ORIGINAL PAPER

Diego L. García-González · Ramón Aparicio

Classification of different quality virgin olive oils by metal-oxide sensors

Abstract The European Union classifies virgin olive oils into three categories, extra virgin, virgin and lampante; lampante being the only oil that cannot be consumed before undergoing a refining process. A mathematical model based on two metal-oxide sensors has been designed and checked in order to detect on-line lampante oils inside the production systems. The model was developed using 114 samples and was successfully tested with an external test set of 55 samples taken from different single varietal olive oils and geographical origins. The model was able to detect 100% of non-lampante virgin olive oils and 89.5% of lampante virgin olive oils.

Keywords Chemometrics · Electronic nose · Metal-oxide sensors · Olive oil · Sensory quality

Introduction

The current EU regulation [1] defines three categories for virgin olive oils (extra virgin, virgin and lampante) because of their diverse sensory and chemical characteristics. The last category also includes the old category "ordinary" [2]. These designations affect their market prices as lampante virgin olive oils cannot be consumed without refining. Thus, the on-line detection of the lowest-quality category, during the olive oil industrial production process, has a notable economical impact.

The sensory analysis plays a crucial role in the classification of virgin olive oils (VOO) into the cited designations. The official method for olive oil sensory analysis is carried out by trained assessors (panel test), a lengthy and costly procedure that cannot be afforded by small enterprises. Besides, the panel test is not an error free procedure since wrong classifications have been detected in international trials due to assessors' subjectivity [3]. A more rapid and cheaper alternative is the analysis of phenols and volatile compounds that are responsible for taste and aroma of virgin olive oil, respectively [4, 5]. Here, the error could be ascribed to the repeatability of whole process without any influence of subjective opinions but the process would not be applied on-line.

The second alternative is the use of sensors for detecting volatile compounds. Sensors do not need any pre-treatment and do not use solvents to detect the presence of volatiles. Furthermore, their main advantages are their low cost and the rapid evaluation of the aroma. Sensors have been applied in many fields of analytical chemistry with success, and they are seen as an emergent technique in food quality and assessment [6, 7]. Thus, metal-oxide semiconductor sensors (MOS) have been used, with varying degree of success, in the classification of edible oils [8], the determination of VOO geographical origin [9] and edible oil shelf-life predictions [10]. On the other hand, previous papers by the authors have stated that metal-oxide sensors are sensitive to certain VOO off-flavors such as rancid, vinegary and fusty [11, 12]. However, VOO sensory classification at the lowest grade (lampante) does not distinguish among the cited off-flavors as samples are so classified. Thus, sensors would be useful in distinguishing the highest quality virgin olive oils from the lowest quality ones if they were able to cluster together all low quality VOO whichever the off-flavor, or combination of off-flavors, was presented in the oil. This paper analyzes the possibilities of metal-oxide semiconductor sensors after working with the samples of a training set and verifying the results with an external test set.

D. L. García-González (✉) · R. Aparicio
Instituto de la Grasa,
Padre García Tejero, 441012 Sevilla, Spain
e-mail: dluisg@cica.es
Tel.: +34-954611550
Fax: +34-954616790

Materials and Methods

Materials

114 samples of virgin olive oil (var. Hojiblanca) were used for training the sensors (training set). The samples were supplied by an association of cooperatives (Hojiblanca SCA, Málaga, Spain) that represents 4% of total Spanish olive oil production. 56 of these samples were qualified as lampante by the assessors of the cited association. The mathematical model (canonical equations) was checked with a test set of 55 samples. 13 of these samples (var. Arbequina, Cornicabra and Picual), collected in different geographical regions, were supplied by Aceites del Sur SA (Sevilla, Spain); four samples were classified as lampante by the trained assessors of the enterprise. The remaining 42 samples originating from 8 producing countries (Spain, Italy, Greece, Tunisia, Morocco, Syria, Algeria and the USA), are comprised of 27 monovarietal virgin olive oils and 15 lampante virgin olive oils. The latter 15 had undergone a rancidity process by subjecting the samples to light, kept inside tubes with a 50% headspace, for three years. All these samples were sensory evaluated by assessors of Instituto de la Grasa. All the samples, training and test sets, were analyzed for 11 months to check the effect of drift in the sensor baseline.

