ELSEVIER

Food Chemistry 76 (2002) 501–507

Food Chemistry

www.elsevier.com/locate/foodchem

Analytical, Nutritional and Clinical Methods Section

Study of oils from Calabrian olive cultivars by chemometric methods

Silvia Lanteri[a,*], Carla Armanino[a], Enzo Perri[b], Annamaria Palopoli[b]

[a]*Dipartimento di Chimica e Tecnologie Farmaceutiche e Alimentari, Università di Genova, Via Brigata Salerno, I-16147 Genova, Italy*

[b]*Istituto Sperimentale per l'Olivicoltura, C.da Li Rocchi, I-87036 Rende (CS), Italy*

Received 15 May 2001; received in revised form 14 October 2001; accepted 14 October 2001

Abstract

A study of characterisation of a typical Italian food is performed by chemometrics. The olive oils from some cultivars of Calabria have been characterised according to their origin and olive genotype using the chemical information mainly provided by 14 chemical parameters of virgin olive oils. In particular, the models of three cultivars of Calabria (Carolea, Cassanese, Dolce di Rossano) were deeply studied. The microclimate of this region has a lower influence on fatty acid composition than the genotype. Using simple and relatively inexpensive analytical parameters as fatty acids and chemometric techniques it has been possible to characterise and classify the olive oils (60–94% prediction rate). © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Olive cultivar; Chemometrics; Classification; Modelling; Protected designation of origin

1. Introduction

The olive oil is one of the most important typical Italian food products. Its characterisation according to its origin and/or genotype is now widely under investigation. In fact, the chemical composition of olive oils may be different due to the influence of agronomic and technological factors, so there is the need to define and to characterise each typical olive oil.

In this work the attention is devoted to the olive oil from some cultivars of Calabria, the second Italian region for olive oil production; the olive oils from the most spread cultivars are studied by chemometric techniques.

In Europe two regulations introduced the Protected Designation of Origin (PDO) of traditional products. The first one protects traditional products (Council Regulation, EEC-N.2081/92) about designations of origin and geographical indication, the second one (Council Regulation, EEC-N.2082/92) defines product types, olive oil included.

The Calabria region in the South of Italy is the second olive growing Italian region with an annual average production of olive oil reaching 100,000 tons. At present it may produce two PDO oils, namely Bruzio (Council Regulation, EEC-N.1065/97) and Lametia (Council Regulation, EEC-N.2107/99), whereas a third olive oil, namely Alto Crotonese, is obtaining the PDO mark. The Calabrian olive tree germplasm is very rich and different (Perri, Sirianni, Godino, & Tartarini, 1999). It comprises several olive cultivars characterised by very interesting physiological (De Nino, Perri, Lombardo, Procopio, Raffaelli, & Sindona, 1997; De Nino, Mazzotti, Morrone, Perri, Raffaelli, & Sindona, 1999), agronomic (Perri, Sirianni, & Pellegrino, 1998) and nutritional features (Perri, Raffaelli, & Sindona, 1999).

The geographic origin of olive oil was studied by chemometric techniques using the data of the chemical composition of olive oils. Good results were obtained by classification methods reported in several papers (Alonso Garcia, & Aparicio Lopez, 1993; Armanino, Leardi, Lanteri, & Modi, 1998; Derde, Coomans, & Massart, 1984; Drava, Forina, Lanteri, & Lupoli, 1994; Forina & Armanino, 1982; Forina & Tiscornia, 1982; Leardi & Paganuzzi, 1987; Montedoro et al., 1995; Perri, Parlati, Palopoli, Pellegrino, & Rizzuti, 1999; Sacchi et al., 1996; Shaw et al., 1997). Fatty acids are extremely useful in characterising olive oil (Parlati, Perri, Palopoli, & Rizzuti, 1994). Their content may be

* Corresponding author. Fax: +39-010-353-2684.
E-mail address: silvia@dictfa.unige.it (S. Lanteri).

PII: S0308-8146(01)00370-3

affected by a number of different factors, such as genotype and ripening stage, so, fatty acid content can be useful to obtain information about these factors.

The main aim of this paper is to evaluate the discriminating ability of the most usually measured parameters (free acidity, peroxide index and fatty acids) in distinguishing among olive oils produced from different cultivars of Calabria. Indeed, the variety factor is undoubtedly one of the most important, but despite this, it is often ignored, either through lack of varietal information, or because the oil has been mixed with oils of different varieties or even because emphasis has been laid only on its place of origin. Moreover, the influence of the environment and other agronomic factors on olive oil composition may be often overestimated, because little notice is taken of the natural geographical distribution of particular cultivars in particular areas.

