Microchemical Journal 206 (2024) 111585

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Employment of MQ gas sensors for the classification of *Cistus ladanifer* essential oils

Francisco Javier Diaz Blasco [a], Sandra Viciano-Tudela [a], Lorena Parra [a], Ali Ahmad [a], Veronika Chaloupková [b], Raquel Bados [b], Luis Saul Esteban Pascual [b], Irene Mediavilla [b], Sandra Sendra [a], Jaime Lloret [a,*]

[a] *Instituto de Investigación para la Gestión Integrada de Zonas Costeras, Universitat Politècnica de València, Gandía C/Paranimf, 1, 46730 Grao de Gandia, Spain*
[b] *Centro de Desarrollo de Energías Renovables (CEDER-CIEMAT). Autovía de Navarra A15, salida 56, 42290 Lubia, SORIA, Spain*

ARTICLE INFO

Keywords:
Metal oxide sensors
Multiclass classification
Viridiflorol, Alpha-pinene
Artificial Intelligence

ABSTRACT

The chemical composition of essential oils (EOs) from *Cistus ladanifer* has a huge variability throughout the year, impacting the oil quality. Nowadays, EO analytic chemistry techniques, which are expensive and destroy the sample, are utilized to measure the chemical composition. In the paper, we propose a combination of low-cost sensors and machine learning based system. As low-cost sensors, seven gas sensors are combined to obtain up to 36 features. Regarding machine learning, 31 multiclass classification algorithms are applied. Data from sensors were collected for 33 samples of EO from *Cistus ladanifer*. The generated dataset was split into training and test datasets, with 75 % of the data for training. The datasets were created to ensure a homogeneous chemical composition distribution on both training and test datasets. There were three target chemical compounds: Alpha-pinene and Viridiflorol as individual compounds and Terpenic Hydrocarbons as a group of chemical compounds. The value of the percentage of each targeted compound is converted into a categoric variable with 5 possible values, 1 being the lowest concentration and 5 being the maximum one. The data of the MQ-sensors were included as the input for the models, and each one of the targeted chemical compounds was selected as an output for different models. The input features were ranged using different algorithms for the feature selection process. The results indicate that there is no valid classification model for Viridiflorol, and limited accuracy is achieved for Alpha-pinene. Meanwhile, for Terpenic Hydrocarbons, an accuracy of 91.6 % is achieved. It is important to highlight that these accuracies were attained when a reduced number of features were included, ranging the number of features from 11 to 13. This is the first case in which MQ-based gas sensors, or other metal oxide sensors, are used to correctly determine the concentration of a chemical compounds in a complex matrix formed by dozens of compounds. This system will provide a cheap method to determine the quality of EOs and confirm the benefits of combining low-cost sensors with machine learning.

1. Introduction

Essential oils (EOs) are volatile, aromatic liquids composed of various secondary metabolites, including terpenes, phenolic compounds, and alcohols. These compounds, synthesized by nearly all plants, are typically extracted from seeds, fruits, flowers, buds, leaves, stems, and roots [10]. Despite their widespread presence in medicinal and aromatic plants, EOs constitute less than 5 % of a plant's total dry matter [34]. They contribute to the distinctive scents of aromatic plants due to their characteristic odors [10,34]. Chemically, EOs are intricate mixtures of bioactive components, including terpenes, terpenoids, and phenolics [37].

Cistus ladanifer, an economically significant shrub, exemplifies the importance of EOs. Characterized by sesquiterpenes (viridiflorol and ledol) and monoterpenoids (bornyl acetate and pinocarveol), the EO of *C. ladanifer* is highly sought after in the perfume industry for its fragrant and fixative properties [30,26]. Recent chemical analyses have identified major compounds such as α-pinene, viridiflorol, (E)-pinocarveol, p-

* Corresponding author.
E-mail addresses: fjdiabla@doctor.upv.es (F.J. Diaz Blasco), svictud@upv.es (S. Viciano-Tudela), loparbo@doctor.upv.es (L. Parra), aahmad1@upv.es (A. Ahmad), veronika.chaloupkova@ciemat.es (V. Chaloupková), raquel.bados@ciemat.es (R. Bados), luis.esteban@ciemat.es (L.S. Esteban Pascual), irene.mediavilla@ciemat.es (I. Mediavilla), sansenco@upv.es (S. Sendra), jlloret@dcom.upv.es (J. Lloret).

https://doi.org/10.1016/j.microc.2024.111585
Received 16 May 2024; Received in revised form 7 August 2024; Accepted 4 September 2024
Available online 7 September 2024
0026-265X/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

cymene, camphene, bornyl acetate, and ledol in *C. ladanifer* EO [33]. Out of these components, α-pinene is often found to be a major constituent ranging from approximately from 30.4 % to 53 %, thus contributing to the oil's characteristic properties [42,12]. Likewise, viridiflorol constitutes about 17.64 % of the essential oil, highlighting its prominence in the composition [42]. Whereas the percentage of terpenic hydrocarbons can vary depending on factors such as extraction methods and growing conditions.

Gas chromatography-mass spectrometry (GC–MS) is the predominant method for detecting and analyzing volatile organic compounds (VOCs), frequently employed alongside headspace solid-phase microextraction. This analytical approach integrates gas chromatography and mass spectrometry through a heated transfer line, facilitating comprehensive analysis [13,22]. VOCs, a broad array of chemicals from sources such as foods, bacteria, and plants [19,21], are detected by the human olfactory system, which, despite its sensitivity, has limitations in discriminating odors. Low-cost metal-oxide gas sensors, or electronic noses (E-noses), have emerged as a promising alternative, offering enhanced sensitivity and selectivity with reduced training requirements [6].

MQ-series gas sensors, known for their cost-effectiveness and high sensitivity, operate on electrochemical principles and modulate resistance in response to specific gases. Widely used in environmental monitoring [9], health [24,29], agriculture [25,28], engineering [18], and the food industry [38], these sensors have proven effective in applications such as detecting spoilage in stored fish [40]and VOCs like acetone, ethanol, and isopropyl alcohol [16]. MQ sensors have also been instrumental in detecting adulteration in EOs. Studies using MQ sensors to detect adulteration in Patchouli Oil and other EOs have achieved significant accuracy, employing methods like the K-nearest neighbors (KNN) algorithm [32,36]. Similarly, Karami et al. [17] demonstrated high accuracy in detecting adulteration in blended edible oils using an E-nose. Roy et al. [31] developed a system utilizing multiple MQ sensors for detecting adulteration in ghee, achieving significant classification accuracy with Principal Component Analysis (PCA) and Discriminant Function Analysis (DFA). In other studies, specific VOCs, including formaldehyde, were successfully differentiated both remotely and on-site, whereas some studies reported utilizing a smart gas sensor system enhanced with machine learning capabilities and integrated within an IoT platform [1,2,4]. In *Cistus ladanifer* EO, MQ sensors have shown significant classification accuracy in assessing adulteration [35,36]. Data from these sensors are often analyzed using advanced statistical and Artificial Intelligence (AI) techniques [3], such as the KNN algorithm and nonlinear autoregressive with exogenous input (NARX) network, demonstrating high accuracy and correlation rates [8]. These methodologies underscore the effectiveness of integrating AI with sensor technology for comprehensive quality assessment and control across various domains.

In consideration of the aforementioned, the present investigation endeavors to discern the variations in the EOs of *C. ladanifer* by analyzing the presence of corresponding organic compounds through the utilization of MQ sensors, specifically MQ2, MQ3, MQ4, MQ5, MQ6, MQ7, and MQ8. On selecting the MQ series sensors (MQ2 − MQ8), we prioritized those with high sensitivity to gases such as methane, propane, butane, and hydrogen, which are indicative of the presence of terpenic hydrocarbons [35,36]. These sensors provide the necessary specificity and sensitivity for our targeted compounds. Conversely, while the MQ135 sensor is well-regarded for detecting a wide range of VOCs, its broader detection range includes many substances irrelevant to our study, potentially complicating the data interpretation for our specific focus on terpenic hydrocarbons. Therefore, the MQ135 sensor was deemed less suitable for this precise application.

Given the significance of EOs as a valuable raw material across various industries, this study aims to introduce a the proposed approach for characterizing EOs, leveraging low-cost sensor networks and machine learning techniques; thereby contributing to the identification of compositional variation in EO obtained throughout the year. Additionally, this non-destructive method for EO characterization holds the potential to inform precise harvesting time for obtaining optimal oil composition. With machine learning classification models, it is possible to find models that predict the class of a test dataset based on the training dataset. This method allows data to perform the predictions even if no clear correlations are found between predictors and output response.

2. Materials and methods

In this section, a detailed description of the EOs sample collection, analysis, and subsequent data processing is provided. Similarly, the sensor and its corresponding node employed in this study are also described comprehensively.

2.1. Sample description

2.1.1. Plant materials

Sample collection was carried out in Bustares and Hiendelaencina, two locations in the Sierra Norte de Guadalajara (Central Spain), a Mediterranean inland climate area [27]were rockrose (*Cistus ladanifer* L.) occupies large areas in a monospecific way. These two locations correspond to different rockrose shrub age (7 and 12 years old respectively) in order to evaluate the influence of the rockrose age on the EO production. The central point UTM coordinates (ETRS 89) of the sampling areas are: Bustares (30 N 495,423 4552862) and Hiendelaencina (30 N 499691, 4546683).

Based on a random sampling of the vegetation, the shrubland in Bustares was 58 % crown cover, 105 cm average height and 92 cm average plant crown diameter. In Hiendelaencina, the shrubland was 70 % crown cover, 115 cm average height and 96 cm crown diameter.

Rockrose was sampled manually using two-handed pruning shears to cut the plants at a height of 20 cm from the ground, simulating the cutting height in a mechanized harvesting. The plants were grouped in bundles of 12–15 kg weight to facilitate handling and transport. Sampling frequency varied throughout the year (once a month from February to August and every 15 days from September to January, corresponding to the period of highest EO production).

