Pattern recognition of acorns from different *Quercus* species based on oil content and fatty acid profile

By **Suzana Ferreira-Dias*[(1)], Dina G. Valente [(1)]** and **José M.F. Abreu [(2)]**

[(1)]*Depart. Agro-Indústrias e Agronomia Tropical, Centro de Microbiologia e Indústrias Agrícolas.
e-mail: suzanafdias@mail.telepac.pt
[(2)] Depart. Produção Agrícola e Animal, Instituto Superior de Agronomia, Tapada da Ajuda,
1349-017 Lisbon; Portugal.

RESUMEN

Patrón de reconocimiento de bellotas procedentes de diferentes especies *Quercus* basado en el contenido de aceite y en el perfil de ácidos grasos.

El objetivo de este estudio fué (i) la caracterización de diferentes especies del género *Quercus* y (ii) la clasificación de las mismas en base al contenido y composición de ácidos grasos del aceite de sus frutos y/o en sus caracteres morfológicos, via técnicas de patrón de reconocimiento (Análisis de Componentes Principales, ACP, Análisis de Cluster, AC, y Análisis Discriminante, AD). Se han estudiado *Quercus rotundifolia* Lam., *Quercus suber* L. y *Quercus pyrenaica* Willd., pertenecientes a la misma zona del centro de Portugal.

Al emplear el contenido de aceite y sus respectivas composiciones de ácidos grasos para caracterizar a las muestras, el ACP reveló grupos bien separados correspondientes a cada especie, los cuales, a su vez, se confirmarón con el AC y el AD. El "ancho" y "longitud" de las bellotas exhibieron un poder discriminante bajo. Las bellotas de *Q. rotundifolia* mostraron el contenido más elevado de aceite, seguidas de las de *Q. suber* y *Q. pyrenaica* (9.1, 5.2 y 3.8%, respectivamente). Los perfiles de ácidos grasos de los aceites de *Q. rotundifolia* y *Q. suber* son similares al del aceite de oliva, mientras que el aceite de las bellotas de *Q.pyrenaica* es más insaturado.

PALABRAS-CLAVE: Aceite - Análisis de cluster - Análisis de componentes principales - Análisis discriminante -- Bellota - Quercus.

SUMMARY

Pattern recognition of acorns from different *Quercus* species based on oil content and fatty acid profile.

The aim of this study was (i) to characterize different species of *Quercus* genus and (ii) to discriminate among them on the basis of the content and fatty acid composition of the oil in their fruits and/or their morphological aspects *via* pattern recognition techniques (Principal Component Analysis, PCA, Cluster Analysis, CA, and Discriminant Analysis, DA). *Quercus rotundifolia* Lam., *Quercus suber* L. and *Quercus pyrenaica* Willd., grown in the same stand in the centre of Portugal, were investigated.

When oil content and respective fatty acid composition were used to characterize samples, well-separated groups corresponding to each of the species were observed by PCA and confirmed by CA and DA. The "width" and "length" of acorns exhibited a low discriminant power. Acorns from *Q. rotundifolia* showed the highest average oil content followed by *Q. suber* and *Q. pyrenaica* acorns (9.1, 5.2 and 3.8%, respectively). Fatty acid profiles of *Q. rotundifolia* and *Q. suber* oils are similar to olive oil while the oil from *Q. pyrenaica* acorns is more unsaturated.

KEY-WORDS: Acorn - Cluster analysis - Discriminant analysis - Principal component analysis -- Oil - Quercus.

1. INTRODUCTION

Quercus is one of the most common genus in the Portuguese Flora where there are 8 different species. All species are acorn producers and each of them is well adapted to diverse soils and climate conditions (Goes, 1991). These fruits were used both for human and animal feed. Until the early seventies, the extraction of the oil from *Q. rotundifolia* acorns was carried out in some oil extraction plants in Portugal, due to the very similar fatty acid composition of this oil to olive oil (Ferrão and Ferrão, 1988). However, the increase in salaries made the harvest too high to make the extraction of this oil viable. Nowadays, the use of olive oil, due to its nutraceutical properties, has drawn attention to edible oils with similar composition.

The aim of this study was (i) to characterize 3 different species of the *Quercus* genus and (ii) to discriminate among them on the basis of the content and fatty acid composition of the oil in their fruits and/or on morphological aspects of the fruits *via* pattern recognition techniques (Principal Component Analysis, Cluster Analysis and Discriminant Analysis). The following different species, grown in the same mixed stand in the center of Portugal (Portalegre) were investigated: *Quercus rotundifolia* Lam., *Quercus suber* L. and *Quercus pyrenaica* Willd.