Finally, the influence of genotype, year of production and fruits ripening has been evaluated.

2. Experimental

The olive oil samples came from different geographic areas of Calabria. For each sample 5–10 kg of olives were picked from trees, which were homogeneous for cultivar, health, and then milled in an laboratory scale hammer mill; after malaxation and pressing, the oil was separated by centrifugation.

Samples (322) of monovarietal olive oil were produced and analysed. Free acidity, peroxide index, and fatty acids were determined on each sample according to the European Official Methods of Analysis (Council Regulation, EEC-N.2568/91). The data set submitted to multivariate data analysis was formed by 322 samples of monovarietal virgin olive oil, each one described by the 14 measured parameters listed in Table 1 together with their code used in figures.

Table 1
The measured variables

No.	Code	Name
1	F.A.	Free acidity
2	P.I.	Peroxide index
3	C16:0	Palmitic acid
4	C16:1	Palmitoleic acid
5	C17:0	Heptadecanoic acid
6	C17:1	Heptadecenoic acid
7	C18:0	Stearic acid
8	C18:1	Oleic acid
9	C18:2	Linoleic acid
10	C18:3	Linolenic acid
11	C20:0	Arachidic acid
12	C20:1	Eicosenoic acid
13	C22:0	Behenic acid
14	C24:0	Lignoceric acid

The data matrix of 322 row (samples) and 14 columns (variables) is horizontally subdivided into 7 classes corresponding to the different varieties listed in Table 2. The olive samples of different cultivars came from particular areas of the region; to distinguish the influence of cultivar and environment, some samples of olives from Cassanese and Carolea were collected in different areas of Calabria.

3. Chemometric methods

Principal component analysis (PCA) (Wold, Esbensen, & Geladi, 1987) is a well-known technique to visualise data and to find the true dimension of a data set. The v parameters (variables), measured for each sample, describe each sample (object) in a v-dimensional space. PCA generates a set of new orthogonal variables (axes), linear combinations of the original ones, so that the maximal amount of variance contained in the data set (information) is concentrated in the first principal components.

The principal components, which have a significance for the interpretation of the studied problem, are called significant components, and can be used in place of the original variables for further treatments or to display the information contained in the data set. The loadings are the coefficients of the original variables defining each principal component. The scores are the co-ordinates of the objects on the new axes.

Unequal class model (Uneq; Derde & Massart, 1986) is the modelling version of the Quadratic Discriminant Analysis, a probabilistic classification method. Class-modelling techniques build a class space whose boundary discriminates between samples fitting the class model and samples that cannot be considered as belonging to the studied class. To define the class boundary that envelops the class space, it is necessary to define the mathematical model of the class and to develop some kind of confidence interval around it.

The model of Uneq is the centroid of the class and the probability distribution of samples in each class is

Table 2
Olive cultivars from Calabria and their main zones of growing

Class index	Cultivar name	Cultivar origin	Number of samples
1	Carolea	Middle Calabria	133
2	Cassanese	North Calabria	46
3	Dolce di Rossano	North East Calabria	45
4	Ottobratica	South West Calabria	36
5	Sinopolese	South West Calabria	28
6	Grossa di Gerace	South East Calabria	21
7	Roggianella	North Calabria	13

hypothesised to be a multivariate normal distribution. The category space is considered as the confidence hyperellipsoid. The class boundary is the estimate of the square of the Mahalanobis distance of an object from the class model. Mahalanobis distance has a χ^2 distribution, but, for objects in the training set, the exact distribution of the estimate of the Mahalanobis distance is a β distribution and χ^2 is an asymptotic approximation, while for the objects in the evaluation set is a Hotelling distribution (Forina, Lanteri, & Sarabia, 1995).

A critical value χ^2_p (at a selected confidence level p per cent) is obtained from the χ^2 distribution with $(V/2, (N_g - V - 1)/2)$ degrees of freedom where N_g is the number of training set objets in category g. This value is used as a measure to judge if a sample falls inside or outside the class boundary.

In this work the boundary of the model of each class was considered as the pP value ($pP = -\log$ probability) at the critical value of variable (at 95% confidence level) obtained from the estimate of the probability density of a multivariate normal distribution.

