Geographical Origin, Cultivar and Harvesting Year Verification of European and Non-European Olive Oils Using Proton Transfer Reaction Mass Spectrometry With Multivariate Data Analysis

Saskia van Ruth[1], Jan Kiers[1], Wies Akkermans[2], Rafa Perez[1], Alex Koot[1], Enzo Perri[3], Massimiliano Pellegrino[3], Jose Manuel Moreno Rojas[4], Claude Guillou[4], Anne Rossignol-Castera[5]

[1]RIKILT – Institute of Food Safety, Wageningen UR, PO Box 230, 6700 AE Wageningen, the Netherlands

[2]Biometris, Wageningen UR, PO Box 16, 6700 AA Wageningen, the Netherlands

[3]CRA Centro di Ricerca per l'Olivicoltura e l'Industria Olearia, C.da Li Rocchi – 87036 Rende (CS), Italy

[4]European Commission – Joint Research Centre, Institute for Health & Consumer Protection, Physical and Chemical Exposure Unit – BEVAPS TP-281, Via Fermi 2, 21020 Ispra (VA), Italy

[5]ITERG, Rue gaspard Monge, Parc Industriel Bersol 2, F 33600 Pessac, France

Abstract. *Volatile profiles of 308 extra virgin olive oil samples originating from four European and five non-European countries were used for oil authentication with respect to their geographical origin, the olive oil cultivar and the year of harvesting. The latter two were expected to provide some insight on the underlying factors of geographical origin classification. The composition of the volatile compounds in the headspace of virgin olive oils were analyzed by proton-transfer-reaction mass spectrometry (PTR-MS). The auto-scaled headspace concentrations of the masses were subjected to partial least square discriminant analysis (PLS-DA) in order to estimate classification models for the olive oil samples. Correct classification rates were assessed by double cross-validation. The models showed varying success rates. The larger the area/region was, the more successful the classification. Eighty-nine percent of the European oils were classified correctly, against 91% of the non-European oils. Success rates of Individual countries varied between 80 and 92. Additional models showed that cultivar determined to some extent origin classification, whereas year of harvesting had a minor influence. The present study shows that PTR-MS is a promising, rapid, non sample preparation requiring, technique for geographical origin determination of food samples.*

Proceedings of the 5th CIGR Section VI International Symposium on Food Processing, Monitoring Technology in Bioprocesses and Food Quality Management, Potsdam, Germany, 31 August - 02 September 2009

792

Introduction

In recent years the popularity of olive oil has steadily increased, mainly because of its associations with a healthy diet (the so-called 'Mediterranean diet'). Its increased popularity lead to a rise in demand of olive oil, resulting in a more widespread production of olive oil in the non-traditional olive oil producing countries, such as Australia and Argentina (Luchetti, 2002).

Olive oil is produced by grinding the fruit of the olive tree (*Olea europaea* L.) into paste and extracting the oil by mechanical or chemical means. Extra virgin olive oil requires no refinement prior to consumption and it retains its characteristic aroma. This aroma, and complimentary flavor, arises from the phenolic content and the large number of volatile constituents of the oil (Morales, Rios & Aparicio, 1997). These volatile compounds can be used for differentiation and authentication of olive oils (Araghipour et al., 2008). Many different factors affect the volatile composition of olive oil, such as fruit genotype, ripening, and processing equipment used in production (Cavaliere et al., 2007; Angerosa et al., 2001; Aparicio & Morales, 1998). Also the climatological conditions and soil type influence the volatile profiles (Kalua et al., 2007). Apparently the geographical origin of an oil is an important quality factor and has led to the introduction of several official regulations for agricultural products such as the protected denomination of origin and protected geographical indication. As a consequence several methods for origin authentication have been tested for reliability, such as GC-MS, GC-FID, HPLC, NIR and more (Araghipour et al., 2008 and references therein).

Recently, PTR-MS was investigated for detection of oxidative alteration of olive oil (Aprea et al., 2006). This technique, originally proposed by Lindinger (Lindinger et al., 1998), looks very promising since it is a fast and sensitive technique for volatile organic compound detection enabling direct on-line headspace analyses of complex samples without the need for sample preparation. Moreover, since the cycle time of an analysis is just under 30 seconds, many samples can be analyzed within a short period. PTR-MS has been evaluated in a preliminary study for geographic classification of European olive oils in combination with a multivariate statistical method (Araghipour et al., 2008). In the present study, PTR-MS is evaluated for origin verification on a set of olive oil of global geographical origin. At the same time various factors which may contribute to correct or incorrect classification (e.g. olive cultivar, harvesting year) were included in the experimental design.