Pattern recognition methods were carried out to determine the set of measurements for sample characterization, *i.e.,* to identify the pattern (Miller and Miller, 1993). Therefore, the available data were used simultaneously rather than sequentially and Principal Component Analysis (PCA), Cluster Analysis (CA) and Discriminant Analysis (DA) were performed on multivariate data.

PCA is an attempt to best describe the shape of a multivariate distribution by considering selected

linear combinations of the original variables rather than the variables themselves (Bolfinger, 1975). In addition, with this technique, the initial m-dimensional space (m variables) may be reduced to n dimensions (n<m) without considerable loss of information (Harman, 1976; Hoffman and Young, 1983). The initial system of m axis is replaced by another system where the new axis are the principal components (Morrison, 1967; Piggott and Sherman, 1986). The first component shows the maximum correlation with all the variables and explains the highest proportion of the global variance (Powers, 1988). This method allows the geometric representation of the original objects (*Quercus acorns*) in a space of reduced dimensions defined by a new set of axis and, consequently the identification of groups of similar objects. It may also provide a particular interpretation of the components and subsequently of the original variables.

The second stage of multivariate data analysis consisted of a Cluster Analysis of the multivariate data in order to confirm the existence of the groups suggested by the plot of the samples in the reduced space defined by the significant principal components. The term Cluster Analysis (CA) covers a number of different clustering concepts and classification algorithms (Mirkin, 1996). In this study, the hierarchic clustering methodology was used, *i.e.*, the clusters of higher levels resulted from aggregations of the clusters of lower levels. The Euclidean distance was used as a measure of similarity or distance between samples. As a rule for the linkage of clusters, the single linkage method (or the nearest neighbor) was used. In this method, the distance between the two closest objects in the different clusters determines the distance between two clusters.

Finally, after the identification of isolated groups of samples by PCA and CA, a Discriminant Analysis (DA) was used to determine which variables discriminate among these groups *a priori* defined (Morrison, 1967; Burgard and Kuznicki, 1990). The basic idea underlying is whether groups differ with regard to the mean of a variable and then use that variable to predict group membership. One can ask whether or not two or more groups are significantly different from each other with respect to the mean of a particular variable. If the means for a variable are significantly different in different groups, then this variable discriminates between the groups. In fact, the procedure is identical to the one-way analysis of variance or to the multivariate analysis of variance if several variables are used.

2. MATERIALS AND METHODS

2.1. Materials

Ripened fruits were obtained, on the same date, from adult trees of a mixed stand (Portalegre, Portugal) with the following *Quercus* species: *Q. rotundifolia*, *Q. suber* and *Q. pyrenaica*. Each sample was collected from one individual tree, i.e., 10 samples of *Q. rotundifolia,* 11 samples of *Q. suber* and 9 samples of *Q. pyrenaica*, were obtained; *n*-hexane *p.a.* was used for oil extraction.

2.2. Methods

Morphological Characterization of the fruits: The average length and width (cm) were measured for 100 fruits from each tree picked at random.

Oil extraction: Before oil extraction, the fruits were dehulled, ground in a household coffee mill (knife cutter type) and heated at 75°C for 90 minutes. The oil from prepared material was extracted by *n*-hexane *p.a.* in a Soxhlet apparatus for 8 hours (Ferreira-Dias *et al.*, 2003). Experiments were carried out in triplicate and the results were expressed on a dry basis.

Chemical characterization of the oil: The fatty acids profile of every oil sample was evaluated as their methyl esters in a gas chromatograph (Carlo Erba, Vega 2000 GC) equipped with a SUPELCO capillary column (SP^{TM}-2380, 0.2μm, 60m x 0.25 mm; fused silica). Both detector and injector (FID) were heated at 250°C. Temperature was programmed as follows: 175°C for 25 minutes, a slope of 5°C/min. from 175°C to 220°C and 220°C for 10 minutes. Hydrogen was the carrier gas at a column head pressure of 60 kPa.

