ACCEPTED VERSION

Fast tool based on electronic nose to predict olive fruit quality after harvest

Diego M. Martínez Gila[a,*], Javier Gámez García[a], Andrea Bellincontro[b], Fabio Mencarelli[b], Juan Gómez Ortega[a]

[a] Robotics, Automation and Computer Vision Group, University of Jaén, Campus Las Lagunillas s/n, 23071, Jaén, Spain
[b] Dipartimento per l'Innovazione nei Sistemi Biologici, Agroalimentari e Forestali, University of Tuscia, Via San Camillo de Lellis s/n, 01100, Viterbo, Italy

ARTICLE INFO

Keywords:
Electronic nose
Harvested olive fruit
Pattern recognition
Olive oil quality
Organoleptic assessment

ABSTRACT

Quality analyses of oil from olive fruit are performed according to regulated procedures and in accredited laboratories that are usually separated from the oil mill. These analytics include organoleptic features involving smelling by human experts. Therefore, oil features depend on the physicochemical conditions of the harvested fruit. An automatic and non-invasive system for monitoring and controlling the process in postharvest stages could optimize the quality of the processed oil. To validate this hypothesis we proposed a methodology based on an electronic nose sensor and pattern recognition algorithms to predict the quality of the oil to be processed from measurements on freshly harvested olive fruit. The pattern recognition algorithms applied were the Naïve Bayes (NB) classifier, the partial least squares discriminant analysis (PLSDA) and a multilayer perceptron (MLP) artificial neural network. Using the measurements performed on 82 samples of olives, the best result was obtained with the MLP network, with 90.2 % success obtained in the classification of the virgin and extra virgin olive oil quality by applying 10-fold cross-validation. Integration of this methodology to virgin olive oil production allows prediction of the quality of the final oil from the olive fruit received from the farmer.

1. Introduction

The production and consumption of olive oil continues to increase; production between 2009 and 2019 seasons increased by 8.4 %, according to the International Olive Oil Council (IOC), 2018. The world consumption of olive oil also increased by 9% over this time, reaching 3 million tons, including consumption of extra virgin oil. This trend is making the implementation of quality control systems in oil mills a necessity. For this purpose, different technologies have been studied in the scientific literature (Aguilera Puerto et al., 2019; Navarro Soto et al., 2018 Martínez Gila et al., 2018; Beltrán Ortega et al., 2016). However, actual implementations in oil mills are still scarce.

The production of olive oil is carried out in the mills through a series of mechanical processes. Basically the process consists of the following phases: harvest phase, reception of the olives at the mill, cleaning of the olives, weighing, grinding of the fruit to make a paste, shaking the paste, decanting the oil by centrifugation, cleaning of the decanted oil, and then storage in stainless steel tanks. Each time a tank is filled, the oil mill master takes a sample of oil and sends it to an accredited chemical laboratory for quality analysis. The product is labeled from lower to higher quality such as lampante, virgin or extra virgin based on the results.

The criteria that establish the quality of the oil are regulated at European level (Reg. CE 1348/2013) and the parameters obtained from each oil sample are physical-chemical and sensory. While the former are related to the oxidative stability of the oil, the latter are based on the sensory perception of a group of 8–10 expert tasters called panelists. The procedure to follow consists basically in smelling the oil and issuing an assessment according to a methodology that is standardized and documented in (IOC, 2018). The highest quality (extra virgin olive oil) is assigned to an oil when there is a total absence of defects (Table 1). The most frequent defects are fusty, muddy, moldy, winey and rancid and most are related to the state of the fruit after harvesting and before being ground (Angerosa et al., 2004; Lerma-García et al., 2010). Moreover other variables can affect to the aroma fingerprint of the olive oil such as the manipulation process of the oil (Reboredo-Rodríguez et al., 2013b, 2014) or the cultivar (Reboredo-Rodríguez et al., 2013a).

For example, injury of olive skin during harvest or during loading and unloading can result in fermentations that result in ethanol, ethyl acetate and acetic acid that is transmitted directly to the oil (Morales et al., 2005). This is also more likely to occur in olives with a high maturity index. Fig. 1 shows an example of fruit of different healthy and maturity conditions. The olives Fig. 1A produce an oil of worse quality than those in Fig. 1B. There may also be lots containing olives of both

* Corresponding author.
E-mail address: dmgila@ujaen.es (D.M. Martínez Gila).

https://doi.org/10.1016/j.postharvbio.2019.111058
Received 2 September 2019; Received in revised form 29 October 2019; Accepted 29 October 2019

0925-5214/ © 2019 Elsevier B.V. All rights reserved.

Table 1
Parameters to classify olive oils according to their quality.

	Parameter	Extra Virgin Olive Oil	Virgin Olive Oil	Lampante Olive Oil
Physiochemical parameters	Acidity	≤0.8	≤2.0	> 2
	Peroxides Index	≤20	≤20	> 20
	K_{232}	≤ 2.50	≤2.60	—
	K_{270}	≤ 0.22	≤0.25	—
	ΔK	≤0.01	≤0.01	—
Organoleptic parameters	Defect Median	0	≤2.50	> 2.5
	Fruity Median	> 0	> 0	—

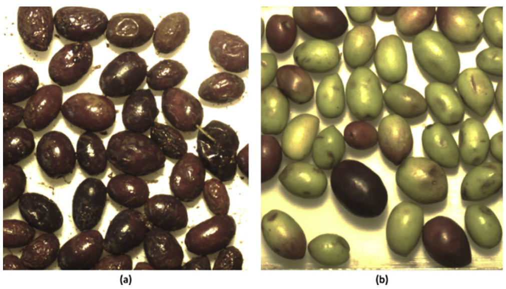

Fig. 1. Olive fruit in different healthy conditions; A. fruit with loss of turgidity and injury, and B. fruit in good condition.

kinds. The number of analysis of this type that the mill performs is limited due to its cost (around 50 € the analysis) and the time to obtain the results may take 1 or 2 days. Therefore, there is interest in sensor technology such as the electronic nose that could be installed in the process line to quickly determine lots of defective olives and thus prevent the oil produced by them from being mixed in the same tank with high quality oil.

