UNIVERSITAT DE BARCELONA

Signal Processing and Machine Learning for Gas Sensors: Gas Source Localization with a Nano-Drone

Javier Burgués Calderón

Signal Processing and Machine Learning for Gas Sensors: Gas Source Localization with a Nano-Drone

Programa de doctorat en Enginyeria i Ciències Aplicades
(Codi: HDK0W)

Autor: Javier Burgués Calderón
Director i tutor: Dr. Santiago Marco Colás

UNIVERSITAT DE BARCELONA

Title: Signal Processing and Machine Learning for Gas Sensors: Gas Source Localization with a Nano-Drone

Signal Processing and Machine Learning for Gas Sensors: Gas Source Localization with a Nano-Drone

Gutta cavat lapidem, non vi, sed saepe cadendo

— Ovid

Abstract

Chemical source localization (CSL) by autonomous robots has been a topic of research since the early 1990s and still today remains elusive beyond simple scenarios. It has numerous potential applications, such as the localization of toxic emissions, malodors, gas leaks and hazardous substances in general, without risking human lives. An intuitive CSL approach is to mimic the known chemo-orientation behaviour of some flying insects, such as moths and mosquitos, which effectively use odor plumes for mating and foraging. However, terrestrial robots are too slow to perform insect-like movements and the response time and limit of detection (LOD) of current odor sensors for key compounds of biological relevance for plume navigation is orders of magnitude higher than in biological chemoreceptors. Instead of using a slow terrestrial robot equipped with complex instrumentation, in this thesis we address the CSL problem with a nano-drone, i.e. a miniaturized aerial robot, equipped with a simple metal oxide semiconductor (MOX) sensor.

Improving key specifications of MOX sensors for this application is one of the core parts of this thesis. Specifically, we introduce novel signal processing methods for estimating and optimizing the LOD, reducing the power consumption and improving the response time. The LOD is a key figure of merit of an odor sensor but difficult to estimate in non-linear devices (e.g. MOX sensors) or sensory systems with multivariate output (e.g. sensor arrays or temperature-modulated MOX sensors). We propose a univariate LOD optimization method based on linearized calibration models and a multivariate approach based on orthogonal partial least squares (O-PLS). To improve the response time, we use high-frequency features extracted from the MOX signal derivative, known as 'bouts' and optimize the algorithm for changing wind conditions and real-time operation. A novel setup consisting on a 3D grid of MOX sensors is proposed for 3D feature selection and real-time visualization of the gas distribution. Two map-based CSL strategies, one based on the instantaneous response and the other one based on bouts, are finally evaluated using the nano-drone in experiments performed in a large indoor environment (160 m^2) where a chemical source is placed in challenging positions.

The experimental results demonstrate that the proposed platform can be used to quickly (< 3 min) build a rough gas distribution map (3D) of the environment and localize the main chemical source within it with small errors. Future research lines in this field may focus on improving the selectivity of MOX sensors and the self-location of nano-drones, to perform fully autonomous exploration of complex environments.

Keywords: Mobile Robotics Olfaction, Gas Sensors, MOX Sensors, Gas Source Localization, Gas Distribution Mapping, UAV, Nano Drone, Signal Processing, Limit of Detection, Low Power Consumption, Temperature-modulation.

Acknowledgments

I would like to thank my supervisor Dr. Santiago Marco for giving me the opportunity to join the Signal and Information Processing for Sensing Systems group. He always had the door open for any question, which he answered effectively. Thanks to my lab teammates for their help (in alphabetical order): Agustin, Ariadna, Ana, Jordi, Juanma, Luciana, Luis, Marta, Rafael, Raqueles, Rudys, Sara, Sergi, Silvia and Toni.

This experience would not have been the same if our research group was not part of the Institute of BioEngineering of Catalonia (IBEC). Here, I found an international and multidisciplinary center in which I had the opportunity to attend more than 50 seminars, workshops, PhD discussions, social activities and much more. I thank to everyone at IBEC that work hard every day to make this possible. Thanks to the PhD committee, for organizing wonderful social activities such as calcotadas, volleyball, lasertag, bowling, ping-pong, etc. As a result of going through IBEC, I feel a much more prepared and open-minded person.

I must also thank Fran Palacio and Andrés Romero, from the University of Barcelona, for their valuable help and advice in soldering electronic components.

During Fall 2017, I did a research stay at Örebro University. I am grateful to the entire AASS team for the warm welcome since the first day, but especially to Achim and Victor for helping a lot. To Per, for helping with electronics and IT issues. Also to Erik, Han, Asif, Rafa. I will always remember our nice meals at Pizza Hat discussing about the problems in the Middle East.

A Loli y Sebas, porque me habéis demostrado que la distancia no importa cuando estáis al otro lado de la pantalla. MDC, he dedicado un especial esfuerzo en escribir con cariño la parte que más te gusta: La Introducción. A Mar, Felipe, Carla, Paula y Blanca. Para las abuelas Maria, Dori y Josefina.

Raquel, a ti te vull escriura aquestas paraulas en catalá per que veuras lo que vaig millorá en aquestos anys merci a ti. Gracies per la teva acogida fins el primer dia que et vaig coneixa, per haber estat al meu costat en lo mal y en lo bo, per compartir las alegrias i las tristeixas. T'vull no només per com ets, sinó per com sóc jo quan estic amb tu.

List of Publications

This thesis is presented as a compilation of papers. The results and contributions presented are reported in the following peer-reviewed articles and conference proceedings.

Peer-reviewed articles in indexed journals:

Paper I. J. Burgués, J.M. Jiménez-Soto, and S. Marco. **Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models**. *Analytica Chimica Acta* (8/80 (Q1) in Analytical Chemistry), 2018, vol. 1013, pp. 13-25. https://doi.org/10.1016/j.aca.2018.01.062

Paper II. J. Burgués, J.M. Jiménez-Soto, and S. Marco. **Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors**. *Analytica Chimica Acta* (8/80 (Q1) in Analytical Chemistry), 2018, vol. 1019, pp. 49-64.
https://doi.org/10.1016/j.aca.2018.03.005

Paper III. J. Burgués and S. Marco. **Low power operation of temperature-modulated Metal Oxide Semiconductor gas sensors**. *Sensors* (16/61 (Q2) in Instruments & Instrumentation), 2018, vol. 18, no. 2, p. 339. https://doi.org/10.3390/s18020339

Paper V. J. Burgués and S. Marco. **Feature Extraction for Transient Chemical Sensor Signals in Response to Turbulent Plumes: Application to Chemical Source Distance Prediction**. *Sensors and Actuators B: Chemical* (2/61 (Q1) in Instruments & Instrumentation), 2019, Submitted.

Paper VII. J. Burgués, V. Hernández, A. J. Lilienthal and S. Marco. **Gas distribution mapping and source localization using a 3D grid of metal oxide semiconductor sensors**. *Sensors and Actuators B: Chemical* (2/61 (Q1) in Instruments & Instrumentation), 2019, Under Revision.

Paper VIII. J. Burgués, V. Hernández, A. J. Lilienthal and S. Marco. **Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping**. *Sensors* (16/61 (Q2) in Instruments & Instrumentation), 2019, vol. 19, no. 3, p. 478.
https://doi.org/10.3390/s19030478

Conference proceedings:

Paper IV. J. Burgués, L. Valdez and S. Marco. **High-bandwidth e-nose for rapid tracking of turbulent plumes**. *International Symposium on Olfaction and Electronic Nose (ISOEN)*, 2019, Accepted.

Paper VI. J. Burgués and S. Marco. **Feature extraction of gas sensor signals for gas source localization**. *International Symposium on Olfaction and Electronic Nose (ISOEN)*, 2019, Accepted.

The following publications are not in the core contributions of this dissertation. However, they correspond to work performed during this thesis, mostly in the form of data collection and co-authoring of the articles.

Peer-reviewed articles in indexed journals:

- J. Palacín, E. Clotet, T. Pallejà, D. Martínez, J. Burgués, J. Fonollosa, A. Pardo and S. Marco. **Application of an array of Metal Oxide Semiconductor gas sensors in an Assistant Personal Robot for early gas leak detection**. *Sensors* (16/61 (Q2) in Instruments & Instrumentation), 2019, Under Revision.

- V. Pomareda, R. Magrans, J.M. Jiménez-Soto, D. Martínez, M. Tresánchez, J. Burgués, J. Palacín and S. Marco. **Chemical source localization fusing concentration information in the presence of chemical background noise**. *Sensors* (16/61 (Q2) in Instruments & Instrumentation), 2017, vol. 17, no. 4, p. 904. https://doi.org/10.3390/s17040904

- L. Fernández, J. Yan, J. Fonollosa, J. Burgués, A. Gutiérrez and S. Marco. **A Practical Method to Estimate the Resolving Power of a Chemical Sensor Array: Application to Feature Selection**. *Frontiers in Chemistry* (44/170 (Q2) in Multidisciplinary Chemistry), 2018, vol. 6, p. 209. https://doi.org/10.3389/fchem.2018.00209

Conference proceedings:

- F. Palacio, J.M. Gómez, J. Burgués, R. Pruna, M. López, A. Scorzoni, S. Zampolli and S. Marco. **Evaluation of MOX Sensor Characteristics in Ultra-Low Power Operation Modes: Application to a Semi-Passive RFID Tag for Food Logistics**. *MDPI Proceedings*, 2017, vol. 1, no. 4, p. 459. https://doi.org/10.3390/proceedings1040459

Contents

Chapter 1

Introduction

Many of the gases and volatile organic compounds (VOCs) present in indoor or outdoor air have toxic or flammable properties that can potentially cause harm to people, environmental damage and financial losses. Such gases may be released from different sources: as exhaust gases from traffic or industry, chemical leaks, flue gases from fires, geochemically active regions, domestic incidents, explosive compounds, foundries, agricultural plants, refineries and petrochemical plants (Figure 1.1(a)), among others. In places such as compost facilities, landfill sites (Figure 1.1(b)) or waste water treatment plants, the emitted gases are not only an environmental threat but can also be olfactory annoying to urban communities located near the emission source [1].