SIMCA (Soft Independent Modelling of Class Analogy; Wold & Sjostrom, 1997) is the first class modelling method introduced in chemistry, its model may be linear, piece-wise linear, planar or a multidimensional figure. The significant principal components of each category build the class model, which is computed after a separate scaling for each category (autoscaling is used in this work). The boundaries of the model are computed assuming the range of the scores on the components, normal model, which can be expanded to obtain the classic expanded range, usually when the number of objects is very small. In this paper the number of objects for each category is relatively high, so a reduced model is more suitable (Forina & Lanteri, 1994). The number of significant principal components of each category used to build the class model is that for which the explained variance for each class is 95%.

Non-error-rate, sensitivity and specificity (Frank & Todeschini, 1994) of the models were the criteria adopted for measuring the classification and modelling performance: non-error-rate is the percentage of correctly classified objects; sensitivity is the non-error-rate for a class, that is the percentage of objects of a class accepted by the class model; specificity indicates the purity of a class, that is the percentage of objects of other class rejected by the class model.

Data matrix was divided into training and evaluation set; the training set was used to build the model, and the evaluation set to test its performance; the leave-one-out validation method was used in this work to test the computed models.

Multivariate data evaluation was performed by QPARVUS 3.0 (Forina, Lanteri, & Armanino, 2000).

4. Results and discussion

As first step, principal component analysis was applied to obtain the best display of this multivariate data set submitted to autoscaling to assign the same numerical weight to each variable.

The scores of the samples (displayed by their class index, Table 1) and the loadings of the variables (displayed by the codes listed in Table 1) on the two first principal components are plotted in Fig. 1: the information retained is the 42% of the total one. Classes are delimited by the class polygons. The cultivar Cassanese, has lowest scores on component one (high contents of C18:3), Carolea, has highest scores on the same component (high contents of C17:0, C17:1, C20:0, C22:0). The second component, mainly due to variations of C18:1, C18:2, C16:0, partly divides Roggianella, Ottobratica and Grossa di Gerace, the last two classes have the highest levels of C18:2. The third principal component (15% of the total explained variance) brings the information of variables C20:1, C24:0, the fourth one (8% of the total explained variance) is prevalently formed by F.A. and C18:0; but these two principal components are not useful to distinguish some cultivars. The information due to F.A. is contained in the fourth and fifth principal component, while that of peroxide index is mainly brought by the seventh component (4% of total variance). The remaining components have almost similar loadings for all the variables, so it is difficult to individuate a significant direction, while the last components retain a very low useless percentage of total information. After a display study of the data set, the second step was the characterisation of the olive oils by classification-modelling methods.

It is difficult to characterise seven categories at the same time considering that the number of samples in each class is very different and that the ratio of objects/variables is lower than three in the most of categories.

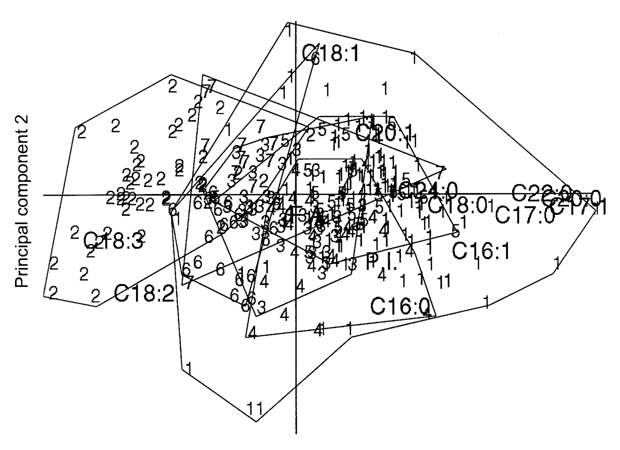

Fig. 1. Score and loading plot. Samples are plotted by the class indexes reported in Table 2. Polygons are drawn to delimit the classes.

However, SIMCA, working in the space of category components, was able to build the models of the seven cultivars, the results are shown in Table 3. Model prediction of Carolea cultivar is very high (92%), that one of Cassanese (80) and Ottobratica cultivar (67%) are high while the other ones decrease from 60 to 19%.