Electronic noses are electromechanical systems formed by a chamber or array of gas sensors and a pneumatic circuit that conducts the air including volatiles, from the product to be analyzed to the sensor chamber (Zhang et al., 2018). The sensors generate an electrical signal in the presence of certain molecular chains. In this work the electronic nose PEN3 (Airsense) was used to detect defects in oils and olives. In the scientific literature we can find the use of this same nose to classify kiwifruit (Ma et al., 2014), and plants (Wang et al., 2019a,2019b) according to their geographical origin, for detecting the freshness of the strawberry during storage (Xing et al., 2018), classifying apples based on their cultivars (Wu et al., 2017), detection of adulterants (Wang et al., 2019a,2019b), and contaminations (Concina et al., 2009). Electronic noses also have been applied to the detection of fungal contaminations and infestations (Liu et al., 2019; Wen et al., 2019), to assess quality (Jiang and Wang, 2016; Raigar et al., 2017) and of fruit maturity (Chen et al., 2019; Huang et al., 2017).

This type of sensor has been used for classification by quality, determination of geographical origin, detection of adulterations and analysis of olive oil deterioration caused by external factors (Majchrzak et al., 2018). Several studies have used sensors to classify high and low quality oils (González Martín et al., 2001, 1999; Guadarrama et al., 2001a, 2001b; James et al., 2004).

The integration of this non-invasive technology in the production process of virgin olive oil can be useful to predict potential quality from the fresh olive fruit. Our hypothesis is that an electronic nose can identify two qualities of olive oils (extra virgin and virgin) on the basis of measurements on olive fruit and on the oil produced from the same olive fruit. To validate this hypothesis, a methodology based on machine learning is proposed in this work to evaluate differences in aromatic profiles between both classes. The result would be an automatic system that can predict the quality of the oil from the analysis of the incoming olives.

2. Materials and methods

2.1. Olive fruit samples

The sampling of olives was carried out in Jaén, Spain, in the facilities of the Picualia mill (www.picualia.com) during the olive harvest in November 2017 to February 2018. 82 lots of olives of 1 kg per lot were collected from a farm at approximately one day intervals. The quality of the olive lots at the beginning of the campaign was higher than the end of campaign according to the maturity index of the fruit. More advanced maturation indexes involve soft olives whose tissue can be broken in harvest or in transport phase and cause loss of quality in the lot. The flow of samples is shown in Fig. 2.

The olive lots were divided in three batches of 300 g and the electronic nose measurements were performed three times per batch. Once the sensory response of the olive batch was obtained, it was transformed into oil using the Abencor system, a small-scale of the olive oil production process, consisting of an olive grinder, a small blender and a small centrifuge. The beating time was 10 min and the beating temperature was 30 °C. The oil was also introduced into the electronic nose system to obtain the sensory response of the oil from each batch. In addition, this oil was sent to the accredited CM Europe laboratory (www.cmeuropa.com) to carry out the official quality analysis. Of the 82 oil samples analyzed, 57 were classified as extra virgin and 25 as virgin. The results of this analysis are shown in Table 2.

2.2. Experimental setup

Analyses were performed using the electronic nose PEN3 (Airsense Analytics GmbH, Schwerin, Germany), a compact (92 × 190 × 255 mm) and lightweight (2.3 kg) portable olfactory system, consisting of a gas sampling unit and a sensor array. Using a software (Win Muster v. 1.6.2), the PEN3 allows acquisition, visualization and analysis of the data and works using filtered ambient air during the cleaning step or dilute gas during sampling with a maximum flow rate of 600 mL/min.

The integrated sensor array is composed of 10 different thermos-regulated (200–500 °C) metal oxide thick film sensors (MOS) positioned in a very small stainless steel chamber (volume: 1.8 mL, temperature: 110 °C) and sensitive to different classes of chemical compounds (Table 3). The selectivity of the sensors is influenced by the sensing and the dopant materials employed, and by the working temperature and sensor geometry.

Each measurement process involved two stages. The first stage was the cleaning process where the air passes through an activated carbon filter before entering the sensor array. In the second stage, air also passes through an activated carbon filter and through the sample, and then through a moisture and particle filter, finally arriving at the sensor arrays. When the volatile compounds of the sample react with the sensing film of the sensor surface, there is an oxygen exchange with a decrease in electrical conductivity, detectable by a transducer element (electrode) attached to each sensor.

Fig. 3 shows the experimental setup for olive fruit batches and oil samples respectively. For each batch a sample of 250 g of fruit or 5 g of