2.3. Statistical analysis

The experimental results, concerning the average "length" and "width" of the fruits from each of the 30 trees studied, the average yield in oil and its fatty acid composition, were put in a matrix form (Matrix A). Samples were presented in rows and variables in columns (matrix 30x8). A matrix 30 x 6 was considered for statistical analysis by removing the columns corresponding to the "length" and "width" of the fruits (Matrix B).

Concerning pattern recognition, a Principal Component Analysis (PCA) was first carried out on the experimental data (matrices A and B). The second stage of multivariate data analysis consisted of a Cluster Analysis (CA) of the data matrices A and B, in order to confirm the existence of the groups suggested by the plot of the samples in the reduced space defined by the significant principal components.

Finally, a Discriminant Analysis (DA) was used on data from matrix B, to determine which variables discriminate between these groups *a priori* defined (Morrison, 1967; Burgard and Kuznicki, 1990). The model of discrimination was built step-by-step and a forward stepwise analysis was followed. At each

step, it was evaluated which variable would contribute most to the discrimination between groups. That variable would then be included in the model, beginning the next step. The maximum number of discriminant functions will be equal to the number of groups minus one or to the number of variables in the analysis, whichever is smaller. The best combination of variables for discriminant analysis includes variables that represent independent measures of product similarities and differences.

In addition, the Classification Functions can be used to determine to which group each case most likely belongs. The classification matrix shows the number of cases that were correctly classified and those that were misclassified.

PCA, CA and DA were performed by using the software "Statistica™", version 5, from Statsoft, USA.

3. RESULTS AND DISCUSSION

3.1. Sample Characterization by Principal Components Analysis

The average values of the oil content, respective fatty acid profile, average length and width of the fruit

Table I

Matrix A containing the average values of oil content (%, w/w, dry basis), respective fatty acid profile (%),average length and width of *Quercus* fruit samples

Sample	Species	Oil	C16:0	Fatty C18:0	Acid C18:1	Profile C18:2	C18:3	Fruit Length (cm)	Fruit Width (cm)
s1	*suber*	4.4	14.8	1.0	63.8	18.3	1.7	3.91	1.86
s2	*suber*	4.6	15.6	1.0	62.7	18.0	1.7	3.73	1.72
s3	*suber*	5.7	14.8	0.95	62.6	19.1	1.7	3.47	1.77
s4	*suber*	5.1	14.9	1.0	62.1	19.6	2.1	3.42	1.69
s5	*suber*	4.7	18.2	1.0	61.5	17.8	1.2	3.50	1.69
s6	*suber*	5.7	18.2	0.9	61.3	17.5	1.7	3.44	1.71
s7	*suber*	5.9	14.6	1.0	63.7	18.1	2.0	3.41	1.69
s8	*suber*	6.6	16.6	1.1	62.0	18.6	1.9	3.49	1.69
s9	*suber*	5.4	14.6	1.1	64.4	17.6	2.1	3.12	1.61
s10	*suber*	4.7	14.9	1.0	62.1	19.6	2.1	3.16	1.58
s11	*suber*	4.7	18.2	1.0	61.5	17.8	1.2	3.16	1.62
i1	*rotundifolia*	8.0	18.0	3.5	65.0	14.0	1.0	3.54	1.53
i2	*rotundifolia*	7.3	15.3	3.0	65.8	14.4	1.0	3.72	1.62
i3	*rotundifolia*	8.6	16.6	3.9	64.4	13.6	1.1	3.41	1.55
i4	*rotundifolia*	9.4	17.2	3.0	63.1	15.2	1.1	3.21	1.65
i5	*rotundifolia*	10.3	18.1	2.6	62.0	16.6	0.8	3.69	1.85
i6	*rotundifolia*	10.2	15.0	2.8	65.9	14.6	1.1	3.72	1.86
i7	*rotundifolia*	10.9	18.1	2.6	61.9	16.6	0.8	3.49	1.71
i8	*rotundifolia*	8.0	19.2	2.4	62.3	14.4	1.1	3.47	1.78
i9	*rotundifolia*	9.6	17.0	2.5	63.2	16.2	1.1	3.55	1.76
i10	*rotundifolia*	9.0	18.3	2.5	62.8	14.8	1.1	3.52	1.74
p1	*pyrenaica*	4.1	15.6	1.5	50.0	29.9	2.0	3.64	1.86
p2	*pyrenaica*	3.8	17.4	1.5	48.7	29.7	1.9	3.77	1.92
p3	*pyrenaica*	4.4	15.4	1.7	48.0	32.0	2.1	3.77	1.83
p4	*pyrenaica*	3.8	16.7	1.3	46.3	32.7	2.3	4.16	1.96
p5	*pyrenaica*	4.2	18.4	1.3	43.4	34.3	1.9	4.00	1.96
p6	*pyrenaica*	2.6	15.6	1.8	52.2	28.2	1.8	3.63	2.08
p7	*pyrenaica*	2.7	16.2	1.6	51.4	28.3	1.4	3.46	2.12
p8	*pyrenaica*	4.2	15.1	1.4	48.5	32.2	2.0	3.94	1.96
p9	*pyrenaica*	4.0	13.6	1.8	49.4	25.1	1.6	3.47	1.82

