Single-point calibration process based integrated electrical impedance analyzer for multi-selective gas detection

Louis Routier, Alexandre Westrelin, Anthyme Cerveaux, Gaël Louis, Thomas Horlach, Pierre Foulon, Kamal Lmimouni, Sébastien Pecqueur, Bilel Hafsi

Discover Applied Sciences

Research

Single-point calibration process based integrated electrical impedance analyzer for multi-selective gas detection

Louis Routier[1] · Alexandre Westrelin[1] · Anthyme Cerveaux[2] · Gaël Louis[2] · Thomas Horlach[2] · Pierre Foulon[2] · Kamal Lmimouni[1] · Sébastien Pecqueur[1] · Bilel Hafsi[1,2]

Received: 23 April 2024 / Accepted: 17 July 2024
Published online: 26 July 2024

Abstract
Impedance analysis is a powerful technique that has become increasingly important in various applications, it represents a leap forward in the field of electronic measurments and diagnostics. In this work, we present the development of miniaturized, multiplexed, and connected platform for impedance spectroscopy. Designed for online measurements and adapted to wireless network architectures, our platform has been tested and optimized to be used for multi-selective chemical organic sensor nodes. This compact and versatile circuit is built from low cost and low power consumption (250 mW) microelectronics components that achieve long duration operability (5 days and 16 h) without compromising on sensor measurement accuracy and precision. We used the well-known impedance network analyzer AD5933 (Analog Devices, Norwood, MA, USA) chip which can measure a spectrum of impedances in the range 5 kHz to 100 kHz. The proposed system is based on ESP32-C3 Microcontroller enabling the management of the AD5933 through its I^2C interface. Our system benefits from two multiplexer components CD74HC4067 allowing calibration process and the interface of 15 conductimetric sensors with real time acquisition (less than 90 ms per acquisition). The system is capable of relaying information through the network for data analysis and storage. The paper describes the microelectronics design, the impedance response over time, the measurement's sensitivity and accuracy and the testing of the platform with embedded chemical sensors for gas classification and recognition.

Article Highlights

- Single calibration process: The main contribution of this paper is to propose a portable multiplexed impedance measurement system with a unique self-calibration capability. This capability is facilitated through the use of a single calibration resistor, which significantly enhances the speed of data acquisition.
- Materials study: To validate our system, a materials study was conducted, focusing on the use of a conductive polymer poly(3hexylthiophene) P3HT doped with various triflate metals. This research aims to assess the system's effective-

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42452-024-06102-x.

✉ Bilel Hafsi, bilel.hafsi@icam.fr; Louis Routier, louis.routier@iemn.fr; Alexandre Westrelin, alexandre.westrelin@iemn.fr; Anthyme Cerveaux, anthyme.cerveaux@icam.fr; Gaël Louis, gael.louis@icam.fr; Thomas Horlach, thomas.horlach@icam.fr; Pierre Foulon, pierre.foulon@icam.fr; Kamal Lmimouni, kamal.lmimouni@iemn.fr; Sébastien Pecqueur, sebastien.pecqueur@iemn.fr | [1]CNRS, Centrale Lille, University Polytechnique Hauts-de-France, UMR 8520 - IEMN, University of Lille, 59000 Lille, France. [2]ICAM School of Engineering, 06 Rue Auber, 59800 Lille, France.

ness in gas recognition applications, demonstrating its potential for accuracy and speed crucial attributes for modern detection technologies.

- Gas recognition: Our developed system has been designed to be suitable to assess volatile samples by their quality, to systematically expose them to the miniaturized array under a static blow (the PCA of the partial dataset of resistance modulation published in our former studies quantifies on the systematicity of the exposures by the organization of the data clusters).These labels are indeed crucial to quantify the materials response with environmental physical variables which may greatly vary within an environment.

Keywords Polymer sensors · Organic semiconductor · Impedance measurements · Environmental monitoring · E-nose

1 Introduction

Recent years have witnessed an increasing interest in smart multi-sensing technologies for various emerging applications: from machine diagnostics in smart industries and connected agriculture to health-monitoring and numerous other Internet-of-Things (IoT) appliances. As the technological landscape is dynamically evolving, continuous development in materials and technologies is required, enabling sensors to become more sensitive, accurate, and versatile [1]. Therefore, considerable attention has been directed toward enhancing the ability to process, analyze, and respond to data, often through embedded electronics and analysis techniques, enabling the screening reproducibility of aerial classes of pollution, threats, diseases, or pleasance.

The identification of volatile organic compounds (VOCs) requires the measurement of different species of gases and particles which have proven to be of great constitutive complexity. Thus, the design of chemical detection platforms with a high number of specific sensors is crucial to identify them as pollutant classes. Chemical, electrochemical and / or biochemical electronic noses (e-noses) [2] are becoming increasingly popular as rapid, specific, and cost-effective detection devices. Despite their remarkable development over the past few years in various fields such as environmental monitoring [3, 4], the food industry [5], health [6], public and industrial safety [7], the potential of integrating them in microsystems as part of the IoTs remains underused. These devices consist of a multiplexed sensitive array combined to analog front-end and a microcontroller unit that enables detection, identification, and quantification of VOCs released from samples. In such array, sensors with different specificities, sensitivities, and stabilities, must be well chosen and co-integrated to promote the identification of a wide range of chemical features within an environmental pattern, allowing the recognition of distinct and specific odor using electronic hardware.

Electrochemical Impedance Spectroscopy (EIS) is a generic technique to measure sensitive materials as voltage-linear electrical elements in real-time. EIS studies are usually performed under laboratory condition, with voluminous and expensive measuring bench system-based impedance spectrometers [8, 9]. If such an approach is well suited for metrology of specific environments, low data throughput within such framework hinders the understanding of collective contributions of multi-material specificity for complex environments pattern recognition. One of the major challenges is the ability to screen multiple sensitive materials co-integrated in an array, which itself is co-integrated as microsensors on a readout platform. This platform allows online measurements and pattern recognition with enough computational resources and the smallest size features for IoT integration into portable systems. This research enables the merging of both the fundamental studies of the chemical interactions between target species and selective materials, and the practical application of developing usable field detection systems for chemical elements, or biomarkers [10]. Our study aims to create a smart and self-sufficient tool capable of providing initial solutions for detecting and recognizing molecular issues where current solutions are lacking, such as in beverage quality control. Besides producing and evaluating sensors, we will also address signal processing, and data collection, storage, transfer, and retrieval using embedded systems based on microcontrollers that can integrate a secure communication solution. Commonly used in various applications such as in bio-impedance analysis, sensing, and electrochemistry, the AD5933 is a highly integrated and low-cost solution for portable applications. Developed by Analog Devices [11], this integrated circuit can perform frequency-domain impedance measurements and can operate in the frequency range of 10 Hz to 100 kHz. AD5933-based systems have been used in a variety of applications, including biomedical [12, 13], agriculture [14], and textile-enabled [15] applications, as well as for the characterization such as the study of metal corrosion [16], and the human skin bioimpedance measurements [17].

Regarding these applications, the AD5933 has been implemented in various works [12, 14], obtaining portable systems suitable for different detection applications across a range of domains, such as renal disease detection and fruit quality monitoring, etc. An exhaustive technical data review of some of the most representative above-mentioned portable impedance meters is presented in Table 1.

In this paper, we demonstrate a low-cost, portable, and miniaturized multi-sensing platform for non-invasive impedance measurement of electrochemical sensors. The device, being compact, provides safe and accurate readings, ideal for monitoring bioimpedance sensors. The paper outlines a circuit that runs under an ESP 32-C3 microcontroller, making it suitable for remote monitoring environment air quality or gas detection. Linear discriminant analysis (LDA) and principal component analysis (PCA) based feature reduction algorithms have been used to analyze data obtained from chemical multi-sensor measurements.

2 Electrochemical multi-sensor system design

2.1 Impedance analyser chip AD5933

The AD5933 is a highly accurate impedance converter system that communicates with microcontrollers via an I^2C. This Integrated Circuit features an internal conversion system and a DDS frequency generator, which allows an external complex impedance to be excited with at specified frequencies. The impedance response is then amplified, filtered, and sampled by a 12-bit, 1 MSPS ADC. Calibration of the system is essential and involves precise control of several factors: the excitation voltage (ranging from 198 $mV_{p\text{-}p}$ to 1.98 $V_{p\text{-}p}$), the amplification setting of internal Programmable Gain Amplifier (PGA) (with selectable gains of 1 × or 5 ×), and the selection of a calibration resistance approximately matching the unknown impedance. Additionally, the circuit also includes a 1024-point DFT Processor that executes a Discrete Fourier Transform, providing real and imaginary impedance values at each frequency step of the sweep.