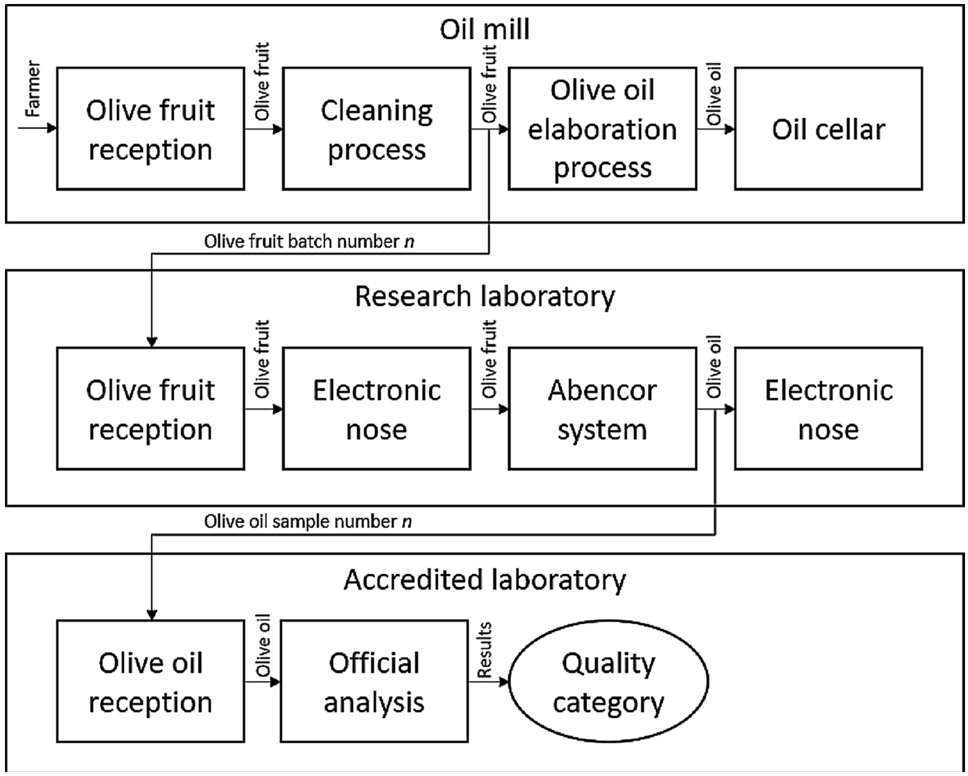

Fig. 2. Flow of samples from the mill to the laboratory.

Table 2
The results of the analyses in an accredited laboratory and according to the IOC regulations.

Number of samples	Quality	Predominant defect	Default medium
57	Extra Virgin	No defect	–
18	Virgin	Moldy	2.25
7	Virgin	Winey	2.61

oil was selected. The samples were placed in a glass container (1 L), which was then covered with parafilm. The tubes were kept at 25 °C for 8 min. For oil samples the procedure was similar, except that they were heated to 30 °C.

The electronic nose operating conditions during the measurements were established as follows: the pre-sampling time was 10 s, the measurement time was 70 s, the sampling time was 1 s, the flush time was 50 s, the zero point time was 10 s and finally the chamber and injection flow were set at 400 mL / min.

3. Results and discussion

3.1. Features extraction

Measurement of each sample made with the electronic nose is a transient time curve reflecting the electrical resistance value for each of the MOS included in the chamber. Each of the transient responses obtained was normalized based on the final resistance value obtained from each sensor after the cleaning process of the sensor chamber. The normalized transient response was formalized in the hypermatrix $x_{n,s,t}^{fruit}$ (Eq. 1) for the olive samples and in the hypermatrix $x_{n,s,t}^{oil}$ (Eq. 2) for oil samples.

$$x_{n,s,t}^{fruit} = \frac{R_{n,s,t}^{fruit}}{R_{n,s}^{0}} \tag{1}$$

$$x_{n,s,t}^{oil} = \frac{R_{n,s,t}^{oil}}{R_{n,s}^{0}} \tag{2}$$

where $R_{n,s,t}^{fruit}$ y $R_{n,s,t}^{oil}$ is the absolute resistance value read from the sensor s for olive and oil samples respectively, for the sample n and at the

Table 3
Metal oxide thick film sensors (MOS) in the sensor array used in the experiment and sensitivities and detection limits for specific organic and inorganic gases. MOS: Metal oxide thick film sensors. n.d.: not detected.

Sensor number	Sensor name	Sensor description and sensitivities	Detection limits
1	W1C	Aromatic organic compounds.	Toluene, 10 mg/kg
2	W5S	Very sensitive, broad range sensitivity, reacts to nitrogen oxides, very sensitive with negative signals.	$N0_2$, 1 mg/kg
3	W3C	Ammonia, also used as sensor for aromatic compounds.	Benzene, 10 mg/kg
4	W6S	Detects mainly hydrogen gas.	H_2, 0.1 mg/kg
5	W5C	Alkanes, aromatic compounds, and non-polar organic compounds.	Propane, 1 mg/kg
6	W1S	Sensitive to methane. Broad range of organic compounds detected.	CH_4, 100 mg/kg
7	W1W	Detects inorganic sulfur compounds. Also sensitive to many terpenes and sulfur containing organic compounds.	H_2S, 1 mg/kg
8	W2S	Detects alcohol, partially sensitive to aromatic compounds, broad range.	CO, 100 mg/kg
9	W2W	Aromatic compounds, inorganic sulfur and organic compounds.	H_2S, 1 mg/kg
10	W3S	Reacts to high concentrations of methane and aliphatic organic compounds.	n.d.

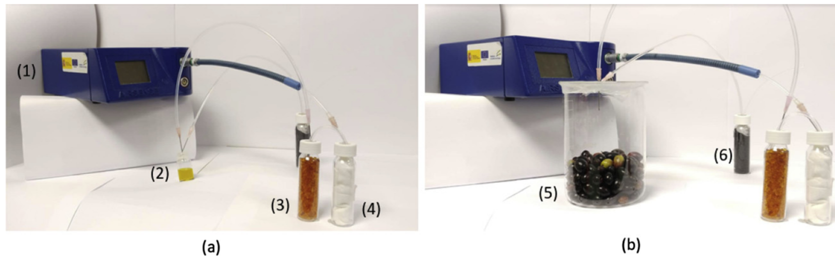

Fig. 3. Experimental setup for measurements on A. oil samples and B. olive fruit, where 1. presents the electronic nose device, 2. denotes olive oil, 3. shows the moisture filter, 4. is the particle filter, 5. denotes olive fruit container and 6. is the charcoal filter.