(a) (b)

Figure 1.1: Examples of gas emitting sites. (a) Petrochemical plant[1]. (b) Landfill[2].

Human's ability to detect and find the source of most dangerous gases is limited because some gases are invisible and, in most cases, have poor olfactory warning properties (i.e. odor thresholds) relative to their dangerous concentrations [2]. For example, methane (CH_4) is a colorless and odorless gas that easily forms underground leading to explosions

[1] `https://www.hpi-llc.com/industries-and-expertise/oil-gas-turbines/refinery-petrochemical-plant-services`

[2] `https://bosstek.com/odor-emissions-a-new-social-and-legal-dynamic`

and mine fires. Dreadful disasters, such as the one of the Courrières mine (France, 1906) or the Benxihu Colliery (China, 1942) that caused the death of 1099 and 1549 miners, respectively, motivated the development of artificial gas detection systems. The modern era of electronic gas sensing started in the late 1920s with the invention of the catalytic combustible gas sensor, to prevent explosions in fuel storage tanks [3]. Since then, many electronic technologies have been developed to detect, quantify and analyze a wide array of gases.

Gas sensors can be classified as in-situ or remote, according to whether they require direct interaction with the gas to produce a measurement or not [4]. In-situ gas sensors, also called point sensors, are based on the change of the properties of a material (e.g. conductivity, resonant frequency, etc.) when a gas interacts with it. Common in-situ sensors include semiconductor sensors, electrochemical sensors, combustible gas sensors, photoionization detectors, infrared point sensors and ultrasonic sensors. On the other hand, remote optical sensors such as tunable diode laser absorption spectroscopers (TDLAS) [5] or optical gas imaging (OGI) cameras measure a fundamental physical property of the target molecule (infrared absorption) without directly interacting with the gas. This prevents some failure mechanisms associated to in-situ gas sensors, such as poisoning and batch-to-batch variation in responsivity. Optical gas analyzers can be very selective for chemical species with unique absorption spectra (e.g. CO_2) but are not able to detect gases that do not absorb infrared energy (e.g. H_2, O_2 or N_2). With a relatively high price, optical remote instruments fill the gap between low-cost in-situ sensors with inferior performance and expensive high-end laboratory equipment, such as the mass spectrometer (MS).

If the potential source of a gas leak is known or a small area needs to be protected against a harmful gas, a fixed gas detection system might be an effective solution (Figure 1.2(a)). Portable gas detectors are commonly utilized in large areas where fixed systems are impractical or cost prohibitive, in search and rescue operations or in environments where people are only present to perform specific tasks on a limited basis (Figure 1.2(b)). A portable gas detector usually integrates one or more point-like sensors, a pump to bring the sample to the sensors and the necessary conditioning electronics. Over the past decade, some remote optical technologies (e.g. TDLAS) have evolved from specialized laboratory equipment to commercial portable devices, such as the remote methane leak detector (RMLD) [6].

The idea of installing a portable gas detector in a remote controlled mobile robot was first reported in 1984 [7], for the localization of gas leaks in nuclear plants minimizing human exposure to radiation. Similar scenarios were devised later for autonomous robots in charge of firefighting, prospecting, demining, environmental monitoring, search and rescue, contraband interception and others [8–11]. Mobile robots are effective in replacing humans dealing with simple but laborious repetitive tasks (e.g. routine patrols), can work in hostile atmospheres (e.g. radioactive, oxygen-deficient, extreme temperatures) and navigate through confined spaces that preclude human entry. This makes them suitable platforms for Chemical, Biological, Radiological, Nuclear, and Explosive (CBRNE) incident response applications [12, 13].

[3] `http://www.gassonic.com/products/gassonic_observer_i`

[4] `https://www.tech4g.com/gas-detection-becoming-mandatory-at-workplace`

(a) (b)

Figure 1.2: (a) Ultrasonic detector used to detect gases escaping from pressurized leaks[3]; (b) Human operator monitoring gases in a sewer tunnel using a PID[4].

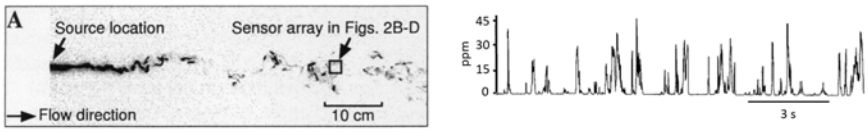
A mobile olfactory robot can perform a number of olfactory related tasks, including chemical trail following (CTF), gas source localization (GSL) and gas distribution mapping (GDM) [14]. Chemical trails are odor markers that a mobile robot navigating an environment can lay on the floor to find its way back to the starting point, to mark parts of the environment which have been already explored or to communicate with other robots. A chemical trail is relatively easy to detect by placing a gas sensor facing downwards in front of the robot and attaching a fan to bring the chemicals closer to the sensor (Figure 1.3(a)). Then, the robot follows a zig-zag trajectory to detect the edges of the trail [15, 16].

(a) (b) (c)

Figure 1.3: (a) Robot with one MOX sensor (1994) [17]; (b) Robot equipped with an electronic nose and an anemometer (2009) [18]; (c) Terrestrial robot equipped with remote optical sensors for monitoring methane leaks in industrial plants (2012) [19].

GSL is a much more challenging task that consists on searching for the point of release of a gaseous substance, e.g., in applications such as gas leak detection or localization of explosives. The first experiments using mobile robots for GSL were reported in the early 1990s [17, 20–23] and still nowadays is an open research problem. GSL robots typically include several gas sensors to compute the chemical gradient and anemometers to determine the wind direction (Figure 1.3(b)). The main challenges are related to the chaotic propagation of chemicals in turbulent fluid flows, the known limitations of in-situ gas sen-

sors and the mobility constraints of mobile robots. Chemicals released in air are advected downstream by wind, in the form of a plume. In most natural indoor and outdoor environments, air flow is in the turbulent regime [24]. Turbulence breaks the plume into patches or eddies (Figure 1.4(a)). Consequently, the instantaneous concentration measured in the centerline of a turbulent plume strongly fluctuates with time (Figure 1.4(b)). High concentration peaks are interspersed with "blanks" of zero or near-zero concentration, which explains why simple orientation to a concentration gradient is unlikely to lead a robot towards the source [21].



(a) (b)

Figure 1.4: (a) Image of a turbulent dye plume captured using planar laser-induced fluorescence (PLIF) [25]; (b) Instantaneous concentration measured by a fast PID (330 Hz) in the centerline of a turbulent plume (the trace is 15 s) [26].

To cope with the intermittent structure of turbulent plumes, researchers developed sophisticated navigation algorithms inspired by the excellent odor tracking behaviour of some insects [27]. These algorithms were mostly demonstrated in simplified test environments with controlled air streams, no obstacles, no chemical interferences and with the robot placed relatively close to the gas source [28]. At this time, it is unclear whether these bioinspired reactive behaviours have better performance than other approaches based on statistical inference from cumulative readings [28–30]. According to Hernandez et al. [28], the bioinspired reactive behaviors that researchers have implemented on mobile robots are modelled too simple to cope with complex environments, gas sensors used are too slow to resolve plume features in a milliseconds scale and mobile robots are not agile enough for performing insect-like reactive movements. Later on, researchers investigated the use of terrestrial robots equipped with optical remote sensors (Figure 1.3(c)) for gas leak detection in complex environments, such as industrial facilities [19] and landfills and biogas production sites [31].

GDM is the task of building a map of the gas distribution based on spatially and temporally distributed concentration measurements. A gas distribution map can highlight areas in which high concentrations of a harmful gas are to be expected (Figure 1.5), providing a means of addressing the GSL problem. Although fixed measurement stations are primarily used for this purpose due to the social problems associated to industrial contamination and dispersion of annoying odors, GDM is an ideal task for mobile robots because they can provide higher spatial resolution than a network of fixed stations, at a much lower cost. Mobile robots represent a trade-off between spatial and temporal resolution, since they can measure in a spatially dense grid but they cannot measure at several times simultaneously. A way to overcome this limitation is to use multiple mobile robots that collaboratively explore the target area [9] or integrating the information provided by mobile robots with data captured by fixed networks [32]. For instance, terrestrial robots have been used for

mapping pollution in public areas [33] and methane emissions in landfill sites [28], among other scenarios [34, 35].

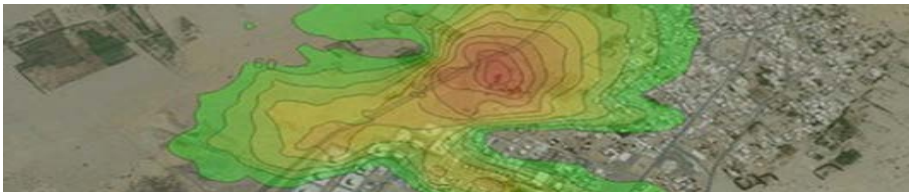


Figure 1.5: Gas distribution map of an industrial site, built using measurements of an aerial vehicle carrying a portable gas detector[5].