The on-board ADC sampling process and the on-board DSP engine perform a Discrete Fourier Transform (DFT) algorithm. The real (the resistance R) and imaginary (the reactance X) components of the unknown impedance are then stored in two 16-bit registers. The impedance of the device being tested is determined through a calibration process that compares the amplitude with a known resistance (see Fig. 1). The impedance magnitude and phase are then easily calculated using the following equations:

$$Magnitude = \sqrt{R^2 + X^2} \tag{1}$$

$$Phase = \tan^{-1}(X/R) \tag{2}$$

The system requires two external components: the device under test, represented as an impedance $Z(\omega) = R + jX(\omega)$, and a reference resistor named R_{FB} placed at the input of amplifier stage. This resistor, used as a feedback resistor, must have a known value that is incorporated into the control code of the AD5933. The AD5933 requires a calibration resistor using a known impedance to determine the gain factor, which sets the system's gain. This is calculated using the following equation:

$$Gain\ Factor = \frac{\left(\frac{1}{R_{CAL}}\right)}{\sqrt{R^2 + X^2}} \tag{3}$$

where R_{CAL} is the calibration resistor value while R is the real component, and X is the imaginary component of the measured impedance.

There are some limitations, mainly in the calibration process, that must be considered when designing new systems, particularly if one needs to perform precise and repeatable measurements. Indeed, the AD5933 system must be calibrated for a known impedance range to determine the gain factor before any valid measurements can take place. The gain factor is determined by placing a known impedance between the input (Pin5) and the output (Pin6) of the AD5933 and measuring the resulting magnitude. However, if the unknown impedance falls outside the calibration range over a wide frequency range, this may lead to incorrect or noisy data. To overcome this, multiple methods have been suggested

Table 1 List of portable impedance meters based on the AD5933 chip

Refs.	Frequency range	Excitation voltage	Maximum error	Impedance range	Data communication	Size
[14]	10 Hz–100 kHz	200 mV, 400 mV, 1 V, 2 V	Magnitude 1% Phase 4.3%	1 Ω–280 kΩ	STM32L486 (USB—BT)	75×40 mm
[18]	1 Hz–100 kHz	200 mV, 400 mV, 1 V, 2 V	3.5%	10 Ω–1 MΩ	Atmel ATmega32	N/A
[20]	4–100 kHz	20–200 mV	2%	100 Ω–10 kΩ	N/A	N/A
[17]	5 Hz–100 kHz	25 mV, 50 mV, 200 mV, 400 mV, 1 V, 2 V, 3 V	6%	180 Ω–165 kΩ	Arduino nano (USB—BT) HC-06 BT	N/A
[21]	0.01 Hz–100 kHz	1 mV–1 V	3%	10 Ω–10 GΩ	Zigbee (USB)	65×120 mm
[22]	0.01 Hz–100 kHz	10 mV–2 V	5%	100 Ω–10 GΩ	Arduino Uno (USB)	74×53 mm
[12] E-NOSE*	1–30 kHz	N/A	0.5%	N/A	Raspberry Pi (USB)	N/A
[23]	10 Hz–100 kHz	200 mV, 400 mV, 1 V, 2 V	5.6%	1 kΩ–10MΩ	EVAL-AD5933EBZ board	100 mm × 90 mm
[24]	1 Hz–100 kHz	189 mV–1.98 V	Magnitude 0.27%	100 Ω–1 kΩ	STM32f407VGT06	N/A
[25]	100 Hz–200 kHz	N/A	4.5%	10 Ω–1 kΩ	MSP430F2274 with ZigBee transceiver (RF/ USB)	35×37×8 mm
Our platform	5–100 kHz	200 mV, 400 mV, 1 V, 2 V	4.8%	1 kΩ–1 MΩ	ESP32-C3 (USB—BT—WIFI)	66×35 mm

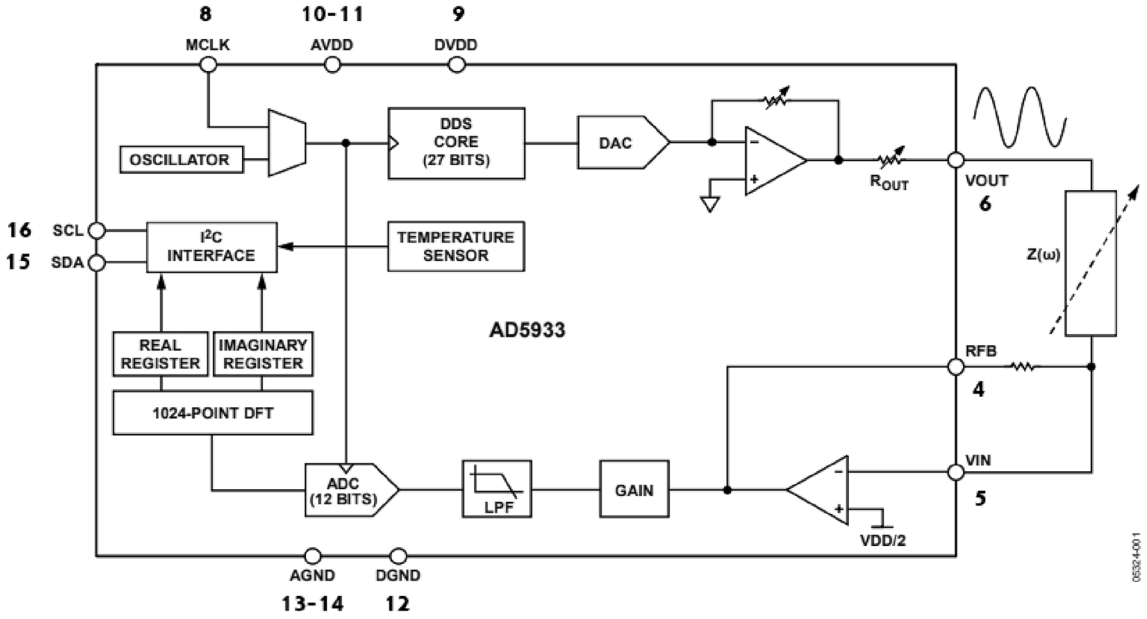


Fig. 1 Block overview of the AD5933 Impedance analyzer [11]

in literature, including using different resistors to cover measurement impedance range and reducing the frequency spectrum limits [14]. Furthermore, the AD5933 has a fixed user-selectable feedback resistance connected between Pin 4 (R_{FB}) and Pin 5 (V_{IN}). It is important for the user to choose a feedback resistance value that, in conjunction with the selected gain of the PGA stage, maintains the signal within the linear range of the ADC.

2.2 Single-point calibration procedure

The AD5933 is a low-cost solution for impedance measurement, but it also has several limitations that obstruct its suitability for EIS measurements. Previous research [18] has employed an Analog Front End (AFE) to enhance the AD5933's capability for Bio Impedance measurements. When calculating the gain factor, it is important that the receiver stage operates within its linear region. This requires careful selection of the system gain settings (Eq. 4). To calibrate the system accurately, the user needs to know the impedance limits. One must then select a resistor that matches the I–V gain setting of the calibration impedance. The gain through the system is defined as:

$$\text{Gain} = V_{OUT} \times \frac{R_{FB}}{Z_{unknown}} \times \text{PGA Gain} \tag{4}$$

Differently from what is suggested in the datasheet [11] and literature [14], the originality of our work, consists of implementing a complete single-point calibration in which the complete frequency range 5–80 kHz is calibrated and data is saved on specific output files. Our calibration method considers the non-idealities of the system and the parasitic contribution of the internal resistance of the analog multiplexers used to interface 15 sensors. The key principle behind this approach is that the application is supported by machine learning rather than being purely metrological: the accuracy range of the measured physical loads does not directly impact the quality of the data clustering. Furthermore, the multiplexing approach required for e-nose application offers a more flexible hardware architecture, which could benefit the calibration process, described in more detail in Sect. 3.

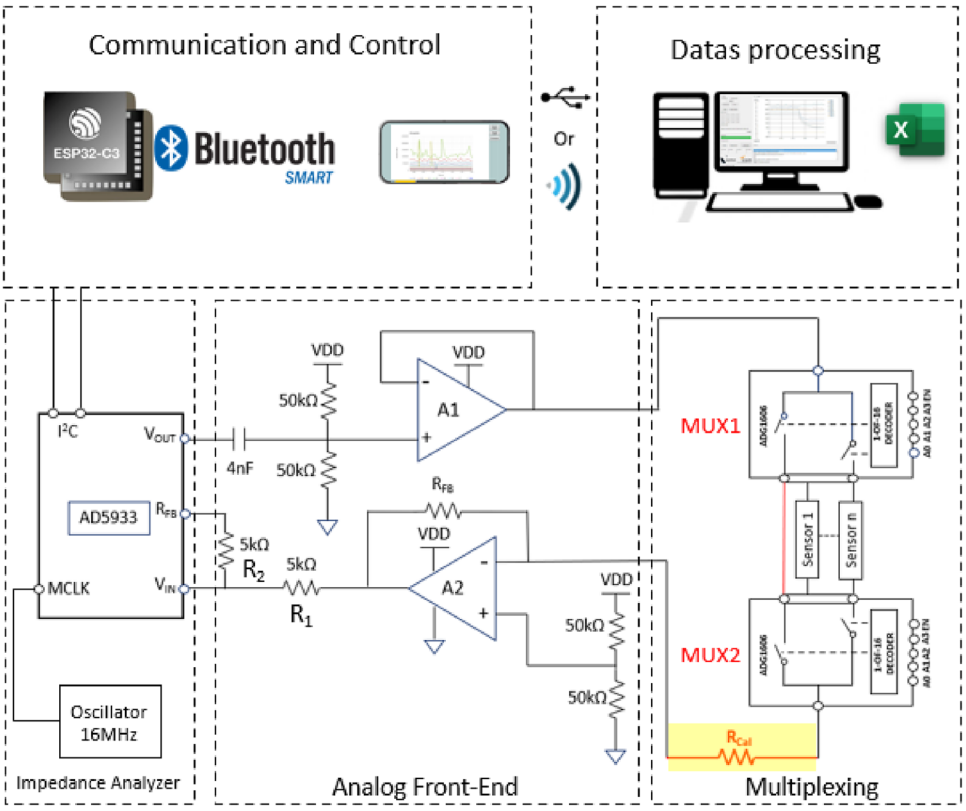


Fig. 2 Block overview with details of electronic components [19]

