Research Article

Classification of monovarietal Sardinian Extra Virgin Olive Oils by ^{1}H NMR Metabolomic†

Running Title ^{1}H NMR Characterization of Sardinian EVOO

Nicola Culeddu[1*], Matilde Chessa[1], Giovanni Bandino[2], Piergiorgio Sedda[2], Roberto Zurru[2], Roberto Anedda[3], Andrea Motroni[4], Maria Giovanna Molinu[5], Sandro Dettori[6], Mario Santona[6]

[1]Institute of Biological Chemistry – National Council of Research, Italy;

[2]AGRIS Sardegna, Agricultural Research Agency of Sardinia Department for research on Wood and Fruit tree, Cagliari, Italy;

[3]Porto Conte Ricerche Srl, SP 55 Porto Conte/Capo Caccia, Tramariglio-Alghero, SS, Italy

[4]ARPAS - Sardinia Environmental Protection Agency Department of Meteorology and Climatology, Sassari, Italy;

[5]Institute of Sciences of Food Production – National Council of Research, Italy

[6]University of Sassari, DIPNET-Department of Science for Nature and Environmental Resources, Sassari, Italy

[*]**Corresponding author**

Dott. Nicola Culeddu

CNR-ICB UOS Sassari, Traversa La Crucca, 3, I-7100 Sassari

Phone: +39 0792841211

Fax: +39 0792841299

E-mail: nicola.culeddu@icb.cnr.it

†This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [**10.1002/ejlt.201700035**].

Received: January 23, 2017 / Revised: July 26, 2017 / Accepted: August 14, 2017

Abstract

The chemical composition affects the sensory properties and quality of Extra Virgin Olive Oil (EVOO). In addition, its knowledge can supply valuable information about the cultivar and its geographical origin. The goal of this study is to obtain a protocol in order to be able to recognize the composition of various Sardinian oils and to consequently correlate them with their production areas. High-resolution ^{1}H NMR spectroscopy was used to analyse 100 (82 training + 18 test sets) samples of EVOO from the Bosana cultivar, collected from different growing areas in Sardinia (Italy). Growth areas were classified on the basis of FAO-UNEP aridity index. NMR data were processed with multivariate statistical analysis. NMR profiling presented a connection between environmental factors of Sardinian cultivation areas and the chemical composition of EVOO. An NMR-based metabolomic approach that used six "one-to-one" OPLS-DA models allowed us to discriminate the different influence of evapotranspiration, solar exposure and altitude on the chemical composition of Bosana EVOO.

Practical applications: Detailed knowledge of NMR spectra pattern variations could have a potential impact on olive oil market. The application of the ^{1}H NMR metabolomic, based on chemometric models, can be a useful tool in order to certificate the geographical origin of EVOO.

Keywords
Composition, Geographic Origin, Protected Denomination of Origin, Food Security, OPLS-DA.

Abbreviations

NMR - nuclear magnetic resonance; PCA - Principal Component analysis; OPLS-DA - Orthogonal Projections to Latent Structures Discriminant Analysis; VIP - Variable Influence on Projection; PDO - Protected Denomination of Origin; ETo - Potential Evapotranspiration; TG – triglycerides; DG– diglycerides.

Introduction

Olive oil production is one of the main agricultural activities in almost all the countries belonging to the Mediterranean area; its presence produces several environmental, social and economic effects.
The olive oil classification concerns quality grade, genuineness (extra virgin, virgin, and lamp); it is essentially based on aroma and taste, peroxide number, spectrophotometric constants, free fatty acid and other chemical constituents, as coded by European Community L248 Regulation [1].
Extra Virgin Olive Oil (EVOO) is obtained through the cold-pressing of the olive paste in order to obtain some peculiar physiochemical properties and its quality is influenced by several factors including genetic, technological, agronomic and environmental conditions [2]. Oil profile of olives grown at different altitudes shows divergences in oleic, linoleic and palmitic acid percentage [3]. Furthermore, temperature and humidity of growth area atmosphere affect the harvest date and consequently the fat composition in EVOO [4, 5] as well as the concentration of 1,2 diacylglicerol esters [6], squalene and sterols content [7, 8] and volatiles profile [9-11]. Higher production of unsaturated fats might be interpreted as a mechanism of cold adaptation involving lipids storage [12]. Various studies have recently been published on the determination of chemical parameters able to discriminate olive cultivars and growth areas native to the Mediterranean basin [13-16].
Extra virgin olive oil characterization has been carried out using multivariate statistics applied to chemical parameters resulting from many analytical techniques (e.g. NMR, Chromatography and Mass Spectrometry) [17-19]. The combination of gas chromatography and statistical multivariate techniques allowed us to discriminate olive oils produced in the North-central areas of Sardinia (which is an island located in the middle of the Mediterranean Sea) from the ones produced in South-eastern areas [20, 21]. High resolution NMR spectroscopy has been successfully applied to the geographical classification of EVOO, through the analysis of the NMR data using chemometric techniques [22-31]. The metabolomic technique implicates the use of different tools (analytical determinations and statistical analysis) for the simultaneous study of several metabolites (biomarkers) of less than 2 kDa in complex biological matrices. Selected metabolites create a network which defines the biochemical basis of the physiological process we were taking into account [32].
In this study, Bosana EVOOs from Sardinia (Italy) were studied through an NMR-based metabolomics approach, in order to distinguish the samples based on the olive growing areas. Bosana is the most widespread olive cultivar in Sardinia, and Bosana EVOO is protected by the European Union Protected Designation of Origin, “PDO Sardegna” regulation (alone or in combination with other Sardinian cultivars for at least 80 %) [1]. Multivariate classification model can actually represent a helpful instrument to strengthen the definition of PDO Sardegna EVOO.