Underground, underwater and aerial applications of gas-sensitive robots have been reported too. Odor robots equipped with ground-penetrating probes (Figure 1.6(a)) have been proposed for searching buried chemical sources, such as truffles, gas/fuel leaks from buried pipes and storage facilities, land mines or victims of avalanches or earthquakes [36–38]. Autonomous underwater vehicles (AUVs) have been used for monitoring natural waters [35, 39] (Figure 1.6(b)) or identification of seafloor structures such as fissures, bubble plumes, exposed hydrates, carbonates, etc. [40, 41].

Unmanned aerial vehicles (UAVs), commonly known as drones (Figure 1.6(c)), equipped with gas detection systems and/or sampling bags have been used for environmental monitoring [35, 42–49], volcanic gas sampling [50–55], localization of fugitive emissions [56, 57], early fire detection [58,59], precision agriculture [60–62], landfill monitoring [63–65], disaster response [66] and mine blasting [67], among others [68, 69].

(a) (b) (c)

Figure 1.6: (a) Ground-penetrating mobile robot, equipping a probe with a MOX gas sensor [37]; (b) Underwater vehicle equipped with a mass spectrometer for monitoring of natural waters [39]; (c) Unmanned aerial vehicle equipped with in-situ gas sensors and a sampling bag for pollution monitoring in residential neighbourhoods[6].

The growing civilian interest in UAVs for gas sensing applications is demonstrated by the rapid emergence of companies selling air quality sampling units for UAVs or UAVs

[5] http://scentroid.com/photo-gallery-4

[6] http://scentroid.com/police-using-scentroid-dr1000-flying-lab-to-combat-smog

with an integrated gas detector. Some examples are provided in Figure 1.7. Scentroid DR1000 is an air quality sampling unit that can be attached to any UAV that can carry a payload heavier than 3.5 kg (Figure 1.7(a)). The DR1000 can host up to five chemical sensors for real-time monitoring of pollutants such as particulate PM1-10, ethanol, formaldehyde or ammonia. It can also include a sampling bag for sample collection in potentially dangerous sites, such as chimneys or flares. This platform and similar ones have been already used by police departments in some European countries for law enforcement purposes; e.g. to detect illegal emissions from chimneys[6] or to find illegal gardens of cannabis[7].

The GasFinder2, from Boreal Laser, is a custom TDLAS device to provide fixed-wing UAVs with the ability to localize fugitive emissions of gases such as methane, carbon dioxide, and ammonia, among others. The analyzer is installed in the body of the aircraft and fibre optic cable carries the laser light from the analyzer to the transceiver (Tx/Rx) placed on one of the UAV's winglets (Figure 1.7(b)). The laser light emitted from the Tx/Rx travels over the wingspan of the UAV and is reflected back by the retro-reflector, forming an active measurement path. The detector within the Tx/Rx converts the received laser light into an electrical signal which is carried back to the analyzer via coaxial cable. If the UAV flies through a plume containing methane concentration above a preprogrammed threshold, the remote controller will display an audible alarm so the UAV pilot can circle around the potential fugitive gas emission to survey the leak further.

Other companies sell UAVs already equipped with optical remote sensors, mainly for detection of methane. For example, the mdTector (Figure 1.7(c)), from Microdrones, consists of a laser absorption spectrophotometer (Pergam Suisse AG, Switzerland) integrated with a Microdrones md4-1000 UAV. Its purpose is to aid operators in the inspection of methane gas infrastructure, natural gas line surveys, landfill emission monitoring, and plant safety, among others. The HS-5000 (Hesai Ltd) is a similar system in which a TDLAS detector is mounted in a small UAV for autonomous natural gas inspection in high-rise buildings. The M600 PRO OGI (Figure 1.7(d)), from Viper Systems, is a Matrice 600 UAV (DJI) equipped with an OGI camera (FLIR Systems) for aerial gas leak detection.

A potential limitation of UAVs for gas sensing tasks is the interaction between the air flows produced by the propulsion system and the existing gas distribution. Rotorcrafts, which are the most popular type of UAV due to its ability to hover, generate strong air flows beneath the rotors (i.e. downwash) that strongly disturb the local gas distribution near the drone and might have negative consequences for the utility of the gas sensor data [70–72]. As can be seen in Figure 1.8, a smoke plume becomes more spread and less concentrated after a drone flies above it. As a result, the on-board gas sensors register less concentration than if the same measurements were taken with a hand-held detector [70]. Estimating the direction of the gas source becomes thus more challenging [71, 72]. UAVs mounting remote optical technologies are not necessarily affected by this problem because

[7] https://www.trendhunter.com/trends/high-tech-aerial-drug-sniffing-dutch-\
police-canna-chopper-cannabis

[8] http://scentroid.com/scentroid-dr1000

[9] https://www.microdrones.com/en/integrated-systems/mdtector/mdtector1000ch4

[10] https://viper-drones.com/systems/dji-m600-pro-optical-gas-imaging-gas-leak-\
detection-system

[11] http://www.boreal-laser.com/products/uav-based-gas-detector

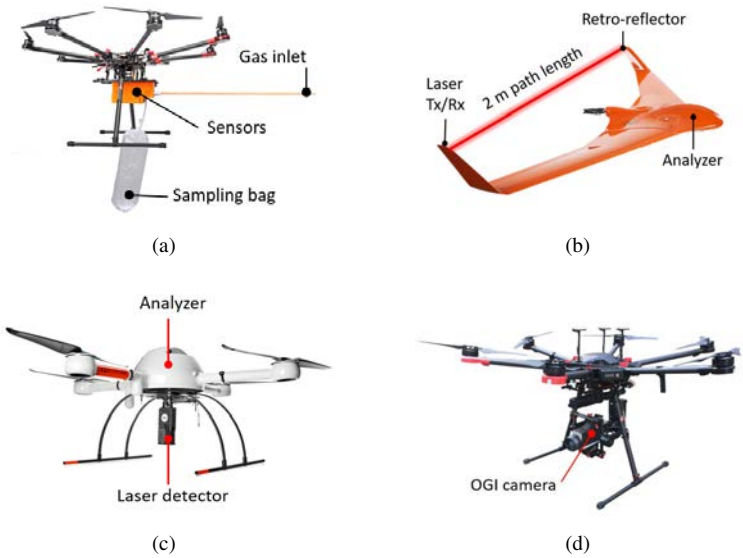


Figure 1.7: Commercial UAVs equipped with gas detectors. (a) Scentroid DR1000 air quality sampling unit attached to a DJI S1000 octocopter[8]; (b) Boreal GasFinder2 laser detector mounted on a Bramor ppX fixed-wing UAV (C-Astral Aerospace Ltd.)[9](c) Microdrones mdTector for methane infrastructure inspection[10]; (d) DJI M600 PRO with OGI camera for gas leak detection[11];

the UAV does not need to be in the plume for measuring. Fixed-wing UAVs produce less disturbance to the plume than rotorcrafts as they generate force behind the aircraft (i.e. wake turbulence).

Figure 1.8: Smoke column (a) before, (b) during and (c) after a DJI M600 hexacopter flies over it (Source: Unpublished results).

Thanks to recent advances in micro-technology, manufacturers of drones have been able to develop miniaturized drones; with insect-sized aircrafts reportedly expected in the future (Figure 1.9). A nano air vehicle (NAV) or nano-drone is extremely small, with a wingspan lower than 15 cm, and weighs less than 50 g [73]. The tiny form-factor and maneuverability of nano-drones allow remote sensing of hazardous environments inaccessible to bigger drones, can fly over areas being unobserved and the propellers introduce

less disturbance to the surrounding gas distribution than larger drones. For example, a nano-drone equipped with gas sensors could be used to search for victims and hazardous gas leaks inside damaged buildings or pockets that form in the aftermath of an earthquake or explosion. Compared to terrestrial robots, a nano-drone could navigate such environments (and many others) much faster, evading obstacles or large gaps and sampling the space in 3D.

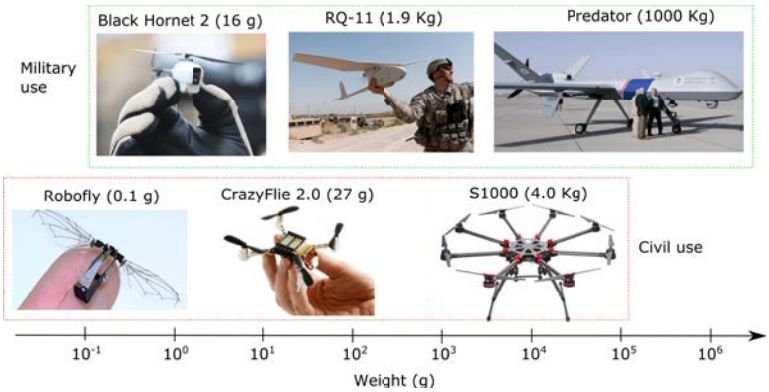


Figure 1.9: Classification of drones according to their approximate weight (and size). The graphic shows the large diversity of drones in which the possible sizes spans seven orders of magnitude (from [74]).