Table II
Principal Component Analysis of matrix A- eigenvalues and variances explained by the principal components extracted from the data matrix A, considering the yield in oil, fatty acid composition and the "length" and "width" of the fruits (see text for details)

Principal Component	Eigenvalue	Total Variance (%)	Cumulative Variance (%)
1	4.35	54.31	54.31
2	1.73	21.63	75.94
3	0.78	9.75	85.69
4	0.47	5.82	91.51
5	0.38	4.73	96.24
6	0.20	2.48	98.72
7	0.08	1.12	99.84
8	0.01	0.16	100.00

samples of *Q. rotundifolia, Q. suber* and *Q. pyrenaica* are shown in Table I, displayed in a matrix form (Matrix A). These data were first submitted to a PCA and the eigenvalues and respective variances of the extracted new axis (principal components) are shown in Table II. According to the criterion proposed by Kaiser, only principal components with eigenvalues greater than 1 must be retained for the analysis, since they explain more than the average variability accounted for by one original variable as meaningful (Dagneli, 1977; Burgard and Kuznicki, 1990). Therefore, the initial 8-dimensional space (defined by 8 variables) can be reduced to a plane, F_1F_2, defined by the first two principal components since they have eigenvalues greater than one. This plane accounts for about 76% of the variance explained by the original data matrix A.

The correlations between the original variables and the first two principal components, *i.e.* the loadings of variables are in Fig. 1-A. The first axis is positively correlated with the concentrations of linoleic and linolenic acids (C18:2 and C18:3, respectively) and negatively with oleic acid (C18:1). The amounts of saturated fatty acids (palmitic, C16:0, and stearic acids, C18:0) are better correlated with the second principal component, increasing along it. They are also well correlated with the negative part of the first principal component, which plots both acids very close to the diagonal of the 2nd quadrant. The presence of the variables "Width" and "Length" of the fruits near to the diagonal of the first quadrant indicates that they are equally correlated with both axes. In order to redistribute the weightings of the variables so as to make them
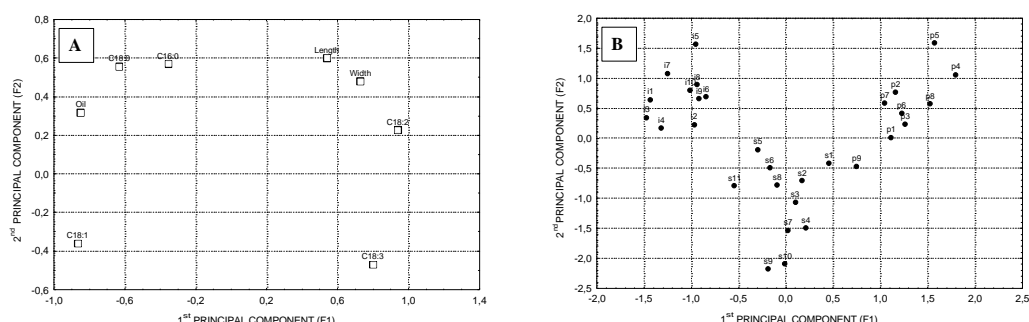

Figure 1
Principal Component Analysis of data from Matrix A- Loadings of morphological characteristics of Quercus fruit samples, oil content and respective fatty acid composition on the first and second principal components (fig.1-A); Plot of the *Quercus* fruit samples (from *Q. suber* (s), *Q. rotundifolia* (i) and *Q. pyrenaica* (p)) on the plane defined by first and second principal components (fig.1-B).

