.21
January	DM	-0.06	0.61	0.25
	ASH	-0.78	1.60	0.92
	CP	0.41	1.61	0.40
	PrA	-32.49	63.00	0.30
February	DM	1.41	1.74	0.07
	ASH	-0.17	1.86	0.90
	CP	0.22	2.21	0.14
	PrA	-17.70	84.76	0.17
March	ASH	0.91	2.68	0.78
	CP	-0.13	2.16	0.20
	PrA	13.21	86.23	0.15
April	ASH	-1.40	1.96	0.89
	CP	0.22	2.23	0.74
	PrA	-67.18	107.18	0.02

Table 2.2.7 Prediction Statistics for November, January and February using the December Equation

Predicting		BIAS	SEP(C)	R^2
November	DM	0.53	0.65	0.46
	ASH	0.44	1.16	0.95
	CP	-0.90	1.67	0.65
	PrA	5.93	92.94	0.17
January	DM	0.18	0.37	0.66
	ASH	-0.97	1.59	0.93
	CP	0.30	1.58	0.50
	PrA	-11.17	73.51	0.64
February	DM	1.77	1.72	0.09
	ASH	-0.69	1.58	0.93
	CP	-0.10	2.25	0.23
	PrA	-14.87	99.29	0.22
March	ASH	-0.70	2.33	0.85
	CP	-0.88	2.78	0.11
	PrA	64.80	127.08	0.12
April	ASH	-3.99	2.97	0.79
	CP	-0.34	2.69	0.47

Table 2.2.8 Prediction Statistics for November, December and February using the January Equation

Predicting		BIAS	SEP(C)	R^2
November	DM	0.32	0.87	0.23
	ASH	-0.06	2.27	0.76
	CP	-1.17	1.20	0.63
	PrA	37.17	86.54	0.21
December	DM	-0.18	0.82	0.15
	ASH	0.31	2.16	0.86
	CP	-0.82	0.87	0.87
	PrA	17.43	73.39	0.44
February	DM	1.62	1.95	-
	ASH	1.60	1.73	0.90
	CP	-0.48	1.27	0.65
	PrA	20.46	92.85	0.10
March	ASH	3.76	2.56	0.82
	CP	-2.34	1.54	0.57
	PrA	93.89	115.34	0.05
April	ASH	2.26	1.92	0.90
	CP	-2.58	1.70	0.82
	PrA	17.73	99.95	0.17

Table 2.2.9 Prediction Statistics for November, December and January using the February Equation

Predicting		BIAS	SEP(C)	R^2
November	DM	-0.93	1.72	0.05
	ASH	0.17	1.71	0.88
	CP	0.17	1.46	0.46
	PrA	12.98	118.97	0.01
December	DM	-1.20	1.54	0.18
	ASH	0.03	1.85	0.90
	CP	0.40	1.77	0.44
	PrA	12.64	97.93	0.07
January	DM	-1.58	1.74	-
	ASH	-1.17	1.46	0.92
	CP	0.23	1.10	0.71
	PrA	-5.69	75.49	0.11
March	ASH	0.66	2.89	0.75
	CP	-1.98	1.42	0.64
	PrA	-76.63	85.02	0.21
April	ASH	-1.40	2.61	0.81
	CP	-2.35	2.40	0.58
	PrA	-252.03	93.40	0.25