Two experimental works [72, 75] already explored the viability of nano-drones for gas sensing tasks. Rossi et al. [72] performed preliminary indoor experiments using a CrazyFlie 2.0 nano-drone equipped with a MOX sensor. Surprisingly, the authors found that the air drawn down around the airframe strongly affected the sensor response, resulting in useless signals. They evaluated several mechanical solutions to hold the sensor out of the radius of turbulence, but the drone became uncontrollable because of inertia problems. The adopted solution was to operate the drone in the so-called "butterfly" mode, in which a human pilot lands the drone in the proximity of the source and halts the motors to take a measurement. In this way, the sensor signals are not affected by the rotors but, at the same time, the 3D sensing capabilities of the drone are not used, and the approach might not scale well to large scenarios.

Fahad et al. [75] equipped the same nano-drone with a chemically sensitive field effect transistor (CS-FET) sensor for hydrogen (H_2) detection. The test environment was a chemical hood in which H_2 was released from the upper part of the hood and the drone ascended from the bottom of the hood to the area near the source ($h = 60$ cm), aided by high tension strings. The sensor response increased as the drone approached the source, reaching its maximum value after hovering (i.e., levitating) near the source for 40 s. The drone was merely used as a proof-of-concept demonstration of the proprietary gas sensor developed by the authors. The above works suggest that a nano-drone might be used for gas source localization, however the experimental scenarios were extremely simple.

The main problems for performing large-scale experiments in complex environments using nano-drones are related to the limited on-board resources and difficulty to control

the platform due to inertia and stability issues. Taking as an example the CrazyFlie 2.0 (CF2) quadcopter, the tiny 240 mAh battery delivers power for up to 7 min of flight and 15 grams of payload, which means that only lightweight and power-efficient sensors can be mounted on board. Self-localization and obstacle avoidance—required for autonomous navigation—are hard to accomplish because laser scanners, for example, are too heavy for the nano-drone payload. Autonomous navigation using RGB cameras is still immature for a real application [76]. The Global Positioning System (GPS) can be used for localization outdoors where, however, nano-drones can often not be controlled stably due to their low inertia and strong wind. Semi-autonomous navigation in indoor areas can be achieved through motion capture systems (MOCAPs) [77] or radio frequency (RF)-based systems [78]. Deploying an external localization system might not be possible in many realistic scenarios (e.g., in a disaster situation) and the drone would have to navigate autonomously or remotely controlled by an operator.

Gas sensing tasks using a nano-drone are also subject to additional constraints, as they must be executed in the short time limited by the battery capacity and relying exclusively on one or two in-situ gas sensors. It should be noted that most research on GSL and GDM is based on terrestrial robots, which can run for hours, perform long measurements of 1–2 min at each sampling location and possibly use selective sensors (e.g., TDLAS, OGI cameras, e-noses) and anemometers. Large drones can fly for 20–25 min and be equipped with the same technology as terrestrial robots. Nano-drones are therefore subject to unprecedented constraints because a stop-sense-go strategy would only allow for at most 10 measurements (of 30 s each), the limited number of sensors hinder rejecting chemical interferences and the absence of anemometry prevents assessing the wind direction, which is a key parameter for GSL. During its operation, the drone can also not fly too fast or the relevant structures of the chemical plume may become blurredor shifted due to the slow response time of the sensors [14].

1.1 Gas dispersion

To select appropriate sensors and algorithms for odor robots, it is of practical interest to understand the dispersion dynamics of the target gas in the target environment. In this section, we explain the basics mechanisms of gas dispersion and present the main gas dispersion models that can be used in certain conditions to predict the spatial concentration of a released gas.

1.1.1 Gas dispersion mechanisms

A chemical released in a fluid (e.g., air) disperses according to two mechanisms, diffusion and advection. Molecular diffusion is the motion of gas particles due to temperature, which explains the movement of particles from regions of high concentration to regions of lower concentration. In the absence of external forces, diffusion would lead to a dynamic equilibrium where the distribution of gas molecules in the environment is uniform. Advection is the movement of gas particles due to the motion of the fluid in which they are suspended, usually creating a "plume" (Figure **??**). Plumes that develop in slowly moving viscous fluids (e.g., liquids) usually show a smooth laminar behavior, whereas plumes in

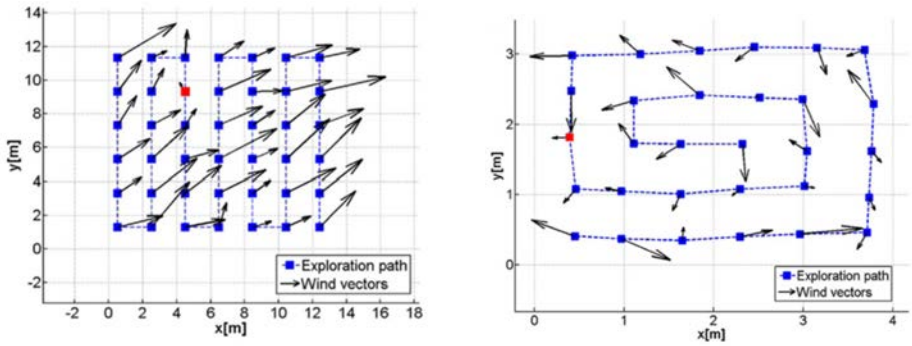
fluids with high flow rate or low viscosity (e.g., gases) typically present turbulent characteristics [79]. Turbulent plumes are characterized by unsteady aperiodic motion, random spatial variations (3D) and intermittent temporal structure.

The thermodynamics of the gas being released and the characteristics of the gas source also affect gas dispersion. Near the release point, the plume can be influenced by the thermodynamics of the gas (e.g., pressure, temperature, density, etc.). Substances which are lighter than air (e.g., hydrogen and helium) and hot gases (e.g., those emitted during a fire) generally show buoyant plumes near the release point. Alternatively, heavy gases, such as propane or chlorine, and cold gases (e.g., Methane is liquefied at –162 ºC for transportation and storage) will generate downward plumes near the leak point. Gases with similar density to air (e.g., ethanol or CO) will be neutrally buoyant, which means that they do not have any intrinsic movement of either up or down. As the plume moves away from the release point, the concentration of the gas decreases and its density becomes similar to that of air. At this point, the plume becomes fully dominated by the air flow of the environment. This is not the case for gases released at high velocity (e.g. due to punctures in pipes, vent stacks or safety valves), since they will cause jets. The velocity inside a jet is usually much higher than the wind speed, so that dispersion in this case is mostly affected by the properties of the jet itself.

Dispersion mechanisms indoors (e.g., inside a building or a cave) are completely different than outdoors, due to boundary effects of the building walls, weak air flow regime, thermal sources (e.g., hot equipment, radiators), forced ventilation systems, etc. In outdoor scenarios, the wind speed is usually strong and its direction can be considered relatively homogeneous across the experimental area for limited periods of time (Figure 1.48), yielding relatively well-defined plumes. Indoor environments present as a defining characteristic a lack of strong or persistent mean flow (Figure 1.48), except in corridors or areas with forced ventilation (either by artificial means or through open windows). The low energy turbulent mixing produces a chaotic patchy gas distribution of the released chemicals and the gas concentration rapidly increases in case of a leak due to boundary effects of the walls [80]. Local maxima in concentration have been observed at some distance to the source if the source has been active for some time [81, 82]. In most natural indoor and outdoor environments, air flow is in the turbulent regime [24].

1.1.2 Gas dispersion models

Gas dispersion models enable the prediction of gas concentrations (as a function of time) at any location surrounding the gas release (up to several kilometres). They are normally used by air quality management agencies to estimate outdoor air quality by predicting the downwind concentration of hazardous gases emitted from industrial plants, vehicular traffic or accidental chemical releases. Although several indoor dispersion models exist [83], the literature in this field is relatively sparse compared to outdoor pollutant dispersion. In this section we describe three popular gas dispersion models: the Gaussian plume model, the filament/particle model and computational fluid dynamics models.



(a) (b)

Figure 1.10: Air flow map of (a) outdoor environment; (b) indoor closed room. At each location, the arrow's length and orientation represent the average wind speed and direction, respectively, computed over 30 s. The air speed varied within the following limits: (a) 0.3 to 1.8 m/s, (b) 0 to 0.08 m/s (from [28]).

Gaussian plume model

Assuming a steady homogeneous fluid flow and a continuous point-like source, the solution to the convection-diffusion equation is the Gaussian plume model (GPM) [84])

$$c(x,y,z) = \frac{q}{2\pi U \sigma_y \sigma_z} \exp\left(-\frac{y^2}{2{\sigma_y}^2}\right) \cdot \left\{ \exp\left(-\frac{(z-H)^2}{2{\sigma_z}^2}\right) + \exp\left(-\frac{(z+H)^2}{2{\sigma_z}^2}\right) \right\} \quad (1.1)$$

where c is the average concentration (g/m^2), x,y,z are the downwind, crosswind and vertical distance from the source (m), respectively, q the release rate (g/s), H the height (m) of plume centerline and U the average wind speed (m/s) along the plume centerline. Equation 1.1 describes a model where the crosswind and vertical dispersion follow a Gaussian distribution with standard devation σ_y and σ_z, respectively, and the source is located at the origin of coordinates (Figure 1.11). The quantities σ_y and σ_z can be computed from charts [85] based on the atmospheric stability class, the mean wind speed (U) and the downwind distance (x). There is physical evidence that the model is valid in many outdoor scenarios if the wind conditions are stable and spatially uniform, there are no obstacles and measurements are averaged for at least 10 minutes [85–88]. GPMs are not adequate when the wind speed is low or unsteady. There exist other formulations of the GPM [85, 89].