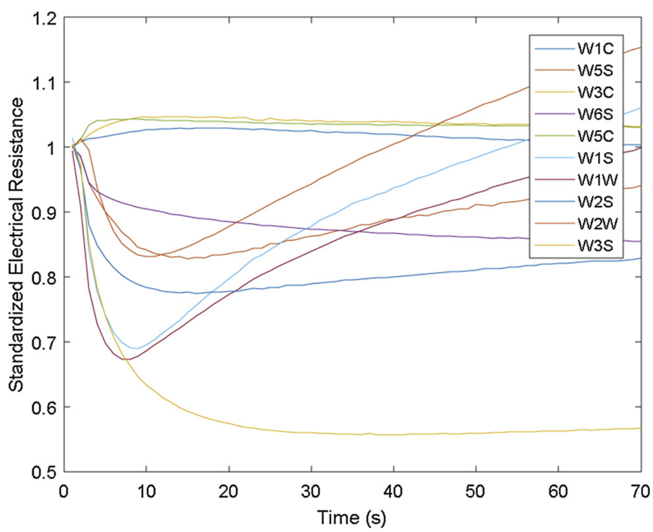

Fig. 4. Transient sensory response of the array of electronic nose sensors for an olive sample over 70 s.

instant t, and $R^0_{n,s}$ is the resistance of the sensor s after the cleaning process and before making the sample measurement n. There is a n to n correspondence between the olive samples and the oil samples made from the olive samples. Fig. 4 shows an example of the normalized response of each sensor during the process of measuring an olive sample.

The second stage of the feature extraction process was to compress the transient response of the sensor array to form a feature vector or olfactory fingerprint. According to the procedure used to generate the olfactory fingerprint, transient compression methods can be grouped into three methods (Gutierrez-Osuna, 2002; Pearce et al., 2002): sub-sampling, parameter extraction and systems identification. The first method is to consider only part of the answer, the second is to extract descriptors of the response and the third is based on identifying the transient response with a mathematical model. In our case, the second method was applied, the method of extracting descriptors. The first derivative was applied to the transient response of each sensor and the extracted descriptors appear in Table 4 where $v^{fruit}_{n,s,f}$ represents the vector of characteristics of the olive samples y $v^{oil}_{n,s,f}$ is the feature vector for oil samples.

Then the dimension of the three-dimensional matrices was reduced $v^{fruit}_{n,s,f}$ and $v^{oil}_{n,s,f}$ two-dimensional multiplying the second dimension and the third dimension, that is, $v^{fruit}_{n,s\times f}$ and $v^{oil}_{n,s\times f}$ respectively. The result is two two-dimensional matrices where the number of rows corresponds to the number of samples ($n = 82$) and the number of columns corresponds to the number of characteristics ($s \times f = 110$).

3.2. Classification algorithms

To evaluate the possibility of using the electronic nose to classify olive and oil samples according to their quality, three supervised classification algorithms were tested: NB, a PLSDA and MLP. The output vector with the class to which each sample belongs was formalized as $y_{n,1}$ where each row n contains the value of its class c. So for each sample n, $y_{n,1}$ is equal to $c_1 = 1$ it the class is “virgin” or it is equal to $c_2 = 2$ if the class is “extra virgin”. Below is a brief description of each of the classification algorithms used.

The NB classifier (John and Langley, 2013) is a probabilistic

Table 4
Descriptive characteristics of the transient response of the sensor array where $v^{fruit}_{n,s,f}$ is the vector of characteristics for olive samples, $v^{oil}_{n,s,f}$ is the feature vector for oil samples, n denotes the sample number (from 1 to 82), s is the sensor number (from 1 to 10) and f is the feature number (from 1 to 11).

Fruit	Oil	Description
$v^{fruit}_{n,s,1} = \max(x^{fruit}_{n,s})$	$v^{oil}_{n,s,1} = \max(x^{oil}_{n,s})$	Maximum value of the response without derivation.
$v^{fruit}_{n,s,2} = \min(x^{fruit}_{n,s})$	$v^{oil}_{n,s,2} = \min(x^{oil}_{n,s})$	Minimum value of the response without derivation.
$v^{fruit}_{n,s,3} = \sum_{t=1}^{70} x^{fruit}_{n,s,t}$	$v^{oil}_{n,s,3} = \sum_{t=1}^{70} x^{fruit}_{n,s,t}$	Sum of all the values of the response without derivation.
$v^{fruit}_{n,s,4} = \frac{v^{fruit}_{n,s,3}}{70}$	$v^{fruit}_{n,s,4} = \frac{v^{fruit}_{n,s,3}}{70}$	Average value of the response without derivation.
$v^{fruit}_{n,s,5} = \sqrt{\frac{\sum_{t=1}^{70} \lvert x^{fruit}_{n,s,t} - v^{fruit}_{n,s,4} \rvert^2}{69}}$	$v^{fruit}_{n,s,5} = \sqrt{\frac{\sum_{t=1}^{70} \lvert x^{fruit}_{n,s,t} - v^{fruit}_{n,s,4} \rvert^2}{69}}$	Standard deviation of the response without derivation.
$v^{fruit}_{n,s,6} = x^{fruit}_{n,s,70}$	$v^{fruit}_{n,s,6} = x^{fruit}_{n,s,70}$	Final value of the response without derivation.
$v^{fruit}_{n,s,7} = \max(x'^{fruit}_{n,s})$	$v^{oil}_{n,s,7} = \max(x'^{oil}_{n,s})$	Maximum value of the first derivative of the response.
$v^{fruit}_{n,s,8} = \min(x'^{fruit}_{n,s})$	$v^{oil}_{n,s,8} = \min(x'^{oil}_{n,s})$	Minimum value of the first derivative of the response.
$v^{fruit}_{n,s,9} = \sum_{t=1}^{70} x'^{fruit}_{n,s,t}$	$v^{oil}_{n,s,9} = \sum_{t=1}^{70} x'^{fruit}_{n,s,t}$	Sum of all the values of the first derivative of the response.
$v^{fruit}_{n,s,10} = \frac{v^{fruit}_{n,s,9}}{70}$	$v^{oil}_{n,s,10} = \frac{v^{oil}_{n,s,9}}{70}$	Average value of the first derivative of the response.
$v^{fruit}_{n,s,11} = \sqrt{\frac{\sum_{t=1}^{70} \lvert x'^{fruit}_{n,s,t} - v^{fruit}_{n,s,10} \rvert^2}{69}}$	$v^{oil}_{n,s,11} = \sqrt{\frac{\sum_{t=1}^{70} \lvert x'^{oil}_{n,s,t} - v^{oil}_{n,s,10} \rvert^2}{69}}$	Standard deviation of the first derivative of the response.