Equipment

A Fox 4000 with ACU500 humidifier supplied by AlphaMOS SA (Toulouse, France) was used. This instrument is equipped with 18 metal oxide sensors, inside three chambers, 6 of them being undoped metal oxide sensors, and 12 being metal oxide sensors doped with noble catalytic metals in order to shift the selectivity spectrum towards different chemical compounds. The temporary and reversible adsorption of volatile reducing compounds at the sensor surface changes its electrical resistance in a non-linear manner [13]. The response is characteristic of each sensor and depends on the concentration and the profile of the volatile compounds.

The air conditioning unit (ACU 500) consists of a thermostat tank containing distilled water through which the carrier gas bubbles continuously. When a valve is opened during the injection time, a controlled mixture of dry and humid industrial air sweeps the headspace of the sampling chamber whose temperature is controlled automatically.

Industrial air, from an air compressor, was used as the carrier gas after being filtered through two columns. The first column was filled with a molecular sieve with a 8/12 mesh (Supelco, Bellefonte PA, USA) to remove the moisture, while the second column was filled with activated carbon (Supelco, Bellefonte PA, USA) to remove hydrocarbons and other undesirable volatile compounds.

Measuring setup

The analytical parameters (sample amount, headspace generation time, sample temperature, flow rate and injection time) were determined following the optimization process described in [11].

A 5-g amount of each sample – enough to cover the bottom of the 100 mL flasks – and these were heated at 34 °C inside a controlled thermostat-sampling chamber for a headspace generation time of 600 s. The responses of the sensors started to be collected immediately after the headspace generation time for a further 600 s; 90 s for the injection time and 510 s for the desorption time. Volatile compounds were pumped into the sensor chambers by the carrier gas (air) at a flow-rate of 100mL/min. After the injection time, a valve was switched and only carrier gas was blown into the sensor chambers at the same flowrate of 100mL/min. After collecting the sensor responses, 900 s of non-measurement time remained. The flowrate of the carrier gas was kept at 500mL/min to ensure that the baseline had indeed recovered before performing the next analysis.

Samples were analyzed in duplicate. Standards for calibration of the sensor array were measured at programmed times to check that the aging of the sensors did not affect the measurements.

Measurements of repeatability

The repeatability studies, either between-days (for 6 months) or during the day, were investigated by consecutively collecting the sensor results of the same sample of virgin olive oil (cv. Farga spiked with 60 ppm of acetic acid) [11]. The maximum %RSD (relative standard deviation) of the repeatability study carried out during the day was 12.0%, the mean being 6.1%. While the mean results of the between-days repeatability study was 11.7%,with a maximum (22.2%) being far higher.

Data pre-treatment

The response of the sensors yields an exponential-like shape but not all this information is useful. After different methods of data preprocessing had been tested, raw data (non-preprocessed data) were selected because they showed the best differential properties [12]. Windowed time slicing (WTS) [14] was used to reduce the information to a small data set. The number of windowing functions was 4, each one applied to a different region of the sensor response [11].

A standard was analyzed before and after each series of analyses with the objective of minimizing the effect of sensor aging and environmental conditions. This information was used to standardize WTS data.

The detection of multivariate outliers was carried out by applying principal components analysis (PCA) [15]. Mahalanobis distance, evaluated as χ^2, was used to discover outliers among samples, while outliers among variables (WTS) were detected by the squared multiple correlation.

Stepwise linear discriminant analysis (SLDA) was applied under the strictest conditions for the selection of the variables in order to diminish over-optimistic models. Thus, tolerance was fixed at 0.01 while the F-to-Enter value (8.6) was obtained from the F-distribution table at F(F)=0.995 for the number of groups (m=2) and the group with the minimum number of samples (n=53) [15].

Statistica (Statsoft Inc., Tulsa, USA) release 5.5 [16], was used to perform the data processing and to implement multivariate data analyses.

Results and Discussion

The hypothesis that the sensor response depends on the amount and composition of volatile compounds had already been demonstrated by the authors when analyzing some of the negative attributes (fusty, rancid and vinegary) by canonical correlation [11]. Based on these promising results, four steps were planned: (i) the detection of multivariate outliers by PCA, (ii) the training process of the supervised procedure SLDA, (iii) the implementation of the canonical equation in a discriminating model, and (iv) the validation of the results with the samples of an external test set.