Filament/particle models

Farrell [90] developed a computationally-efficient plume model that not only accounts for long-term plume characteristics but also for short-term features called filaments. The plume is represented as a sequence of 'puffs' [91] that are released sequentially by the

[12] https://en.wikipedia.org/wiki/Atmospheric_dispersion_modeling

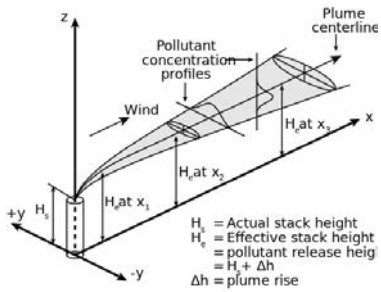


Figure 1.11: Schematic diagram of a Gaussian plume[12]

source, and each puff is composed of multiple filaments (Figure 1.12). It is assumed that filaments are dispersed by wind, so that the position $\mathbf{p}_i(t) = (p_x, p_y, p_z)$ of the i-th filament at time t can be computed from its previous position $\mathbf{p}_i(t-1)$ and the wind vector. The instantaneous concentration C (ng/m^3) at location $\mathbf{x} = (x, y, z)$ and time t is the sum of the concentrations contributed by the N filaments of the plume

$$C(\mathbf{x}, t) = \sum_{i=1}^{N} \frac{Q}{\sqrt{8\pi^3 R_i^3(t)}} \exp\left(-\frac{\|\mathbf{x} - \mathbf{p}_i(t)\|^2}{R_i^2(t)}\right) \tag{1.2}$$

where Q is the release rate (ng/s) and $R_i(t)$ the size (m^2) of the i-th filament at time t.

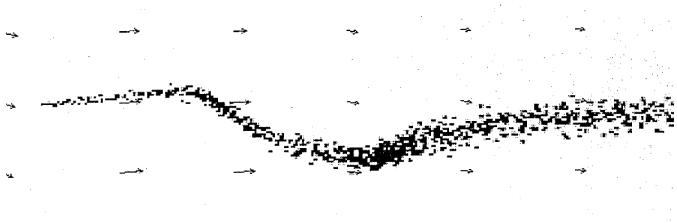


Figure 1.12: Simulated odor plume using the filament-based model (from [90]). The arrows indicate the wind direction.

Computational fluid dynamics (CFD)

Computational fluid dynamics (CFD) [92] dispersion models aim at numerically solving the analytically intractable Navier-Stokes (NS) equations [93] in 3 dimensions. These equations are a set of nonlinear partial differential equations that model the pressure (p), density (r), temperature (T) and velocity of the fluid (v_x, v_y and v_z) as a function of the spatial coordinates (x, y, z) and the time (t). Given some initial conditions, the NS equations could in theory predict the properties of the fluid (independent variables) at each point of space and time (dependent variables), which would enable predicting the gas concentration $c(x, y, z, t)$.

A direct numerical simulation (DNS), i.e. solving the whole range of spatial and temporal scales of the turbulence, requires extremely high time resolution and fine three-dimensional grids, which makes it computationally infeasible [94]. Two methods can be employed to simplify the Navier–Stokes equations in such a way that small-scale fluctuations do not have to be directly simulated: large eddy simulation (LES) and Reynolds averaged Navier Stokes (RANS). LES filters out eddies smaller than a certain size and only simulates the largest eddies. RANS considers all the scales of turbulence but only models the time-averaged airflow in the environment, using turbulence models [95]. CFD results strongly depend on the ability of the chosen turbulence model (there are hundreds of different turbulent models) and the specified initial conditions to predict all physical phenomena of the target environment. CFD simulations of gas releases inside buildings tend to be over simplistic, due to difficult-to-model convective flows (e.g., heat conduction through the building enclosure, heat from objects, solar radiation through the windows) and mechanical movements (e.g., forced ventilation systems and occupant movement) [94].

CFD models have been developed for simulating dense gas dispersion events [96], atmoshperic dispersion in urban areas [97] or gas dispersion in indoor scenarios [94,98,99]. GADEN [99] combines a CFD wind simulator and a filament-based dispersion model to produce a 3D representation of gas dispersion in realistic indoor scenarios, e.g., including walls and obstacles. An advantage over other CFD simulators is that multiple gas sources and different chemical substances can be simulated simultaneously (Figure 1.13).

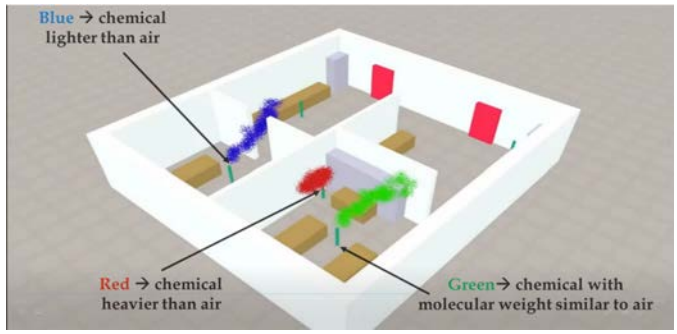


Figure 1.13: Snapshot of GADEN simulation[13] of an indoor environment with three simultaneous gas sources with different molecular weights.

1.2 Gas source localization algorithms

GSL algorithms are usually classified according to the characteristics of the fluid in which they are supposed to work: diffusion-dominated flow, strong turbulent flow or weak turbulent flow [27]. GSL strategies in diffusion-dominated environments are usually gradient-based, whereas GSL for turbulent flows can be divided into three groups [27]: reactive plume tracking (bioinspired) algorithms, plume modelling algorithms and map-based approaches (Figure 1.14).

[14] https://www.youtube.com/watch?v=ZPGtk8KLtiE&t=139s

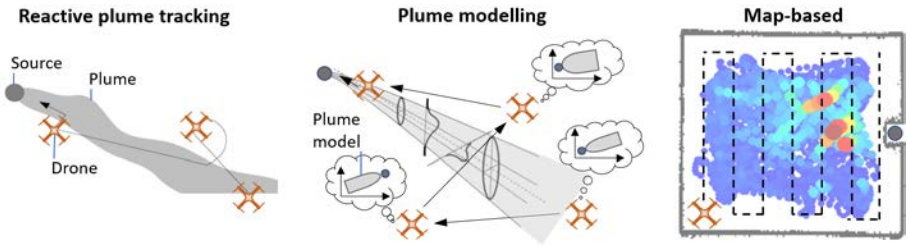


Figure 1.14: Gas source localization strategies. (a) Reactive plume tracking; (b) Plume modelling; (c) Gas distribution mapping

1.2.1 Diffusion dominated flow

When the fluid flow is dominated by viscosity, as in underground or underwater scenarios, smooth chemical distributions with a single peak at the location of the gas source can be found. In this case, simply following the concentration gradient is an effective method to localize the gas source (Figure 1.15). A 2D chemical gradient can be computed using a single sensor [20], two symetrically arranged sensors [21] or three sensors shaped in a triangle [100]. In the single sensor case, acquisition of spatially separated measurements by moving the robot between the sites of the reading is necessary [20, 101]. This is also necessary in the case of multiple sensors if the chemical field is weak or the separation between the sensors small. Using multiple sensors is faster but requires matching the responses of the different sensors and is more prone to sensor drift.

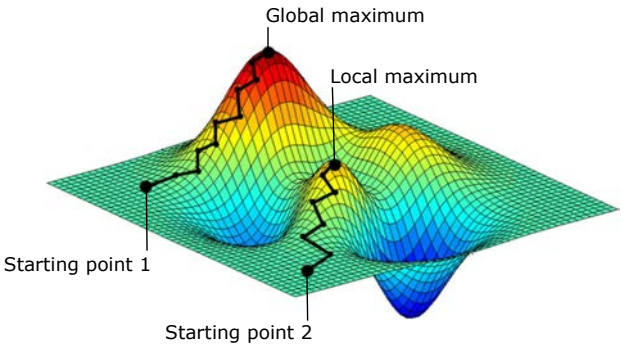


Figure 1.15: Gradient ascent method (2D). The color map represents the concentration. Depending on the starting point the searcher might get trapped in local maxima (adapted from [102]).

Holland and Melhuish [101] derived gradient-based algorithms for robots with a single sensor, based on the behaviour of the E.coli bacteria and the Planarian worm. These algorithms implement a set of simple reactive rules that produce slowly zig-zag movements towards the source. Russell et al. [103] proposed a modification of the Planarian algorithm,

the so-called hex-path algorithm, which replaces the zig-zag pattern by an hexagonal trajectory to increase the robustness of the algorithm. They demonstrated the algorithm in the localization of an ethanol source buried in sand. Algorithms based on gradient ascent have been also proposed for localization of underwater [41] and above-ground [20] sources.

Some atmospheric gas releases can be adequately described by a smooth Gaussian distribution when averaging measurements over temporal windows of about 10 minutes [85, 104]. Although a gradient-based method could theoretically be used to localize a source in this scenario, the exploration becomes extremely slow. Ishida et al. [21] developed an "odor compass", composed of two MOX sensors, to compute the chemical gradient and used the device to localize an ethanol gas source in a small wind tunnel (70 x 80 cm^2). Several hours were required to locate the source because the concentration gradient was negligibly small. Sandini et al. [105] expanded the gradient-based algorithm to multiple cooperating robots to increase the search speed.