3 Measurements and validation

3.1 Multiplexed impedance system

The AD5933 allows the user to perform a frequency sweep with a user-defined start frequency, a frequency resolution, and a number of points in the sweep. In addition, the device enables the user to set a peak-to-peak output signal value of 198 mV, 383 mV, 0.76 V, and 1.98 V with different DC offsets and a known frequency, this signal acts as an excitation voltage for external unknown impedances under test. We choose to set the output voltage at the minimum voltage (V_{PP} = 198 mV) proposed by analog device to avoid electrochemically damaging our conducting polymer sensitive materials in chemical environments. The clock for the DDS is generated from either an external reference clock which is provided by the user at MCLK or by the internal oscillator. In our case, we use a 16 MHz external clock which provides a more stable and accurate frequency source than the integrated internal clock. This stability can affect the quality of the frequency sweep and consequently lead to more accurate and consistent impedance measurements (more details about the printed circuit board in the supporting information). After calibration, the measurement process is performed, the unknown impedance values are calculated by means of this equation:

$$Z = \frac{1}{\sqrt{r^2 + x^2} \times \text{Gain Factor}} \tag{5}$$

where r and x are, respectively, the real and the imaginary values obtained from the internal register during the measurement process.

In order to calibrate the system, one must measure the magnitude produced by a two-terminal cell with a known resistor values. Since we are interested in the relative variation of a reactance-less impedance, neither capacitance nor inductance are involved in the sensing of our transducer model. The proposed solution consists of using a single point calibration based on R_{CAL} = 20 kΩ resistor placed at the output of the second multiplexer, the measurement cover a [1 kΩ, 1 MΩ] impedance range. The current-to-voltage resistor R_{FB} was set at the same resistance values 20 kΩ as shown in Fig. 2.

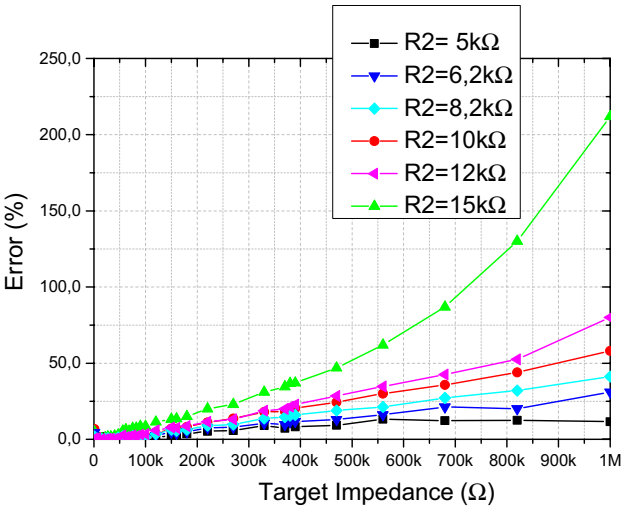


Fig. 3 Percentage error curve of the measured impedance values versus target impedance values

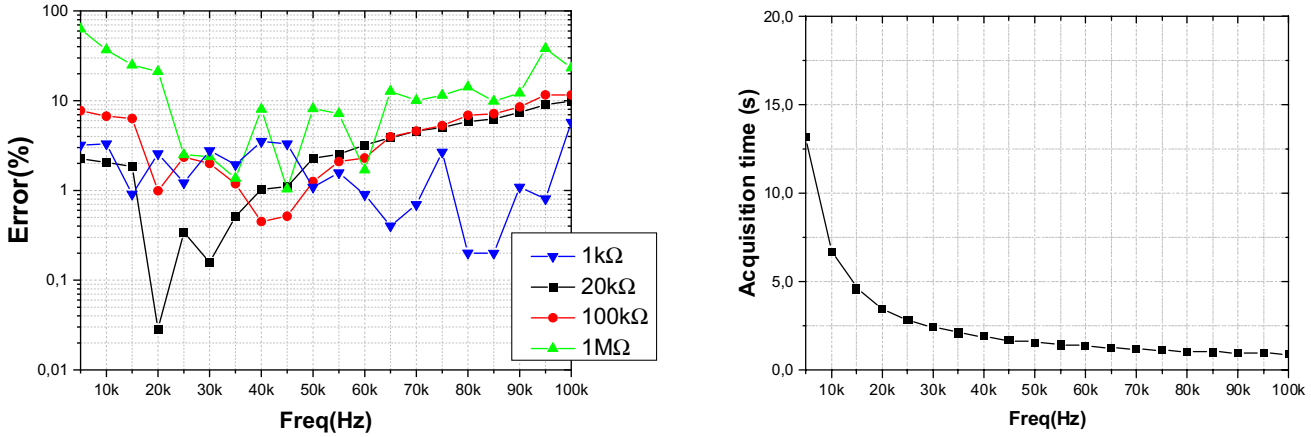


Fig. 4 **a** Absolute impedance error at frequencies between 5 and 100 kHz for different known resistor mean values (1 kΩ, 20 kΩ, 100 kΩ, 1 MΩ) **b** Acquisition time response as function as the frequency variation for 15 multiplexed sensors,

The receive stage consists of a current-to-voltage amplifier, a programmable gain amplifier (PGA). The PGA allows the user to increase the output of the current-to-voltage amplifier by either a factor of 5 or 1, using the control register (located in the Register Map section at Register Address 0×80). The filtered output is then passed through a 12-bit, 1 MSPS ADC, this stage's schematic is depicted in Fig. 1. The unknown impedance is connected between the V_{OUT} and V_{IN} pins. The current-to-voltage amplifier results in a voltage present at the V_{IN} pin acting as a virtual ground with a DC value set at $V_{DD}/2$. The current generated across the unknown impedance flows into the V_{IN} pin and produces a voltage signal at the output of the current-to-voltage converter. As mentioned before, the current-to-voltage amplifier's gain is determined by a selectable R_2 feedback resistor that is connected between Pin 4 (R_{FB}) and Pin 5 (V_{IN}).

It is crucial for the user to choose a feedback resistor value that, along with the PGA stage's selected gain, keeps the signal within the ADC's linear range. Figure 3 shows measurements for feedback resistance values: 5, 6, 8, 10, 12 and 15 kΩ and for target impedance ranging from 1 kΩ to 1 MΩ were conducted. An increase was observed in the percentage error with the successive increments of the feedback resistor value. The response of the circuit was consistent and showed a maximum error value that does not exceed 10% for $R_1 = 5$ kΩ, indicating linear behavior of the gain current to voltage amplifier. Furthermore, the AD5933's frequency response is limited, causing a variation in the gain factor, and resulting in errors in impedance calculations across a range of frequencies. To reduce these errors, it is better to limit the frequency sweep to a narrow range. We choose to study the response of the system to frequency variation. EIS characteristic for a frequency range from 5 to 100 kHz with 5 kHz step was measured using known resistor values (1 kΩ, 20 kΩ, 100 kΩ, 1 MΩ), (Fig. 4a). By analyzing these results, it can be noticed that for extreme frequency values (very low frequencies and very high ones), large errors in the shape of the measured resistance characteristics occurs, the error can be reduced by

adjusting excitation frequency in the range of [30–80 kHz] (Fig. 4b), but still need more adjustment to get better accuracy in the measurments. Moreover, given the requirement of conducting real-time measurement, our system must allow fast acquisition time of 15 chemical conductimetric sensors. The theoretical sampling time is the contribution of three characteristic times: t_{Cal} which is the time for the AD5933 to calculate the impedance values (generally, it is estimated to be 1 ms for a 16 MHz system clock [11]), the time allowed for serial I^2C protocol instruction, such as reading the estimated impedance, value registers, sending the repeat instruction time, and checking the impedance measurement status. And finally, the Settling Time Cycles t_{Cycles}, which is determined by multiplying the total number of settling cycles and the period of the impedance excitation signal. In order to increase accuracy and reduce the time between each acquisition, we choose to fix the frequency at three different ranges as illustrated in the flowchart in Fig. 5a.