2 Material and methods

2.1 Sampling

100 monovarietal samples of EVOO were selected by "AGRIS Sardegna" (Regional Agency for Agricultural Research). Olive (*Olea europaea* L.) cultivar Bosana was obtained from rainfed orchards. All oil samples were produced from olives harvested at optimum level of ripeness for Bosana cultivar (Jaen Maturity Index between 2 and 3). The olive oil production was performed in either two-phase or three-phase mills. Figure 1 shows the four growing areas in Sardinia, namely Sassarese (area 1), Nuorese- Baronia (2), Campidano (3) and Planargia-Montiferru (4). We then randomly selected 18 samples (test set) for validation purposes. Growth areas distribution of samples are reported in Table 1.

2.2 Classification of growing areas

A general climate classification of each area was performed by analysing both the agroclimatic index (FAO-UNEP aridity index) and the Rivas-Martinez bioclimatic classification. The Aridity Index was calculated as the ratio between average annual (or monthly) precipitation amounts (P) and average potential evotranspiration (ETo)*. Based on aridity index, an area is classified as "arid" when average rainfall does not reach 50% of evotranspiration from soil-plant systems [33]. Potential Evapotranspiration (ETo)* was calculated with Hargreaves and Samani equation based on extraterrestrial global radiation. These results are confirmed by the prevalent bioclimatic type of the study areas as classified by Rivas Martinez applied to Sardinian climatic conditions [34]. Temperature and precipitation data are related to the period between 1971 and 2010 (http://www.sar.sardegna.it).

2.3 Chemicals

Formaldehyde, pentanal, n-hexanal, squalene, 2-E-hexenal, formate, β-sitosterol, cycloartenol, , chloroform-d, and DMSO-d6 were purchased from Sigma-Aldrich S.r.l. Milan, Italy.

2.4 NMR analysis

For the NMR experiments, each sample was prepared following a procedure previously published by

Mannina et al. [16]. Briefly, 20 µl of olive oil were placed into a 5mm NMR tube, mixed with 700 µl of chloroform-d; (Aldrich, USA) and 20 µl of DMSO-d6 (Aldrich, USA). ^{1}H-NMR spectra were recorded on a Bruker AVANCE II 600 spectrometer (Bruker, Germany) operating at 600.13 MHz for 1H resonance. A standard ^{1}H one-dimensional (zg sequence) NMR experiment was performed using 256 scans, a 90° pulse (9.5 µsec), power level 63 W. Acquisition time was 2.0 sec followed by a 2.0 sec delay for a total of 18 min, spectral width was 12 ppm and 128K data point for acquisition and processing were used. Prior to applying Fourier transform (FT), the FIDs were multiplied by an exponential function corresponding to a line broadening of 0.3 Hz. Chemical shifts (c.s.) were calibrated respect to $CDCl_3$ (7.27 ppm). All ^{1}H NMR spectra were manually phased and baseline corrected within TOPSPIN 2.0 (Bruker GmbH, Karlsruhe, Germany). Data were reduced to 469 integrated regions corresponding to the chemical shift range of 10-0.5 with a region width of 0.02 ppm using Analysis of MIXtures (AMIX) (Bruker GmbH, Karlsruhe, Germany), $CDCl_3$ solvent signal and the extreme noise zone buckets were excluded. The most important variable was assigned on the basis of literature data [35-38] and by direct addition of standards.

2.5 Statistical Methods

Before any chemometric analysis, data were normalized by setting the total region of each spectrum to 100 and the resulting ASCII file was imported into Microsoft EXCEL for the addition of labels. The bucket matrix was imported into SIMCA-P software version 13.0, (Umetrics AB, Umea, Sweden) for statistical analysis. Data were centered using CRT function in SIMCA-P. Multivariate methods based on projections were applied for data analysis and modelling. PCA (Principal Component Analysis) data analysis was performed for exploratory purposes and outliers recognition, while Projection to Latent Structures (PLS)-based methods were used for discriminant analysis and data set comparison. We used Orthogonal extension of PLS-DA [39] in which the first latent variable took into account only correlated data variations.

OPLS-DA models were evaluated using the goodness-of-fit parameter (R2Y) and the predictive ability parameter (Q2Y), which was calculated by a seven-round internal cross-validation. R2Y represents the proportion of variance explained by a given component in the model, whereas Q2Y is defined as the proportion of variance in the data predictable by the PLS-DA model under cross-validation.

To prove the reliability of our models and avoid over-fitting, a 7-fold full cross-validation was performed, as well as a permutation test on the class response, in accordance with good practice for model validation [40] (see supplementary materials S1). The classification of a test data set using the model-parameters based on the training data set provides information about robustness and

generalization of models [41].

3 Results and discussion

3.1 Climatological classification of olive areas

According to FAO-UNEP Aridity Index results, Sardinia is then divided in 4 main areas of oil production (Figure 1).
Dry sub-humid class is dominant in zone 1 and 2, while zone 3 is mostly characterized by arid conditions. Zone 4 is mainly classified as sub-humid.