The raw information was first clustered into four WTS [14] and later standardized to avoid hypothetical sensor ageing. The collected information was transferred into a Statistica [16] file and the next step was the detection of outliers.

The study of outliers is extremely necessary as they can greatly affect the magnitudes of the decision equation

Fig. 1 Results of applying the canonical equation to the samples of training and first test sets
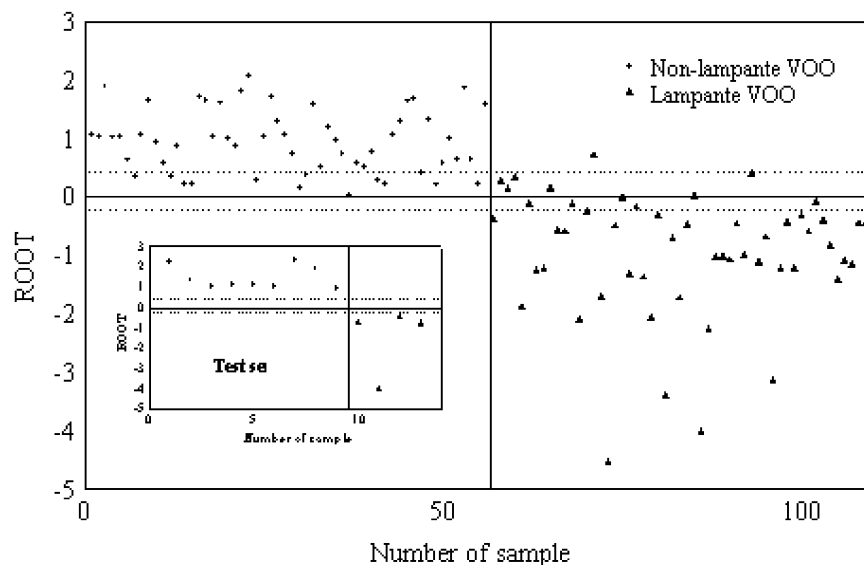
coefficients. This study was carried out by multivariate procedures, as the problems originate mostly from multivariate outliers among variables and cases. Thus, five multivariate outliers among cases (2 non-lampante and 3 lampante virgin olive oils) and four multivariate outliers among variables (sensors 2, 5, 12 and 13) were detected by PCA and removed.

Once the outliers had been removed (5 samples and 4 sensors), the WTS of the remaining sensors were submitted to SLDA under the conditions cited above. This statistical procedure automatically selected 3 variables from the initial set of 56 variables (14 sensors * 4 WTS per sensor). As there were only two groups (lampante vs. non-lampante virgin olive oils) that needed to be distinguished, SLDA produced only one canonical equation based on the fourth WTS of sensors 1 and 18 plus the first WTS of sensor 1:

$$y = 6.96 - 0.25 * \text{Sensor1 (WTS4)}$$
$$+ 0.10 * \text{Sensor1 (WTS1)} + 0.01 * \text{Sensor18 (WTS4)}$$
$$(1)$$

At first sight, the equation 1 contains information from different aspects. Thus, relevant information concerns the processes of adsorption (WTS1) and desorption (WTS4) of volatiles. There is no discrimination concerning the sensor characteristics because sensor 1 is undoped while sensor 18 is doped [6]. No discrimination was detected in terms of the sensor chambers since one sensor is placed in the first chamber (sensor 1) and the other is inside the third chamber (sensor 18).

Figure 1 shows the result of the canonical equation distinguishing lampante virgin olive oil from the other categories. The y-axis indicates the quality level of samples according to the sensor responses as this axis shows the values of applying the canonical equation to each sample of the training set. Positive values of the canonical equation correspond to non-lampante virgin

olive oils while negative values indicate that samples are lampante virgin olive oil. The x-axis indicates the sample number. The procedure was able to classify 100% of non-lampante virgin olive oils versus only 83% of lampante virgin olive oils. The diversity of possible off-flavors (rancid, vinegary, fusty, winey, mould sediment, cucumber, etc) and their intensity range might explain the difficulty that sensors had in order to cluster all of them into only one group.