Table III
Principal Component Analysis of matrix B- eigenvalues and variances explained by the principal components extracted from the data matrix B, corresponding to the data of matrix A concerning the yield in oil and its fatty acid composition

Principal Component	Eigenvalue	Total Variance (%)	Cumulative Variance (%)
1	3.73	62.17	62.17
2	1.20	20.06	82.23
3	0.69	11.52	93.75
4	0.23	3.82	97.57
5	0.13	2.19	99.76
6	0.02	0.24	100.00

nearer to or farther away from each of the axes, rotation of principal components by Varimax technique was also performed (Burgard and Kuznicki, 1990). However, no improvement on interpreting principal components and/or clustering the samples was achieved.

When the samples were plot on the plane formed by the first and the second principal components, F_1F_2 (Fig. 1-B), the different samples seem to be grouped according to the species. The *Q. rotundifolia* fruit samples show the highest oil content and their oil is richer in saturated fatty acids (palmitic and stearic acids). Larger and longer fruits are observed for *Q. pyrenaica*. The lowest oil content is obtained for *Q. pyrenaica* acorns. This oil is richer in unsaturated fatty acids (mainly linoleic acid). The *Q. suber* fruits are in an intermediate position in the plane F_1F_2. This suggests that for *Q. suber*, both oil content and

composition are between the values observed for *Q. rotundifolia* and *Q. pyrenaica* fruits.

As an attempt to better separate samples by species, the variables concerning the morphological characterization of the fruits were ignored on Matrix A (Table I.) and a second PCA on these smaller set of data (Matrix B) was performed.

Similarly as obtained for Matrix A, the original information can be displayed on the plane F_1F_2, corresponding to the eigenvalues higher than the unity (Table III). This plane explains about 82% of the initial information contained in the matrix B.

The loadings of the initial variables on first and second components (Fig. 2-A) are similar to those observed for Matrix A. With respect to the projection of the samples on the plane F_1F_2 (Fig. 2-B), a better separation of the clusters according to the species is obtained. Therefore, the length and width of the fruits showed not to be adequate parameters to
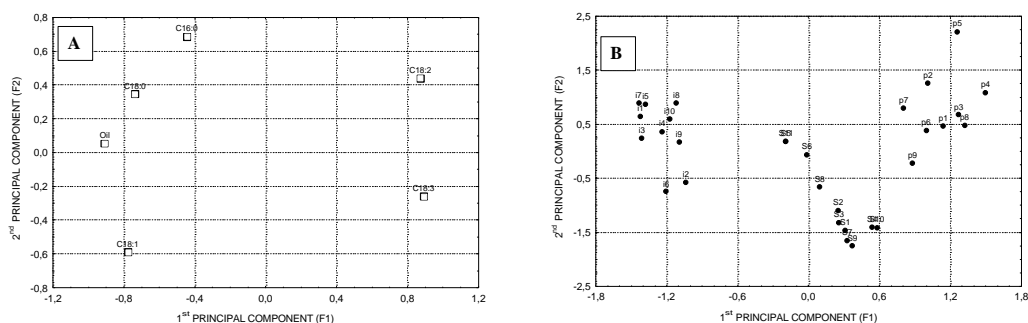

Figure 2
Principal Component Analysis of data from Matrix B- Loadings of oil content of Quercus fruit samples and respective fatty acid composition on the first and second principal components (fig.2-A); Plot of the *Quercus* fruit samples (from *Q. suber* (s), *Q. rotundifolia* (i) and *Q. pyrenaica* (p)) on the plane defined by first and second principal components (fig.2-B).
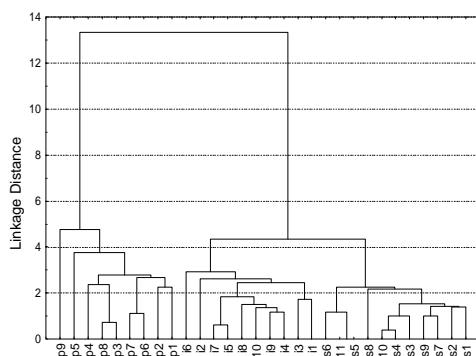
Figure 3
Dendrogram of the *Quercus* fruit samples (from *Q. suber* (s),
Q. rotundifolia (i) and *Q. pyrenaica* (p)) based on their oil content
and fatty acid composition.