Materials and Methods

Three hundred and eight olive oil samples were obtained courtesy of the EC-JRC Ispra, Centro di Ricerca per l'Oliviocoltura e l'industria Olearia Rende, and ITERG Pessac. The samples originated from 4 continents and in total 9 countries: Europe: 239 (27 from France, 36 from Greece, 119 from Italy and 57 from Spain); Africa: 8 (all originated from South Africa); Australia: 32 (29 from Australia and 3 from New Zealand); South America: 29 (6 from Argentine and 23 from Chile).

Measurements were performed using a commercial HR PTR-MS system (Ionicon GmbH, Innsbruck, Austria). Samples were equilibrated at 30°C for at least 30 minutes (5 ml sample in a 250 ml screw cap glass bottle) and the headspace of the samples was delivered directly to the inlet of the PTR-MS system. The temperature of the inlet and drift chamber were both 60°C. Measurements were carried out in the mass scan mode and the mass spectra were collected in the range of 20-150 amu. A dwell time of 0.2 s mass^{-1} was used, resulting in a cycle time just under 30 s. Analyses were carried out in triplicate. The data were background and transmission corrected.

PTR-MS spectra yield fingerprints containing hundreds of masses and are therefore high-dimensional vectors. The dimensionality of these data was reduced by PLS-DA. The analyses were carried out in Matlab using the PLS toolbox (Wise et al., 2006). The average of triplicate measurements was subjected to PLS-DA in order to estimate classification models for the olive oil samples. The raw data were auto-scaled. A double cross validation procedure was performed to obtain both the number of factors and the performance: 10 % of the samples were randomly removed from the data set, and a model built with the remaining samples was used to classify these left out samples. In each cycle, the best number of components was estimated in an inner loop. After assessing the performance, the double cross validation was repeated 100 times to get insight on the repeatability of the classification results. These 100 replications can be summarized and visualized in boxplots (for example Figure 1). For each sample i there is a posterior probability p_i (p_{i1}, p_{i2}, ...,p_{ik}) for membership in each of the k classes. The sample is assigned to the class with the highest probability. Let p_{im} be the probability for this class, and let p_{im2} be the probability for the class with the second-largest probability, then an ad-hoc coefficient can be calculated: $Ci = \dfrac{Pit}{Pim + Pim2}$ (with $p_{it} = p_{im}$ if the classification is correct, and $p_{it} = p_{im2}$ if the classification is incorrect). If sample i is correctly classified and the difference between p_{im} and p_{im2} is large, C will be near 1, indicating high confidence. Incorrect classification and large difference between p_{im} and p_{im2} yields a coefficient near 0 (also indicating high confidence). Small differences between p_{im} and p_{im2} give intermediate coefficients which indicate low confidence. By doing this it is possible to investigate if there are objects that are consistently classified in the wrong class. Lastly 100 random permutations of the data were performed, resulting in 100 data sets that give a 'baseline' against which the prediction obtained in the first step can be compared, also visualized in a box plot (for example Figure 2). The model based on random permutations is expected to have low confidence in prediction results, with coefficients around 0.5.

Results and Discussion

European versus non-European

PLS-DA classification of European versus non-European olive oils using their PTR-MS mass spectra resulted in 89% correctly classified European oils (213 correct classifications out of 239 European oils), and 91% correctly classified non-European oils (63 out of 69 non-European oils).