The sampling methodology at both locations was as follows: Randomly, a central point is chosen to start the sample collection, covering each day a different area within the total area to be sampled (10 ha in Bustares and 3 ha in Hiendelaencina). At a distance of 20 m in a random direction, the next sampling point is chosen. Around it, rockrose is cut to form another bundle. Then, in a perpendicular direction and at a distance of 20 m, another collection point is established, and so on until a daily rockrose sample of 80 kg is completed. In addition, the following dasometric parameters were measured each sampling day of ten plants to characterize the scrubland area: shrub crown cover and total height and crown diameter. About 60 kg of fresh wild plants were collected between 18th October 2021 and 26th September 2022 per location and were stored under shelter between 24 and 48 h.

2.1.2. Steam distillation and EO analysis

After collection and storage, the biomass was crushed with a screen size of 20 mm using a shredder (18.5 kW, slow-rotating single-shaft type) and then, the biomass was distilled in a steam distillation pilot plant, performing three distillation tests per sample (Fig. 1 A). This pilot plant is composed by a stainless steel still (50 L) where 15 kg of fresh ground biomass were loaded. A steam flow of 11.5 kg/h (0.5 barg) produced by an electric boiler extracted the EO from the plant and was condensed in a cooling coil. Finally, the separation of the EO and the hydrolate (aqueous phase) was performed in a glass Florentine flask (Fig. 1 B). The steam distillation tests had a duration of 1 h per test and, once the EO sample was separated, it was weighted and stored in a refrigerator at 4 °C.

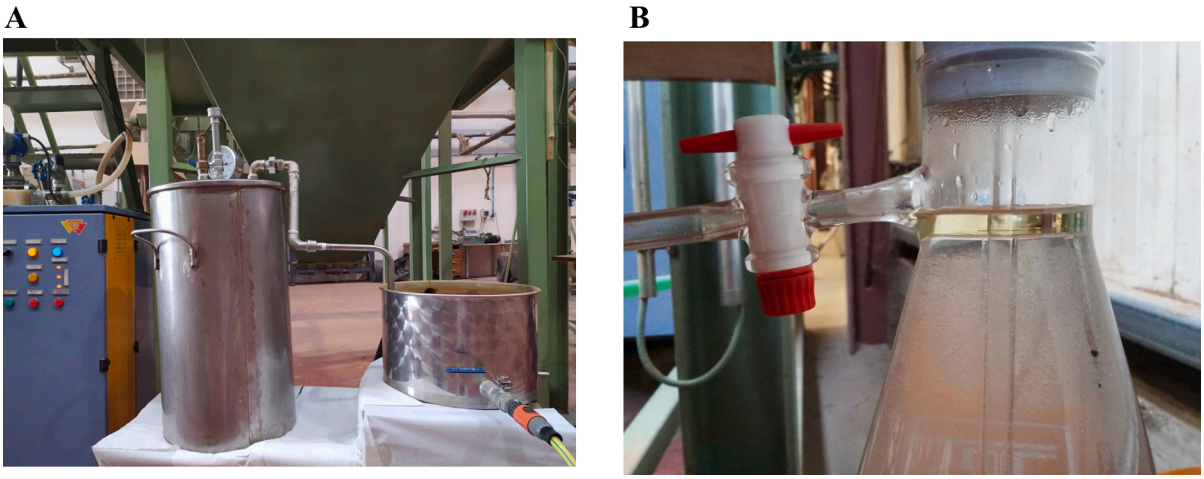

Fig. 1. Pilot plant for EO extraction. A) Steam distillation pilot plant B) EO and hydrolate detail.

2.1.3. Eos samples

Following the collection and analysis of the EOs, a detailed description of the samples is provided. Some of the EOs samples used for VOC measurement are shown in Fig. 2. The EOs were analysed by GC (FID)-MS using the same procedure as described in [23].

A total of 33 samples were obtained from *C. ladanifer* plants grown between February 28, 2021, and September 26, 2022. Subsequently, all the compounds extracted from the previous distillation and analysis of the EO's were studied. Thus, obtaining that Alpha-pinene, Viridiflorol and a combination of Terpenic Hydrocarbons, were the most remarkable. Whereas the concentration ranges for Alpha-pinene were found to be from 15.81 to 61.07 %; for Viridiflorol between 2.76 and 18.37 %; and for Terpenic Hydrocarbons from 22.48 to 71.49 %. These concentration ranges are shown in Fig. 3.

2.2. Sensors description

MQ gas sensors constitute a range of devices meticulously crafted to identify diverse chemical components within the air. The MQ sensor lineup encompasses a diverse array, with each model tailored to identify specific substances for designated purposes. These applications may include detecting flammable gases, assessing air quality, or identifying alcohol presence in exhaled air (Table 1). Thus, a total of 35 different sensor values were obtained using the proposed sensors. Each sensor value corresponds to the signal of a specific sensor and the application of calibration equations provided by the manufacturer. These are the 35 features used for machine learning. Besides these 35 features, an additional feature is added, the timestamp. There are a total of 7 MQ gas sensors; those sensors are relatively inexpensive devices. The price varies from one model to another but, in general, ranges between 1 and 1.5 €, from international sellers on eBay [11]and AliExpress [7].

An integral component of MQ sensors is the electrochemical sensor, which undergoes resistance changes upon contact with various substances. Common to all MQ models is a heater essential for elevating the sensor's temperature, rendering its materials sensitive. It is crucial to note that until the heater attains the operational temperature, the sensor's readings cannot be deemed reliable.

The proposed model consists of two different circuits inside the sensor (Fig. 4). One circuit, with heating function, is situated between points (c) on the diagram. The second circuit spanning points (a) to (b) serves the purpose of quantifying the resistive alterations in the sensor induced by the presence of gas. A diverse array of such sensors exhibits sensitivity to various gases that are inherently present in our environment.

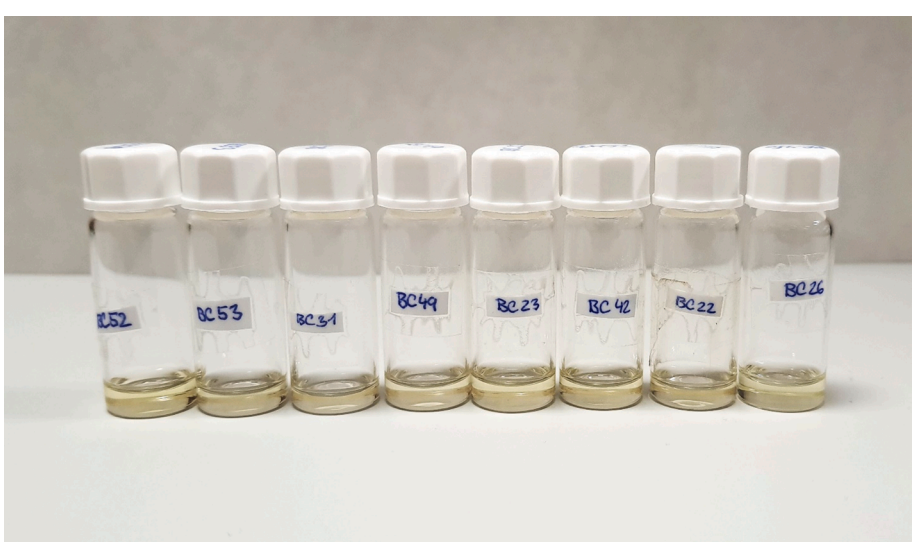

Fig. 2. Image of some of the oil samples utilized for sensor-based data collection.

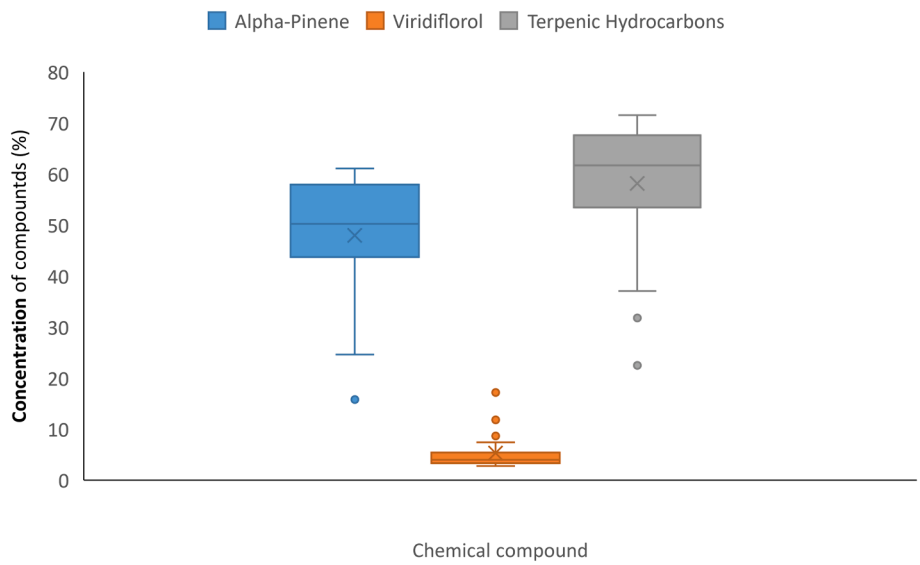

Fig. 3. Maximum and minimum concentration ranges for the studied organic compounds in *Cistus ladanifer* EOs.

Table 1
Summary of compounds detected by each MQ sensor.

Type of MQ	Compound Detection									
	Methane	Butane	Liquified Petroleum Gas (LPG)	Smoke	Alcohol	Ethanol	Compressed Natural Gas (CNG)	Natural Gas	Carbon Monoxide (CO)	Hydrogen (H)
MQ2	X	X	X	X						
MQ3				X	X	X				
MQ4	X						X			
MQ5			X					X		
MQ6		X	X							
MQ7									X	
MQ8										X

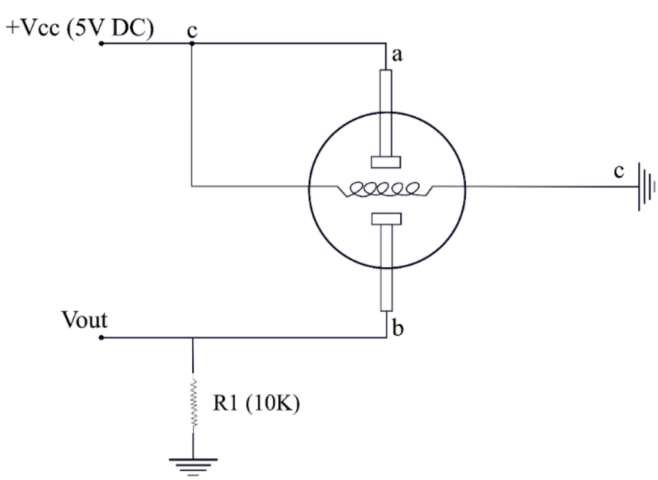

Fig. 4. Pictorial description of the MQ sensor system.