Table IV
**Summary of Stepwise Discriminant Analysis by
forward method (*** at a significance level,
p <0.001; NS, not significant, p>0.05)**

Step	Variable to Enter	F to enter
1	C18:1	201.83***
2	C18:0	121.47***
3	Oil	23.89***
4	C16:0	9.93***
5	C18:3	1.39 (NS)
6	C18:2	1.24 (NS)

discriminate between samples from different species of the *Quercus* genus, as previously suggested by the results from the analysis of variance.

3.2. Sample Characterization by Cluster Analysis

In a second stage, a CA was carried out to investigate the feasibility of the clusters suggested by PCA. The hierarchical tree diagram (dendrogram) for the 30 *Quercus* samples described by the "oil content" and fatty acid composition (6 variables; Matrix B) is shown in Fig. 3. Similar dendrogram is obtained when data from Matrix A is used (not shown). This confirms the small power of "fruit length" and "fruit width" on discrimination and characterization of *Quercus* samples. For a linkage distance higher than 5, only two clusters can be defined: a Cluster corresponding to *Q. pyrenaica* fruits, and another Cluster where the sub-cluster of *Q. rotundifolia* fruits and the sub-cluster of *Q. suber* are joined together. For a linkage distance of 4, three clusters, corresponding to the different species, can be identified. Only an outlier from *Q. pyrenaica* (P9) was observed at this linkage level (Fig. 3).

3.3. Sample Characterization by Discriminant Analysis

After confirming by Cluster Analysis the existence of isolated groups of samples, a Discriminant Analysis was used to determine which variables discriminate between the groups *a priori* defined, corresponding to the 3 species. Due to the low discriminant power exhibited by the variables "Length" and "Width" of the *Quercus* fruits on the previous data analysis (PCA and CA) only the oil

content of the fruits and their fatty acid composition (Matrix B) were used in Discriminant Analysis. Table IV summarizes the successive steps of the forward stepwise analysis. The highest the F-value of a variable, the highest is the discriminating power of that variable. Therefore, the most important variable to discriminate between *Quercus* species was oleic acid content followed by stearic acid, oil content and palmitic acid level.

The existence of the 3 entirely distinct *Quercus* species can be confirmed by the plot of the *Quercus* samples onto the plane formed by the two Discriminant Functions (canonical roots) found by canonical analysis (Fig. 4).

In addition, the following Classification Functions were established to define every group (species) and can be used to determine to which group each case most likely belongs:
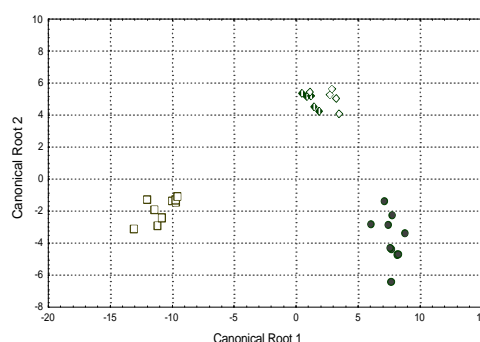

Figure 4
Discriminant Analysis- Plot of 30 *Quercus* samples from 3 different species (◇, *Q. suber;* •, *Q. rotundifolia*; , *Q. pyrenaica*) on the plane defined by 2 canonical roots (discriminant functions).

Table V
Classification Matrix- The diagonal shows the number of cases correctly classified

Observed	% Correct	Predicted Classifications		
		Q.rotundifolia	Q. pyrenaica	Q. suber
Q. rotundifolia	100.00	10	-	-
Q. pyrenaica	100.00	-	9	-
Q. suber	100.00	-	-	11
TOTAL	100.00	10	9	11

Q. rotundifolia = (-4104.60) + 88.88[C18:1] + 49.76[C18:0] + 53.93[OIL] + 71.58[C16:0] + 48.81[C18:3] + 41.76[C18:2]   (eq. 1)

Q. suber = (-3834.57) + 87.35[C18:1] + 19.15[C18:0] + 44.04[OIL] + 68.62[C16:0] + 59.61[C18:3] + 41.25[C18:2]   (eq. 2)

Q. pyrenaica = (-3176.34) + 77.41[C18:1] + 41.07[C18:0] + 34.67[OIL] + 59.36[C16:0] + 34.91[C18:3] + 45.42[C18:2]   (eq. 3)

Where [C16:0], [C18:0], [C18:1], [C18:2] and [C18:3] are the concentrations (%) of palmitic, stearic, oleic, linoleic and linolenic acids, respectively; [OIL] represents oil content of the acorns (%, w/w), on a dry basis.