1.2.2 Strong turbulent fluid flow

If there is strong wind, gas molecules will be dispersed into the environment by turbulence and simple orientation to a concentration gradient is unlikely to lead a robot towards the source. In this case, the robot must use the intermittent chemosensory cues provided in the gas plume to track it towards its origin.

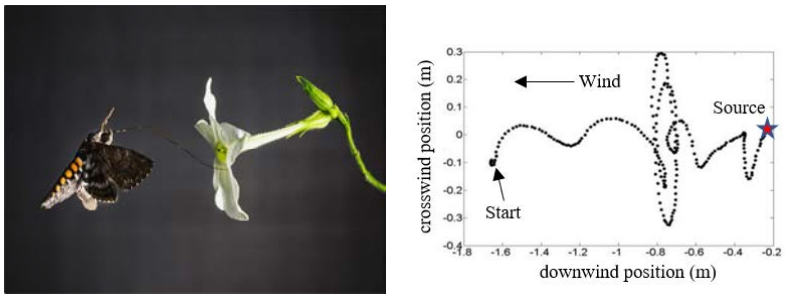
Reactive plume tracking (bioinspired) algorithms

An immediate approach for developing a plume tracking strategy is to mimic the excellent odor plume tracking capabilities of some animals. Insects, for example, use odor plumes for mating, foraging or to detect the presence of a predator. Experimental observations of the male moth tracking pheromone plumes in wind-tunnels have been used to decode key plume tracking behaviors of these insects (Figure 1.16(a)). Contact with individual filaments of the plume or 'odor hits', rather than concentration gradients, seem the most likely features insects use to orient rapidly in turbulent plumes [87,106,107]. Upon contact with the plume, the moth flies upwind upon and oscillates cross-wind to reacquire the plume whenever it is lost [106, 108] (Figure 1.16(b)). Observed behavioral changes as the moth approaches the pheromone source, such as decrease in flight speed, might indicate the existence of a different chemical signature near the source [109].

Reactive plume tracking algorithms imitating the behavior of the silkworm moth, dung beetle and e. coli bacteria, among others, have been implemented in mobile robot [27,111, 112]. These bioinspired implementations usually consist of three phases: plume finding, plume tracking and source declaration [113].

Plume finding In order to establish the initial contact with the plume, either the robot passively waits at the start position until detecting the plume [23,114] or actively searches the plume by moving obliquely upwind [115], randomly [115], cross-wind [17] or spiraling [113]. It was found in simulation that cross-wind searching was more effective than

[13] https://phys.org/news/2016-05-scent-hawk-moths-best-fitting.html



(a) (b)

Figure 1.16: (a) Male moth (from[15]); (b) Male moth tracking a pheromone plume (adapted from [110]).

oblique and random searching [115]. It was also found that sniffing (i.e. dragging air to the sensors with a pump/fan) improves plume detection [116]. Because of the low probability of odor encounter far from the source, active searching is necessary in large exploration areas, especially if UAVs are used due to their limited flight time [70].

Contact with the plume is typically declared when the measured concentration (or the sensor response) surpasses a predefined threshold [17, 113]. For example, Hayes et al. [113] set this threshold at 4σ (σ is the standard deviation of the sensor baseline measured in clean air) to minimize the false positive rate. Most bioinspired algorithms declare that the plume is lost when the measured concentration falls behind the same threshold determined for plume detection. Because a turbulent plume is intermittent and meanders, using a fixed threshold to decide when the robot is "in the plume" and "not in the plume" will produce many false detections/non-detections [117]. Additionally, the slow response time of some gas sensing technologies may delay the detection of the plume (Figure 1.17), making the robot lose the plume [117]. Changes in the background concentration as the robot approaches the gas source may also hinder the application of a fixed threshold (which is estimated in clean air conditions).

To make the plume detection more robust, some authors proposed adaptive thresholds [117, 118] and hysteresis [37, 119]. Ishida et al. [117] defines that the robot enters (resp. exists) the plume when the sensor response changes by more than 10-20% of a recent local maximum (resp. minimum). Li et al. [118] defines a threshold that is computed over the moving average of the signal. Purnamadjaja and Russell [37, 119] propose two separate thresholds for defining 'in' and 'out' of the plume, although it is not specified how these thresholds should be determined. Pashami et al. [120] presented a review of methods to detect changes in MOX sensor signals due to a distant gas source.

Other authors [113,121] consider a turbulent gas plume as a sequence of packets rather than a continuous trail (Figure 1.18). Hayes et al. [113] proposed the "surge spiral" algorithm in which the robot triggers the "out the plume" behaviour when the frequency of packet encounter falls behind a certain threshold. Schmuker et al. [121] recently presented a signal processing method to detect plume patches using features extracted from the derivative of the MOX sensor signals (the so-called 'bouts'). The advantage of us-

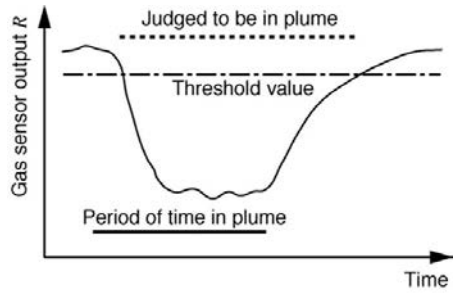


Figure 1.17: Threshold-based plume detection. The thick line shows the period of time when the robot was actually in the plume, whereas the dotted line indicates when the robot is judged to be in the plume according to the sensor response and the chosen threshold (dash-dot line) (from [117]).

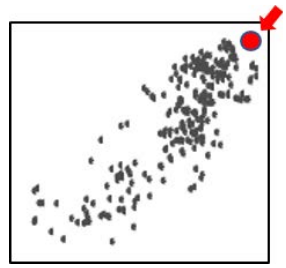


Figure 1.18: Plume hits during one hour in a wind tunnel. The red circle indicates the location of the gas source and the red arrow indicates the air flow direction (adapted from [113]).

ing the derivative instead of the raw response is that its dynamics are much faster and is not affected by changes in the background concentration. It nonetheless requires proper smoothing of the response and the derivative, and a bout amplitude threshold (which the authors set at 3σ) to discard low-amplitude bouts produced by noise (i.e. false positives).

Plume tracking Once the plume has been detected, straight [113], oblique [23] and zig-zag upwind movements [22] can lead the robot towards the source (Figure 1.19(a)). If the robot exits the plume, plume reacquisition strategies such as spiraling [122] (Figure 1.19(b)), zig-zag followed by spiraling [123] and casting [23] have been proposed in the literature. Lochmatter et al. [124] found that algorithms based on upwind surge (surge-cast or surge-spiral) were significantly faster than pure casting, in wind-tunnel experiments using a terrestrial robot. Neumman et al. [70] found that surge-cast with straight upwind movements outperformed zig-zag and gradient-based methods in simulation. They also evaluated these three algorithms in real-world outdoor experiments using a micro-UAV flying at constant height. Although the results do not clarify which algorithm works better, they suggest that plume tracking using a UAV can be successful only under stable wind flow.

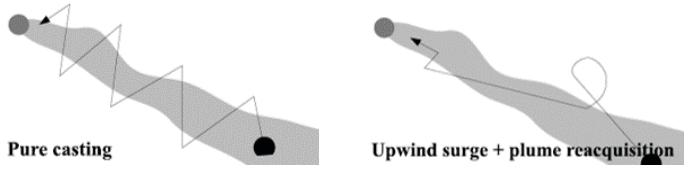


Figure 1.19: Examples of reactive plume tracking strategies (from [29]).(a) Pure casting; (b) Upwind surge combined with spiraling (for plume reacquisiton).

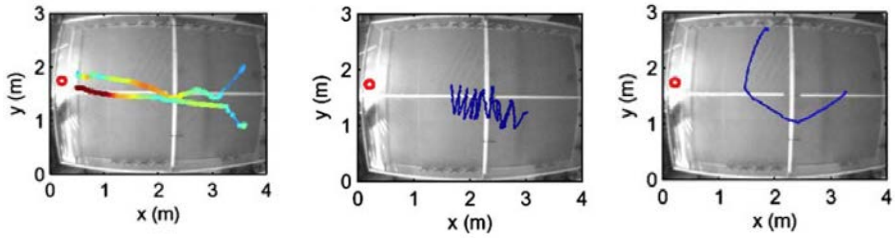
Source declaration A reactive plume tracking algorithm needs a termination condition to stop the search, i.e. declare when the source has been located. For simplicity, in most works the source is declared (and the experiment finished) as soon as the robot gets within a certain distance (typically 20-50 cm) of the true source location [17, 29, 70, 125–127]. This approach presents three drawbacks: (1) it requires knowledge about the true source location, (2) in small-sized arenas, random walking searchers would often attain the boundaries of the gas source, and (3) it is unknown what would be the final position of the robot if the experiment was not stopped at that point.