Figure 2.2.31 SDT Spectra for Fg and Ds Species

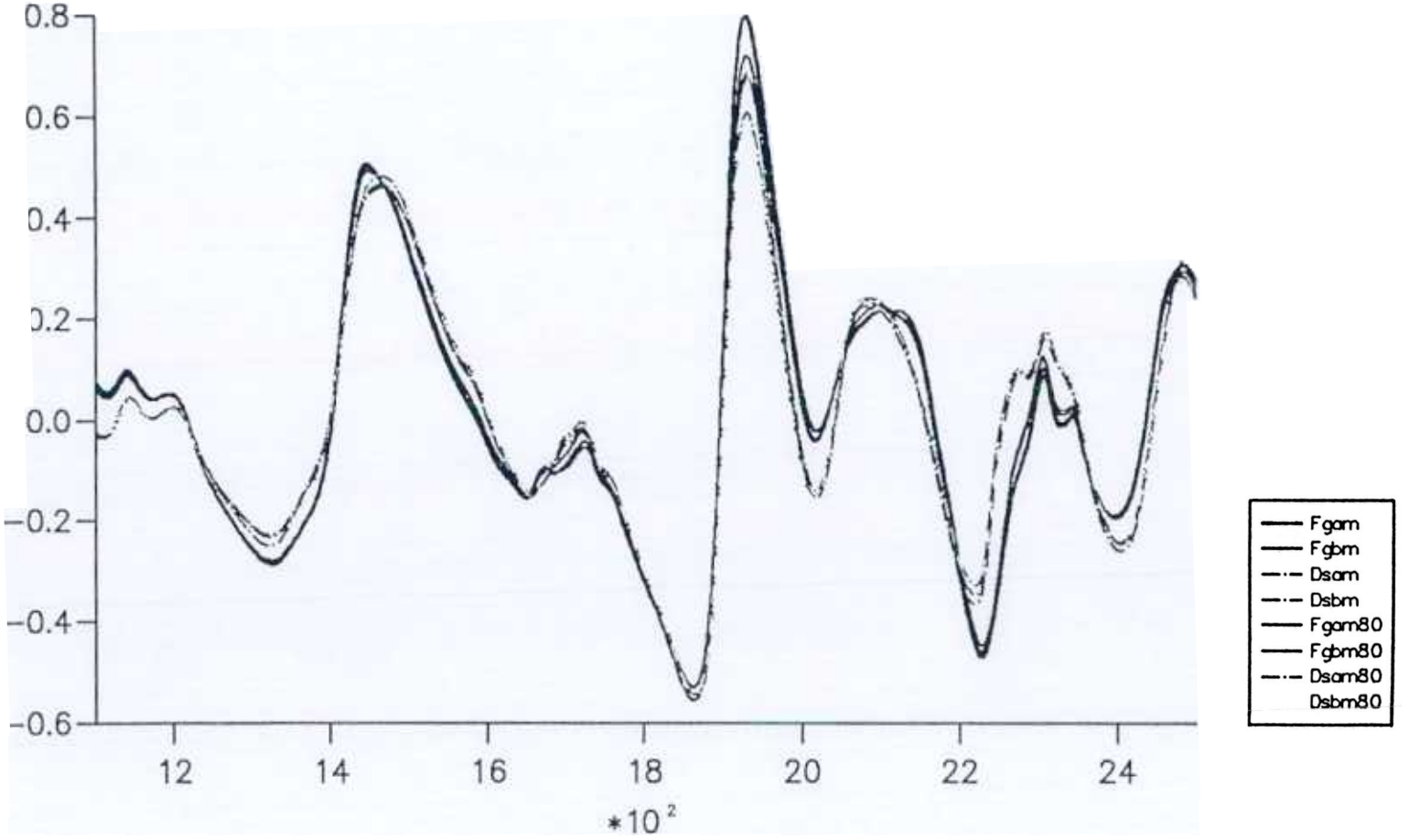