2.3. Node selection

An Arduino Mega 2560 is selected for this application. This node's board is based on the ATmega2560 microcontroller, having more computing power, more memory capacity and more expansion lines compared to previous models. This node is ideal for projects with large numbers of sensors. Arduino Mega 2560 can be powered through USB connection or through an external power supply and can operate with a 6 V to 20 V supply. It has a FLASH memory capacity of 256 KB, for storing program code, of which 8 KB is used for the boot loader, 8 KB of SRAM and 4 KB of EEPROM. This node has 54 digital pins, which can be used for either input or output, operating at 5 V and providing or receiving 20 mA.

The system is based on a series of MQ sensors, connected to the node as can be seen in Fig. 5. These sensors are connected to Raspberry Pi 4 microcontrollers that collect the data and store it in a MariaDB database (DB). The Raspberry Pi sends the data collected to an Access Point to be downloaded by the user when needed.

2.4. Data gathering procedure

In this study, a rigorous experimental protocol, characterized by meticulous sample handling, precise data acquisition, and controlled atmospheric conditions: reinforcing the scientific rigor and integrity of the study, was followed.

Precisely, the experimental procedure commenced with the retrieval of oil samples from the refrigerator, which were previously stored at 4 °C to ensure their preservation and integrity. They were left at room temperature for 5 to 10 min to acclimatize. Subsequently, 0.05 mL of these samples were then introduced into a closed measuring chamber as depicted in Fig. 6, using a Pasteur pipette, where they underwent a process of volatilization at ambient temperature. Within this chamber, MQ gas sensors were strategically fitted to record the corresponding levels of VOCs, ensuring precise and sensitive detection.

Each oil sample was meticulously introduced into the measurement chamber, with a single drop serving as the medium for analysis. The sampling interval for data collection was 2 h facilitating sensor cleansing and ensuring the purity of the atmospheric conditions within the chamber, with readings measured over a 24-hour period to allow for the

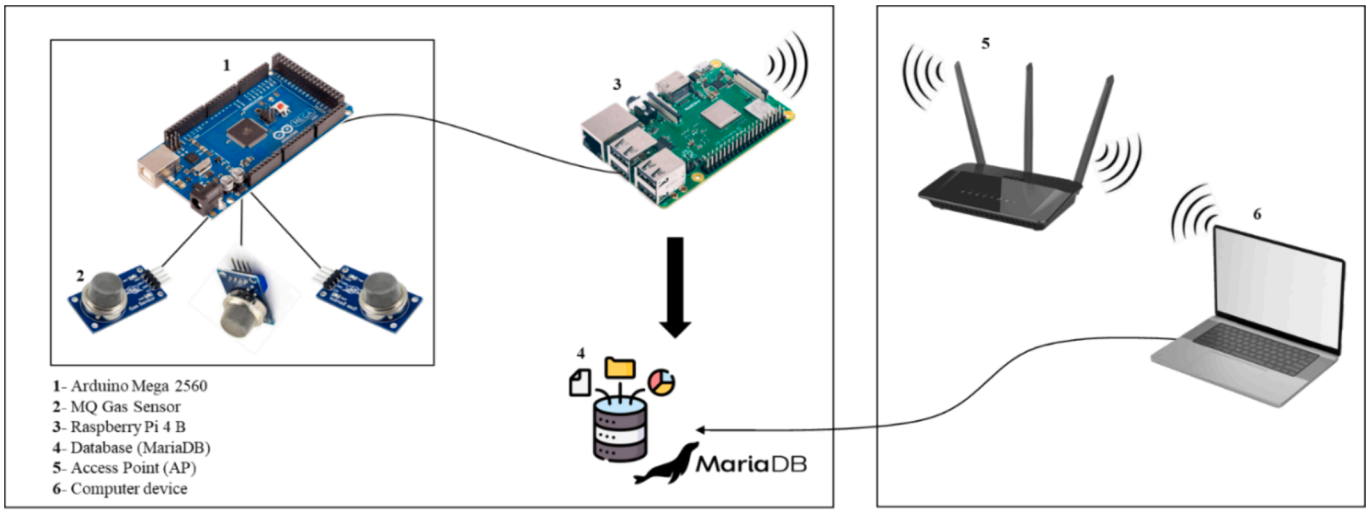

Fig. 5. Illustration of the sensing node. Whereas numbers represent each part of the system.

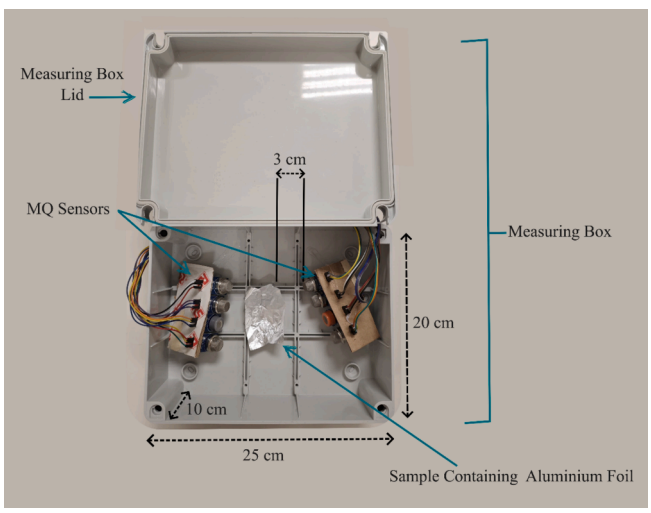

Fig. 6. Detailed view of the Oolatile Organic Compounds (VOCs) measuring box along with its dimensions and MQ gas sensors.

complete volatilization of compounds at an ambient temperature oscillating between 25 °C to 34 °C and humidity levels ranging from 22 % to 38 %. The gas sensors collected a comprehensive dataset, comprising a minimum of 1000 readings, capturing nuanced variations in VOC levels. A stepwise description of the process is shown in Fig. 7.

2.5. Data processing

2.5.1. Pre-processing of data

The sensor mediated collected data was processed and filtered for analysis. Subsequently, the data from 2 h 53 min to 8 h 40 min after adding the EO sample was filtered. Data below 200 lectures, 2 h 53 min, were discarded in order to avoid errors related to the stabilization of the box conditions, such as sensor heating or allowing the evaporation of the samples. On another note, data above 600 lectures, 8 h 40 min, was also excluded to set a limit on data filter. The average of every 10 readings was calculated. It was decided to perform these averages to obtain homogeneous results, as well as to avoid data errors taken by the sensor or physical errors, such as electronic failures.

The data extracted from the EOs was processed and filtered for analysis. Initially, the data was analyzed without any preprocessing, leading to inconclusive results after conducting over 800 classification models and more than 400 regression models. Since the most predominant compounds in *C. ladanifer* oils are Alpha-pinene, an attempt was made to perform a specific classification for these compounds. Additionally, it was decided to perform a classification of all the Terpenic Hydrocarbons as a whole. Alpha-pinene is more volatile due to its lower molecular weight and simpler structure, evaporating easily at room temperature. To evaluate the effectiveness of this method with other compounds, one with different characteristics has been selected. In contrast to alpha-pinene, viridiflorol is less volatile due to its higher molecular weight and more complex structure. Therefore, to facilitate the analysis, 5 different classes were created according to their quantity, ranging from very low to very high (from 1 to 5). The classification for the three chemical compounds on the five output categories has been conducted, ensuring a homogeneous distribution of samples among groups. The thresholds for each group are displayed in Table 2.

2.5.2. Performed analysis

The methodology employed in this study involved the utilization of machine learning-based classification models implemented through the MATLAB (R2022b) programming environment (Fig. 8). The hardware used to conduct the test has an Intel(R) Core(TM) i7-4770 K CPU at 3.50 GHz with 32 GB of RAM. The primary goal of this classification task was to discern differences among organic compounds within the EOs profile of *C. ladanifer*. For ensuring model robustness, 75 % of the entire dataset comprising of 1275 registers, 955 registers were allocated for training and validation, while the remaining 25 % was dedicated to testing the generated models. Thus, we have a training dataset matrix with a dimension of 955x36 and a test dataset with a matrix dimension of 36x32. Specifically, a “Multiclass” classification approach was adopted, incorporating 10-fold cross-validation, to ensure model robustness and prevent overfitting.

A comprehensive evaluation was conducted using a total of 31 models, which were subsequently compared to determine their efficacy. Initially, the utilized models incorporated 36 features as input, derived from a compilation of data, including readings from MQ sensors and temporal information. Subsequently, feature selection was conducted using feature ranking algorithms. The following algorithms have been used, namely Maximum Relevance — Minimum Redundancy (MRMR) and Chi-square (Chi2) algorithms. These algorithms identified the most relevant features, respectively, based on optimal values produced during the selection process. These processes have been conducted independently for each of the predicted compounds (alpha-pinene, viridiflorol, and terpenic hydrocarbons).