3.2 NMR spectra

NMR spectra of olive oil have been previously described in detail [11, 16]. NMR signals refer to organic compounds such as fatty acids, glycerol, aldehydes responsible for olfactory properties (fruity, artichoke, freshly mown grass and almond), polyphenols (responsible for bitter and pungent taste) and other alcohols such as β-sitosterol and cycloarthenol with nutritional interest. A typical NMR spectrum is shown with expansions of aliphatic and aromatic regions (figure 2).

3.3 Statistical analysis using Orthogonal Projections to Latent Structures (OPLS-DA)

Regarding the mill technology, no statistically significant differences were found between “two-phase” and “three-phase” samples (data not shown). The discriminant functions have been generated using the stepwise method. In this case, the analysis has been carried out on all the available oils to test the efficiency of this method in comparison with PCA, with the aim of authenticating the area of oils origin. Four-class model shows a good separation of samples (see supplementary materials S2).
Usually, projection methods for classification, such as PLSDA or O2PLS-DA, are able to produce efficient classification models for not more than 4-5 classes of samples. It is not the easiest task to identify specific putative biomarkers for each class, mostly due to the simultaneous influence of all the four classes on the global model. That is why we agreed on preferring “one-to-one” strategy in this paper by comparing pairwise models. Six OPLS-DA models of NMR data (Figure 3) were then applied to maximize the discrimination of growing areas and to focus on metabolic variation. The class separation is concentrated into the first predictive component; it is sufficient to interpret its loadings to find out which descriptors have the best discriminatory power.

Table 3 summarizes the results of OPLS-DA for each pairwise model. The first five leading separation metabolites (in terms of Variable Influence on Projection, [39]) in pairwise models are summarized in Table 4.

3.4 Pairwise models

All models show a good discrimination among classes; we examined each model, obtaining different characterizing metabolites.

OPLS-DA pairwise model between Zone 1 and Zone 2 (i.e. Model 1,2), distinguishes EVOOs by higher fatty acid contents from Zone 1. The high fat content in oil for Zone 1 derives from the effect of high temperature and low humidity [3].

Model 1,3 separate the EVOO samples on the basis of fat content and higher squalene in Zone 3 than in Zone 1. In previous studies, Romero et al. [42] report as volatile composition is mainly affected by pedoclimatic conditions. Furthermore Salvo et al. [7] report that the variation in the squalene contents in EVOO from ten different countries may be affected by geographic and climatic conditions; Fernàndez- Cuesta et al. [12] show instead that squalene content in oil becomes constant at optimum level of olive ripeness.

In model 1,4 we found bis-allylic group and mono-unsaturated fatty acids as putative biomarkers of growth zone. The synthesis of certain fatty acid classes during the different ripening stages can be explained by the conversion between different classes by means of specific enzymes and it is reported as influenced by bioclimatic conditions [4]. Thus, the higher production of unsaturated fats in the Zone 4 might be interpreted as a mechanism of adaptation mainly involving storage lipids, as reported by Matteucci et al. [5].

Model 2,3 stand out thanks to higher fatty acid contents in EVOO from Zone 3. The high fat content in oil for Zone 3 derives from the effect of high temperature [3].

In model 2,4, the Zone 4 is characterized by high production of unsaturated fats, as discussed in model 1,4.

Model 3,4 is characterized by aridity (high temperature and low humidity) effect on EVOO from zone 3, as indicated by the higher fat content [3].

Based on our results, we can describe the effect of different climatic areas (classified by FAO UNEP index) on composition EVOO from Bosana cultivar growth in Sardinia. The differences between the classes could be explained as a combination of availability of water (evapotranspiration), solar exposure and altitude. Evapotranspiration slows down the ripening process of olive resulting in a (sometimes even partial) limitation of TG synthesis. Solar exposure accelerates the ripening process

of olive resulting in high levels of saturated fat and content of oil in drupes.

Drupes grown at higher altitudes present higher monounsaturated fatty acids amounts than those coming from lower altitudes [43]. Similar results were reported by Osman et a1. [44] and Deidda et al. [45].

4 Conclusions

Many adaptive growing and physiological mechanisms allow plants to efficiently respond to environmental changes of every cultivation zone. High resolution ^{1}H NMR allowed to recognize extra virgin olive oils produced in different growing areas of Sardinia on the base of chemical information. Determination of fatty acids, volatile compounds and acylglycerols compositions, in combination with statistical methods of data processing, can lead to characterize the most important monocultivar oils in Sardinia, Bosana, on the base of their production area. Although, it is still difficult to decompose the influence of each environmental parameter on the composition of olive oil; the data obtained rationalize dependence of the Bosana EVOO composition on water availability (evapotranspiration), solar exposure and altitude. These results are the first step towards the definition of more precise analytical profiles characterizing the various monovarietal Sardinian oils.

This methodology proposes the construction of chemometric models for the division into production areas, which could be a key tool for the allocation of the EU Protected Designation of Origin and in order to safeguard valuable products.

The authors have declared no conflict of interest

Acknowledgements

We wish to thank Prof. Mario Falchi for his scientific support and helpful discussions.

References

[1] A.VV.: EC 1991 Regulation European Community. Regulation, *Off. J. European Community.* 1991, 2568/9, L248.

[2] El Riachy, M., Priego-Capote, F., León, L., Rallo, L., de Castro, L., Dolores, M., Hydrophilic antioxidants of virgin olive oil. Part 2: Biosynthesis and biotransformation of phenolic compounds in virgin olive oil as affected by agronomic and processing factors. *Eur. J. Lipid Sci. Technol.* 2011, *113*, 692-707.