Figure 2.2.32 Spectra for Fg and Ds species dried at 60

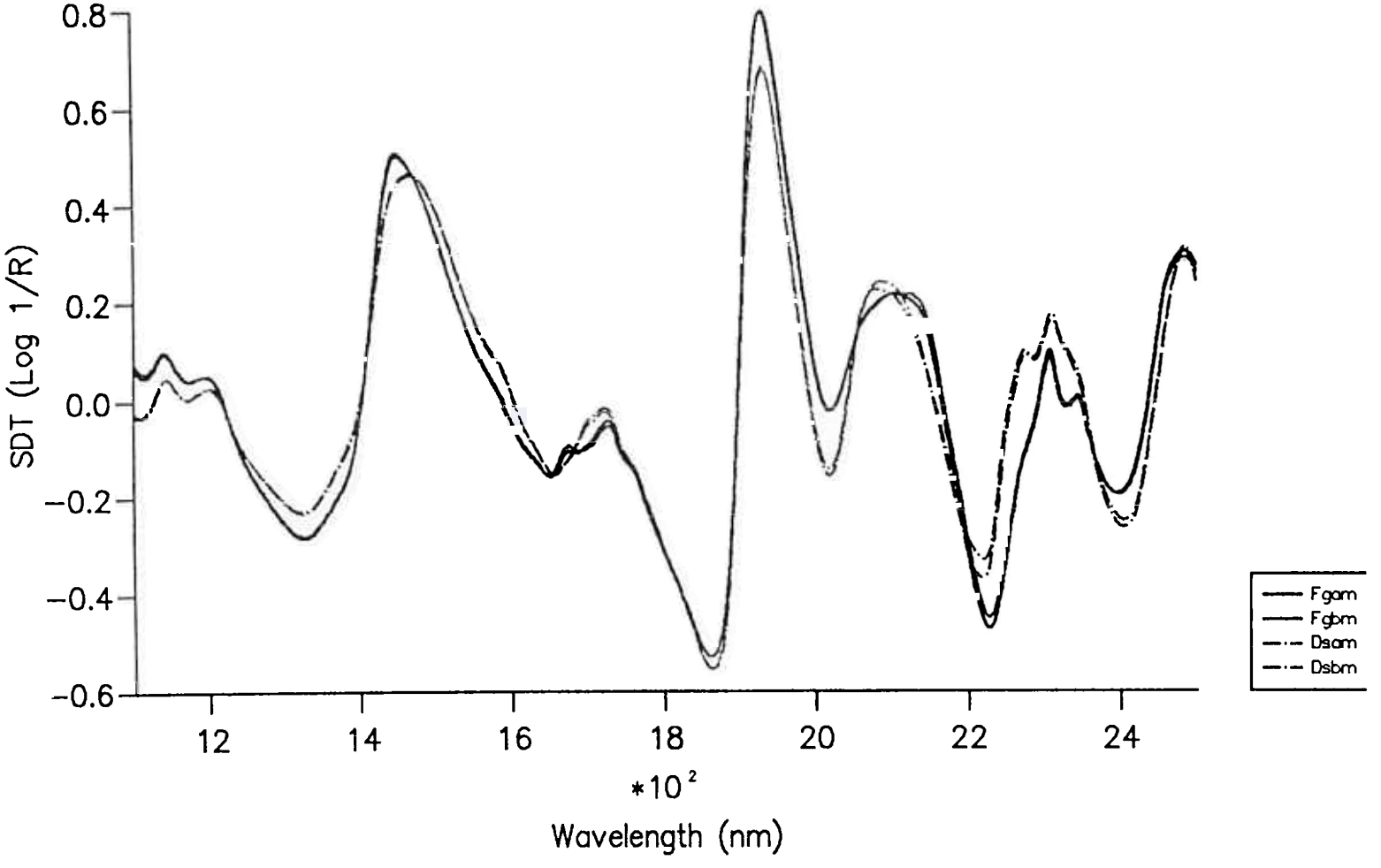


Figure 2.2.33 Spectra for Fg and Ds species dried at 80