A total of 31 classification models were executed, encompassing

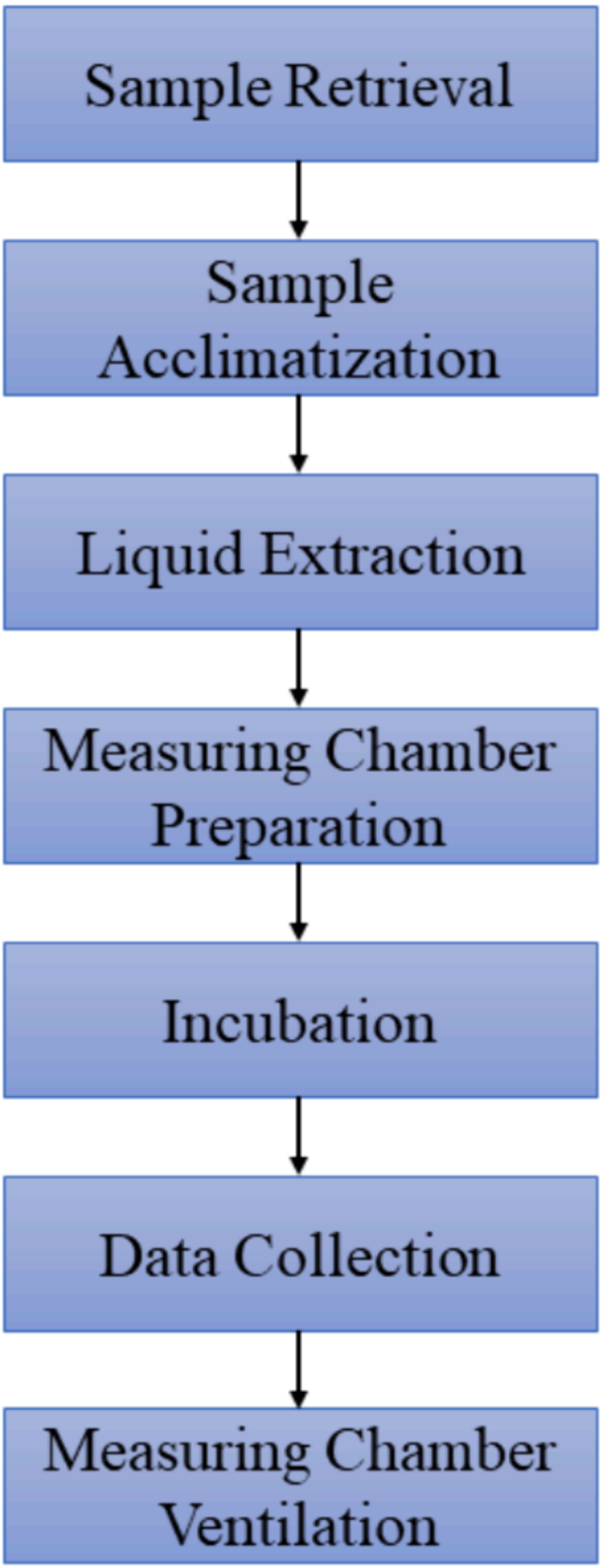

Fig. 7. Stepwise illustration of data gathering procedure.

Table 2
Details of organic compounds classification.

Class	Tag	Alpha-pinene (%)	Viridiflorol (%)	Terpenic Hydrocarbons (%)
1	Very Low	0–30	0 – 2.92	0–55
2	Low	30–45	2.92 – 3.35	55–60
3	Medium	45–50	3.35 – 4.1	60–65
4	High	50–55	4.1 – 5.4	65–70
5	Very High	>55	> 5.4	<70

diverse algorithms such as Linear Regression (LR), Support Vector Machine (SVM), Gaussian Process Regression (GPR), Logistic Regression Kernel, and Neural Networks (NN), among others, as shown in Table 3. Some hyperparameters of the main models are tuned, considering different values for considered neighbours, kernel scales or numbers or layers in the network. All this information linked to the hyperparameters is summarized in Table 3. The comprehensive evaluation of these models involved assessing their performance based on the selected features, providing a nuanced understanding of their effectiveness in distinguishing organic compounds within the EOs profile.

Three metrics are used to evaluate the performance of the generated models: accuracy, sensitivity, and specificity. For the calculation of these metrics, the number of True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) are considered. The formulas of these metrics can be seen in equations (1)–(3).

$$Accuracy = \frac{TP + TN}{Allpredictions} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

3. Results

In this section, the obtained results are described. First of all, the general results of all the tests that were conducted are summarized. Then, the details of the performance of the generated models for Terpenic Hydrocarbons.

3.1. General results

First of all, the general results in terms of maximum achieved accuracies for a summarised number of features are included in Table 4. The model with better accuracy is identified according to the model id provided in Table 3. It can be seen that the accuracies in the training and validation dataset are very high compared with the ones achieved in the test dataset. Moreover, it can be seen that while the highest accuracies are attained in the training and validation dataset when all the features are included, this trend is not observed for the test dataset. Furthermore, we can highlight that the accuracy of the test dataset for the viridiflorol is extremely low compared with the ones for the Alpha-pinene and Terpenic Hydrocarbons. Thus, the results of viridiflorol data will not be included in the subsequent sections. For all the modelled compounds, classification models, including the PCA, which explains the 95 % variance, offered worse results than for other combinations of features. The features (36 to 10) are detailed in Table 5. In this table, the name of features refers to the sensor and the calibration equation employed.

3.2. Performance for the different models of Terpenic Hydrocarbons

To start with, we are going to evaluate the obtained performance with the different conducted classification models, see Fig. 9a–g for the different number of features. The different colours represent the

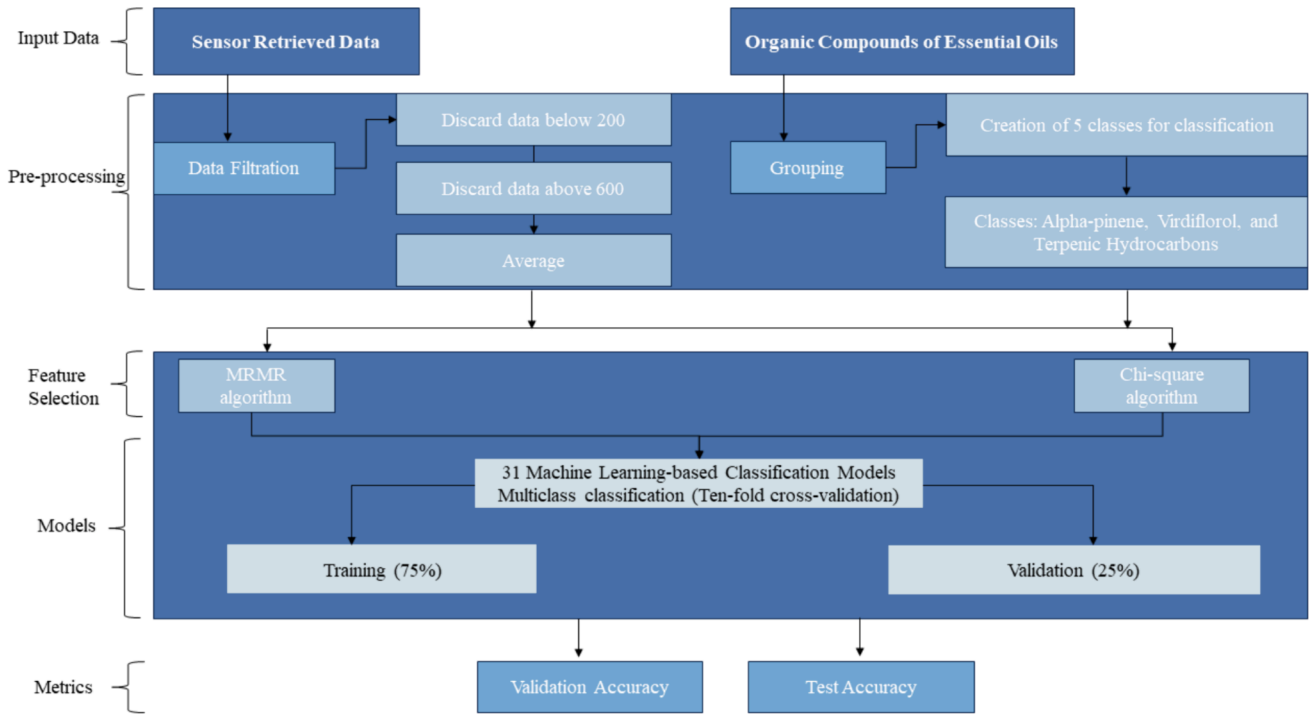

Fig. 8. Schematic illustration of the data processing procedure.

sensitivity for the training and validation dataset in orange and for the test dataset in blue. Models are represented as indicated in Table 4, and the multiple bars indicate the number of included parameters. There are more results in the next subsection about the effect of the different number of parameters.

According to the obtained results, we can identify that there are 2 models which failed to classify the datasets: the Quadratic Discriminant and the Gaussian Naive Bayes. There are a group of models in which the attained accuracies in the training and validation dataset are below 80 % regardless of the number of selected parameters: Coarse Tree, Linear Discriminant, Coarse Gaussian SVM, Subspace Discriminant, and Coarse KNN. Then, there are a group of models whose best accuracies range from 80 to 95 %: Kernel Naive Bayes, Linear SVM, SVM Kernel, and Logistic Regression Kernel. The rest of the models achieved accuracies higher than 95 %. The model that reached the highest sensitivity in the validation dataset is the Weighted KNN, which had an sensitivity of 100 %. Regarding the accuracies of the test dataset, the results are considerably lower than in the validation dataset. There are 6 models which do not surpass the 40 % sensitivity in the test dataset (Fine Tree, Medium Tree, Coarse Tree, Boosted Trees, Bagged Trees, and RUSBoosted Trees). Following, a group of 8 models are characterized by their best accuracies from 40 to 60 %, Fine KNN, Medium KNN, Cosine KNN, Cubic KNN, Weighted KNN, Subspace Discriminant, Subspace KNN, and Trilayered Neural Network. A total of 7 models has their best accuracies ranging from 60 to 70 %: Linear Discriminant, Kernel Naive Bayes, Linear SVM, Coarse Gaussian SVM, Coarse KNN, Narrow Neural Network, and Medium Neural Network. Finally, there are 8 models which have sensitivities above 70 %, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Wide Neural Network, Bilayered Neural Network, SVM Kernel, Logistic Regression Kernel. Among these models, the one with the highest sensitivity, 79.06 %, is Medium Gaussian SVM.

Most of the models have similar sensitivity in the training and validation dataset regardless of the number of included parameters. Nonetheless, it should be noted that some models are strongly affected by the input data. Particularly, on the one hand, the kernel-based models (Kernel Naive Bayes, SVM Kernel, Logistic Regression Kernel), have the best performance in the training dataset when the number of included parameters is low. The highest sensitivity is about 90 % and decreases to approximately 70 %. Nevertheless, for the test dataset, this trend is not followed. On the other hand, the opposite trend is observed in the Linear SVM, Cubic SVM, and Subspace Discriminant models for the training and validation dataset. Considering the sensitivity of the test dataset, there is no clear trend. The effect of the included features is analyzed in the following section.