3.2 Software corrections

To go further in accuracy assessment, our circuit was then tested with different loads from the 1 kΩ resistor 1 MΩ. To perform online calibration and update it if needed, we used a 1:16 multiplexer that switches between the known value of a calibration resistor to 15 unknown impedances to measure. Results for the measured loads show the difference between the measured values and the real ones for loads higher than 400 kΩ (Fig. 5b), which represents an increasing error with the increase in target resistance values. The originality of our study lies in its unique calibration approach, which is based on a fixed R_{CAL}. To enable ease of calibration and to improve measurement accuracy, we have included correction equations that can be selected at firmware level. This functionality enables a large degree of freedom in the impedance exploration as the user can choose to select the operating frequency, the flowchart of Fig. 5a shows a simplified pseudocode for both calibration and measurement loop. A calibration flag set to "False" is employed to monitor

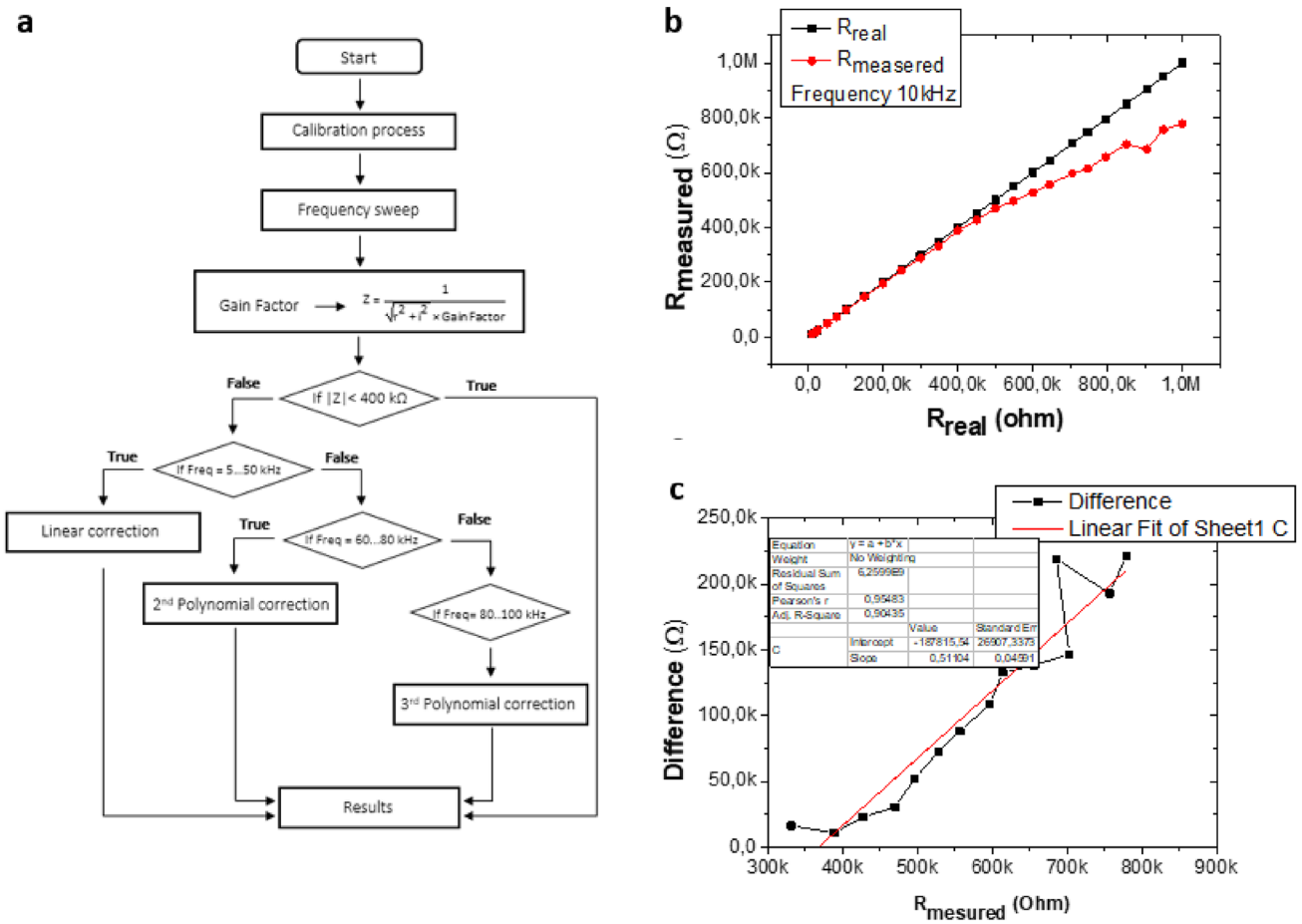


Fig. 5 **a** Flowchart for data acquisition, **b** Results for the measured resistors compared to with different loads from the 1 kΩ resistor 1 MΩ, for a 20 kHz frequency excitation voltage **c** Correction equation extracted from the linear fit of the real and measured impedance difference curve

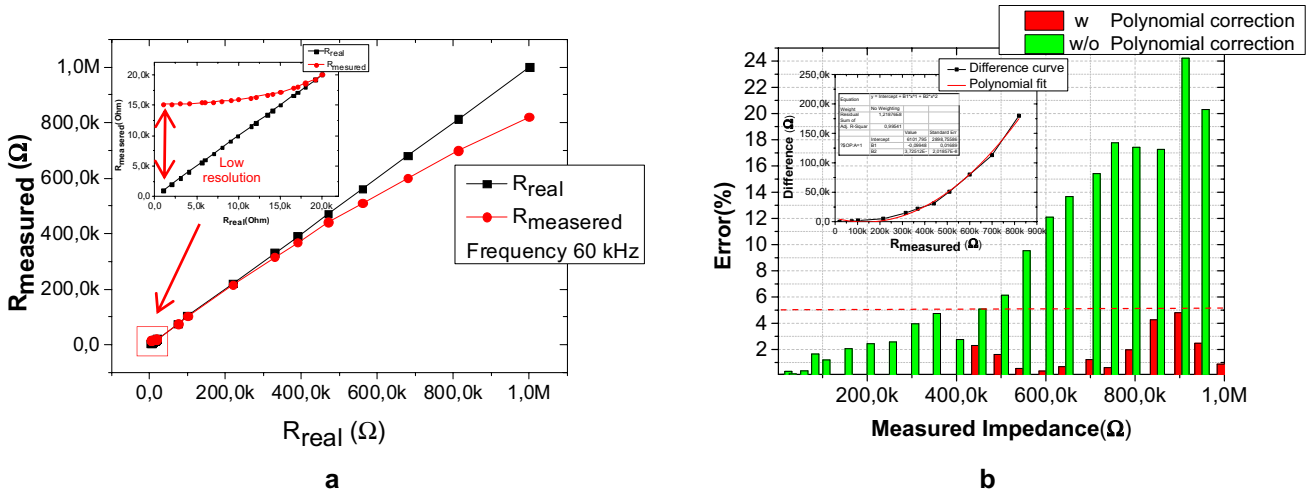


Fig. 6 **a** Results for the measured resistors compared to with different loads from the 1 kΩ resistor 1 MΩ, for a 60 kHz frequency excitation voltage, **b** correction equation extracted from a second-order polynomial fitting function of the real and measured impedance difference curve

whether the calibration has been successfully carried out. Once the calibration has been executed at least once, this flag is set to "True". When calibration is performed, the second step consists of checking whether gain and phase factors are available prior to running measurements, otherwise leading to the returning of an error message. The AD5933 performs then a Discrete Fourier Transform (DFT) on the acquired measurements and sets precise constraints in terms of used frequency. Unlike conventionally calibration process that involve multiplexing of different R_{CAL}, our methodology is tailored specifically for biological sensors. These sensors respond to gases by altering their impedance values across a wide range, spanning up three decades, making the conventional method incompatible. All the obtained measurements were processed several times with a rinsing, falling and random manner. The correction process was applied independently from the measurement process. To enable greater flexibility, in our implementation each frequency comprises a corresponding correction equation extracted by the user to avoid incurring in saturation of the circuit, as the impedance accuracy may be significantly affected in the complete 5–100 kHz frequency range. The correction process was applied using the extracted equation, which represents the difference between the measured resistance curve and the target one, as shown in Fig. 5c. The next decision point evaluates the impedance magnitude |*Z*|. If |Z| is less than 400 kΩ, the process moves to unknown load calculation. In the case where impedance values are greater than 400 kΩ, a correction procedure is required. The decision criteria are based on the target frequency defined by the user. If the operating frequency is fixed in the [5–50 kHz] range values, curve deviation is compensated using a correction equation extracted from the linear fit of the real and measured impedance difference curve. Figure 5b shows an example of the extracted correction equation for 10 kHz excitation frequency.

In the case of excitation frequency ranging between 60 and 80 kHz, the user must implement a second order polynomial correction, otherwise, another a third order polynomial correction equation should be implemented, (for more details refer to supplementary information).

We conducted the same study by testing the impedance response with various loads ranging from a 1 kΩ to 1 MΩ at 60 kHz (target frequency for e-nose application). The results showed a gap between the measured values and real ones for loads above 400 kΩ (as seen in Fig. 6a), resulting in an error rate exceeding 20% for the highest impedance case. We extract the equation that manages the difference between the measured resistance and the target ones, the curve presented in the inset of Fig. 6b exhibits second-order polynomial behavior. To decrease the error observed above 400 kΩ, we implemented a correction equation in the C + + code [26].