Afterward, to study the influence of the production area, only three cultivars out of the seven ones were selected: Cassanese, Carolea and Dolce di Rossano, for the reasons. The samples of Carolea and Cassanese were grown in several areas spread all over Calabria, so these two cultivars were selected to detect the differences related to their diffusion on the whole region. On the contrary, the samples of Dolce di Rossano come from an unique narrow area of production with similar microclimatic conditions, and the main reason to select this cultivar is that the olive oils belonging to Colline Ioniche Presilane submention of Bruzio PDO are obtained by olives of Dolce di Rossano cultivar for the most part. It is to be underlined also that the ratio objects/variables within this class is acceptable to use Uneq method.

So, in order to verify whether, with the measured chemical parameters, the discriminant information hold by variety is higher than the differentiation due to growing area, the three classes Carolea, Cassanese and Dolce di Rossano are studied by the two modelling methods, Uneq and SIMCA. The results were confirmed by two methods based on different theoretical models: the 14 dimension centroid of Uneq and the SIMCA box of the class components.

Table 3
SIMCA results for the seven cultivars

Cultivar	Classification rate (%)	Prediction rate (%)
Carolea	97.7	92.5
Cassanese	97.8	80.4
Dolce di Rossano	80.0	60.0
Ottobratica	77.8	66.7
Sinopolese	75.0	42.9
Grossa di Gerace	61.9	19.1
Roggianella	76.9	53.8

The classification and prediction rates are reported in Table 4, they were obtained with the leave-one-out method of validation. The classes Cassanese and Carolea have high classification and prediction rates, upper than 80% by Uneq and SIMCA, while the prediction rate of the models of Dolce di Rossano is lower than the other ones, particularly for the model built by SIMCA. It is noteworthy that the models of the two cultivars coming from the whole Region has a good prediction, it means that the 14 measured parameters can differentiate them; the genotype of olives characterises the obtained product. The differences among the oils depend on numerous factors, not only on the geographical origin. Beyond the prediction results, it is interesting to study more in detail the built models considering their sensitivity and their specificity.

The sensitivity of models (Table 4), computed with all the samples in the training set, are similar for the three classes, i.e. their models well classify the samples of the class; SIMCA models classify better the samples used to compute each category.

In Table 4 the specificity of the class models is reported too, while the models of the three classes correctly classify an high percentage of their samples, differences became evident considering specificities.

In the case of Uneq method the very high specificity of Cassanese towards Carolea and towards Dolce di Rossano means that the class space built by the 14 measured parameters, is well identified, that is Cassanese samples can be almost perfectly distinguished from samples of the other two classes. SIMCA specificity towards Dolce di Rossano is very low. But considering that the Cassanese model have a sufficiently good prediction rate, it is possible conclude that this cultivar is well characterised by 14 measured parameters, in fact Cassanese samples are produced on the whole region but the geographical origin does not influence them.

The specificity of Carolea model is high toward Cassanese, i.e. its samples are very different from those of Cassanese; at the same time the Carolea models have a good prediction rate, so the conclusions are the similar to those ones for Cassanese model. Instead, the specificity of Carolea model towards Dolce di Rossano is low

Table 4
Classification and modelling results for the three selected cultivars

	Classification rate (%)		Prediction rate (%)		Sensitivity (%)		Specificity(%)			
	Uneq	SIMCA	Uneq	SIMCA	Uneq	SIMCA	Uneq		SIMCA	
Carolea	97.0	98.5	93.9	94.7	85.7	91.0	vs 2	100.0	vs 2	93.5
							vs 3	46.7	vs 3	11.9
Cassanese	97.8	97.8	86.7	80.4	76.1	93.5	vs 1	98.5	vs 1	99.2
							vs 3	93.3	vs 3	30.1
Dolce di Rossano	95.6	80.0	77.8	60.0	80.0	100.0	vs 1	69.2	vs 1	99.2
							vs 2	89.1	vs 2	97.8

Specificity is reported for each class versus the other ones identified by their class index.

with both methods, so, in spite of its high prediction rate, this model accepts the samples of Dolce di Rossano too.

The specificity of Dolce di Rossano is enough high also if its model has a bit lower specificity with Uneq model for Carolea. Coomans plot (Figs. 2 and 3) was used to visualise both the sensitivity and the specificity of the models of the studied classes both their specificity towards the other four cultivars considered as test set. Some of the test set samples were classified into the three class models according to the following percentages: 35% into Carolea, 4% into Cassanese and 11% into Dolce di Rossano; while the 50% of the test set samples were correctly classified as outliers.