The samples placed close to the value zero of the y-axis, indicate that their quality is near to lampante virgin olive oils, which means "ordinary virgin olive oil" according to the previous EU regulation [2]. It is a zone where the risk of wrong classifications is high enough due to the absence of discontinuity in the sensory evaluation, and hence it could be seen as a transition zone between both large groups (lampante vs. non-lampante virgin olive oils). The confidence limits of this zone correspond to $\bar{y} \pm SD$; $\bar{y}$ and SD being the mean and standard deviation of the values of the canonical equation when applied to samples classified as ordinary virgin olive oils by the panel test of Hojiblanca SCA. All the wrongly classified samples, with one exception, were located inside the zone for ordinary virgin olive oils. The exception, classified with a median of defects (Md) 6.4, was re-evaluated by assessors of Instituto de la Grasa. A new score punctuation of Md=5.8 was given which means that the sample would be inside the category "ordinary". Assessors also detected a light vinegary note in this oil that was not probably noted by the metal-oxide sensors.

The next objective was to check the canonical equation with the external test set of 55 samples. First of all, the equation was applied to a set of 13 different single varietal olive oil samples (Arbequina, Cornicabra and Picual) from different Spanish geographical origins. The objective was to check if the varieties and the sensory evaluation, now carried out by the panel test of a multinational enterprise (Aceites del Sur SA), affected

the model. All the samples (100%) were correctly classified thus pointing out the validity of the proposed mathematical model (Fig. 1).

Finally, 42 single varietal olive oil samples were submitted to the mathematical model. Assessors of Instituto de la Grasa classified 27 of them as extra-virgin olive oil and the remaining 15 as lampante-virgin olive oils. The objective of this test was to check the model with a large set of single varietal extra-virgin olive oils from different geographical origins (Spain, Italy, Greece, Tunisia, Morocco, Syria, Algeria and the USA), and with a set of single varietal olive oils that had undergone an extreme rancidity process by subjecting the samples to light for three years. All the extra virgin olive oils were correctly classified (100%) while the classification was lower for lampante virgin olive oils (87%). The misclassified samples belonged to varieties Cornicabra and Lechin.

In conclusion, the model, based on exclusively a canonical equation, was able to distinguish lampante virgin olive oils from the other categories with only two misclassifications when analyzing the test set (oxidized varieties). The other misclassified sample (training set) showed the difficulty of a complete agreement between panel tests when evaluating olive oils [3].

References

1. E.C. Off. J. Eur. Communities, 2002, May 15, Regulation 796/2002
2. E.C. Off. J. Eur. Communities, 1991, July 11, Regulation 2568/1991
3. Ranzani C (1994) Grasas Aceites 45:1-4
4. Aparicio R, Morales MT, Alonso MV (1996) J Am Oil Chem Soc 73:1253–1264
5. Morales MT, Tsimidou M (2000) The Role of Volatile Compounds and Polyphenols in Olive Oil Sensory Quality. In Harwood J, Aparicio R (eds) Handbook of Olive Oil: Analysis and properties, Aspen, Gaithersburg, Maryland, pp 393–458
6. García-González DL, Aparicio R (2002) Grasas Aceites 53:96–114
7. Shurmer HV, Gardner JW (1992) Sens Actuator B 8:1-11
8. González Y, Pérez JL, Moreno B, García C (1999) Anal Chim Acta 384:83–94
9. Pardo M, Sberveglieri G, Taroni A, Masulli F, Valentini G (2001) Anal Chim Acta 446:223–232
10. Shiers V, Adechy M (1998) Semin Food Anal 3:43–52
11. García-González DL, Aparicio R (2002) J Agric Food Chem 50:1809–1814
12. García-González DL, Aparicio R (2002) Eur Food Res Technol 215:118–123.DOI 10.1007/s00217–002–0527–9
13. Mielle P (1996) Trends Food Sci Technol 7:432–438
14. Gutiérrez-Osuna R, Nagle HT (1999) IEEE Trans Syst Man Cyber 29:626–632
15. Tabachnick BG, Fidell LS (1983) Using Multivariate Statistics. Harper & Row Publishers, New York
16. Statsoft (2000) STATISTICA, release 5.5. Statsoft, Tulsa, OK