The application of a second correction equation was also considered to increase the accuracy response in the 1–400 kΩ load range. The issue is that for low resistance values (typically below 20 kΩ), we have a very small resolution window (Fig. 6a inset). To enable ease of calibration, we conducted the calibration process with only one known resistance (R_{CAL} = 20 kΩ). Subsequently, the measurement of all the sensor values is carried out. Conventional method involves connecting the calibration resistor between the two multiplexers [14], our strategy involves shifting the calibration resistor position at the of the second multiplexer. Hence, the R_{CAL} value is subtracted from the measured unknown impedance

value in the software part. By Implementing the equivalent circuit model fitting, the resulted percentage error of our system does not exceed 5% (Fig. 6b). We should note that these results were obtained by averaging ten measurements cycles related to passive loads that cover [1 kΩ–1 MΩ] range values.

The performances of our system were compared to with similar state-of-the-art portable impedance analyzers, (see Table 1). An analysis of the available information on instruments reported in the literature shows that our system offers comparable excitation voltage and impedance range values. Notably, our system can be upgraded to cover a frequency range of 5–100 kHz (see Supplementary Information); users can implement a correction equation tailored to the specific type of application. It is important to highlight that our analysis aims to underscore the strategic position and capabilities of our system within this broader context, while recognizing that each technology may be optimized to meet different needs. Furthermore, our system demonstrates similar accuracy compared to some results in literature. In fact, the maximum impedance error resulted to be 4.8%, which is comparable to those is reported in [22, 25]. On the other side, system reported in [14] shows an average Root Mean Square Error (RMSE) of 1.0% and 4.3% for the passive components, for impedance magnitude and phase, respectively. These results cover [1–280 kΩ] impedance range. We should note that in this work a different approach was used in the calibration process, which consist of using multiple calibration resistances. Recently, Hu et al. [24] proposed an auto-calibrated measurement system for 1D matrices of impedimetric sensors, in this work the observed maximum impedance deviation value is 0.27%, this study was evaluated with 10 kHz working frequency for within the 100-1 kΩ range.

The major originality of our work lies in the capability to cover an impedance range of [1 kΩ–1 MΩ] with only one calibration resistance, which is not possible for other impedance systems presented in the Table 1. While error appears to be already optimized to ensure good performances, our system can be easily used for multivariate detection applications, such as electronic nose (e-nose) technologies.

4 Chemical sensing application

An electronic nose (or e-nose) is an electronic interface that reads out chemically sensitive transducers to recognize the environment with on-board multivariate data analysis techniques to Classify information features. Among them, some consist of an array of sensitive gas sensors, usually result in a complex preparation process, high-power consumption, and redundant sensing signals. This section outlines the study of the electrical responses of various conductimetric cells. Each cell is tuned using a drop-casted conducting polymer, namely poly(3-hexylthiophene) P3HT, and different triflate doping salts. Based on the obtained data, the recognition of various gas species has been conducted using Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA)-based feature reduction algorithms. Sensors consist of 15 interdigital electrodes, each 1 mm wide and 3 mm long, with a separated gap of 0.15 mm between them. We should note that the current consumption of our system was evaluated under different operating conditions, acquired using a DC power probe combined with Rohde & Schwarz RTB2004 oscilloscope. While connected through Bluetooth, the system consumes 35 mA with established connection, and 50 mA while calibration and measuring process. Our system presents an onboard 3.7 V-3400 mAh two Li–Po battery, providing an autonomy of ~ 136 h (5 days and 16 h) (see supporting information). A 3D view of the electrodes is shown in Fig. 7.

4.1 Clustering using principal component analysis (PCA) and linear discriminant analysis (LDA)

PCA is a widely used technique that reduces the complexity of data, offering better visualization. It operates by projecting the high-dimensional data matrix onto a lower-dimensional space defined by principal components. The resulting score and loading matrices allow clustering of classes without the need of labeling them [27]. The number of dimensions in these matrices is smaller than the original data matrix. The columns of the score matrix, known as Principal Components (PCs), are orthogonal to each other and are organized to preserve most of the data variance post-transformation. The scores are then displayed as scatters in two- or three-dimensional spaces, where different (PC) are plotted against one another. The first principal component take into account the largest amount of variation in the data, while each subsequent component explains as much of the remaining variation as possible. Objects that were similar in the original space are depicted close to each other in the score plots. The primary benefit of this approach is its ability to transform multidimensional data into easily interpretable two- or three-dimensional graphs, all while retaining most of the original data variance. While PCA provides a powerful framework for reducing data complexity without regard to class labels, the next stage of our analysis involves Linear Discriminant Analysis (LDA), a supervised technique that incorporates class

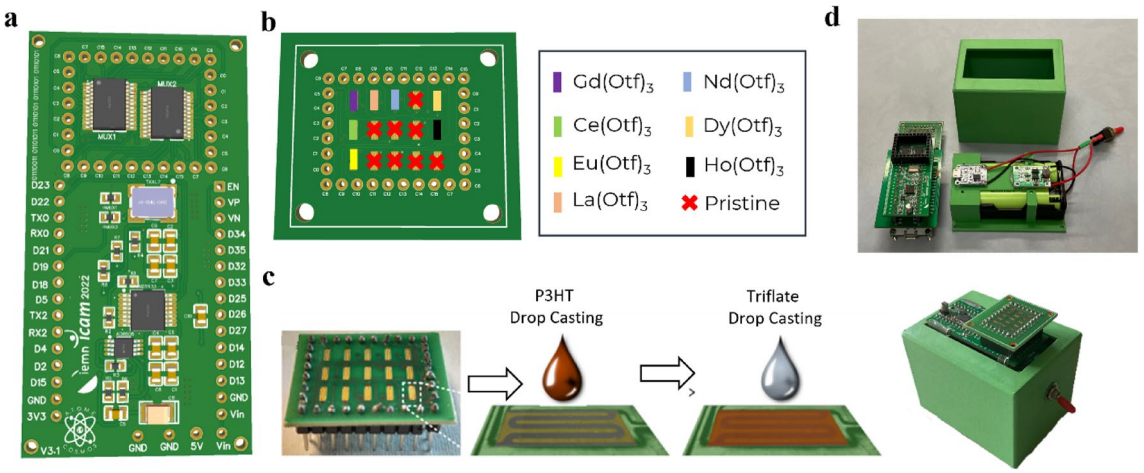
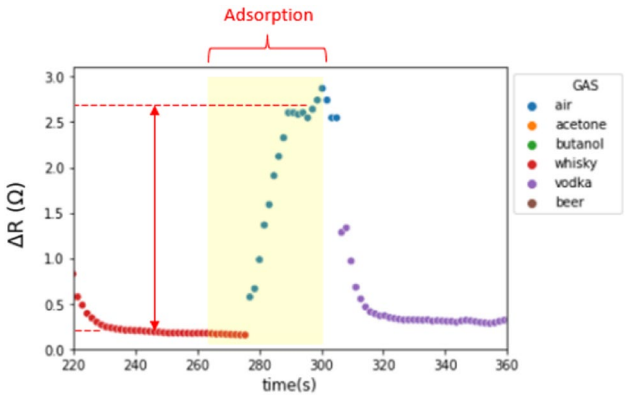


Fig. 7 **a** 3D view of the printed circuit board of our electrochemical impedance analyzer, **b** 3D view of our proposed interdigital sensors **c** drop-casting process of the sensitive materials: the conducting polymer (electrotransducer), and the doping salt (chemical sensitizer). The volatile organic compound (VOC) response analysis was carried out using two approaches: unsupervised and supervised classification. Among these methods: Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA), **d** e-Nose prototype packaging, system is based on two 3400 mAh Li-ion batteries

Fig. 8 Gas sensor response with feature extracted



labels to enhance the discrimination power of the feature reduction process [28]. The features are selected through LDA, which performs by maximizing the variance between species and minimizing the variance within species, thereby resulting in improved separability in the dataset. Unlike PCA, which seeks directions with the greatest variation in the dataset and ranks components based on the variation they explain, LDA specifically ranks features based on the variation they explain among species. Most of feature reduction algorithms aims to reduce the size of data while retaining sufficient information to ensure discrimination between the classes. These algorithms are implemented using a post-processing approach. In our case, we have developed all these algorithms based on a Python code to meet future real-time processing requirements.

4.2 Data preprocessing

The study was based on collected data of eight different sensors based on seven different materials and a reference cell. The processing of Sensors responses processing is performed using a post-processing algorithm using Python code. The first step involves filtering the curves using a moving average equation to reduce noise and remove any incoherent measures. Subsequently, the resistance is calculated from the real and imaginary values provided by our system. For each sensor, features are extracted during the adsorption phases, as illustrated in Fig. 8.

Typically, the analysis has been performed on a dataset of dynamical relative resistance modulation ΔR/R extracted from the curve. Normalization was carried out by dividing the resistance of the sensing materials by their baseline resistance values.