According to Fig. 10, where the results are divided according to the number of used features, we can see that there are no apparently high differences between the number of parameters. The most significant factor in this case is the type of model, which has been described in the subsection above. There is one significant difference: when the number of parameters is 36, the Linear Discriminant model fails. This does not occur when the number of included features is lower.

Following, we analyze the results for each feature combination. For 10 features, the average in the training and validation dataset is 90.86 %. The sensitivities range from 60.73 % for the Coarse KNN to 99.71 % for the Fine KNN and Weighted KNN. The average sensitivity in this dataset for 11 features is 90.63 %, ranging from 61.88 % for the Coarse KNN to 99.89 % for the Weighted KNN. Regarding the performance when 12 features are included, the average sensitivity is 90.27 % and ranges from 60.54 % for the Coarse KNN to 99.89 % for the Weighted KNN. For 13 features, the sensitivity is characterized by an average value of 90.11 %, with maximum and minimum values of 100 % for Weighted KNN and 59.89 % for Coarse KNN. For 14 features, the average sensitivity in the training and validation datasets stands at 89.90 %. sensitivity levels vary from 59.89 % for the Coarse KNN to 100 % for the Weighted KNN classifiers. With 15 features, the dataset's average sensitivity is 89.89 %, with sensitivity rates spanning from 61.88 % for the Coarse KNN to 99.89 % for the Weighted KNN. Finally, when 36 features are incorporated, the average sensitivity increases to 91.15 %, with sensitivity levels ranging from 65.13 % for the Logistic Regression Kernel to 99.89 % for the Weighted KNN classifier.

Subsequently, we examine the outcomes for various combinations of features. With 10 features, the mean sensitivity for the test dataset reaches 44.55 %. The precision ranges from 12.19 % for the Coarse Tree to 78.44 % for the Medium Gaussian SVM classifiers. For 11 features, the

Table 3
Summary of all the classification models employed.

Id	Preset	Hyperparameters
1	Fine Tree	Maximum n° of splits: 100;
2	Medium Tree	Maximum n° of splits: 20
3	Coarse Tree	Maximum n° of splits: 4
4	Linear Discriminant	
5	Quadratic Discriminant	
6	Gaussian Naive Bayes	Distribution name for numeric predictors: Gaussian
7	Kernel Naive Bayes	Distribution name for numeric predictors: Kernel
8	Linear SVM	Kernel function: Linear; Kernel scale: Automatic
9	Quadratic SVM	Kernel function: Quadratic; Kernel scale: Automatic
10	Cubic SVM	Kernel function: Cubic; Kernel scale: Automatic
11	Fine Gaussian SVM	Kernel function: Gaussian; Kernel scale: 1.5
12	Medium Gaussian SVM	Kernel function: Gaussian; Kernel scale: 6
13	Coarse Gaussian SVM	Kernel function: Gaussian; Kernel scale: 24;
14	Fine KNN	N° of neighbors: 1; Distance metric: Euclidean; Distance weight: Equal;
15	Medium KNN	N° of neighbors: 10; Distance metric: Euclidean; Distance weight: Equal;
16	Coarse KNN	N° of neighbors: 100; Distance metric: Euclidean; Distance weight: Equal;
17	Cosine KNN	N° of neighbors: 10; Distance metric: Cosine; Distance weight: Equal;
18	Cubic KNN	N° of neighbors: 10; Distance metric: Minkowski; Distance weight: Equal;
19	Weighted KNN	N° of neighbors: 10; Distance metric: Euclidean; Distance weight: Squared inverse;
20	Boosted Trees (BT)	Ensemble method: AdaBoost; Learner type: Decision tree; Maximum n° of splits: 20; N° of learners: 30; Learning rate: 0.1; N° of predictors to sample: Select All
21	Bagged Trees	Ensemble method: Bag; Learner type: Decision tree; Maximum n° of splits: 954; N° of learners: 30; N° of predictors to sample: Select All
22	Subspace Discriminant	Ensemble method: Subspace; Learner type: Discriminant; N° of learners: 30; Subspace dimension: 18
23	Subspace KNN	Ensemble method: Subspace; Learner type: Nearest neighbors; N° of learners: 30; Subspace dimension: 18
24	RUSBoosted Trees	Ensemble method: RUSBoost; Learner type: Decision tree; Maximum N° of splits: 20; Number of learners: 30; Learning rate: 0.1; Number of predictors to sample: Select All
25	Narrow NN	N° of fully connected layers: 1; First layer size: 10;
26	Medium NN	N° of fully connected layers: 1; First layer size: 25;
27	Wide NN	N° of fully connected layers: 1; First layer size: 100;
28	Bilayered NN	N° of fully connected layers: 2; First layer size: 10; Second layer size: 10;
29	Trilayered NN	N° of fully connected layers: 3; First layer size: 10; Second layer size: 10; Third layer size: 10;
30	SVM Kernel	Learner: SVM
31	Logistic Regression Kernel	Learner: Logistic Regression

average sensitivity in this dataset is 44.17 %, spanning from 12.19 % for the Coarse Tree to 79.06 % for the Medium Gaussian SVM classifiers. Evaluating performance with 12 features, the average sensitivity records at 45.52 %, varying from 12.19 % for the Coarse Tree to 79.06 % for the Medium Gaussian SVM classifiers. With 13 features, the sensitivity is characterized by an average of 53.31 %, showcasing a high of 79.06 % for Medium Gaussian SVM classifiers and a low of 12.19 % for Coarse Tree. The dataset featuring 14 parameters exhibits an average sensitivity of 51.27 % for the training and validation dataset, fluctuating from 12.19 % for the Coarse Tree to 77.19 % for the Cubic SVM classifier. For 15 features, the dataset's mean sensitivity stands at 51.44 %, with sensitivity rates ranging from 12.19 % for the Coarse Tree to 75.94 % for the Cubic SVM classifier. Finally, upon incorporating 36 features, the average sensitivity decreases to 44.46 %, with precision levels spanning from 12.19 % for the Coarse Tree to 74.06 % for the Medium Gaussian SVM classifiers.

Considering the averaged sensitivities for each number of parameters, the highest sensitivity for the training and validation dataset is attained with 36 features. Nevertheless, the highest sensitivity in the test dataset is attained with the 13 features. After evaluating this data, we can confirm that the best results in terms of sensitivity in the test are attained when the number of features is between 11 and 13. These results are not coincident with the training and validation sensitivities, which achieve the highest sensitivities, equal to 100 %, when the number of parameters is equal to 13 or 14.

Now, the results of the confusion matrices in both datasets are compared. These results for 11features can be seen in Fig. 11 a and 11b. The confusion matrices for 12 features are depicted in Fig. 11 c and Fig. 11 d. Finally, Fig. 11 e and Fig. 11 f present the output of the model when 13 features are included. There are no relevant differences between the obtained results. The results for the test dataset in the models generated with 11 and 12 models are the same, and the differences in the

Table 5
Top selected features for the different VOCs.

Predicted VOC	Alpha-pinene	Viridiflorol	Terpenic Hydrocarbons
Top 15 features	MQ5_H2, MQ5_LPG, MQ2_Alcohol, MQ2: Propane, MQ2_LPG MQ7_H2, MQ7_CO, MQ6_CO, MQ2_CO, MQ5_CO, MQ5_CH4. MQ6_Alcohol, MQ5_Alcohol, MQ2_H2, MQ6_CH4	MQ5_H2, MQ5_CH4, MQ5_CO, MQ5_Alcohol, MQ8_H2, MQ6_H2, MQ6_CH4, MQ6_Alcohol, MQ6_CH4, MQ6_CO, MQ8_LPG, MQ8_Alcohol, MQ7_H2, MQ7_CO, MQ7_Alcohol	MQ8_H2, MQ4_LPG, MQ3_Alcohol, MQ6_Alcohol, MQ3_CO, MQ8_Alcohol, MQ6_CH4, MQ2_LPG, MQ5_Alcohol, MQ6_H2, MQ4_CO, MQ6_LPG, MQ8_CH4, MQ8_CO, MQ4_Smoke

Table 4
Summary of the achieved accuracy in conducted tests.

Chemical compound	No. of features	Model	Accuracy Training-Validation	Test	Sensitivity Training-Validation	Test	Specificity Training-Validation	Test
Alpha-pinene	36 (all)	SVM (12)	99.2	79.0	97.9	47.5	99.5	86.9
Alpha-pinene	10	NB (7)	89.7	84.9	74.1	62.3	93.5	90.6
Alpha-pinene	15	NB (6)	84.5	80.0	61.2	50.0	90.3	87.5
Alpha-pinene	– (PCA)	KNN (14)	91.1	73.9	77.7	34.7	94.4	83.7
Viridiflorol	36 (all)	NB (7)	89.4	74.0	73.6	35.0	93.4	83.8
Viridiflorol	10	NN (26)	99.6	72.3	99.0	30.6	99.7	82.7
Viridiflorol	15	DA (4)	90.2	75.5	75.6	38.8	93.9	84.7
Viridiflorol	– (PCA)	BT (20)	89.5	65.5	73.8	13.8	93.5	78.4
Terpenic Hydrocarbons	36 (all)	SVM (12)	99.4	89.6	98.4	74.1	99.6	93.5
Terpenic Hydrocarbons	10	SVM (11)	98.7	91.4	96.6	78.4	99.2	94.6
Terpenic Hydrocarbons	15	SVM (10)	99.7	90.4	99.3	75.9	99.8	94.0
Terpenic Hydrocarbons	– (PCA)	SVM (9)	88.1	77.6	70.2	44.1	92.5	86.0