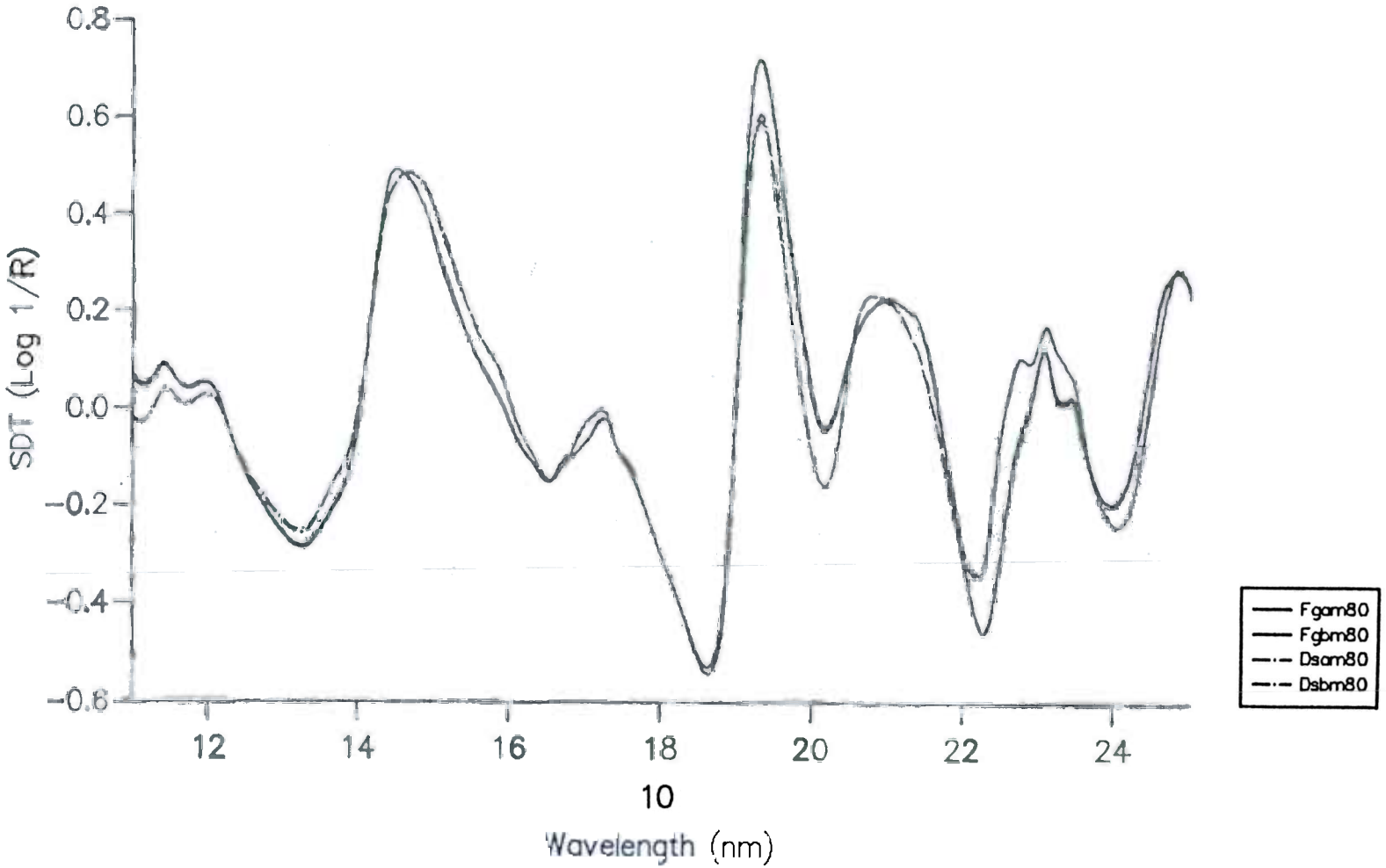
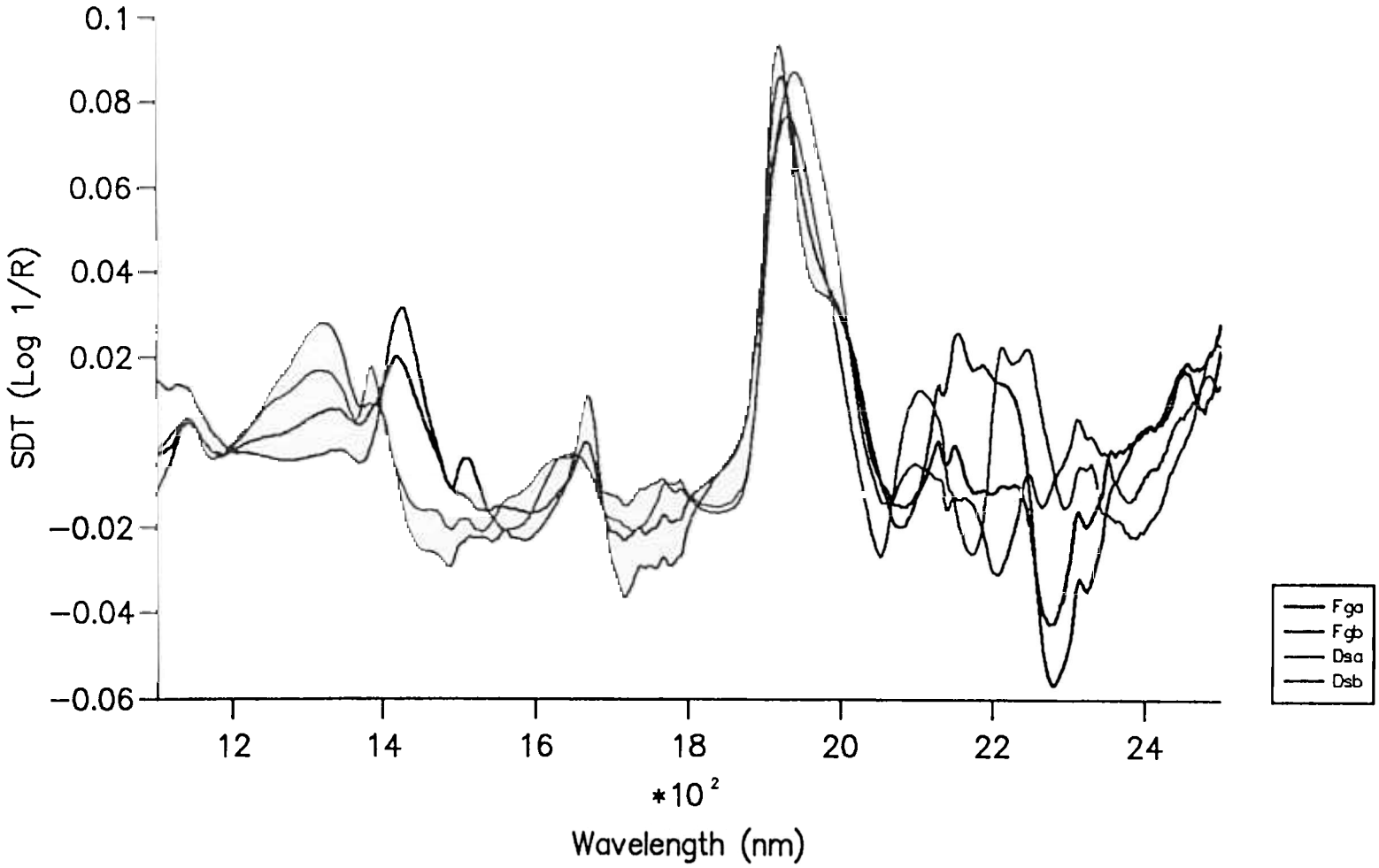