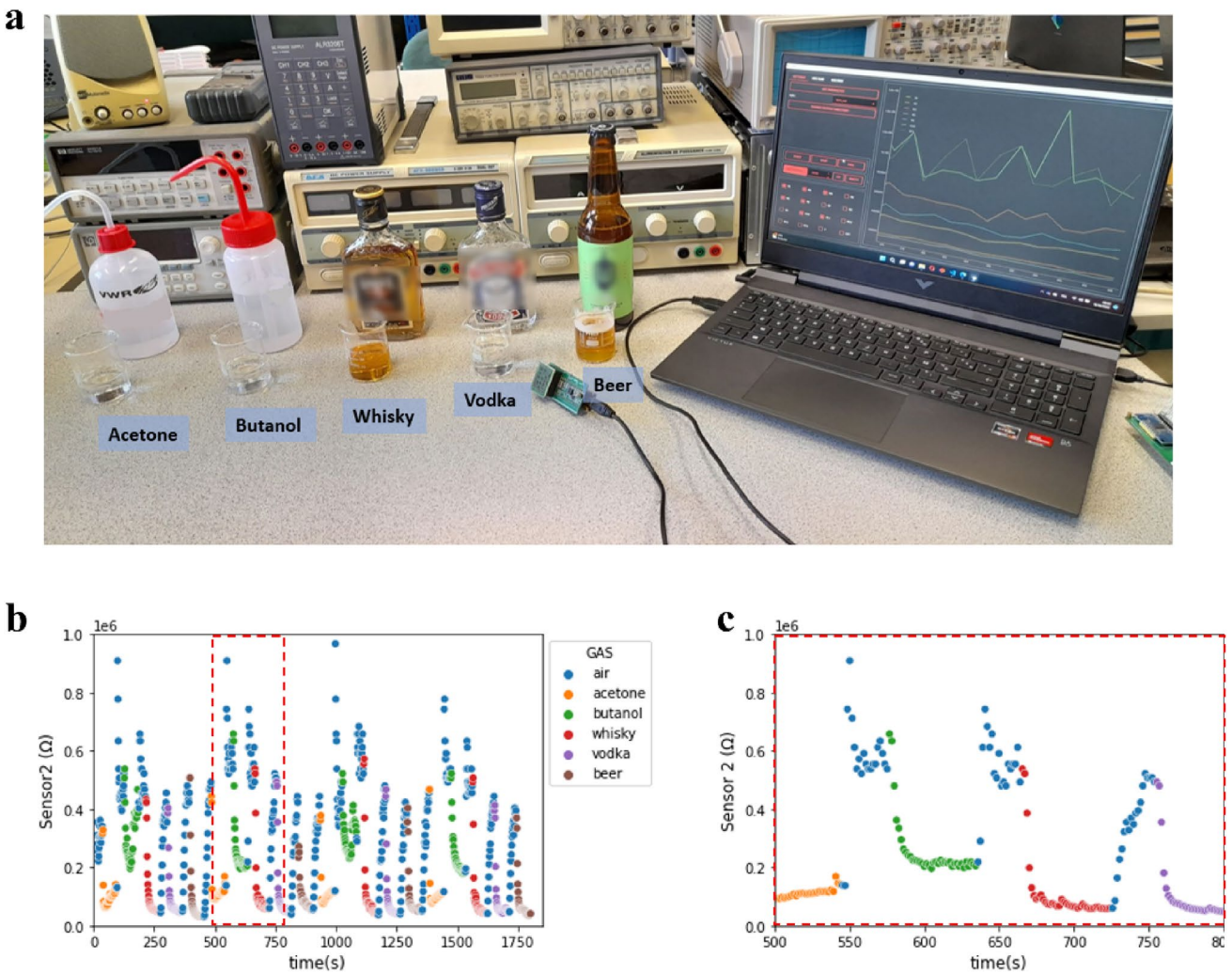


Fig. 9 **a** The setup of electronic-nose base AD5933 impedance analyser **b** response over time of sensor number 2 for different exposed solutions (acetone, butanol, whisky, vodka, beer), **c** Zoom of the sensor 2 impedance response

4.3 Classification results

The electrical performances of seven commercially available triflate salts ($Gd(OTf)_3$, $La(OTf)_3$, $Nd(OTf)_3$, $Dy(OTf)_3$, $Ce(OTf)_3$, $Ho(OTf)_3$, and $Eu(OTf)_3$) were studied by drop casting them onto poly(3-hexylthiophene) (P3HT), one of the most studied conductive polymers due to its excellent electrical properties and ease of processing [29]. This polymer is widely used in gas sensor applications. Zhanbo Cao [30] demonstrated that Organic Thin Film Transistors (OTFTs) with blended P3HT and n-type small molecule dicyanomethylene-terminated quinoid compound (TFT-CN), exhibit enhanced sensing performance for detecting H_2S gas. Additionally, the dynamic response of chemoresistive gas sensors based on P3HT nanofibers (NFs) to gaseous acetone was assessed using a flow-injection analysis setup, designed to emulate actual breath exhalations. In this study, we utilized P3HT as an active layer deposited on top of interdigital gold fingers. The experiment aims to mimic the olfactory system, which identifies odor patterns by chemisorbing elementary components of a volatome on various nodes that gather different VOCs, each reacting distinctly with their environment in a chemospicific manner [31]. Simultaneous multidimensional data processing was employed to recognize a variety of vapor saturation solutions. Figure 9a shows the experiment protocol, including sensors, system for acquisition, pre-processing, and feature extraction.

Data processing tasks such as noise filtering, normalization, and post-classification of the collected data were performed using Python code based on Scikit-learn, an open-source library for machine learning [32, 33]. Five solutions, including acetone, alcohol, whisky, vodka, and beer, were selected for this experiment. These solutions were selected to

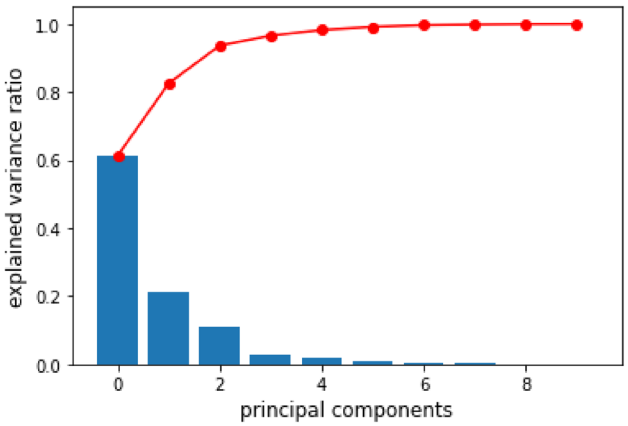


Fig. 10 Explained variance ratio plot showing that 90% of variance is included into the first three principal components

include target molecules with different chemical compositions, facilitating the study the response of our e-nose system. In our case, responses from all six sensors were the correlated attributes, which were transformed into a set of uncorrelated principal components. We exposed our e-nose to each vapor saturation for 90 s, followed by a purge time of 30 s in air, the transient changes in resistance of the triflated polythiophene coatings in response to vapor saturation were recorded. (see supporting information).

We should note that the peak resistance values on the graph represent the response of sensors to pure air, while the minimum values correspond to response of the sensors to each specific gas (Fig. 9b, c). Subsequently,the six responses from the sensor matrix were used as input parameters in PCA analysis.

By examining the six-loading data of the principal components, one can determine which combination of sensors in the array is most effective for distinguishing the target gases. Figure 11a, b display the responses and the correlation of each sensor within the array to all the target gases. Each response value plotted is the average of two consecutive resistance readings. The heat map clearly shows that all the individual triflated polythiophene coatings exhibited high sensitivity and dynamics to butanol and acetone vapor saturation. In contrast, the relative variation is smaller in the case of beer, whisky and vodka which can be attributed to the fact that these solutions are less volatiles than organic solvents–as the beverages mainly contains water– which clearly affects the relative variation of resistance values of each sensor. Previous research has reported that the sensitivity and selectivity can be improved and optimized by introducing dopants that alter the energy band structure and morphology of the semiconductor. Data analysis reveals that sensors coated with the same polymer but doped by different triflate dopants exhibit different resistance modulations when exposed to substances such as acetone, water, and isopropanol [34]. The observed changes in polymer resistance due to gas exposure are attributed to a decrease of the charge carrier's, which is correlates with a molecular gas's ability to transfer electrons with an electrophilic Lewis acid dopant [31, 34]. Using a sensor array rather than a single sensor, enables the effective exploitation of the collective effect of several triflated polythiophene coatings on a single tool to extract multiselectivity. This is because each gas produces a distinct pattern imprint on the sensing array, which can be further analyzed with multivariate data analysis techniques and machine learning algorithms. These analytical tools make it possible to detect and identify specific gases with greater precision. The component loadings in a principal component analysis (PCA) represent the correlation coefficients between the sensors in the array and the principal components, as demonstrated in the data. New datasets are created as linear combinations of the original ones, these datasets are highly correlated with each other. The features extracted from the training data using PCA are then applied to the test dataset for gas identification. Responses from the array are resolved by a smaller number of (PCs) that account for most of the variance in the observed attributes. The first three principal components (PC1–PC2–PC3) capture the maximum of the variance from the original data set. Figure 10 shows explained PCA ratio plot, where the relative contribution of the principal components PC1, PC2 and PC3 reflects 60%, 20% and 10%, respectively, resulting in 90% of the total variance. The principal components in PCA are orthogonal to each other, meaning they are at right angles and independent of each other. This orthogonality ensures that each principal component captures a unique aspect of the data's variance. The rest of the principal components captured components typically capture a much smaller portion of the total variance, and they might not contribute significantly to understanding the main patterns in the data.