In fact, when the observed classifications were compared to the predicted by these classification functions no case was misclassified, *i.e.* all the cases were correctly classified (Table V).

3.4. Final Remarks

The results concerning oil content of *Q. suber* and *Q. rotundifolia* fruits and respective fatty acid profile are in agreement with those obtained by Ferrão and Ferrão (1988) for the same fruit species harvested from 1984 to 1986 in the Centre and South of Portugal. In fact, data from matrix A (Table I) show that acorns from *Q. rotundifolia* have an average oil content of 9.1% and *Q. suber,* of 5.2%. The lowest values were observed for *Q. pyrenaica* fruits (average value of 3.8%).

Both for oils from *Q. rotundifolia* and *Q. suber* fruits, fatty acid profiles are within the variation range exhibited by olive oil fatty acids (Commission of the Codex Alimentarius, 1993). However, *Q. rotundifolia* fruit oil is richer in saturated fatty acids (palmitic and stearic acids) than the oil from *Q. suber* fruits. In addition, a slightly higher linolenic acid content was observed in *Q. suber* fruit oil (1.2-2.1%) when

compared to its content in olive oil (maximum of 1.5%). Concerning *Q. pyrenaica* acorns, its oil is rather unsaturated and quite different from olive oil. The highest levels of linoleic (25.1-34.3%) and linolenic acids (1.4-2.3%) and the lowest oleic acid content (43.4-52.2%) were observed in the oil from this fruit.

As a main conclusion, the knowledge of oil content and respective fatty acid profile of *Q. rotundifolia*, *Q. suber* and *Q. pyrenaica* fruits allows the identification of acorns from these species via pattern recognition techniques (PCA coupled with CA and DA).

ACKNOWLEDGMENTS

REFERENCES

Bolfinger, V.J. (1975). An Introduction to Some Multivariate Techniques with Applications in Field Experiments, in *Developments in Field Experiment Design and Analysis*, pp. 73-83. V.J. Bolfinger and J.L. Wheeler (Eds.). Commonwealth Agricultural Bureaux.

Burgard, D.R., Kuznicki, J.T. (1990) Chemometrics: Chemical and Sensory Data, pp. 135-185. CRC Press, Boston.

Commission of the Codex Alimentarius (1993). Revised Norm for Olive Oil, Cl 1993/15-FO.

Dagneli, P. (1977) Analyse Statistique à Plusieurs Variables, pp. 167-224, Presses Agronomiques de Gembloux, A.S.B.L., Gembloux.

Ferrão, J.E.M., Ferrão, A.M.B.C. (1988). A bolota e a glande, potencialidades em óleo e farinha, *Revista de Ciências Agrárias*, **2,** 15-29.

Ferreira-Dias S., Valente, D.G., Abreu, J.M.F. (2003), Comparison between ethanol and hexane for oil extraction from *Quercus suber* L. fruits. *Grasas y Aceites,* **4**, 378-383 .

Goes E. (1991). A Floresta Portuguesa: Sua Importância e Descrição das Espécies de Maior Interesse, pp. 127-144, Portucel, Portugal.

Harman, H.H. (1976). Modern Factor Analysis, pp. 133-135, The University of Chicago Press, Chicago and London.

Hoffman, D.L., Young, F.W. (1983) Quantitative Analysis of Qualitative Data: Applications in Food Research in *Food Research and Data Analysis*, pp. 69-93, H. Martens and H. Russwurm Jr., (Eds.). Applied Science Publishers, London and New York.

Miller, J.C., Miller, J.N. (1993) Statistical Analysis for Analytical Chemistry, pp. 169-203, Ellis Horwood and Prentice Hall.

Mirkin, B. (1996) Mathematical Classification and Clustering, pp. 109-158, Kluwer Academic Publishers, Dordrecht, Boston and London.

Morrison, D.F. (1967) Multivariate Statistical Methods, pp. 226-302, Kogakusha International Student Edition, McGraw-Hill, New York.

Piggott, J.R., Sherman, K. (1986) Methods to Aid Interpretation of Multidimensional Data, in *Statistical Procedures in Food Research*, J.R. Piggott, (Ed.), Elsevier Applied Science, London and New York.

Powers, P.P. (1988) Current Practices and Application of Descriptive Methods, in *Sensory Analysis of Foods*, J.R. Piggott (Ed.), Elsevier Applied Science, London and New York.