Figure 2.2.34 Difference spectra between samples dried at 60 and 80

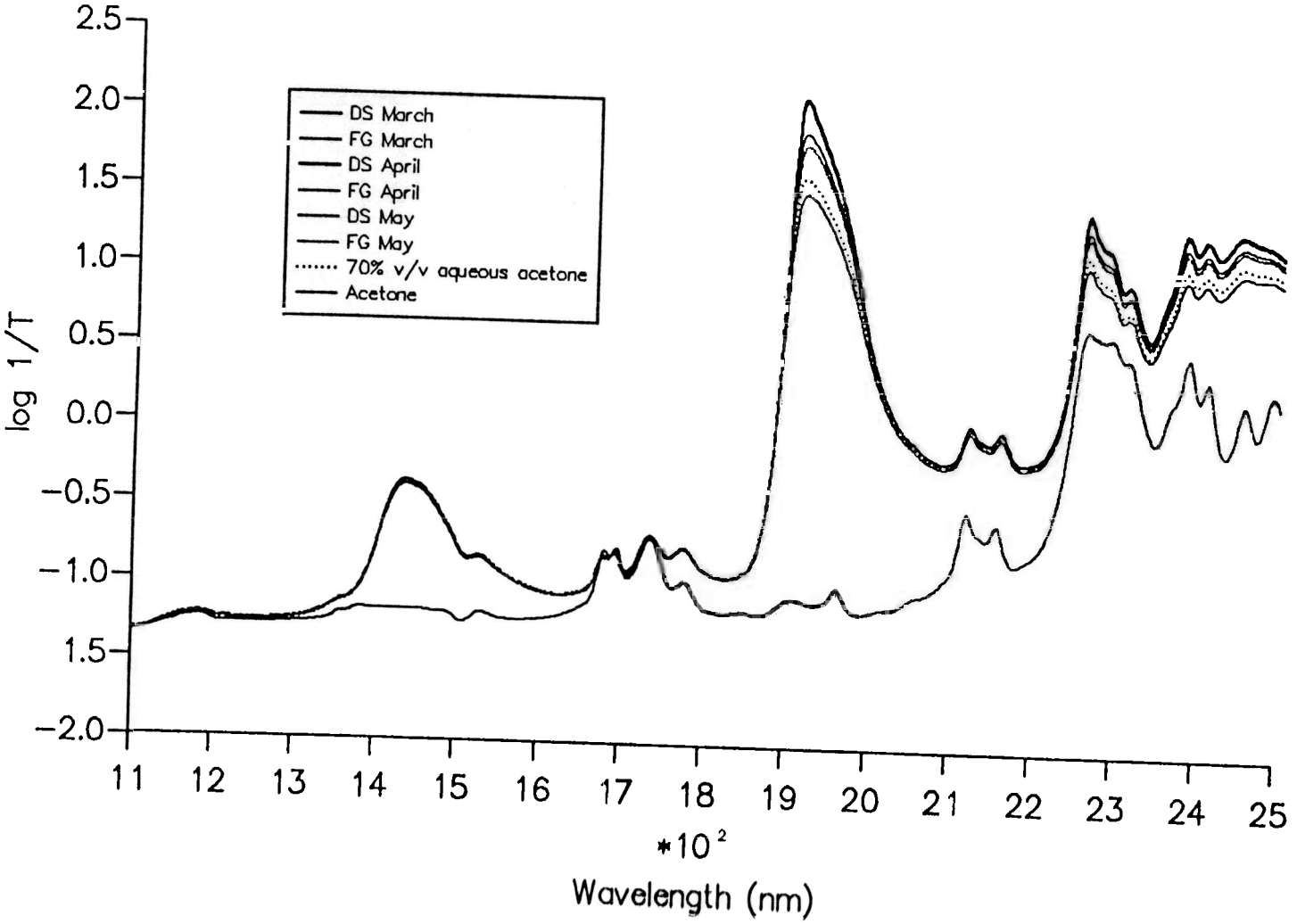


for only the Fg species. These are regions of the NIR spectrum known to be associated with tannins (13,14,15), suggesting that differences in tannin content may be detected using this technique. However, it must be remembered that drying at the higher temperature may have brought about other changes in the samples, not simply differences in the tannin content.

2.2.4 Transmission Spectroscopy

Near Infrared Transmission (NIT) spectroscopy allows the study of transparent liquids. Near infrared radiation is transmitted through the liquid sample and the data is collected in log (1/T), where T=Transmission. A preliminary investigation was performed using NIT to examine the extracts used in the protein precipitation assay (12). Duplicate samples from the Fg and Ds species were extracted using 70% aqueous acetone and the liquid extracts scanned. Figure 2.2.35 shows the Log 1/T spectra for acetone, 70% aqueous acetone and the mean spectra for the Fg and Ds species from months March, April and May. The spectra of the extracts for the two species appeared very similar, suggesting that the tannins in the extract are not present in a sufficiently high concentration to be detected. Further research is therefore required to assess this and other extract techniques for the routine analysis of tannins.

Figure 2.2.35 Mean for each tree in each month



3. IMPLICATIONS OF THE RESULTS

The initial aim of this work was to assess the use of NIR spectroscopy for identifying differences between provenances. Wet chemistry has not been able to differentiate between provenances. In the absence of chemical data in the current work, information has been obtained about the *Gliricidia* provenances using only the NIR spectral data. PCA and cluster analysis have proved useful in identifying differences between and within provenances. These statistical techniques and others employed in this study have enabled provenances showing the greatest spectral differences to be identified, also regions of the NIR spectra which may be associated with these differences have been identified. Further, this work has shown that calibration equations can be developed to predict DM, ash, CP, DOMD, DMD, OMD, TGP-C, TGP for *Gliricidia* samples. However, only a small representation of the population of *Gliricidia* provenances was available in this study. For this type of work to be more successful, more samples need to be included in the calibration population, with access to accurate laboratory data.