It should be noted that a higher number of sensors results in more features for each data point, increasing the dimensionality of the dataset. This leads to an increase in computational complexity and time required to perform PCA but provides more information, helping in better distinguishing between different data points. Standardizing features is a

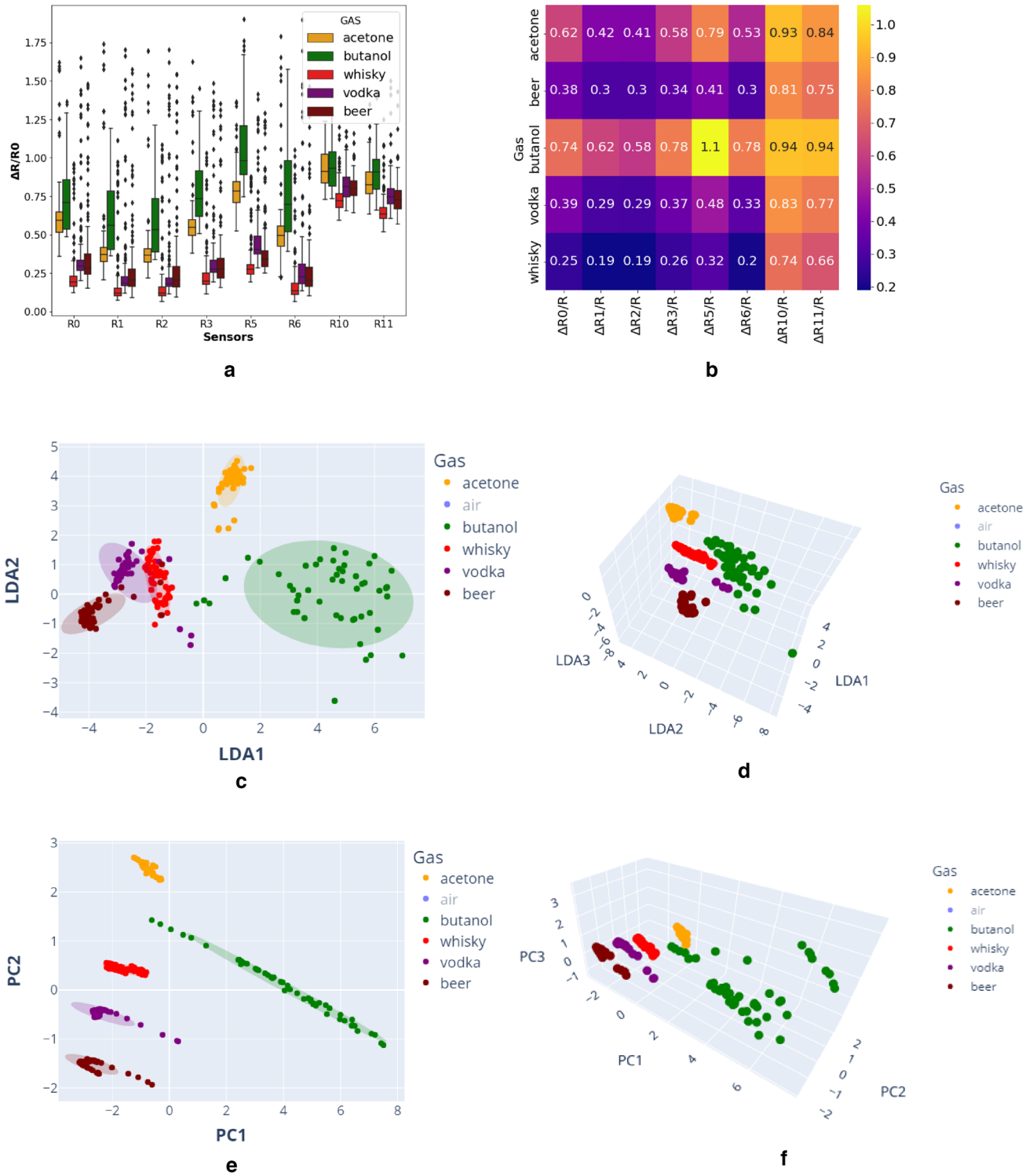


Fig. 11 **a** Statistics of sensor impedance responses toward vapor exposures **b** Correlation heatmaps showing the connection coefficients between gas and the relative variation of each sensor impedance **c** 2D LDA plot showing the discrimination of different vapor saturation based on normalized from different sensor data. **d** 3D LDA representation **e** 2D PCA plot showing the clustering of different vapor saturation based on normalized data acquired from different sensors **f** 3D PCA representation

crucial pre-processing step that involves scaling the input data to have a zero mean and unit standard deviation. This process is typically performed before conducting PCA to prevent bias resulting from differences in input scale. In fact, if the input data have very different scales, variables with the largest scales dominate the analysis, regardless of their contributions as the principal component loadings. The 2D and 3D PCA plots shown in Fig. 11e, f are based on normalized data. It is evident from these plots that the data for each solution are distinctly separated from each other, which is an accurate result for classification. However, we also observe the presence of sensor drift in the PCA projection dataset. This tendency of the test set to drift by a gradual shift of samples in one direction over time. Such phenomenon lead to the overlapping of the different classes and, consequently, to a loss of classification performance as time goes on. Drift effect is a widely addressed topic in the literature [35], and typically stems from unknown dynamic processes within the sensor, often related to changes in the sensing material. These modifications are usually caused by irreversible phenomena such as poisoning, aging, environmental conditions, thermo-mechanical degradation, or exposure to humidity. Sensor drift is defined as the temporal shift of sensors' response under stable environmental (physical and chemical) conditions. Several strategies have been proposed to tackle the problem of drift effect. Current solutions include the use of "adaptive" models that continuously update the classifier in real time. These adaptive models aim to adjust the classifier online taking into account pattern changes. Neural networks, such as self-organizing maps (SOMs) [35] or adaptive resonance theory (ART) networks [36], are frequently employed to overcome this effect. Nevertheless, this feature can also be addressed by combining multiple measurements taken over a certain period of time, rather than depending on a single series of measurements for training. The LDA classification algorithm generally considered to be a better choice for classification tasks compared to PCA, since it incorporates class labels and is specifically designed to optimize the separation between classes [37]. However, it is important to note that each algorithm serves different purposes and processes distinct properties. Whether one is "better" than the other depends on the specific context and objectives of a given problem. For unsupervised dimensionality reduction or when class labels are not available, PCA is the adequate algorithm. Conversely, for classification tasks where maximizing the interclass separation is crucial, LDA may be more suitable. In some scenarios, it might be advantageous to first apply PCA to reduce noise and then apply LDA to achieve further dimensionality reduction with a focus on class separation. Figure 11c, d show the result of the 2D and 3D discrimination of the different studied solutions. We should note that all extracted features have been normalized. Discrimination of the five-vapor saturation, regardless of their nature, is clearly achievable; the first two discriminant functions mainly separate butanol from the other gases. This separation can be attributed to the sensors' heightened sensitivity towards this particular solution. Additionally, in the case of ethanol-based beverages, the clustering groups are located in the same region, as these solutions are water-rich. Additionally, the same LDA plot also facilitates the direct correlation between the groups and the nature of the ethanol concentration in these solutions.

5 Conclusion

In this paper, we focused on the development of a compact, connected, and portable impedance analyzer. The AD5933-based system operates efficiently in [5–100 kHz] frequency range with high accuracy impedance measurement (error less than 5% for [1 kΩ–1 MΩ] impedance range) and this using only one calibration resistance. To go further in sensor application, we demonstrate a proof-of-concept for a low-cost, and portable e-nose application. Our platform has been tested and optimized to be used for multi-selective chemical organic sensor nodes. The multiplexed chemical sensor array is built upon functionalized chemiresistive nanomaterials to target various VOCs. The electrical performance of eight triflate salts, combined with an organic semiconductor poly(3-hexylthiophene) and used as an active layer, were studied. Sensor's impedances were used as input data of two classifications algorithms (PCA and LDA). The goal of this research is to provide a cost-effective system combined with feature reduction approach such as LDA and PCA, which can be used for software implementation for real-time applications. Nevertheless, given the better classification provided by fewer components LDA than PCA classification, this integrated chemiresistive gas sensor platform could be a powerful alternative to other diagnostic tools available for long-term monitoring.

6 Future work

Our system shows a good accuracy of the measurement compared to a benchtop impedance analyzer. Hence, it is important to study the impedance response by changing some parameters such as time exposer, excitation frequency, which can impact the classification algorithms. Other data analytics approaches will be investigated within the project with the purpose of deriving insights, predictions, or recommendations from the collected data including physical sensor properties gathered from the sensors. In particular, a specific data analysis service based on a combination of neural network and clustering analysis will be investigated.