Countries one-versus-all

For the olive oils worldwide a number of two-class classifications was performed, contrasting each country with an 'Other' class containing all samples from all other countries. In these analyses, only countries with eight or more samples were included (Table 1). The model predicted the origin of Spanish and non-Spanish samples for 92% of the samples correctly. Figure 1 and 2 present the classifications of Spain versus all countries in boxplots. For the other countries, success rates varied between 80 and 86%.
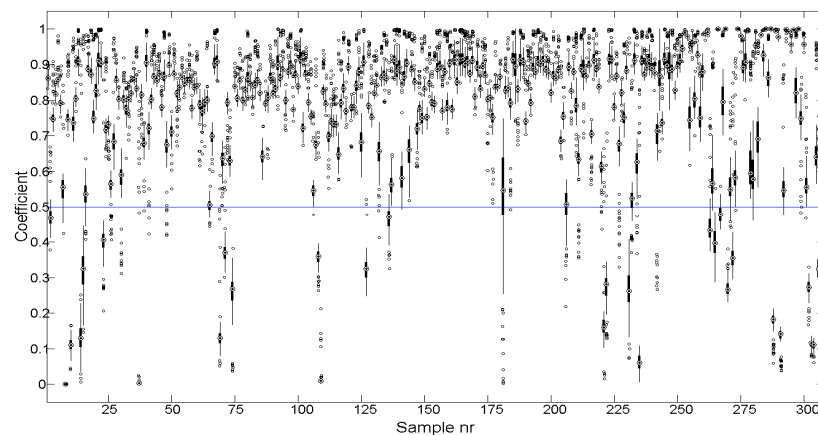
Figure 1. Boxplot of classification of Spanish samples versus all other samples with a PLS-DA model on PTR-MS headspace data of 308 olive oil samples. Values of coefficient C were obtained with 100 replications of double cross validations. Samples 252 to 308 originated from Spain and samples 1 to 251 from elsewhere.
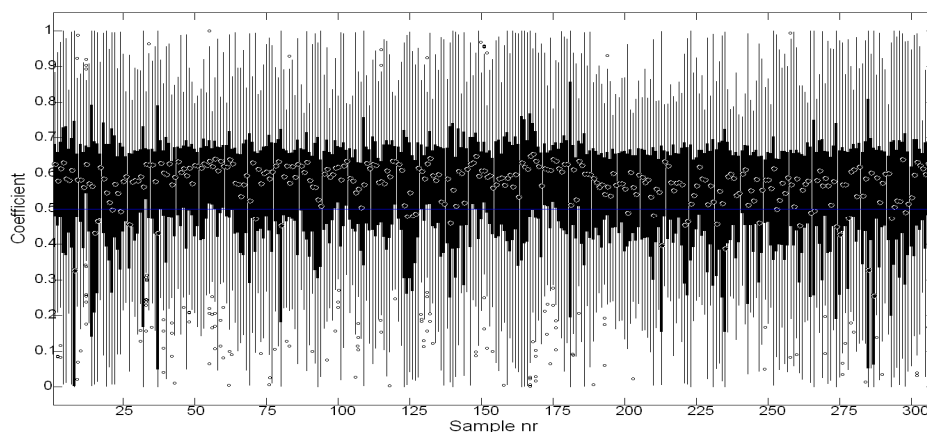

Figure 2. Boxplot of classification of Spanish samples versus all other samples with a PLS-DA model on PTR-MS headspace data of 308 olive oil samples. Values of coefficient C were obtained by performing 100 random permutations of the double cross validations.

Table 1. Numbers and (percentages) of correct predictions (countries one-versus-all) of 299 olive oil samples.

Origin	Numbers	Correctly classified		Total correctly
		Country	Rest	classified
France	27	14 (52%)	248 (88%)	262 (85%)
Greece	36	29 (81%)	229 (84%)	258 (84%)
Italy	119	97 (82%)	148 (78%)	245 (80%)
Spain	57	47 (82%)	238 (95%)	285 (92%)
Australia	29	27 (93%)	237 (85%)	264 (86%)
South Africa	8	6 (75%)	260 (87%)	266 (86%)
Chile	23	19 (83%)	244 (86%)	263 (85%)

Year of harvesting

Two different harvesting years per country were examined, for the Northern hemisphere harvests 2004/2005 (148 samples) and 2006/2007 (73 samples), and for the Southern hemisphere harvests 2006 (36 samples) and 2007 (33 samples). Some samples (18) of unknown growing year were removed from the set of 308 samples. The harvesting year was predicted from the PTR-MS spectra. For the Northern hemisphere harvest 2004/2005 80% of the samples were correctly classified (118 out of 148 samples) and for the harvest of 2006-2007 58% (42 out of 73 samples). For the Southern hemisphere harvest 2006 53% (19 out of 36 samples) and of harvest 2007 61% were correctly classified (20 out of 33 samples).

Taking into account Northern and Southern hemisphere, overall the year of harvesting was correctly predicted for 69% of the samples (199 out of 290 samples). Considering a 50% chance, this is not very high. It seems, therefore, that in comparison with geographical origin, harvesting year plays a less important role in the development of the volatile profiles of extra virgin olive oil. At least for the years tested.

Cultivar

The olive cultivar of the European extra virgin olive oils was predicted from the PTR-MS spectra (volatile profiles) (Table 2). The mean classification success rate was 62% overall, but varied considerably between cultivars. Some cultivars consisted of fairly small sample groups (6-7 samples) which may explain their poor results.

Table 2. European varieties: Classification of European olive oils into 7 varieties by their PTR-MS spectral data and PLS-DA: absolute numbers, percentages in brackets, the correctly classified samples in bold[a].