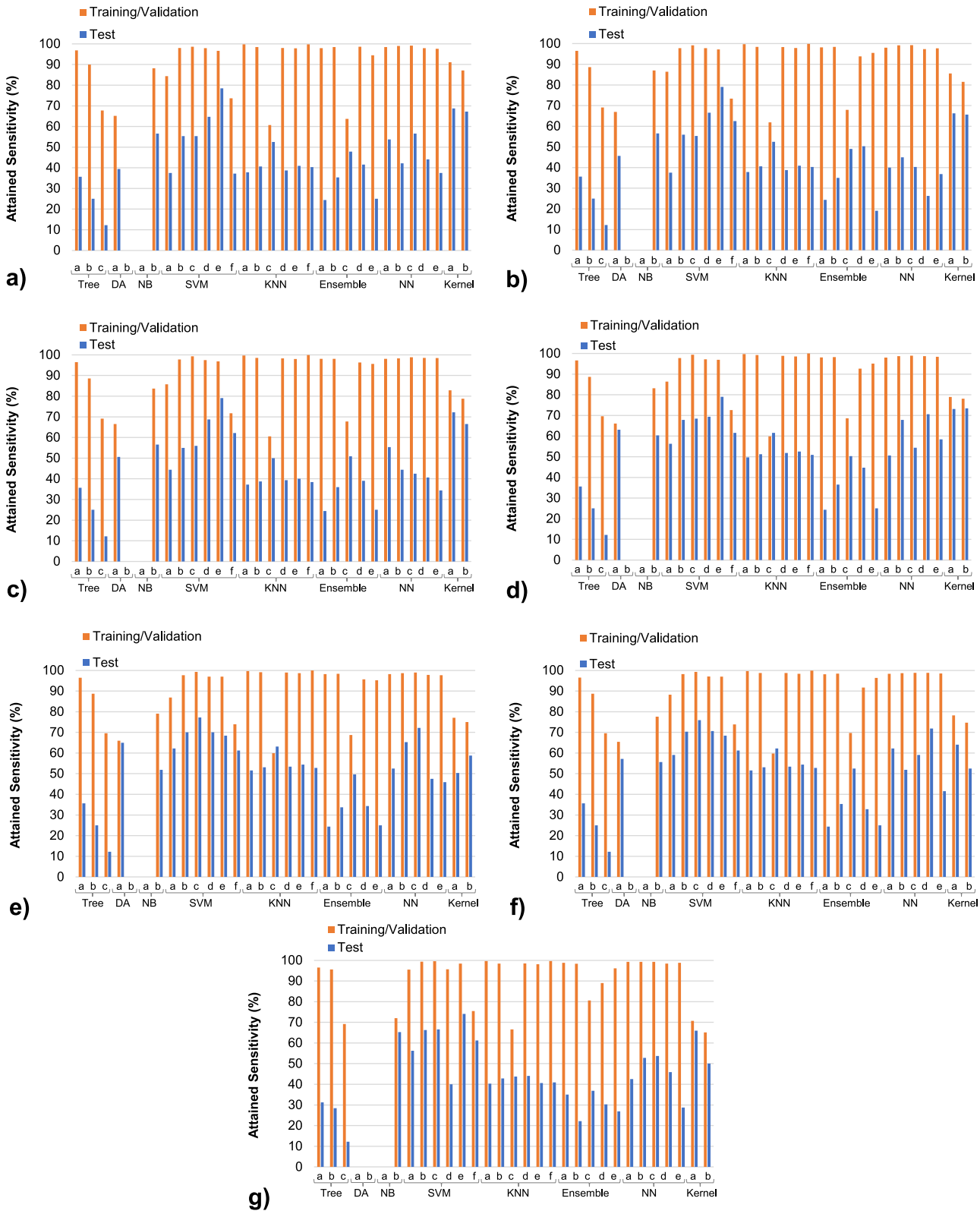

Fig. 9. Attained sensitivities for different tested models for Terpenic Hydrocarbons classification for a) 10 features, b) 11 features, c) 12 features, d) 13 features, e) 14 features, f) 15 features, and g) all features.

training and validation datasets are minimal. The greater changes are linked to errors in the second with 13 features. Considering the low differences in the results achieved and with the aim of optimizing energy consumption, a model with a lower number of features has been selected.

Train: 98.87 96.82 99.27.
Test: 91.63 77.3 94.48.

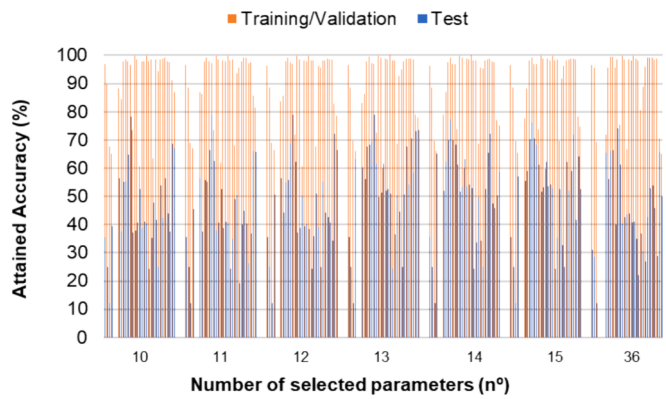

Fig. 10. Attained sensitivities for different number of selected features for Terpenic Hydrocarbons classification.

4. Discussion

Among the different studies in which authors used gas sensors and E-Noses, most of them are applied for a very specific case (detection of ripening of products, growth stages of fruits, determining the origin of products). Although most of these applications have offered high accuracy in the results, and the measures of the gas sensor models have been proven to be related to the presence of chemical compounds [39,6], the quantification of these compounds is not provided.

Concerning the paper, which detects and identifies individual compounds, we highlight the following ones. On the one hand, [41] evaluated the response of a sensing array (composed of 36 fluorescent polymers), to differentiate a total of 30 analytes and a control. The chemical compounds were tested individually with the E-Nose, and the results were promising. Nevertheless, when compounds are tested in pairs, the performance of the system is reduced. The authors did not include details of the results. On the other hand, [20] used MQ-based and TGS-based gas sensors to differentiate chemical compounds. Finally, Liu et al. in 2019, used an array of 9 functionalized reduced graphene oxide to identify individual organic volatile compounds. Nevertheless, even the high accuracy rate in determining the origin of the substances, none of this paper has quantified with regression or classification models the concentration of the given compounds, as it is intended in this paper.

The rest of the paper focused on using gas sensors to classify different products according to their features. We are going to detail the results obtained from papers that analyzed EOs. Some studies have been conducted on *Rosa damascena* [14,15]to classify the EOs according to the most common chemical compound s [14]and according to their genotype [15]. In 2016, data of 7 sensors (TGS822, TGS842, SP15A, SP32, SP53, TGS2610, TGS2620) have been used with Fuzzy ARTMAP and LDA to differentiate among three different classes. The achieved sensitivities range from 82 to 99 % when 10 features are included and 62 to 95 when all features are added. Nonetheless, the authors do not provide information about the quantity of selected chemical compounds for the classification nor give clear data about the range in each clase. Moreover, with gathered data only training/validation has been performed. No information about the test is included in the paper, which precludes the fair assessment of the truthfulness of their system in real conditions. Meanwhile, in 2017, EOs were measured using 7 sensors (MQ3, MQ5, MQ135, MQ138, TGS822, TGS832) to identify the genotypes of 3 classes of the *R. damascena*. The authors applied LDA and SVM as ML algorithms for multiclass classification. Achieved accuracies reached 95 to 99 % in the validation dataset. Nevertheless, as in the previous case, the authors do not inform about the chemical compounds contained in the EOs of each genotype. Moreover, no test has been conducted precluding a comparison with the proposed method.

In 2020, there are two examples of using similar systems for EOs obtained from trees. On the one hand, Graboski et al. 2020 quantified the chemical compounds in clove EO. For that purpose, the authors created their own gas sensors based on carbon nanocomposite, creating 6 different sensors. The generated data were used to train and validate the model using PCA and IDMAP. In this paper the authors achieved R^2 of 0.99, which supposes a very good performance. Nevertheless, this performance was obtained with sensors characterized by a high cost and which are not found commercially. On the other hand, Aghoutane et al. [5] used ML and gas sensors to differentiate between EOs from two species: *Aucoumea klaineana* and *Canarium schweinfurthii*. The authors used 6 low-cost sensors (TGS815, TGS821, TGS822, TGS224, TGS825, TGS42), and 3 ML algorithms (PCA, DFA, HCA). However, the authors do not provide information about the achieved accuracies nor the detection quantification of specific chemical components of the EOs.

From 2021, several examples of using EOs from herbs and fruits are found. The first examples consist of different sources of EOs (*Thymus vulgaris*, *Artemisia dracunculus*, *Mentha arvensis*, *Citrus sinensis*, *Citrus limon*, and *Mangifera indica*) by Rasekh et al. (2021). A total of 9 gas sensors (MQ9, MQ4, MQ135, MQ8, TGS2620, MQ135, TGS813, TGS822, MQ3) were used. The obtained data was included as an input for 4 ML algorithms: LDA, QDA, sigmoid SVM, or radial SVM. The accuracy in the validation reached 96.7 to 100 %. Even though the achieved accuracies are high, no tests were conducted. In addition, the authors do not identify or quantify chemical compounds. Another example of application in herbs, *Mentha piperita*, is found in 2022 by Zorpeykar et al. [44]. The authors used 8 sensors (MQ3, MQ9, MQ135, MQ136, TGS2620, TGS2602, TGS813, TGS822) to differentiate the obtained EO from distilled water extracts. For the classification 3 ML algorithms were used (PCA, LDA, ANN), averaged achieved accuracies from validation and test ranged from 85.6 to 100 %. Even though tests were conducted, averaged data of validation and test precludes the correct comparison. Furthermore, in this paper, no quantification of chemical compounds was conducted. In the same year, Liu et al. (2022), proposed the use of gas sensors for recognition and classification of aroma oils. The authors used up to 30 different aromas and measured them with 24 printable chemiresistive odor sensors created for this application. The obtained data were analyzed with N-way k-shot *meta*-learning. The achieved accuracy was 98.70. Though the accurate results, no information about the test dataset is given. In 2023, [43] used the gas sensors to classify the quality of commercial EOs from Lavander. The authors have used 4 MOS sensors (MQ3, MQ4, MQ6, MQ135) and 3 ML algorithms (KNN, NB, DA), attaining accuracy in the validation of 65 %. It must be noted that in this paper, no quantification of the chemical compounds of the tested EOs was carried out. In the same year, Sudarmaji et al. [32] used gas sensors to detect adulteration of patchouli oil with candlenut oil. The authors selected 9 gas sensors for this purpose, which include TGS-2602, TGS2620, TGS2600, MQ5, MQ135, MQ138, FIS12A, FIS30SB, FISAQ1. The sensed data was analyzed with PCA. According to their results, the best results were achieved with 6 of the 9 sensors. Nevertheless, the authors did not provide accuracy in the paper. As in previous cases, no quantification of chemical compounds was conducted.