A less common approach is to declare the source when the measured concentration exceeds a predefined threshold [128, 129]. This requires prior knowledge about the source intensity, which is typically unknown. Alternative source declaration methods have been proposed based on the relative variance of the response [130], mean and variability [122], inter-hit distance [40, 122, 131], particle filter [70, 118], divergence operator [132] or using vision [128]

Experimental evaluation Reactive plume tracking algorithms have been demonstrated mostly in simulation and simplified experimental scenarios in which the robots travel small distances across a constant unidirectional airflow in an arena free of obstacles [110, 112, 124, 127] (Figure 1.20(a)). Plumes in such scenarios are relatively thin and well-defined, extending uninterruptedly several meters downwind of the source. Bioinspired algorithms are extremely inefficient in natural scenarios where the air speed is too weak (indoor scenarios) [23] or the wind direction is constantly shifting (outdoor scenarios) [70]. This was demonstrated by ceasing the air flow in wind tunnel experiments [112] (Figure **??**) or by introducing sudden changes in wind direction outdoors [70].

Odor compass

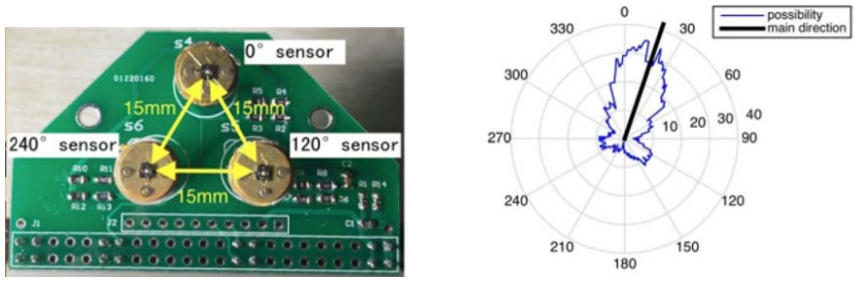
The purpose of an odor compass is to determine the direction towards the odor source. Wei et al. [133] recently developed a 2D odor compass using 3 naked MOX sensors shaped in a triangle (Figure 1.21(a)). The direction of the odor is computed after 2.8 s of measurement from the lag between matched features in the band-pass filtered sensor signals. A difference of Gaussians (DoG) band-pass filter is used to highlight areas of the signal where the rate of change is maximum. Features with maximum or minimum DoG values are



(a) (b) (c)

Figure 1.20: Evaluation of a reactive plume tracking algorithm (surge-cast) in a wind tunnel. (a) Strong unidirectional air flow; (b-c) Absence of air flow. The red circle indicates the location of the gas source. Adapted from [112].

matched using a scale-invariant feature transform (SIFT) algorithm. The direction of the odor and the confidence in the estimated direction are represented in a wind rose diagram (Figure 1.21(b)).



(a) (b)

Figure 1.21: (a) Odor compass made of 3 MOX sensors; (b) Wind rose representation of the output of the system (the thick black line indicates the estimated source direction). From [133].

The device was used by a human operator to localize a gas source in small indoor (4×5 m^2) and outdoor (6×7 m^2) scenarios by following the estimated odor direction. In both cases, a fan was used to induce a moderate directional airflow in the environment because the odor compass is not reliable when the wind speed is either too low or too high, or the wind shifts direction. Success rates between 75% and 90% were achieved, with a mean distance to the source of 0.2-0.5 m.

Plume modelling algorithms

An alternative to reactive plume tracking is to use a mathematical model of the plume as a heuristic to guide the search process. Plume modelling algorithms assume that the spatial distribution of gas concentrations can be explained by an underlying mathematical model,

such as Gaussian shaped plumes [84] or filament/particle models [90, 134, 135], among others. Instead of reactively tracking the plume, the robot alternates periods of exploration of the environment (i.e. collect measurements of concentration and wind) with periods of exploitation (i.e. move towards the source). This adaptative behaviour, that involves learning and spatial memory, can increase the speed and accuracy in gas source localization when plumes are highly intermittent [136]. According to Lilienthal et al. [14], algorithms that store information of previous encounters with a plume should in principle outperform methods that discard such information.

Gaussian plume model. The Gaussian plume model (GPM) has been used by several authors [126, 137–140] to localize a source in indoor environments with strong unidirectional airflow. Ishida et al. [137] originally developed an algorithm to estimate the location of a point source placed on the floor (i.e. $H = 0$), by fitting a GPM [89] with local measurements of wind and concentration obtained by a terrestrial robot (i.e. $x = 0$). The gas distribution at location (x, y) is obtained by substituting $H = 0$ and $x = 0$ into the equivalent of Equation 1.1 in [89] and transforming the coordinate system to include the source (x_s, y_s) as a parameter of the model (Figure 1.22). The gas distribution on the ground is given by

$$c(x,y) = \frac{q}{2\pi K d_s} \exp\left(-\frac{U}{2K}(d_s - \Delta x)\right) \tag{1.3}$$

where K is the turbulent diffusion coefficient (cm^2/s), d_s is the Euclidean distance (m) from (x, y) to (x_s, y_s), and Δx is given by

$$\Delta x = (x_s - x)cos\theta + (y_s - y)sin\theta \tag{1.4}$$

where θ is the angle from the x-axis to the upwind direction. The algorithm determines x_s and y_s from Equation 1.3 using a gradient descent method and Kalman filtering, given multiple measurements of c, x, y and U (collected by the robot at different locations).

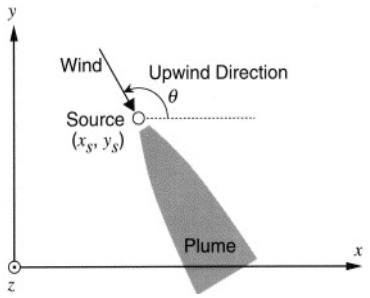


Figure 1.22: Gaussian plume model with coordinate system used by Ishida et al. [137] in which the source is not at the origin of coordinates.

A theoretical advantage of this model is that the source can be found without moving towards it. Its practical applicability is however limited because (i) values for K and

q must be given a priori (e.g., $K = 30$ cm^2/s and $q = 200$ cm^3/min in [138, 139], (ii) the environmental conditions are not usually as stable as it is assumed by the model and (iii) the long time-averaging required to observe a Gaussian shaped plume might not be efficient for rapid GSL using a single robot. Indeed, Sanchez et al. [140] recently used a GPM approach in a small room of 15 m^2 and 10 minutes were required to localize the source. Marques et al. [126] found similar performance between this algorithm and a bio-inspired algorithm (silkworm moth) across 20 tests in a ventilated indoor arena.

Filament/particle models Farrell et al. [141] developed a probabilistic algorithm to locate a gas source in an intermittent plume, assuming a particle model and constant, uniform and known wind. The local density of filaments is used as a heuristic to move towards the source, since the probability of filament encounter is higher near the source. Using binary chemical detections and wind measurements, the robot recursively builds a maximum likelihood map that indicates which regions are likely to contain the odor source. At each time step, the robot is directed to areas that maximize the likelihood of the detection of odor or the likelihood of finding the source. The source is declared when the last three odor encounters are separated by less than 3-5 m. Pang and Farrell [142] improved the efficiency of the algorithm by estimating the likelihood map using Bayesian methods. The algorithm of Pang and Farrell was recently revised by Pomareda et al. [143] to use quantitative concentration measurements instead of binary detections.In practice, estimating the density of filaments in a certain region requires temporal averaging, which makes this method similar to algorithms based on the Gaussian plume model.

In Infotaxis [144], the plume is represented by a set of detectable "particles" emitted by the source at a constant rate and transported by a turbulent flow [134,135] (Figure 1.23). Assuming that the probability of particle encounter as a function of distance to the source is known, the algorithm builds a posterior probability map where the value of each cell represents the probability of containing the source. A particle encounter is declared when the measured concentration exceeds a certain threshold. At each time step, the searcher measures the gas concentration for a certain period of time at the current location, counts the number of particle 'hits' and updates the map. Then, it visits the cell that locally maximizes the expected rate of information acquisition. The intuitive idea is that particles arrive at a higher rate close to the source, hence tracking the maximum rate of information acquisition will guide the searcher to the source.

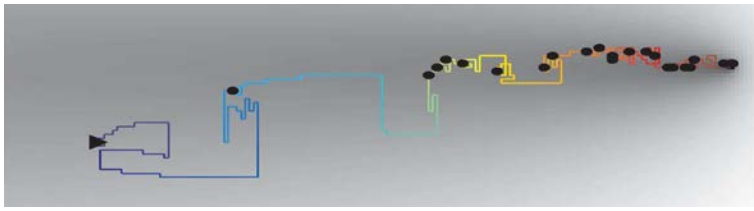


Figure 1.23: Typical infotactic trajectory. The black dots represent the odor detection events (from [144]).

Recently, Voges et al. [123] found slightly higher success rates of infotaxis when compared to three bioinspired algorithms (93% versus 84-89%) in locating a pheromone source in wind tunnel experiments. Recordings from the antenna of a moth mounted on a mobile robot were used to avoid the long response time of MOX sensors. The authors claim that cognitive searching (e.g., infotaxis) might be more appropriate than reactive plume tracking when the robot's starting location is far from the odor plume. The main practical limitations of infotaxis are that it requires prior knowledge of the spatial probability of odor hits, it assumes that the emission rate is constant and it is unclear how the binary concentration threshold should be estimated.

Vuka et al. [145] proposed a GSL algorithm for moderate air flow environments (air speed < 0.3 m/s) in which the goal is to build a map of the amplitudes of detected 'bouts' [121]. The 'bouts' are specific features extracted from the derivative of the sensor response, which are supposedly caused by contact with gas patches. The underlying assumptions of Vuka's algorithm is that the bout amplitude increases with proximity to the source and that cells in upwind direction from the current measurement point have a similar bout amplitude (to avoid exploring those cells). The algorithm divides the search area into cells of equal size and the robots visits each cell to measure the bout amplitude and wind direction. After the map has been sufficiently explored, the cell with the highest bout amplitude is declared as the source (Figure 1.24).