[3] Beltràn, G., Del Rio, C., Sanchez, S., Martìnez, L., Influence of Harvest Date and Crop Yield on the Fatty Acid Composition of Virgin Olive Oils from Cv. Picual. *J. Agric. Food Chem*. 2004, *52*, 3434-3440.

[4] Esmaeili, A., Shaykhmoradi, F., Naseri, R., Comparison of oil content and fatty acid composition of native olive genotypes in different region of Liam, Iran. *Intl. J. Agri. Crop Sci.* 2012, *4*, 434-438.

[5] Matteucci, M., D'Angeli, S., Errico, S., Lamanna, R., Perrotta, G., Altamura, M., Cold affects the transcription of fatty acid desaturases and oil quality in the fruit of Olea europaea L. genotypes with different cold hardiness.

J. Exp. Bot. 2011, *62*, 3403-3420.

[6] Salvo, A., Rotondo, A., La Torre, G. L., Cicero, N., Dugo, G., Determination of 1,2/1,3-diglycerides in Sicilian extra-virgin olive oils by 1H-NMR over a one-year storage period. Nat. Prod. Res. 2017, 31, 822-828.

[7] Salvo, A., La Torre, G. L., Rotondo, A., Mangano, V., Casale, K. E., Pellizzeri V., Clodoveo, M. L., Corbo, F., Cicero, N., Dugo, G., Determination of Squalene in Organic Extra Virgin Olive Oils (EVOOs) by UPLC/PDA Using a Single-Step SPE Sample Preparation. Food Anal. Methods. 2017, 5, 1377-1385.

[8] Tranchida, P.Q., Salivo, S., Bonaccorsi, I., Rotondo A., Dugo P., Mondello, L., Analysis of the unsaponifiable fraction of lipids belonging to various milk-types by using comprehensive two-dimensional gas chromatography with dual mass spectrometry/flame ionization detection and with the support of high resolution time-of-flight mass spectrometry for structural elucidation. *J. Chromatogr. A*, 2013, *1313*, 194–201.

[9] Flamini, G., Cioni, P. L., Morelli, I., Volatiles from Leaves, Fruits, and Virgin Oil from Olea europaea Cv. Olivastra Seggianese from Italy. *J. Agric. Food Chem.* 2003, *51*, 1382-1386.

[10] Rotondo, A., Salvo, A., Giuffrida, D., Dugo, G., Rotondo, E., NMR analysis of aldehydes in sicilian extra-virgin olive oils by DPFGSE techniques. *AAPP Atti della Accademia Peloritana dei Pericolanti, Classe di Scienze Fisiche, Matematiche e Naturali.* 2011, *89*, C1A8901002-2–8.

[11] Dugo, G., Rotondo, A., Mallamace, D., Cicero, N., Salvo, A., Rotondo, E., Corsaro, C., Enhanced detection of aldehydes in extra-virgin olive oil by means of band selective NMR spectroscopy. *Physica A: Statistical Mechanics and its Applications.* 2015, *420*, 258-264.

[12] Fernández-Cuesta, A., León, L., Velascoa, L., De la Rosa, R., Changes in squalene and sterols associated with olive maturation. *Food Res. Int.* 2013, *54*, 1885-1889.

[13] Aparicio, R., Morales, M. T., Alonso, V., Authentication of European virgin olive oils by their chemical compounds, sensory attributes, and consumers' attitudes. *J. Agr. Food Chem.* 1997, *45*, 1076-1083.

[14] Dugo, G., Lo Curto S., Alfa M., Di Bella G., Lo Turco V., Mavrogeni E., Saitta M., Salvo F., Pizzimenti G., R., M., Caratterizzazione di oli di oliva vergini siciliani. *Riv. Ital. Sostanze Gr.* 2000, *77*, 803-833.

[15] Boggia, R., Zunin, P., Lanteri, S., Rossi, N., Evangelisti, F., Classification and class-modeling of "Riviera Ligure" extra-virgin olive oil using chemical-physical parameters. *J. Agr. Food Chem.* 2002, *50*, 2444-2449.

[16] Mannina, L., Patumi, M., Proietti, N., Bassi, D., Segre, A. L., Geographical characterization of Italian extra virgin olive oils using high-field 1H NMR spectroscopy. *J. Agr. Food Chem.* 2001, *49*, 2687-2696.

[17] Mincione B., Giuffrè A. M., Leuzzi U., Lombardo N., Ricerche sugli oli di oliva monovarietali. Nota I. Caratterizzazione della produzione della Calabria. *Riv. Ital. Sostanze Gr.* 1994, *71*, 335-339.

[18] Armanino, C., Leardi, R., Lanteri, S., Modi, G., Chemometric analysis of Tuscan olive oils. *Chemometr. Intell. Lab.* 1989, *5*, 343-354.

[19] Boschelle, O., Giomo, A., Conte, L., Lercker, G., Caratterizzazione delle cultivar di olivo del Golfo di Trieste mediante metodi chemiometrici applicati ai dati chimico-fisici. *Riv. Ital. Sostanze Gr.* 1994, *71*, 57-65.

[20] Scano P, Lai A, Bandino G, Deiana, M, Dessi, M. A., Sardinian Olive Oils - A territorial investigation based on the fatty acids composition. *Riv. Ital. Sostanze Gr.* 2001, *78*, 629-632.