Statistical analysis of the *Gliricidia* samples has shown that the NIR data can be used to perceive some inter- and intra-provenance 'differences'. Using NIR spectroscopy it is possible to detect qualitative and quantitative differences between the provenances, however, further research is required to ascertain what exactly these 'differences' are.

Calibrations were developed successfully for the Nepal samples. There were obviously differences between samples from the different months which will need to be taken into account when predicting future samples using these equations. It is not clear at this point whether individual equations for each month would be required or whether broad-based equations representing all samples and species may be acceptable for the prediction of future samples. More samples are required to validate the calibration equations developed to date.

The statistical techniques employed in this study have been useful in identifying differences between species and looking at the distribution of the different species with respect to each other using only the spectral data. More information is required about the samples to assist in the identification of these differences.

Examination of the Nepal tree samples over the growing season revealed differences between species. It is now necessary to ascertain what these differences are, in particular

what is causing the apparent change in the months of February and April.

Transmission spectroscopy offers the possibility of a fast reliable technique for the analysis of tannins in liquid extracts, but is reliant on the extraction procedure at the moment.

4. PRIORITY TASKS

Two papers are being prepared for submission to refereed scientific journals and at least one more is expected. Preliminary results on the *Gliricidia* were presented at an NIR Users Meeting (MAFF, Starcross, UK).

5. BIBLIOGRAPHY

(1) GENSTAT 5 Procedure Library Manual (1989) Numerical Algorithms Group Ltd., (Editors: Payne, R.W. & Arnold, G.M).

(2) EMC X0162, Theodorou, M.K. & Brooks, A.E. (1990) Evaluation of a new laboratory procedure for estimating the fermentation kinetics of tropical feeds.

(3) GENSTAT 5 (1988) GENSTAT 5 Reference Manual. Clarendon Press, Oxford.

(4) Cowe, I.A. & McNicol, J.W. (1985) The use of principal components in the analysis of near-infrared spectra. *Applied Spectroscopy*, **39(2)**: 257-265.

(5) Gabriel, K.R. (1971) The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, **58**: 453.

(6) Mark, H.L. & Tunnel, D. (1985) Qualitative Near-Infrared Reflectance Analysis using Mahalanobis Distances. *Analytical Chemistry*, **57**: 1449-1456.

(7) Tilley, J.M.A. & Terry, R.A. (1963) A two stage technique for the *in-vitro* digestion of forage crops. *Journal of British Grassland Society*, **18**: 104-111.

(8) Naes, T., Irgens, C. & Martens, H. (1986) Comparison of Linear Statistical Methods for Calibration of NIR Instruments. *Applied Statistics*, **35**, No. 2, 195-206.

(9) Infrasoft International (ISI) (1991) NIRS 2 manual.

(10) Stone, M. (1974) Cross-validatory Choice and Assessment of Statistical Prediction. *Journal Royal Statistical Society (B)*, **36**, 111-133.

(11) Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1991) Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, **43(5)**: 772-777.

(12) Wood, C. (1991) Personal Communication.

(13) Windham, W.R., Fales, S.L. & Hoveland, C.S. (1988) Analysis of Tannin Concentration in *Sericea Lespedeza* by Near Infrared Spectroscopy. *Crop Science*, **28**:705-708.

(14) Mueller-Harvey, I., Dhanoa, M.S. & Barnes, R.J. (1990) NIRS of Sorghum Residues. *Proceedings of the 3rd International NIRS Conference*. Brussels, Belgium.

(15) EMC X0093, McAllan, A.B. (1991) Evaluation of cereal crop residues: influence of species, variety and environment on nutritive value.

6 ACKNOWLEDGMENTS

The author would like to thank M S Dhanoa for assistance with statistical analyses and interpretation of the data, Dr M K Theodorou and A E Brooks for the gas production analyses, J Stewart and T Simons (OFI) for the information regarding the *Gliricidia* provenances, and Dr C Wood (NRI) for the data on the Nepal samples.

7. NAME OF REPORT AUTHOR

Name: Susan J Lister

Signature: Susan J. Lister Date: 15-4-92