Author contributions All authors contributed to the study. Conception and design were performed by Bilel HAFSI, data collection and analysis were performed by Gaël LOUIS, Anthyme CERVEAUX and Thomas HORLACH. The C++ code and incorporation of correction equations were the contribution of Louis ROUTIER. Human machine interface was developed by Pierre FOULON for android application and Alexandre WESTRELIN for python application. Semiconductor and triflate salts formulation and drop casting were performed by Sebastien PECQUEUR. The first draft of the manuscript was written by Bilel HAFSI, and all authors including Kamal LMIMOUNI commented on previous versions of the manuscript. All authors read and approved the final manuscript. The project was supervised by Bilel HAFSI.

Funding The authors gratefully acknowledge financial support from the COSMOS internal project of the Institute of Electronic, Microelectronic and Nanotechnology (IEMN), Institut Catholique des Arts et Métiers' ATOME project and French National Nanofabrication Network RENATECH for financial support of the IEMN technological platforms. Special thanks to Digilent for providing us the PMODIA based AD5933.

Availability of data and materials The authors declare that the data supporting this study are available within the paper, its supplementary information files. Copies of the data can be obtained via GitHub link: https://github.com/Bilel-HAFSI/ATOME-project-E-Nose-based-AD5933.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Kumar, et al. Internet of Things is a revolutionary approach for future technology enhancement: a review. J Big Data. 2019;6:111.
2. Anisimov DS, Chekusova VP, Trul AA, et al. Fully integrated ultra-sensitive electronic nose based on organic field-effect transistors. Sci Rep. 2021;11:10683.
3. Ramya M, et al. A recent advancement on the applications of nanomaterials in electrochemical sensors and biosensors. Chemosphere. 2022;308:136416.
4. Hara TO, Singh B. Electrochemical biosensors for detection of pesticides and heavy metal toxicants in water: recent trends and progress. ACS EST Water. 2021;1(3):462–78.
5. Xing L, Zhang W, Fu L, Lorenzo JM, Hao Y. Fabrication and application of electrochemical sensor for analyzing hydrogen peroxide in food system and biological samples. Food Chem. 2022;385:132555.
6. Chung M, Fortunato G, Radacsi N. Wearable flexible sweat sensors for healthcare monitoring: a review. J R Soc Interface. 2019;16(159):20190217.
7. Dodevska T, Hadzhiev D, Shterev I. Electrochemical sensors for the safety and quality control of cosmetics: an overview of achievements and challenges: review paper. J Electrochem Sci Eng. 2022;14(1):3–35.
8. Iurilli P, Brivio C, Wood V. On the use of electrochemical impedance spectroscopy to characterize and model the aging phenomena of lithium-ion batteries: a critical review. J Power Sources. 2021;505:229860.
9. Kashyap D, et al. Application of electrochemical impedance spectroscopy in bio-fuel cell characterization: a review. Int J Hydrog Energy. 2014;39(35):20159–70.
10. Attallah O, Morsi I. An electronic nose for identifying multiple combustible/harmful gases and their concentration levels via artificial intelligence. Measurement. 2022;199:111458.

11. Analog Devices. 1 MSPS, 12-bit impedance converter network analyzer. Analog Devices, Norwood, MA, USA, Datasheet AD5933, 2010. p. 40.
12. Le Maout P et al. A low cost, handheld E-nose for renal diseases early diagnosis. In: 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Honolulu, HI, 2018. p. 2817–2820.
13. Burgos-Flórez F, Rodríguez A, Cervera E, Zucolotto V, Sanjuán M, Villalba PJ. TBISTAT: an open-source, wireless portable, electrochemical impedance spectroscopy capable potentiostat for the point-of-care detection of S100B in plasma samples. PLoS ONE. 2022;17(2):e0263738.
14. Ibba P, et al. Design and validation of a portable AD5933–based impedance analyzer for smart agriculture. IEEE Access. 2021;9:63656–75.
15. Ferreira J, Seoane F, Lindecrantz K. Portable bioimpedance monitor evaluation for continuous impedance measurements. Towards wearable plethysmography applications. In: 35th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), Osaka, 2013. p. 559–562.
16. Yu R, et al. Impedance-based wireless sensor network for metal-protective coating evaluation. Int J Meas Technol Instrum Eng (IJMTIE). 2011;1(3):25–39.
17. Robert Harvey J, Mendelson Y. A portable sensor for skin bioimpedance measurements. Int J Sens Sens Netw. 2019;7(1):1.
18. Chabowski K, Piasecki T, Dzierka A, Nitsch K. Simple wide frequency range impedance meter based on AD5933 integrated circuit. Metrol Meas Syst. 2015;22(1):13–24.
19. Routier L, et al. Portable multiplexed system-based AD5933 impedance analyzer: toward multiselective gas recognition. IEEE Sens Lett. 2024;8(7):1–4. https://doi.org/10.1109/LSENS.2024.3415789.
20. Margo C, Katrib J, Nadi M, Rouane A. A four-electrode low frequency impedance spectroscopy measurement system using the AD5933 measurement chip. Physiol Meas. 2013;34(4):391–405.
21. Hoja J, Lentka G. A family of new generation miniaturized impedance analyzers for technical object diagnostics. Metrol Meas Syst. 2013;20(1):43–52.
22. Grassini S, Corbellini S, Parvis M, Angelini E, Zucchi F. A simple Arduino-based EIS system for in situ corrosion monitoring of metallic works of art. Measurement. 2018;114:508–14.
23. Placidi P, Vergini CVD, Papini N, Cecconi M, Mezzanotte P, Scorzoni A. Low-cost and low-frequency impedance meter for soil water content measurement in the precision agriculture scenario. IEEE Trans Instrum Meas. 2023;72:1–13.
24. Hu Z, Kallel AY, Lu T, Al-Hamry A, Kanoun O. An auto-calibrated measurement system for 1-D matrices of impedimetric sensors. IEEE Sens J. 2024;24(9):14970–6.
25. Bogónez-Franco P, Bayés-Genís A, Rosell J, Bragós R. Performance of an implantable impedance spectroscopy monitor using ZigBee. J Phys Conf Ser. 2010;224:012163. https://doi.org/10.1088/1742-6596/224/1/012163.
26. https://github.com/mjmeli/arduino-ad5933/blob/master/examples/ad5933-test/ad5933-test.ino.
27. Khan MAH, Thomson B, Debnath R, Motayed A, Rao MV. Nanowire-based sensor array for detection of cross-sensitive gases using PCA and machine learning algorithms. IEEE Sens J. 2020;20(11):6020–8.
28. Akbar MA, et al. An empirical study for PCA- and LDA-based feature reduction for gas identification. IEEE Sens J. 2016;16(14):5734–46.
29. Han S, Zhuang X, Shi W, Yang X, Li L, Yu J. Poly(3-hexylthiophene)/polystyrene (P3HT/PS) blends based organic field-effect transistor ammonia gas sensor. Sens Actuators B Chem. 2016;225:10–5. https://doi.org/10.1016/j.snb.2015.11.005.
30. Cao Z, Huo X, Ma Q, Song J, Pan Q, Chen L, Lai J, Shan X, Gao J. TFT-CN/P3HT blending active layer based two-component organic field-effect transistor for improved H2S gas detection. Sens Actuators B Chem. 2023;385:133685.
31. Boujnah A, Boubaker A, Kalboussi A, Lmimouni K, Pecqueur S. Mildly doped polythiophene with triflates for molecular recognition. Synth Met. 2021;280:116890.
32. Hwang C, Chen M-S, Shih C-M, Chen H-Y, Liu WK. Apply Scikit-learn in python to analyze driver behavior based on OBD data. In: 32nd International Conference on advanced information networking and applications workshops (WAINA), Krakow, 2018. p. 636–639.
33. Saraf E, Pradhan S, Joshi S, Sountharrajan S. Behavioral segmentation with product estimation using K-means clustering and seasonal ARIMA. In: 6th international conference on trends in electronics and informatics (ICOEI), Tirunelveli, India, 2022. p. 1641–1648.
34. Boujnah A, Boubaker A, Pecqueur S, Lmimouni K, Kalboussi A. An electronic nose using conductometric gas sensors based on P3HT doped with triflates for gas detection using computational techniques (PCA, LDA, and kNN). J Mater Sci Mater Electron. 2022;33:27132–46.
35. Padilla M, Perera A, Montoliu I, Chaudry A, Persaud K, Marco S. Drift compensation of gas sensor array data by Orthogonal Signal Correction. Chemom Intell Lab Syst. 2010;100:28–35.
36. Holmberg M, Winquist F, Lundstrom I, Davide F, DiNatale C, Damico A. Drift counteraction for an electronic nose. Sens Actuators B Chem. 1996;36:528.
37. Ding H, Ge HF, Liu JH. High performance of gas identification by wavelet transform-based fast feature extraction from temperature modulated semiconductor gas sensors. Sens Actuators B Chem. 2005;107:749.