[21] Vacca V, Fenu P, Franco M A, Sferlazzo G., Primo approccio alla caratterizzazione dei composti fenolici di oli vergini di oliva della Sardegna: elaborazioni statistiche multivariate dei risultati in HPLC. *Riv. Ital. Sostanze Gr.* 1993, *70*, 595-599.

[22] Sacchi, R., Patumi, M., Fontanazza, G., Barone, P., Fiordiponti, P., Mannina, L., Rossi, E., Segre, A., A high-field ^{1}H nuclear magnetic resonance study of the minor components in virgin olive oils. *J. Am. Oil Chem. Soc.* 1996, *73*, 747-758.

[23] Scano, P., Casu, M., Lai, A., Saba, G., Dessi, M. A., Deiana, M., Corongiu, F. P., Bandino, G., Recognition and quantitation of cis-vaccenic and eicosenoic fatty acids in olive oils by ^{13}C nuclear magnetic resonance spectroscopy. *Lipids.* 1999, *34*, 757-759.

[24] Bentivenga, G., D'auria, M., Viggiani, F., ^{1}H NMR characterization of virgin olive oils from Basilicata. *Riv. Ital. Sostanze Gr.* 2003, *80*, 171-172.

[25] Sacco, A., Brescia, M. A., Liuzzi, V., Reniero, F., Guillou, G., Ghelli, S., Van der Meer, P., Characterization of Italian olive oils based on analytical and nuclear magnetic resonance determinations. *J. Am. Oil Chem. Soc.* 2000, *77*, 619-625.

[26] Mannina, L., Patumi, M., Proietti, M. N., Segre, A. L., P.O.D. (Protected Designation of Origin): Geographical characterization of Tuscan extra virgin olive oils using High Field ^{1}H NMR spectroscopy. *Ital. J. Food Sci.* 2001, *13*, 53-63.

[27] Karkoula, E., Skantzari, A., Melliou, E., Magiatis, P., Quantitative measurement of major secoiridoid derivatives in olive oil using qNMR. Proof of the artificial formation of aldehydic oleuropein and ligstroside aglycon isomers. *J. Agr. Food Chem.* 2014, *62*, 600-607.

[28] Del Coco, L., Pascali, D., Angelica, S., Iacovelli, V., Cesari, G., Schena, F. P., Fanizzi, F. P., Following the olive oil production chain: 1D and 2D NMR study of olive paste, pomace, and oil. *Eur. J. Lipid Sci. Technol.* 2014, *116*, 1513-1521.

[29] Piccinonna, S., Ragone, R., Stocchero, M., Del Coco, L., De Pascali, S. A., Schena, F. P., & Fanizzi, F. P., Robustness of NMR-based metabolomics to generate comparable data sets for olive oil cultivar classification. An inter-laboratory study on Apulian olive oils. *Food Chem*. 2016, *199*, 675-683.

[30] Del Coco, L., Mondelli, D., Mezzapesa, G. N., Miano, T., De Pascali, S. A., Girelli, C. R., Fanizzi, F. P., Protected designation of origin extra virgin olive oils assessment by nuclear magnetic resonance and multivariate statistical analysis: "Terra di Bari", an apulian (southeast Italy) case study. *J. Am. Oil Chem. Soc.* 2016, *93*, 373-381.

[31] Mannina, L., Dugo, G., Salvo, F., Cicero, L., Ansanelli, G., Calcagni, C., Segre, A., Study of the cultivar-composition relationship in Sicilian olive oils by GC, NMR, and statistical methods. *J. Agr. Food Chem.* 2003, *51*, 120-127.

[32] Hong, J., Yang, L., Zhang, D., Shi, J., Plant Metabolomics: An Indispensable System Biology Tool for Plant Science. *Int. J. Mol. Sci.* 2016, *17*, 767.

[33] Motroni, A., Canu, S., Bianco, G., Loj, G., Monitoring sensitive areas to desertification in Sardinia: The contribute of the Regional Agrometeorological Service. In: Desertification and risk analysis using high and medium resolution satellite data, Springer Netherlands, 2009, pp. 117-128.

[34] Canu, S., Rosati, L., Fiori, M., Motroni, A., Filigheddu, R., Farris, E., Bioclimate map of Sardinia (Italy). *J. Maps.* 2015, *11*, 711-718.

[35] Nieva-Echevarría, B., Goicoechea, E., Manzanos, M. J., Guillén, M. D., A method based on ^{1}H NMR spectral data useful to evaluate the hydrolysis level in complex lipid mixtures. *Food Res. Int.* 2014, *66*, 379-387.

[36] Dais, P., Hatzakis, E., Quality assessment and authentication of virgin olive oil by NMR spectroscopy: A critical review. *Anal. Chim. Acta* 2013, *765*, 1-27.

[37] Alonso-Salces, R. M., Heberger, K., Holland, M. V., Moreno-Rojas, J. M. Mariani, C., Bellan, G., Reniero, F., Guillou, C. Multivariate analysis of NMR fingerprint of the unsaponificable fraction of virgin olive oils for authentication purposes. *Food Chem.* 2010, *118*, 956-965.

[38] D'Imperio, M., Mannina, L., Capitani, D., Bidet, O., Rossi, E., Bucarelli, F. M., Quaglia, G. B., Segre, A., NMR and statistical study of olive oils from Lazio: a geographical, ecological and agronomic characterization. *Food Chem.* 2007, *105*, 1256-1267.