The Coomans plot axes correspond to the SIMCA distance of the first two class models (Fig. 2), the straight lines are drawn at the critical level of the distance of each model. The most of samples of the test set are correctly plotted in the oultier space, they are not accepted by any cultivar. The most of samples of Cassanese are accepted by their class model, few samples are accepted also by the model of Carolea, and few samples are considered as outliers. The most of samples of Carolea lie inside their class model, only one sample is accepted also by Cassanese model, and some samples are considered as outliers. The samples of Dolce di Rossano should be considered as ouliers (not accepted by both models), but a lot of them are accepted by the model of Carolea. In Fig. 3 the same consideration can be made for the model of Cassanese and Dolce di Rossano: the two model have high specificity. Using chemometric techniques it was possible to show the differences among olive oils obtained by the olives of these genotypes using the information contained in the data set.

Since other kinds of information was present in the data set (the distance from the sea and the year of harvest were known for each sample) chemometrics was to used to find the relation among olive oils and microclimate and harvest year. The distance from the sea was taken into account to find some influence of microclimate on sample of olive oils. To eliminate the variation of cultivar only the sample of the same cultivar was elaborated. Therefore, the most numerous samples of Carolea cultivars were submitted to the elaboration. Four new subcategories were built inside Carolea: category 1 consisted of 39 samples with distance from the sea of 0 km, category 2 of 28 samples with distance from the sea of 5 km, category 3 of 49 samples with distance from the sea of 10 km, category 4 of 17 samples with distance from the sea of 20 km.

One-way analysis of variance was performed to check significant differences among classes, the class differences were non significant (at a level $P=0.05$) for the most of variables. Consequently, it can be concluded that the microclimate of Calabria is enough homogeneous and the variations depending on it are negligible respect to those arising from the genotype.

The year of harvest is known for each sample, so the differences due to it were studied. The 133 oil samples of Carolea cultivar were picked in three years: 1993 (47 samples), 1994 (39 samples) and 1995 (47 samples).

The loading and score plot of the first two principal components (Fig. 4; 47% explained variance) shows the differences between year 1994 and the other two years. These differences were verified by the results of classification analysis. Quadratic discriminant analysis gave correct prediction of 73% for the category 1993, 87% for the category 1994 and 83% for the category of 1995 with leave-one-out validation method. The specificity of the models gave further information about the differences among the olive oils of harvest year 1993–1995. The model of year 1993 is highly specific towards that of year 1994, specificity towards year 1994 is 80%, but not for year 1995, the specificity towards year 1995 is 30%.

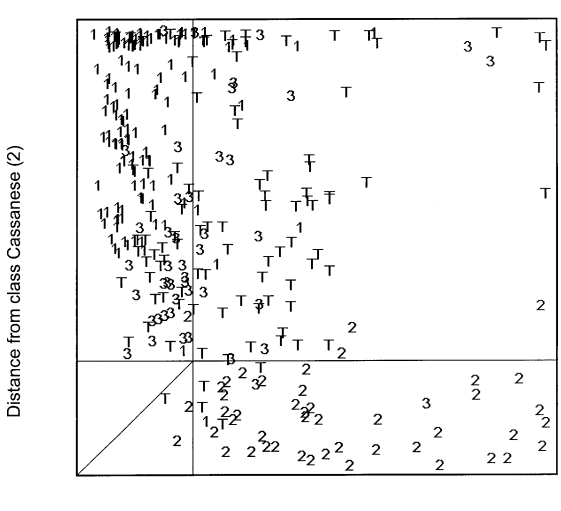

Fig. 2. Coomans plot of Cassanese and Carolea model.

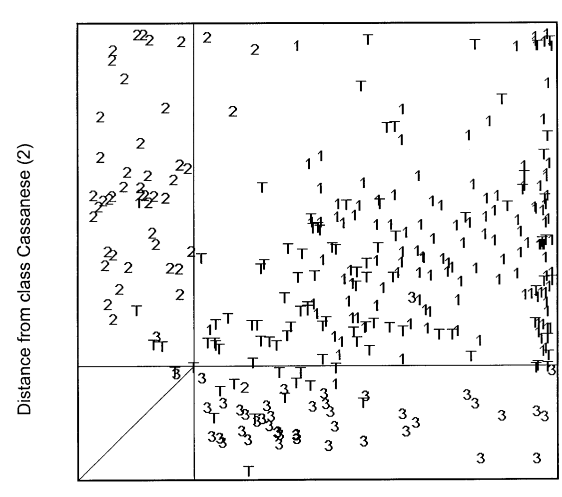

Fig. 3. Coomans plot of Cassanese and Dolce di Rossano model.