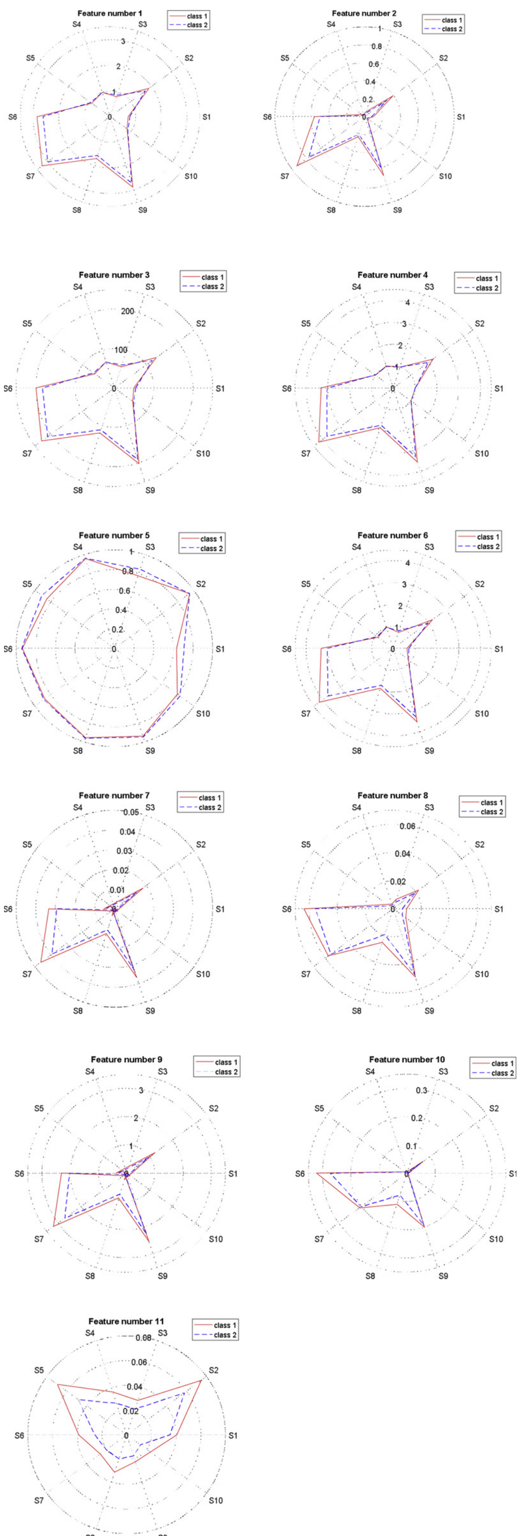

Fig. 5. Average value of the features extracted from the transient response of the electronic nose for olive fruit samples. Class 1 belongs to samples of virgin category and class 2 to samples of extra virgin category.

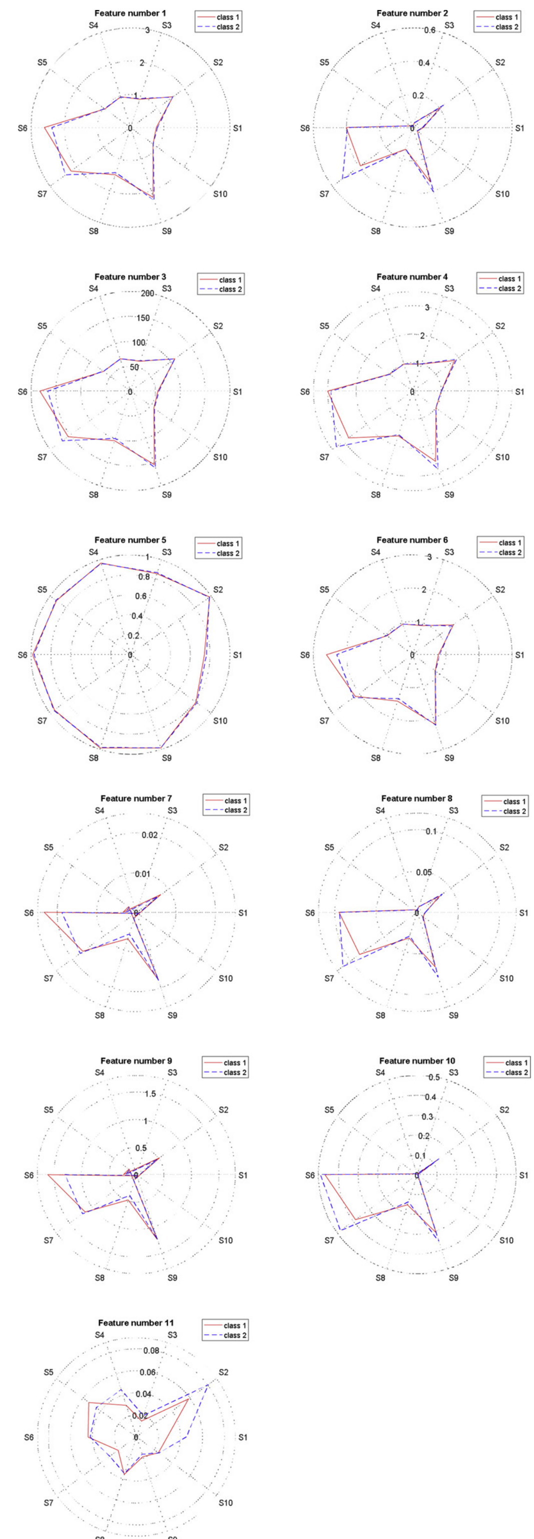

Fig. 6. Average value of the features extracted from the transient response of the electronic nose for oil samples. Class 1 belongs to samples of virgin category and class 2 to samples of extra virgin category.