[39] Trygg, J., Wold, S., Orthogonal projections to latent structures (O-PLS). *J. Chemom.* 2002, *16*, 119-128.

[40] Westerhuis, J. A., Hoefsloot, H. C., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J., van Duijnhoven, J. P., van Dorsten, F. A., Assessment of PLSDA cross validation. *Metabolomics.* 2008, *4*, 81-89.

[41] Cloarec, O., Dumas, M. E., Trygg, J., Craig, A., Barton, R. H., Lindon, J. C., Nicholson, J. K., Holmes, E., Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H NMR spectroscopic metabonomic studies. *Anal. Chem.* 2005, *77*, 517-526.

[42] Romero, N., Saavedra,J., Tapia,F., Sepúlveda,B., Aparicio, R., Influence of agroclimatic parameters on phenolic and volatile compounds of Chilean virgin olive oils and characterization based on geographical origin, cultivar and ripening stage. *J. Sci. Food Agric.* 2016, *96*, 583–592.

[43] Di Vaio, C., Nocerino, S., Paduano, A., Sacchi R., Influence of some environmental factors on drupe maturation and olive oil composition. *J. Sci. Food Agric*. 2013, *93*, 1134-1139.

[44] Osman, M., Metzidakis, I., Gerasopoulos, D., Kiritsakis, A., Qualitative changes in olive oil of fruits collected from trees grown at two altitudes. *Riv. Ital. Sostanze Gr*. 1994, *71*, 187-190.

[45] Deidda P., Nieddu G., Spano D., Bandino G., Orru V., Solinas M., Serraiocco A., Olive oil quality in relation to environmental conditions. *Acta Hortic* "II International Symposium on Olive Growing" 1994, 354-357.

Tables

Area/Samples	Training Set	Test Set
1 Sassarese (SS)	36	6

2 Nuorese - Baronia (NB)	25	6
3 Campidano (CA)	9	2
4 Planargia – Montiferru (PM)	12	4

Tab. 1 Growth Zones distribution of samples.

CLASS	**ZONE 1**	**ZONE 2**	**ZONE 3**	**ZONE 4**
ARID	4,9	3,6	60,9	0,6
DRY SUB-HUMID	75,4	62,3	22,5	38,3
SUB-HUMID	19,7	27,3	15,2	56,5
HUMID	0,0	6,7	1,4	4,7

Tab. 2 Calculated FAO-UNEP Aridity Index for the 4 sampling areas.

	OPLS-DA				PLS-DA	
Model	R2Y	Q2Y	Cross Valid.	Ext. Valid.	R2Y	Q2Y
4 areas (1+3)	0.834	0.567	89%	83%	0.855	0.488
1 vs 2 (1+2)	0.864	0.648	100%	67%	0.864	0.625
1 vs 3(1+5)	0.917	0.568	100%	63%	0.917	0.481
1 vs 4(1+2)	0.989	0.941	100%	80%	0.989	0.921
2 vs 3(1+3)	0.978	0.692	100%	88%	0.905	0.607
2 vs 4(1+4)	0.853	0.623	97%	70%	0.891	0.686
3 vs 4(1+1)	0.982	0.918	100	83%	0.982	0.908

Tab. 3 Autofit results for 4 class model and pairwise models. Cross validation and external validation results. Numbers in bracket are the number of latent variables.

Growth Zone	2 (NB)	3 (CA)	4 (PM)

1 (SS)	Saturated alcohols (3.82) ↑ ^{13}C satellite signal of methylene protons of the glyceryl group; (3.92) ↑ ^{13}C satellites of protons of triacylglycerols ; (3.94) ↑ Unknown (7.82) ↓ Triacylglycerols; (3.98) ↑	Formaldehyde; (8.02) ↓ Pentanal; (1.34) ↓ $(CH_2)_n$; (1.32) ↓ n-hexanal; (9.72) ↓ **CH_2**-O-C=O; (4.28) ↓ **CH_3**-CH2; (0.89) ↑ Squalene; (1.60) ↓ **-$(CH_2)_n$-** (acyl group); (1.36) ↓ 2-E-hexenal; (9.48) ↓ CH-2 of sn-1,2diacylglycerols; (4.30) ↑	CH_3-**CH_2**; (1.18) ↑ **CH_3**-CH_2; (0.82) ↑ CH_2-**CH_2**-COOH (free acids); (2.54) ↓ Allylic protons broad; (1.98) ↑ **H_2C**-CH=CH-**CH_2**; (1.30) ↓ **CH_2**-COOR; (2.22) ↑ CH-2 Glycerol;(5.30) ↑ **CH_2**-CH_2-COOR (esterified); (2.38) ↑ Bis-allylic group (2.70) ↓
2 (NB)		Allylic protons (broad): (1.98) ↑ CH_2-**CH_2**-COOH; (2.54) ↓ Unknown (2.98) ↑ **CH_2**-CH_2-COOH; (1.50) ↓ **CH_2**-CH_2-COOR (esterified); (2.38) ↓ CH-2 Glycerol; (5.30) ↓	**CH_3**-CH_2: (0.82) ↑ CH_3-**CH_2**: (1.18) ↑ **H_2C**-CH=CH-**CH_2**. (1.26) ↑ Allylic protons (broad); (1.98) ↑ **CH_2**-COOH (2.22) ↓ **CH_2**-CH_2-COOH; (1.50) ↓ **CH_2**-CH_2-COOR (esterified); (2.38) ↓ Unknown (3.34) ↑ CH-2 Glycerol; (5.30) ↑
3 (CA)			CH_2 of sn-1,3 diacylglycerols; (4.02) ↑ D-Glucose; (3.82) ↑ **H-C=C-H** olefinic; (3.92) ↑ **CH_2**-O-C=O; (3.86) ↑ ^{13}C satellites of protons of Triacylglycerols; (3.94) ↑ Unknown (7.82) ↓ Saturated alcohols; (3.82) ↑ Triacylglycerol; (4.38) ↑

Tab. 4 Summary (in terms of variable importance) of metabolites (ppm) leading separation between 2 growth areas. Up arrow ↑ indicate if metabolite is higher in first (row) growth area or ↓ in second (column) growth area.