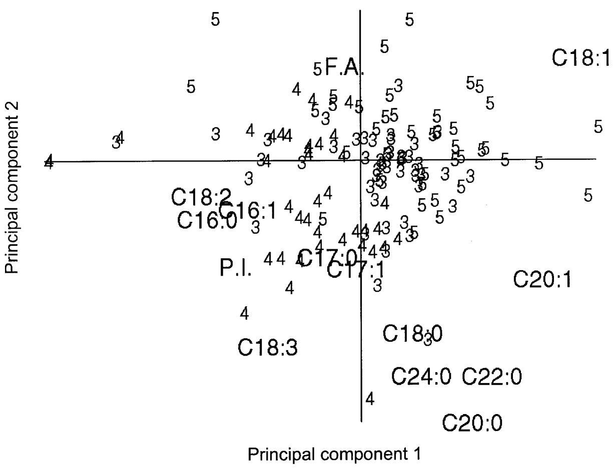

Fig. 4. Score and loading plot of the samples of the three harvest years. Samples are plotted by their harvest years: (3) 1993, (4) 1994, (5) 1995.

The model of year 1995 is highly specific towards that of year 1994, specificity towards year 1994 is 87%, but not for year 1993, the specificity towards year 1993 is 43%. From the values of specificity of the classes it is evident that the two models of years 1993 and 1995 are different from the model of 1994, i.e. the quality of olive oil is similar for the two harvest year.

5. Conclusions

The chemical information provided by the 14 variables measured on virgin olive oil samples from three important cultivars of Calabria enabled us to distinguish mainly the olive oil samples of Cassanese cultivar, traditionally cultivated near Cassano in the north of Calabria. This means that it is possible to distinguish and protect the olive oils belonging to the Sibaritide submention of Bruzio PDO. The differentiation of the olive oils belonging to Colline Ioniche Presilane submention of Bruzio PDO is not so evident. In the case of samples from Carolea only the specificity towards Cassanese cultivar was high, whereas the specificity towards Dolce di Rossano was low, probably because of more close genetic distance among them.

In addition, the results of the present work suggest that since the microclimate of Calabria is enough homogeneous, its influence on fatty acids is negligible with respect to influence of the genotype. Under these conditions any discrimination of olive oil samples based on geographical area of origin is very difficult.

Finally, the statistical analysis also allowed us to highlight the most significant variables for the data set, thus confirming the usefulness of fatty acids to distinguish the monovarietal olive oils belonging to some particular cultivars. Therefore, the combination of simple and relatively inexpensive analytical parameters such as fatty acids and chemometric techniques represents a useful tool to characterise the Calabrian olive oils.

Acknowledgements

This work was supported by research grants from the Italian Ministry of University and Scientific Research, the National Council of Research and the Calabria Region.

References

Alonso Garcia, M. V., & Aparicio Lopez, R. (1993). Characterization of european virgin olive oils using fatty acids. *Grasas y Aceites*, *44*(1), 18–24.

Armanino, C., Leardi, R., Lanteri, S., & Modi, G. (1998). Chemometric analysis of Tuscan olive oils. *Chemometrics and Intelligent Laboratory Systems*, *5*, 343–354.

De Nino, A., Mazzotti, F., Morrone, S. P., Perri, E., Raffaelli, A., & Sindona, G. (1999). Characterisation of Cassanese olive cultivar throug the identification of new trace components by ionspray tandem mass spectrometry. *Journal of Mass Spectrometry*, *34*, 10–16.

De Nino, A., Perri, E., Lombardo, N., Procopio, A., Raffaelli, A., & Sindona, G. (1997). Direct identification of phenolic glucosides from leaves extracts by ionspray ionization tandem mass spectrometry. *Journal of Mass Spectrometry*, *32*, 533–541.

Derde, M. P., Coomans, D., & Massart, D. L. (1984). SIMCA (Soft independent modelling of class analogy). Demonstrated with characterisation and classification of Italian olive oil. *Journal of the Association of Official Analytical Chemists*, *67*, 721–727.