Variety	Classification							Total number
	A-C-S	KORON	TAGG	PINO-T	MO-SLF	ARBEQ	PICU-H	of samples
Aglandau, Cayanne, Salonenque	**6(67%)**	0(0%)	0(0%)	0(0%)	2(22%)	0(0%)	1(11%)	9
Koroneiki	1(3%)	**28(88%)**	1(3%)	0(0%)	1(3%)	1(3%)	0(0%)	32
Taggiasca (= Lagagnina)	0(0%)	1(4%)	**15(56%)**	4(15%)	6(22%)	0(0%)	1(4%)	27
Pinola + Taggiasca	0(0%)	0(0%)	5(71%)	**0(0%)**	2(29%)	0(0%)	0(0%)	7
Moraiolo, S.Fel.-Lecc.-Fran.	0(0%)	1(17%)	0(0%)	0(0%)	**4(67%)**	1(17%)	0(0%)	6
Arbequina	1(14%)	2(29%)	0(0%)	0(0%)	2(29%)	**0(0%)**	2(29%)	7
Picuda / Hojiblanca	0(0%)	0(0%)	0(0%)	0(0%)	0(0%)	2(22%)	**7(78%)**	9

[a] Number correct: 60 out of 97 samples (62%)

Table 3. Classification of European olive oil varieties grouped within countries by PLS-DA with cross-validation by their PTR-MS spectral data: absolute numbers, percentages in brackets, the correctly classified samples in bold[a].

Class	Classification				Total number
	ACS-Fr	AP-Sp	Kor-Gr	MPT-It	of samples
Agla., Caye., Salo. (France)	**4(44%)**	1(11%)	1(11%)	3(33%)	9
Arbe., Picu./Hoji. (Spain)	1(6%)	**12(75%)**	3(19%)	0(0%)	16
Koro. (Greece)	1(3%)	0(0%)	**30(94%)**	1(3%)	32
Mora.+S.Fel.-L-F, P.+T., Tagg. (Italy)	0(0%)	1(3%)	2(5%)	**37(93%)**	40

[a] Number correct: 83 out of 97 samples (86 %)

For the investigated subset of European olive oils the 7 varieties occurred in single countries: Aglandau, Cayenne, Salonenque in France, Koroneiki in Greece, Moraiolo, S.Felice-Leccino-Frantoio, Pinola + Taggiasca and Taggiasca in Italy, and Arbequina and Picudo/Hojiblanca in Spain (Table 3). Eighty-six percent of the samples classified in olive oil varieties grouped within European countries were classified correctly. These results indicate that differences between oils from distinct geographical origin are determined to a considerable extent by the olive cultivars. However, it is not the sole factor, other geographical conditions are important as well, since prediction of the cultivar only was less successful.

Conclusions

Olive cultivar is an important parameter in geographical origin authentication, the year of harvesting plays a minor role. PTR-MS in combination with multivariate data analysis is a promising technique for geographical origin determination of olive oil samples.

References

Angerosa, F., Mostallino, R., Basti, C. & Vito, R. 2001. Influence of malaxation temperature and time on the quality of virgin olive oils. *Food Chemistry* **72**: 19-28.

Aprea, E., Biasioli, F., Sani, G., Cantini, C., Märk, T.D., & Gasperi, F. 2006. Proton transfer reaction-mass spectrometry (PTR-MS) headspace analysis for rapid detection of oxidative alteration of olive oil. *Journal of Agricultural and Food Chemistry* **54**: 7635-7640.

Araghipour, N., Colineau, J., Koot, A., Akkermans, W., Moreno Rojas, J.M., Beauchamp, J., Wisthaler, A., Märk, T.D., Downey, G., Guillou, C., Mannina, L. & Van Ruth, S. 2008. Geographical origin classification of olive oils by PTR-MS. *Food Chemistry* **108**: 374-383.

Cavaliere, B., De Nino, A., Hayet, F., Lazez, A., Macchione, B., Moncef, C., Perri, E., Sindona, G., Tagarelli, A. 2007. A metabolomic approach to the evaluation of the origin of extra virgin olive oil: a conventional statistical treatment of mass spectrometric analytical data. *Journal of Agricultural and Food Chemistry* **55**: 1454-1462.

Kalua, C.M., Allen, M.S., Bedgood, D.R., Jr., Bishop, A.G., Prenzler, P.D., & Robards, K. 2007. Olive oil volatile compounds, flavour development and quality. *Food Chemistry* **100**: 273-286.

Lindinger, W., Hansel, A., Jordan, A. (1998) On-line monitoring of volatile organic compounds at ppt level by means of proton-transfer-reaction mass spectrometry (PTR-MS): medical application, food control and environmental research. *Int. J. Mass Spectrom. Ion Process* **173**: 191-241.

Luchetti, F. (2002) Importance and future of olive oil in the world market – An introduction to olive oil. *European Journal of Lipid Science and Technology* **104**: 559-563.

Morales, M.T., Rios, J.J., & Aparicio, R. (1997). Changes in the volatile composition of virgin olive oil during oxidation – Flavors and off-flavors. *Journal of Agricultural and Food Chemistry*, **45**: 2666-2673.

Wise, B. M., Shaver, J.M., Gallager, N.B., Windig, W., Bro, R., & Koch, R.S. 2006. *PLS-Toolbox Version 4.0 for use with Matlab^{TM}.* Wenatchee, WA, USA: Eigenvector Research Inc.