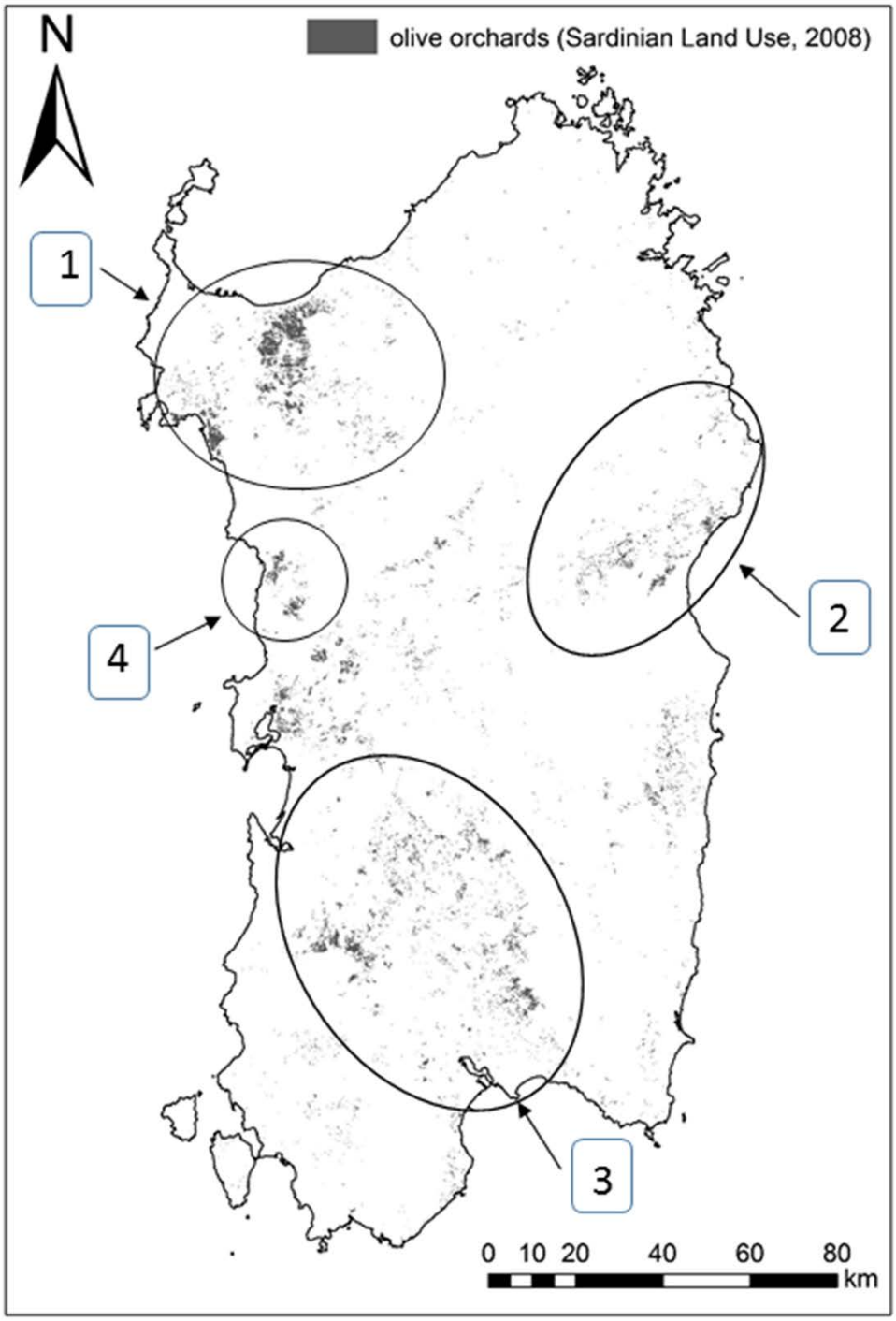

Figure 1

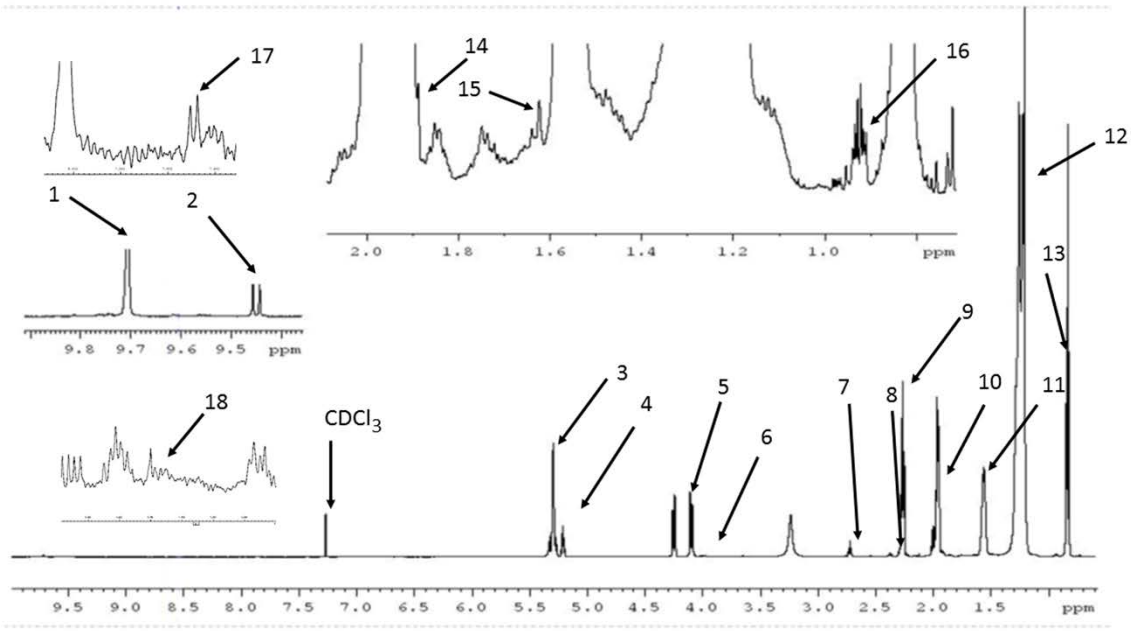