Finally, we find some examples in which the EO of *C. ladanifer* is measured with gas sensors in 2023. In the first paper, Viciano-Tudela et al. (2023) used 7 MOS gas sensors: MQ2, MQ3, MQ4, MQ5, MQ6, MQ7, MQ8 to identify *C. ladanifer* EO adulterated EO with EO *Pinus pinaster*. The authors used ANN to identify the added EOs with achieved accuracy of 99.97 %. The same sensors were used by Viciano-Tudela et al. (2023) to recognize different types of EOs from herbs and fruits. In this case, the used ML models were higher, including ANN, NB, KNN, DA, and SVM, achieving accuracies from 88.74 to 99.82 %. Nonetheless, in those papers, the authors have not used gas sensors to quantify the content of any chemical compound.

A summary of the aforementioned cases of gas sensors used in EOs is summarised in Table 6. As can be seen, the accuracy achieved for the training and validation dataset is similar to the ones reported by other

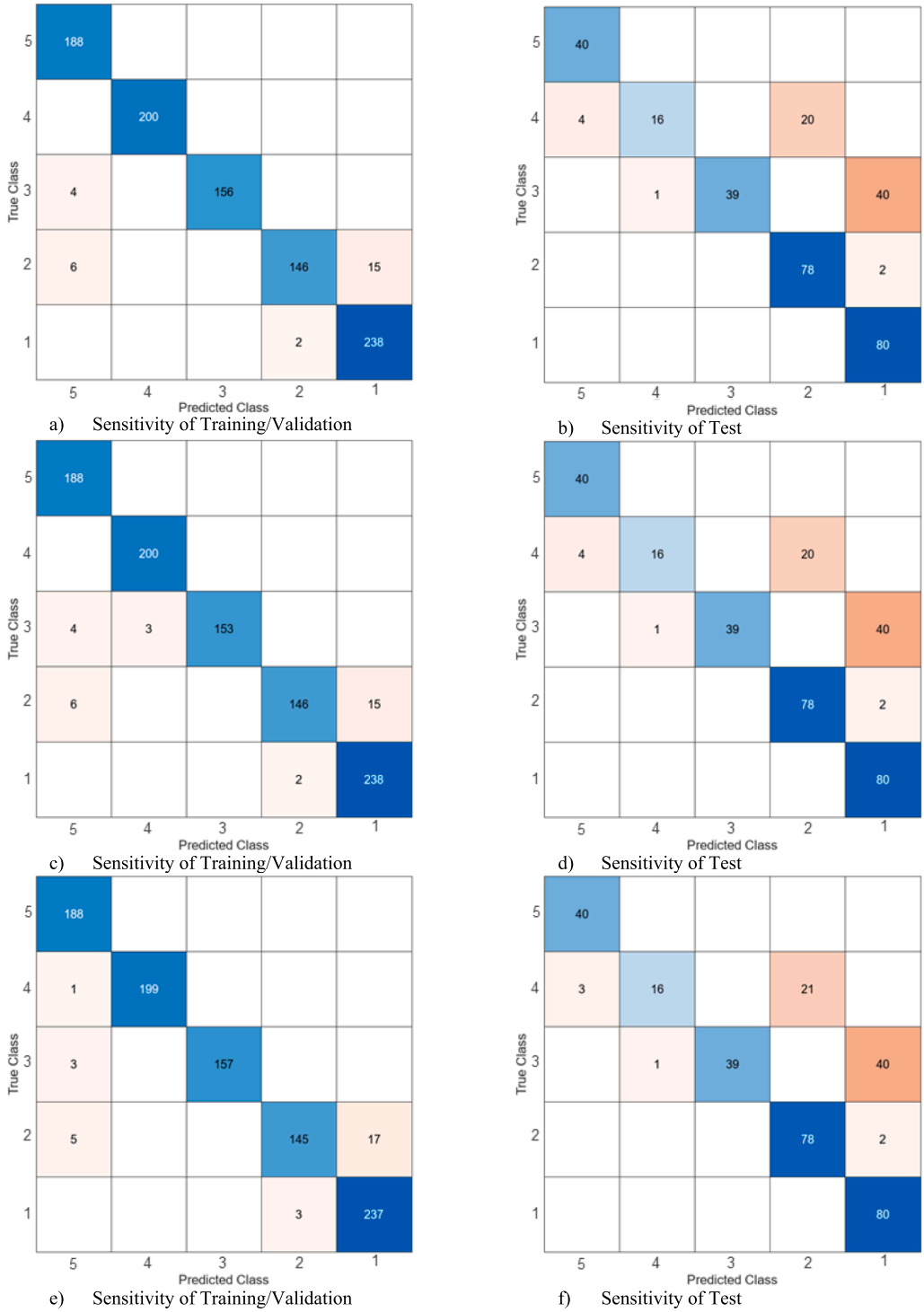

Fig. 11. Confusion matrix for different models in both training/validation and test with Medium Gaussian SVM model. a) training/validation for 11 features, b) test for 11 features, c) training/validation for 12 features, d) test for 12 features, e) raining/validation for 13 features, f) test for 13 features.

authors. Even though the accuracy in the test dataset is lower than the one indicated by [44], it must be noted that test accuracy is not provided in multiple papers.

5. Conclusions

This paper proposes an approach combining low-cost sensors and machine-learning techniques to address this issue. Results reveal the inability to establish a valid classification model for Viridiflorol or Alpha-pinene, while for Terpenic hydrocarbons, an accuracy of 91.6 % is attained. Notably, these accuracies are obtained with a reduced number of features (11 to 13). This study marks the first successful application of MQ-based gas sensors, or similar metal oxide sensors, in accurately determining the concentration of a chemical compound.

In future work, models will be created to classify other chemical compounds or groups of chemical compounds (alcohols, ketones, etc.) which are relevant to the quality of EOs. With the aim of improving accuracy, convolutional neural networks and procedures with variable

Table 6
Summary of existing uses and accuracies in the use of metal oxide sensors for classifying EOs.

Year	EOs of:	N° of sensors	Used sensors	N° of ML algorithm	ML algorithm	Accuracy in validation	Accuracy in test
2016	*Rosa damascena*	7	TGS822, TGS842, SP15A, SP32, SP53, TGS2610, TGS2620	2	Fuzzy ARTMAP and LDA	82–99	–
2017	*Rosa damascena*	7	MQ3, MQ5, MQ135, MQ138, TGS822, TGS832	2	LDA, SVM	95 % to 99 %	
2020	*Clove*	6	Tailored	1	PCA and IDMAP	–	–
2020	*Aucoumea klaineana and Canarium schweinfurthii*	6	TGS815, TGS821, TGS822, TGS224, TGS825, TGS42	3	PCA, DFA, HCA	–	
2021	*Thymus vulgaris, Artemisia dracunculus, Mentha arvensis, Citrus sinensis, Citrus limon, and Mangifera indica*	9	MQ9, MQ4, MQ135, MQ8, TGS2620, MQ135, TGS813, TGS822, MQ3	4	LDA, QDA, sigmoid SVM, or radial SVM	96.7 to 100 %	–
2022	*M. piperita*	8	MQ3, MQ9, MQ135, MQ136, TGS2620, TGS2602, TGS813, TGS822,	3	PCA, LDA, ANN,	85.6 to 100 % as averaged training, validation and test	
2022	30 different sources	24	Tailored	1	N-way k-shot *meta*-learning	98.70 %	
2023	*Commercial Lavander EO*	4	MQ3, MQ4, MQ6, MQ135	3	KNN, NB, DA	65 %	–
2023	*Patchouli oil*	9	TGS-2602, TGS2620, TGS2600, MQ5, MQ135, MQ138, FIS12A, FIS30SB, FISAQ1	1	PCA	–	–
2023	*Cistus ladanifer*	7	MQ2, MQ3, MQ4, MQ5, MQ6, MQ7, MQ8	1	ANN	99.97 %	
2023	*Cistus ladanifer, Pinus pinaster, Cistus ladanifer oil adulterated Pinus pinaster, Melaleuca alternifolia, tea tree, and red*	7	MQ2, MQ3, MQ4, MQ5, MQ6, MQ7, MQ8	5	ANN, NB, KNN, DA, SVM	88.74 to 99.82 %	
2024	*Cistus ladanifer*	7	MQ2, MQ3, MQ4, MQ5, MQ6, MQ7, MQ8	31	See Table 3	98.87 %	91.63 %

selection will be applied to the generated datasets in order to evaluate their ability to improve the current models. Moreover, additional inputs will be considered, such as absorption spectra of ultraviolet, visible and infrared light.

CRediT authorship contribution statement

Francisco Javier Diaz Blasco: Writing – original draft, Investigation, Conceptualization. **Sandra Viciano-Tudela:** Writing – original draft, Investigation, Formal analysis, Data curation. **Lorena Parra:** Writing – original draft, Investigation, Formal analysis, Data curation. **Ali Ahmad:** Writing – original draft, Investigation, Conceptualization. **Veronika Chaloupková:** Writing – original draft, Investigation, Data curation. **Raquel Bados:** Writing – original draft, Investigation, Data curation. **Luis Saul Esteban Pascual:** Writing – review & editing, Supervision, Investigation, Funding acquisition. **Irene Mediavilla:** Writing – review & editing, Supervision, Investigation, Funding acquisition. **Sandra Sendra:** Writing – original draft, Investigation, Formal analysis, Data curation. **Jaime Lloret:** Writing – review & editing, Supervision, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jaime Lloret reports financial support was provided by Spain Ministry of Science and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work has been funded by the “Ministerio de Ciencia e Innovación” through the Project PID2020-114467RR-C33/AEI/10.13039/501100011033, and by the "Ministerio de Economía y Competitividad" through the Project TED2021-131040B-C31. This study also forms part of the ThinkInAzul programme and was supported by MCIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat Valenciana (THINKINAZUL/2021/002). Llamada iniciada 13:02.