classification method and is based on obtaining the probability of belonging to each class. Since the extracted characteristics are continuous variables and are normally distributed, the distribution of each class can be represented as a Gaussian probability density function in terms of its mean μ_c and standard deviation σ_c. In this way the probability of belonging of each sample to each class will be given by Eqs. (3) and (4).

$$p(v_{n,s\times f}c) = g(v_{n,s\times f};\, \mu_c,\, \sigma_c),\ \ where \tag{3}$$

$$g(v_{n,s\times f};\, \mu_c,\, \sigma_c) = \frac{1}{\sigma_c\sqrt{2\pi}}e^{-\frac{(v-\mu_c)^2}{2\sigma_c^2}} \tag{4}$$

where $g(v_{n,s\times f};\, \mu_c,\, \sigma_c)$ s the Gaussian probability density function.

Table 5
Results obtained for the different algorithms evaluated. NB: Naïve Bayes, PLSDA: Partial Least Squares Discriminant Analysis, MLP: Multilayer Perceptron, ROC: Receiver Operating Characteristic.

Data matrix used	Classification algorithm	Correctly classified	Incorrectly classified	SR (%)	ROC	Kappa statistic
NB	$v_{n,s\times f}^{oil}$	65	17	79.27	0.78	0.54
PLSDA		67	15	82.00	0.86	0.59
MLP		69	13	84.14	0.84	0.62
NB	$v_{n,s\times f}^{fruit}$	52	30	63.41	0.71	0.27
PLSDA		71	11	87.00	0.98	0.69
MLP		74	8	90.24	0.88	0.78
NB	$v_{n,s\times f}^{fruit+oil}$	58	24	70.73	0.73	0.38
PLSDA		71	11	87.00	0.86	0.69
MLP		67	15	81.70	0.85	0.56

On the other hand, the PLSDA algorithm (Ballabio and Consonni, 2013) identifies linear combinations of the variables $v_{n,s\times f}$ that maximize their variance using class information, $y_{n,1}$. The multivariate model is defined by the equations Eq. (5) y Eq. (6).

$$V = TP^T + E \tag{5}$$

$$Y = UQ^T + F \tag{6}$$

where T represents the scores of V and P is the loadings matrix of V. The same goes for Eq. (6). The objective of PLSDA is to find the latent variables TP^T y UQ^T that maximize the covariance with the class vector Y.

Finally, the MLP classifier (Duda et al., 2001) was evaluated. The multilayer perceptron neural network is a neural network formed by multiple layers and has the advantage of being able to solve classification problems where the classes are not linearly separable. For our case, the MLP network was configured with three layers. An input layer with 110 nodes (equal to the number of features extracted), a hidden layer with 58 nodes with sigmoidal transfer functions and an output layer with 2 nodes (equal to the number of classes). The training of the network was backpropagation type and the number of epochs used to train the MLP was 500. This limit was established experimentally by checking that a higher number of epochs did not reduce the training error.

The different classification algorithms were applied three times: 1. characteristics of the olive samples $v_{n,s,f}^{fruit}$, 2. characteristics of the oil samples $v_{n,s,f}^{oil}$, and 3. characteristics of the oil samples together with those of olives $v_{n,s,f}^{fruit+oil}$.

The performance of each algorithm was evaluated applying cross-validation type 10-fold. This type of validation involves 10 iterations. In each iteration 90 % of the data set is used to train the model and the remaining 10 % for the test.

As a result, 10 confusion matrices are obtained $MC_{i\times n\times n}$ where in each iteration i the sum of the elements of the diagonal is related to the success rate SR. The total success rate will be obtained according to Eq. (7)

$$SR(\%) = \frac{\sum_{i=1}^{10} MC_{i\times 1\times 1} + \sum_{i=1}^{10} MC_{i\times 2\times 2}}{82} \times 100 \tag{7}$$

Figs. 5 and 6show the average values for each sensor and for each parameter obtained from the transient response of the electronic nose when the measures were performed over olive fruits and olive oils respectively. Two classes are presented where class 1 is "virgin" and class 2 is "extra virgin".

The most discriminating sensor between classes was sensor number 7 (Fig. 5). The values between classes were higher for features number 2, 4, 6, 7 and 9. According to the datasheet of the electronic nose (PEN3) employed in this research, sensor number 7 (W1W) is sensitive to many terpenes and the presence of fermented olive fruit in lower quality batches is related to the percentage of total terpenes (Dabbou et al., 2012). This issue could explain the differences between features for the sensor 7. Moreover, the presence of volatile compounds of terpenes is partly responsible for the aroma of the olive oil and consequently of its quality (Reboredo-Rodríguez et al., 2013a). This relationship may explain that the sensor 7 denotes greater discriminatory power in the majority of features extracted from the measurements made on olive oils (Fig. 6).

When the characteristics of the oil samples were used (matrix $v_{n,s\times f}^{oil}$) the classification algorithm that obtained the best result was the MLP neural network with a success rate of 84.14 % (of 82 samples, 69 were correctly classified). This result can be seen in Table 5 and it confirms the separability of the classes previously documented (Guadarrama et al., 2001b) but without applying classification algorithms. James et al. (2004) found that 99 % success was achieved, although the quality of the samples was not the same as the authors classified virgin olive oil, lampante oil and sunflower oil.

The best result was obtained with the characteristics extracted from the transient sensory response on olives ($v_{n,s\times f}^{fruit}$). The best result was obtained with MLP with a success in classification of 90.2 %. This result confirms the differences between the two kinds of olives seen in Fig. 5 and indicates that the incorporation of an electronic nose at the input of the olive oil production process can be useful to detect defects before the oil is produced and to be able to classify olives based on the quality of the oil that is going to be produced.