Figure 2

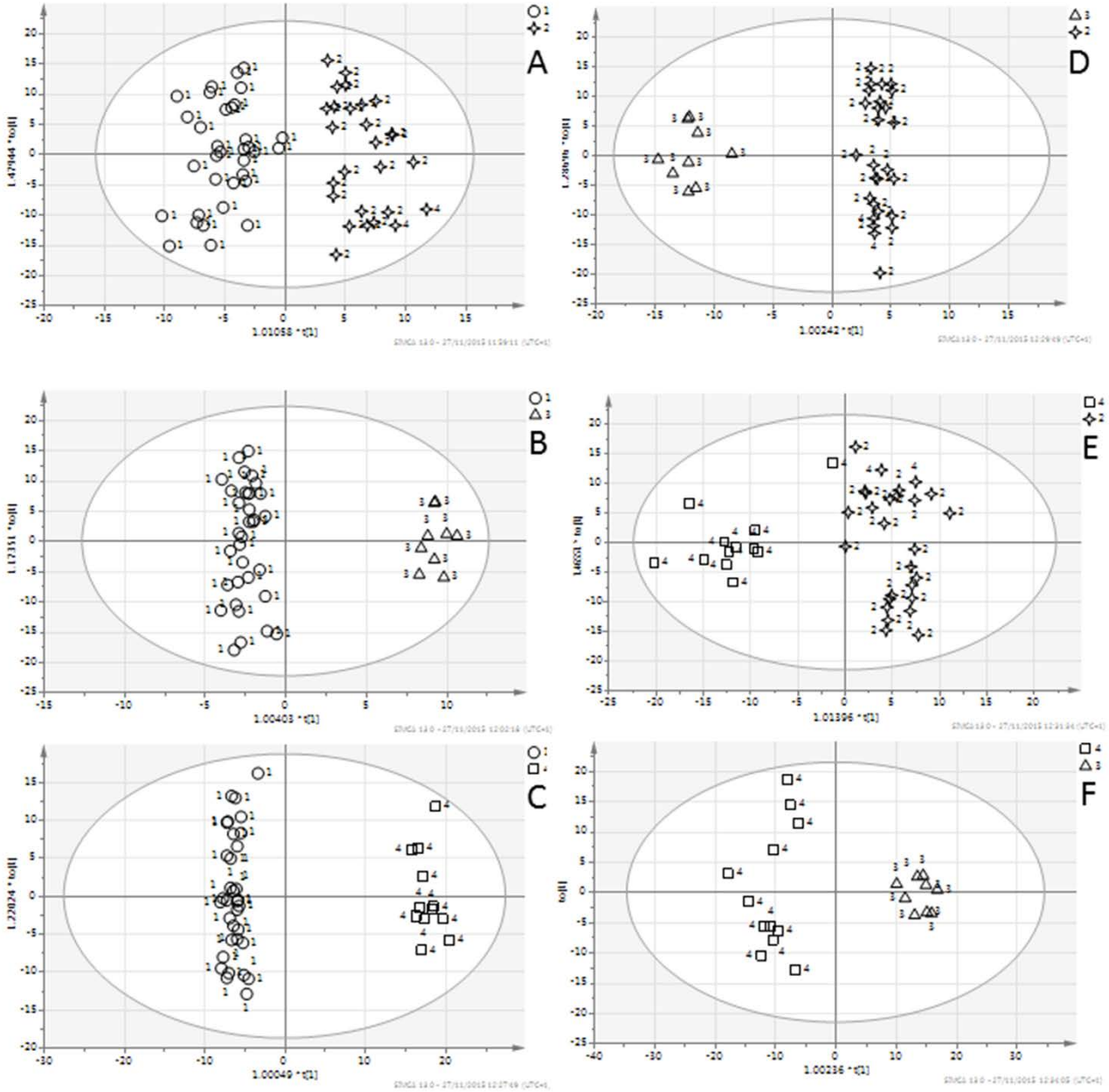

Figure 3