Derde, M. P., & Massart, D. L. (1986). UNEQ: a disjoint modelling technique for pattern recognition based on normal distribution. *Analytica Chimica Acta*, *184*, 33–51.

Drava, G., Forina, M., Lanteri, S., & Lupoli, M. (1994). Development of the chemical model of a typical food product: olive oil from an Italian region (Basilicata). *Journal of Science and Food Agriculture*, *65*, 21–30.

Forina, M., & Armanino, C. (1982). Eigenvector projection and simplified non linear mapping of fatty acid content of Italian olive oils. *Annali di Chimica*, *72*, 127–141.

Forina, M., & Lanteri, S. (1994). Data analysis in food chemistry. In B. R. Kowalski (Ed.), *Chemometrics. Mathematics and statistics in chemistry* (pp. 305–349). Seattle: Reidel.

Forina, M., Lanteri, S., & Armanino, C. (2000). *QPARVUS 3.0: an extendable package of programs for data explorative analysis, classification and regression analysis*. Genova, Italy: Department of Chimica e Tecnologie Farmaceutiche e Alimentari, University of Genova. (Free download available at: http://PARVUS.unige.it).

Forina, M., Lanteri, S., & Sarabia, L. (1995). Distance and class space in the UNEQ class-modeling technique. *Journal of Chemometrics*, *9*, 69–89.

Forina, M., & Tiscornia, E. (1982). Pattern recognition methods in the prediction of Italian olive oil origin by their fatty acid content. *Annali di Chimica*, *72*, 143–155.

Frank, I. E., & Todeschini, R. (1994). *The data analysis handbook*. Amsterdam: Elsevier.

Leardi, R., & Paganuzzi, V. (1987). Caratterizzazione dell'origine di oli di oliva extra vergini mediante metodi chemiometrici applicati alla frazione sterolica. *La Rivista Italiana delle Sostanze Grasse*, *64*, 131–136.

Montedoro, G. F., Servili, M., Baldioli, M., Selvaggini, R., Perretti, G., Magnarini, C., Cossignani, L., & Daminai, P. (1995). Characterization of some Italian virgin olive oils in relation to origin area. *La Rivista Italiana delle Sostanze Grasse*, *72*, 403–414.

Parlati, M. V., Perri, E., Palopoli, A., & Rizzuti, B. (1994). Further observations on oil quality of calabrian olive cultivars. *Acta Horticulturae*, *356*, 327–330.

Perri, E., Parlati, M. V., Palopoli, A., Pellegrino, M., & Rizzuti, B. (1999). Characterization of Italian olive oils using fatty acids. *Acta Horticulturae*, *474*(2), 627–630.

Perri, E., Raffaelli, A., & Sindona, G. (1999). Quantitation of oleuropein in virgin olive oil by ionspray mass spectrometry-selected reaction monitoring. *Journal of Agriculture and Food Chemistry*, *47*(10), 4156–4160.

Perri, E., Sirianni, R., & Pellegrino, M. (1998). Selezione varietale dell'olivo in Calabria. Caratterizzazione della cultivar 'Tonda Dolce'. In *Abstracts of Congress Biodiversità: germoplasma locali e loro valorizzazione*, Alghero, 9–11 September.

Perri, E., Sirianni, R., Godino, G., & Tartarini, S. (1999). Characterization of Italian olive (*Olea europaea* L.) cultivars by statistical and molecular methods. *Acta Horticulturae*, *474*(2), 489–493.

Sacchi, R., Patumi, M., Fontanazza, G., Barone, P., Fiordiponti, P., Mannina, L., Rossi, E., & Segre, A. L. (1996). A high-yield 1H Nuclear Magnetic Resonance study of the minor components in virgin olive oils. *Journal of American Oil Chemist's Society*, *73*(6), 747–758.

Shaw, A. D., Di Camillo, A., Vlahov, G., Jones, A., Bianchi, G., Rowland, J., & Kell, D. B. (1997). Discrimination of the variety and region of origin of extra virgin olive oils using 13C NMR and multivariate calibration with variable reduction. *Analytica Chimica Acta*, *348*, 1–3, 357–374.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, *2*, 37–52.

Wold, S., & Sjostrom, M. (1997). SIMCA: a method for analyzing chemical data in term of similarity and analogy. In B. R. Kowalski (Ed.), *ACS Symposium Series*, No. 52.