Finally, the characteristics extracted from the oil samples were merged with the characteristics extracted from the olive samples ($v_{n,s\times f}^{fruit+oil}$) and the best success rate was 87 % with the PLSDA algorithm. This result confirms the aromatic intensity of the olives has more information than the aromatic intensity of the oil for detection of defects.

4. Conclusions

The parameters and methodology to determine the quality of olive oil are regulated by European regulations and this quality depends mainly on the olive fruit conditions after harvest. This paper proposes a rapid methodology based on an electronic nose sensor and on different pattern recognition algorithms to predict the organoleptic quality of virgin and extra virgin olive oil. The measurements with the electronic nose were made on olives and on the oil made from these fruit. Three classification algorithms were evaluated: NB, PLSDA and MLP neural network. The best classification success ratio was 84.14 % for olive oil samples and 90.2 % for olive fruit samples, both with the MLP algorithm. The proposed methodology could be used in the field to predict the quality of the oil that is going to be made with the freshly harvested fruits. Also the system could be installed at the input of the virgin olive oil production process to classify fruits according to their quality.

4.1. Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been partially supported by the project of the Ministry of Spain with reference DPI2016-78290-R. The authors thank the oil mill Picualia (www.piculia.com) for the olive samples provided to carry out this study. Also thank CM Europa (www.cmeuropa.com) for the organoleptic analysis.

References

Aguilera Puerto, D., Cáceres Moreno, Ó., Martínez Gila, D.M., Gómez Ortega, J., Gámez García, J., 2019. Online system for the identification and classification of olive fruits for the olive oil production process. J. Food Meas. Charact. 13, 716–727. https://doi.org/10.1007/s11694-018-9984-0.

Angerosa, F., Servili, M., Selvaggini, R., Taticchi, A., Esposto, S., Montedoro, G., 2004. Volatile compounds in virgin olive oil: occurrence and their relationship with the quality. J. Chromatogr. A 1054, 17–31. https://doi.org/10.1016/j.chroma.2004.07.093.

Ballabio, D., Consonni, V., 2013. Classification tools in chemistry. Part 1: linear models. PLS-DA. Anal. Methods 5, 3790. https://doi.org/10.1039/c3ay40582f.

Beltrán Ortega, J., Martínez Gila, D.M., Aguilera Puerto, D., Gámez García, J., Gómez Ortega, J., 2016. Novel technologies for monitoring the in-line quality of virgin olive oil during manufacturing and storage. J. Sci. Food Agric. 96, 4644–4662. https://doi.org/10.1002/jsfa.7733.

Chen, H., Zhang, M., Guo, Z., 2019. Discrimination of fresh-cut broccoli freshness by volatiles using electronic nose and gas chromatography-mass spectrometry. Postharvest Biol. Technol. 148, 168–175. https://doi.org/10.1016/J.POSTHARVBIO.2018.10.019.

Concina, I., Falasconi, M., Gobbi, E., Bianchi, F., Musci, M., Mattarozzi, M., Pardo, M., Mangia, A., Careri, M., Sberveglieri, G., 2009. Early detection of microbial contamination in processed tomatoes by electronic nose. Food Control 20, 873–880. https://doi.org/10.1016/J.FOODCONT.2008.11.006.

Dabbou, S., Issaoui, M., Brahmi, F., Nakbi, A., Chehab, H., Mechri, B., Hammami, M., 2012. Changes in volatile compounds during processing of tunisian-style table olives. J. Am. Oil Chem. Soc. 89, 347–354. https://doi.org/10.1007/s11746-011-1907-8.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern classification. Wiley.

González Martín, Y., Cerrato Oliveros, M.C., P'rez Pavón, J.L., García Pinto, C., Moreno Cordero, B., 2001. Electronic nose based on metal oxide semiconductor sensors and pattern recognition techniques: characterisation of vegetable oils. Anal. Chim. Acta 449, 69–80. https://doi.org/10.1016/S0003-2670(01)01355-1.

González Martín, Y., Luis Pérez Pavón, J., Moreno Cordero, B., García Pinto, C., 1999. Classification of vegetable oils by linear discriminant analysis of Electronic. Nose data. Anal. Chim. Acta 384, 83–94. https://doi.org/10.1016/S0003-2670(98)00851-4.

Guadarrama, A., Rodríguez-Méndez, M.L., Sanz, C., Ríos, J.L., De Saja, J.A., 2001a. Electronic nose based on conducting polymers for the quality control of the olive oil aroma. Anal. Chim. Acta 432, 283–292. https://doi.org/10.1016/S0003-2670(00)01383-0.

Guadarrama, A., Rodríguez-Méndez, M.L., Sanz, C., Ríos, J.L., De Saja, J.A., 2001b. Electronic nose based on conducting polymers for the quality control of the olive oil aroma. Anal. Chim. Acta 432, 283–292. https://doi.org/10.1016/S0003-2670(00)01383-0.

Gutierrez-Osuna, R., 2002. Pattern analysis for machine olfaction: a review. IEEE Sens. J. 2, 189–202. https://doi.org/10.1109/JSEN.2002.800688.

Huang, L., Meng, L., Zhu, N., Wu, D., 2017. A primary study on forecasting the days before decay of peach fruit using near-infrared spectroscopy and electronic nose techniques. Postharvest Biol. Technol. 133, 104–112. https://doi.org/10.1016/J.POSTHARVBIO.2017.07.014.

International Olive Oil Council, 2018. Wold Olive Oil Production. http://www.internationaloliveoil.org/documents/viewfile/4244-production1-ang/.

IOC, 2018. Sensory Analysis of Olive Oil – Method for the Organoleptic Assessment of Virgin Olive Oil. http://www.internationaloliveoil.org/documents/viewfile/3685-orga6.

James, D., Scott, S.M., O'hare, W.T., Ali, Z., Rowell, F.J., 2004. Classification of fresh edible oils using a coated piezoelectric sensor array-based electronic nose with soft computing approach for pattern recognition. Trans. Inst. Meas. Control 26, 3–18. https://doi.org/10.1191/0142331204tm0102oa.