References

[1] S. Acharyya, S. Nag, P.K. Guha, Ultra-selective tin oxide-based chemiresistive gas sensor employing signal transform and machine learning techniques, Anal. Chim. Acta 1217 (2022) 339996.
[2] S. Acharyya, P.K. Guha, Hierarchical zinc stannate nanoneedle-based sensitive detection of formaldehyde, ACS Applied Electronic Materials 5 (6) (2023) 3446–3453.
[3] S. Acharyya, A. Ghosh, S. Nag, S.B. Majumder, P.K. Guha, Smart and selective gas sensor system empowered with machine learning over IoT platform, IEEE Internet Things J. (2023).
[4] S. Acharyya, P.K. Guha, Enhanced formaldehyde sensing performance employing plasma-treated hierarchical SnO2 nanosheets through oxygen vacancy modulation, Appl. Surf. Sci. 655 (2024) 159640.
[5] Y. Aghoutane, M. Moufid, S. Motia, G.S. Padzys, L.P. Omouendze, E. Llobet, N. El Bari, Characterization and analysis of Okoume and aiele essential oils from Gabon by gc-ms, electronic nose, and their antibacterial activity assessment, Sensors 20 (23) (2020) 6750.
[6] M.M. Ali, N. Hashim, S. Abd Aziz, O. Lasekan, Principles and recent advances in electronic nose for quality inspection of agricultural and food products, Trends Food Sci. Technol. 99 (2020) 1–10.
[7] AliExpress, 2024. MQ Gas Sensors. Available at: https://shorturl.at/fFIM4. Last Accessed on February 21, 2024.
[8] A. Amkor, N. El Barbri, Artificial intelligence methods for classification and prediction of potatoes harvested from fertilized soil based on a sensor array response, Sens. Actuators, A 349 (2023) 114106.
[9] Y.R. Carrillo-Amado, M.A. Califa-Urquiza, J.A. Ramón-Valencia, Calibration and standardization of air quality measurements using MQ sensors, Respuestas 25 (2020) 70–77.
[10] W. Dhifi, S. Bellili, S. Jazi, N. Bahloul, W. Mnif, Essential oils' chemical characterization and investigation of some biological activities: A critical review, Medicines 3 (2016) 25.
[11] eBay, 2024. MQ Gas Detection Sensor. Available at: https://shorturl.at/cdkCP. Last Accessed on February 21, 2024.
[12] El Hachlafi, N., Kandsi, F., Elbouzidi, A., Lafdil, F. Z., Nouioura, G., Abdallah, E. M., Abdnim, R., Bnouham, M., Al-Mijalli, S. H., Mrabti, H. N., & Fikri-Benbrahim, K. (2024). Extraction of bioactive compound-rich essential oil from Cistus ladanifer L. by microwave-assisted hydrodistillation: GC-MS characterization, in vitro pharmacological activities, and molecular docking. Separations, 11(7), 199.
[13] P.B. Gomes, V.G. Mata, A. Rodrigues, Characterization of the Portuguese-grown Cistus ladanifer essential oil, J. Essent. Oil Res. 17 (2005) 160–165.

[14] A. Gorji-Chakespari, A.M. Nikbakht, F. Sefidkon, M. Ghasemi-Varnamkhasti, J. Brezmes, E. Llobet, Performance comparison of fuzzy ARTMAP and LDA in qualitative classification of iranian rosa damascena essential oils by an electronic nose, Sensors 16 (5) (2016) 636.
[15] A. Gorji-Chakespari, A.M. Nikbakht, F. Sefidkon, M. Ghasemi-Varnamkhasti, E. L. Valero, Classification of essential oil composition in Rosa damascena Mill. genotypes using an electronic nose, J. Appl. Res. Med. Aromat. Plants 4 (2017) 27–34.
[16] J. Huang, J. Wu, Robust and rapid detection of mixed volatile organic compounds in flow through air by a low cost electronic nose, Chemosensors 8 (2020) 73.
[17] H. Karami, M. Rasekh, E. Mirzaee-Ghaleh, Application of the E-nose machine system to detect adulterations in mixed edible oils using chemometrics methods, J. Food Process. Preserv. 44 (2020) e14696.
[18] M.M. Khan, Sensor-based gas leakage detector system, Engineering Proceedings 2 (2020) 28.
[19] M. Khatib, H. Haick, Sensors for volatile organic compounds, ACS Nano 16 (2022) 7080–7115.
[20] Klouček, P., Malina, M., Havlík, J., Maršík, P., & Maršán, L. (2016). Identification and quantification of plant volatiles by low-cost electronic nose. The 2nd IMEKOFOODS, 110-114.
[21] Z. Liu, M. Wang, M. Wu, X. Li, H. Liu, N. Niu, S. Li, L. Chen, Volatile organic compounds (VOCs) from plants: From release to detection, TrAC Trends Anal. Chem. 158 (2023) 116872.
[22] P.J. Marriott, R. Shellie, C. Cornwell, Gas chromatographic technologies for the analysis of essential oils, J. Chromatogr. A 936 (2001) 1–22.
[23] Mediavilla, I., Blázquez, M.A., Ruiz, A., Esteban, L.S., 2021. Influence of the storage of cistus ladanifer l. Bales from mechanised harvesting on the essential oil yield and qualitative composition. Molecules 26, 2379.
[24] T.C. Miller, S.D. Morgera, S.E. Saddow, A. Takshi, M. Palm, Electronic nose with detection method for alcohol, acetone, and carbon monoxide in coronavirus disease 2019 breath simulation model, IEEE Sens. J. 21 (2021) 15935–15943.
[25] M.J. Oates, N. Abu-Khalaf, C. Molina-Cabrera, A. Ruiz-Canales, J. Ramos, B. W. Bahder, Detection of lethal bronzing disease in cabbage palms (Sabal palmetto) using a low-cost electronic nose, Biosensors 10 (2020) 188.
[26] J.L. Oller-López, R. Rodríguez, J.M. Cuerva, J.E. Oltra, B. Bazdi, A. Dahdouh, A. Lamarti, A.I. Mansour, Composition of the essential oils of Cistus ladaniferus and C. monspeliensis from Morocco, J. Essent. Oil Res. 17 (2005) 553–555.
[27] M.C. Peel, B.L. Finlayson, T.A. McMahon, Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci. 11 (2007) 1633–1644.
[28] A. Popa, M. Hnatiuc, M. Paun, O. Geman, D.J. Hemanth, D. Dorcea, L.H. Son, S. Ghita, An intelligent IoT-based food quality monitoring approach using low-cost sensors, Symmetry 11 (2019) 374.
[29] Rahmad, I.F., Nababan, E.B., Tanti, L., Triandi, B., Ekadiansyah, E., Fragastia, V.A., 2019. Application of the alcohol sensor MQ-303A to detect alcohol levels on car driver, 2019 7th International Conference on Cyber and IT Service Management (CITSM). IEEE, pp. 1-5.
[30] J.R. Raimundo, D.F. Frazão, J.L. Domingues, C. Quintela-Sabarís, T.P. Dentinho, O. Anjos, M. Alves, F. Delgado, Neglected Mediterranean plant species are valuable resources: The example of Cistus ladanifer, Planta 248 (2018) 1351–1364.
[31] M. Roy, N. Hariharan, D. Manoj, I. Auddy, S. Shanmugasundaram, Development of metal oxide semiconductor gas sensor based electronic nose system for adulteration detection in ghee, Pharm. Innov. J 10 (2021) 30–38.
[32] A. Sudarmaji, A. Margiwiyatno, S.B. Sulistyo, Characteristics of array MOS gas sensors in detection of adulteration on patchouli oil with candlenut oil. AIP Conference Proceedings, AIP Publishing, 2023.
[33] C.S. Tavares, A. Martins, M.L. Faleiro, M.G. Miguel, L.C. Duarte, J.A. Gameiro, L. B. Roseiro, A.C. Figueiredo, Bioproducts from forest biomass: Essential oils and hydrolates from wastes of Cupressus lusitanica Mill. and Cistus ladanifer L, Ind. Crop. Prod. 144 (2020) 112034.
[34] F. Valderrama, F. Ruiz, An optimal control approach to steam distillation of essential oils from aromatic plants, Comput. Chem. Eng. 117 (2018) 25–31.
[35] S. Viciano-Tudela, L. Parra, P. Navarro-Garcia, S. Sendra, J. Lloret, Proposal of a New System for Essential Oil Classification Based on Low-Cost Gas Sensor and Machine Learning Techniques, Sensors 23 (2023) 5812.
[36] S. Viciano-Tudela, S. Sendra, L. Parra, J.M. Jimenez, J. Lloret, Proposal of a Gas Sensor-Based Device for Detecting Adulteration in Essential Oil of Cistus ladanifer, Sustainability 15 (2023) 3357.
[37] H.C. Voon, R. Bhat, G. Rusul, Flower extracts and their essential oils as potential antimicrobial agents for food uses and pharmaceutical applications, Compr. Rev. Food Sci. Food Saf. 11 (2012) 34–55.
[38] J. Wawrzyniak, Advancements in improving selectivity of metal oxide semiconductor gas sensors opening new perspectives for their application in food industry, Sensors 23 (2023) 9548.
[39] J. Xu, K. Liu, C. Zhang, Electronic nose for volatile organic compounds analysis in rice aging, Trends Food Sci. Technol. 109 (2021) 83–93.
[40] E. Yavuzer, Determination of fish quality parameters with low cost electronic nose, Food Biosci. 41 (2021) 100948.
[41] P. Zhao, Y. Wu, C. Feng, L. Wang, Y. Ding, A. Hu, Conjugated polymer nanoparticles based fluorescent electronic nose for the identification of volatile compounds, Anal. Chem. 90 (7) (2018) 4815–4822.
[42] H. Zidane, M. Elmiz, F. Aouinti, A. Tahani, J. Wathelet, M. Sindic, A. Elbachiri, Chemical composition and antioxidant activity of essential oil, various organic extracts of Cistus ladanifer and Cistus libanotis growing in Eastern Morocco, Afr. J. Biotechnol. 12 (34) (2013).
[43] Z. Zlatev, A. Draganova, Development and research of a measuring device for gas and optical features of essential oils, AIP Conference Proceedings Vol. 2889, No. 1 (2023).
[44] S. Zorpeykar, E. Mirzaee-Ghaleh, H. Karami, Z. Ramedani, A.D. Wilson, Electronic Nose Analysis and Statistical Methods for Investigating Volatile Organic Compounds and Yield of Mint Essential Oils Obtained by Hydrodistillation, Chemosensors 10 (11) (2022) 486.