Jiang, S., Wang, J., 2016. Internal quality detection of Chinese pecans (Carya cathayensis) during storage using electronic nose responses combined with physicochemical methods. Postharvest Biol. Technol. 118, 17–25. https://doi.org/10.1016/J.POSTHARVBIO.2016.03.016.

John, G.H., Langley, P., 2013. Estimating Continuous Distributions in Bayesian Classifiers.

Lerma-García, M.J., Cerretani, L., Cevoli, C., Simó-Alfonso, E.F., Bendini, A., Gallina Toschi, T., 2010. Use of electronic nose to determine defect percentage in oils. Comparison with sensory panel results. Sens. Actuators B Chem. 147, 283–289. https://doi.org/10.1016/j.snb.2010.03.058.

Liu, Q., Sun, K., Zhao, N., Yang, J., Zhang, Y., Ma, C., Pan, L., Tu, K., 2019. Information fusion of hyperspectral imaging and electronic nose for evaluation of fungal contamination in strawberries during decay. Postharvest Biol. Technol. 153, 152–160. https://doi.org/10.1016/J.POSTHARVBIO.2019.03.017.

Ma, Y., Guo, B., Wei, Y., Wei, S., Zhao, H., 2014. The feasibility and stability of distinguishing the kiwi fruit geographical origin based on electronic nose analysis. Food Sci. Technol. Res. 20, 1173–1181. https://doi.org/10.3136/fstr.20.1173.

Majchrzak, T., Wojnowski, W., Dymerski, T., Gębicki, J., Namieśnik, J., 2018. Electronic noses in classification and quality control of edible oils: a review. Food Chem. 246, 192–201. https://doi.org/10.1016/j.foodchem.2017.11.013.

Martínez Gila, D.M., Cano Marchal, P., Gómez Ortega, J., Gámez García, J., 2018. Non-invasive methodology to estimate polyphenol content in extra virgin olive oil based on stepwise multilinear regression. Sensors 18, 975. https://doi.org/10.3390/s18040975.

Morales, M.T., Luna, G., Aparicio, R., 2005. Comparative study of virgin olive oil sensory defects. Food Chem. 91, 293–301. https://doi.org/10.1016/j.foodchem.2004.06.011.

Navarro Soto, J., Satorres Martínez, S., Martínez Gila, D.M., Gómez Ortega, J., Gámez García, J., 2018. Fast and reliable determination of virgin olive oil quality by fruit inspection using computer vision. Sensors 18, 3826. https://doi.org/10.3390/S18113826.

Pearce, T.C., Schiffman, S.S., Nagle, H.T., Gardner, J.W. (Eds.), 2002. Handbook of Machine Olfaction. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG. https://doi.org/10.1002/3527601597.

Raigar, R.K., Upadhyay, R., Mishra, H.N., 2017. Storage quality assessment of shelled peanuts using non-destructive electronic nose combined with fuzzy logic approach. Postharvest Biol. Technol. 132, 43–50. https://doi.org/10.1016/J.POSTHARVBIO.2017.05.016.

Reboredo-Rodríguez, P., González-Barreiro, C., Cancho-Grande, B., Simal-Gándara, J., 2014. Improvements in the malaxation process to enhance the aroma quality of extra virgin olive oils. Food Chem. 158, 534–545. https://doi.org/10.1016/j.foodchem.2014.02.140.

Reboredo-Rodríguez, P., González-Barreiro, C., Cancho-Grande, B., Simal-Gándara, J., 2013a. Concentrations of aroma compounds and odor activity values of odorant series in different olive cultivars and their oils. J. Agric. Food Chem. 61, 5252–5259. https://doi.org/10.1021/jf400804m.

Reboredo-Rodríguez, P., González-Barreiro, C., Cancho-Grande, B., Simal-Gándara, J., 2013b. Effects of sedimentation plus racking process in the extra virgin olive oil aroma fingerprint obtained by DHS–TD/GC–MS. Food Bioprocess Technol. 6, 1290–1301. https://doi.org/10.1007/s11947-011-0751-z.

Wang, Q., Li, L., Ding, W., Zhang, D., Wang, J., Reed, K., Zhang, B., 2019a. Adulterant identification in mutton by electronic nose and gas chromatography-mass spectrometer. Food Control 98, 431–438. https://doi.org/10.1016/J.FOODCONT.2018.11.038.

Wang, Z.-C., Yan, Y., Nisar, T., Sun, L., Zeng, Y., Guo, Y., Wang, H., Fang, Z., 2019b. Multivariate statistical analysis combined with e-nose and e-tongue assays simplifies the tracing of geographical origins of Lycium ruthenicum Murray grown in China. Food Control 98, 457–464. https://doi.org/10.1016/J.FOODCONT.2018.12.012.

Wen, T., Zheng, L., Dong, S., Gong, Z., Sang, M., Long, X., Luo, M., Peng, H., 2019. Rapid detection and classification of citrus fruits infestation by Bactrocera dorsalis (Hendel) based on electronic nose. Postharvest Biol. Technol. 147, 156–165. https://doi.org/10.1016/J.POSTHARVBIO.2018.09.017.

Wu, H., Yue, T., Xu, Z., Zhang, C., 2017. Sensor Array Optimization and Discrimination of Apple Juices According to Variety by an Electronic Nose. https://doi.org/10.1039/c6ay02610a.

Xing, M., Sun, K., Liu, Q., Pan, L., Tu, K., 2018. Development of novel electronic nose applied for strawberry freshness detection during storage. Int. J. Food Eng. 14. https://doi.org/10.1515/ijfe-2018-0111.

Zhang, X., Cheng, J., Wu, L., Mei, Y., Jaffrezic-Renault, N., Guo, Z., 2018. An overview of an artificial nose system. Talanta 184, 93–102. https://doi.org/10.1016/j.talanta.2018.02.113.