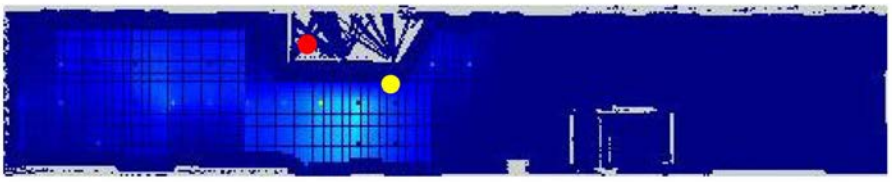


Figure 1.24: Mean bout amplitude map in an indoor corridor (wind was flowing towards the south-east). The dark and light blue regions represent areas with low and high mean bout amplitude, respectively. The red and yellow circles indicate the gas source and the final position of the robot, respectively (from [145]).

Experiments were performed in a large ventilated indoor corridor (22 x 4 m^2), with 8 successful trials out of 12 (67%). An experiment was declared successful if the cell with the highest bout amplitude was the nearest one to the true source location, along the wind direction. The average time required to find the source is not specified and it is unclear how the source shall be declared when its true location is unknown.

3D algorithms

According to Lilienthal et al. [14] and Ishida et al. [34], addressing the gas source tracing problem in 3D is the next step to go for research in this field. Prior to the commercial availability of low-cost UAVs, researchers demonstrated the advantages of measuring the gas concentration in 3D either by manually moving a gas sensing probe to different heights [111, 146, 147] (Figure 1.25a), using a terrestrial robot with an extendable vane [148] (Figure 1.25b) or with sensors mounted at multiple heights [129, 149] (Fig-

ure 1.25c), using a blimp tethered to a ground robot [150] (Figure 1.25d) or using a 3D traversing system [151] (Figure 1.25e).

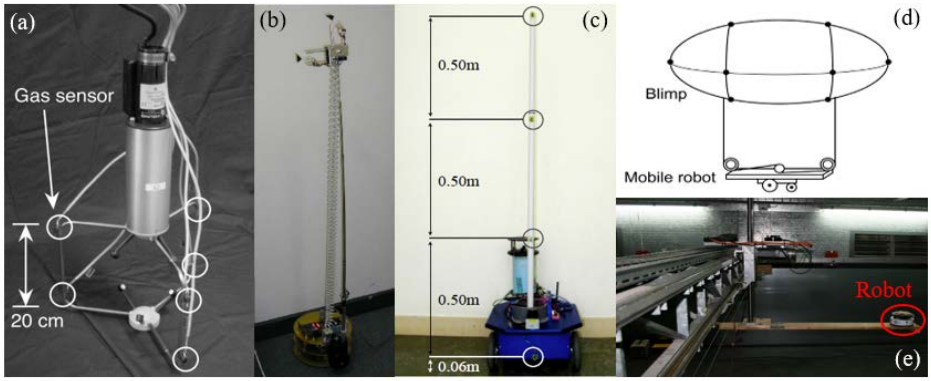


Figure 1.25: Early 3D gas sensing instruments. (a) Manual sensing probe [147]; (b) Terrestrial robot with MOX sensors mounted on an extendable vane [148]; (c) Terrestrial robot with MOX sensors mounted at different heights [129]; (d) Blimp tethered to a ground robot [150]; (e) 3D traversing arm carrying a terrestrial robot [151]

;

Soares et al. [151] emulated an aerial robot using a 3-axis traversing system in a wind tunnel. They compared the readings of a terrestrial robot (Figure 1.26(a)) and the pseudo-aerial robot (Figure 1.26(b)), with the gas source placed slightly higher than the height of the terrestrial robots. In general, the aerial robot sensed much more concentration than terrestrial robots. As the aerial robot approaches the source, the fluctuations of the sensor response increase, indicating proximity to the source. The opposite behavior was seen in the terrestrial robot. Moving towards the source reduced both the average and the variance of the measurements, because the vertical width of the plume of a point gas source is minimal near the source. This highlights a clear advantage for those robots capable of measuring at the same height of the source.

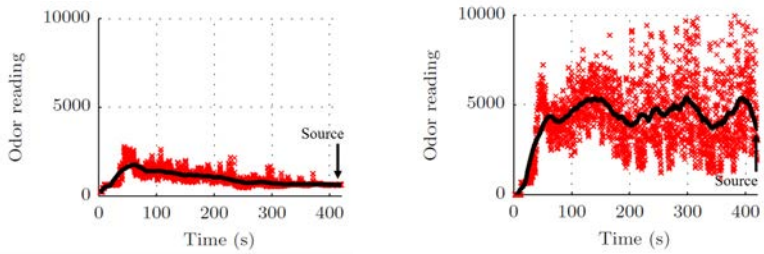
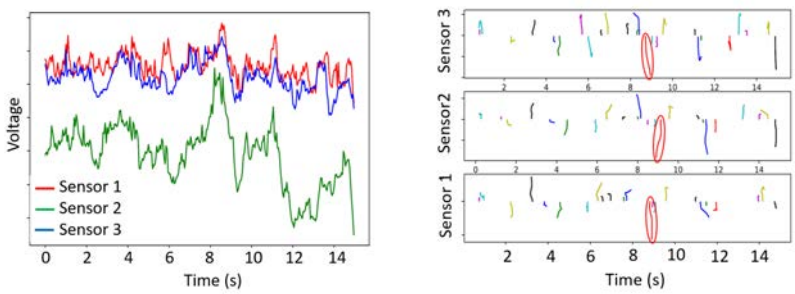


Figure 1.26: Comparison of sensor readings of (a) terrestrial and (b) pseudo-aerial robots in a wind tunnel (adapted from [151]).

Osorio et al. [129] used a robot with three MOX sensors attached at different heights to compare the performance of two bioinspired algorithms (surge-cast and cast) and a gradient-based one, when data of the three sensors is fused versus using only data from the bottom sensor. The data fusion consists in taking the maximum concentration of the three sensors. Experiments in a large indoor corridor with weak air flow showed that using data from the three sensors improved the success rate from 56% to 70%, due to a better reacquisition of the plume when sampling at multiple heights.

Ishida et al. [150] used a small blimp with ten MOX sensors distributed around its surface, which was tethered to a ground robot (a wrench was used to regulate the height of the blimp). Experiments performed in a strongly ventilated indoor arena (5.7 x 3.5 m^2) suggest that the variability of the sensor signals is more useful for source localization than the average response. Although a blimp is advantageous because it can fly with little disturbance to the surrounding gas distribution, its limited maneuverability hinder the practical application to realistic indoor scenarios or environments with strong wind.

Odor source orientation inference (OSOI). Luo et al. [71] developed the so-called "flying odor compass", which is a commercial quadrotor equipped with 3 MOX sensors shaped in a triangle (inspired by the design of the 2D odor compass presented by Wei et al. [133]). The proposed OSOI algorithm computes the odor direction (e_{odor}) by fusing the gas flow direction (e_{g}) and the wind direction (e_{w}). For computing e_{g}, the raw sensor signals (Figure 1.27(a)) are filtered with a wavelet transform modulus maxima method (WTMM), i.e. a band-pass filter that identifies singularities in the variability of the sensor signals (Figure 1.27(b)). The authors claim that singularities with a wavelet coefficient higher than a threshold are produced by contact with odor patches, so those with a small coefficient are filtered out. The time difference between the same event across the three sensors is used to estimate e_{g}.



(a) (b)

Figure 1.27: OSOI algorithm. (a) Raw sensor signals; (b) Wavelet transform of the three sensors, with modulus lines colored by grouping. The red ellipse highlights the same modulus line detected by different sensors. Adapted from [71].

The system was used to locate a source in a small indoor environment (4 x 3 m^2) with strong natural airflow, achieving a mean success rate of 95% (across 60 trials) and a mean

time to reach the source slightly higher than 2 min. The sensitivity of the algorithm to the threshold used to filter out small events is not specified.

3D odor compass. Ishida et al. [146] developed an odor compass made of 4 MOX sensors and a small fan (to draw air towards the sensors), which are mounted on a platform that can rotate in 3D. The odor direction is estimated from the delay of the sensor responses sampled during 24 s of rotation of the sensor probe. The utility of the compass to locate an elevated gas source was demonstrated in a extremely small indoor arena (≈ 2 m^2) with two air supply openings. The probe was manually moved in steps of 20 cm in the odor direction provided by the device and 5-6 minutes were required to localize the source. The low exploration efficiency (≈ 2.5 min/m^2) might limit the applicability of this system to larger scenarios.

Ishida et al. [147] improved its previous odor compass by developing a new one consisting of a 3D ultrasonic anemometer surrounded by 6 MOX sensors. The direction towards the gas source $\mathbf{d} = (d_x, d_y, d_z)$ is determined by a linear combination of the upwind vector $\mathbf{v} = (v_x, v_y, v_z)$ and the gas concentration gradient $\mathbf{g} = (g_x, g_y, g_z)$

$$\mathbf{d} = \mathbf{v} + k\mathbf{g} \tag{1.5}$$

where k is a constant. Equation 1.5 assumes that $\mathbf{g}$ always points towards the centre of the plume (Figure 1.34(a)), which is a very strong assumption that may not be satisfied in many realistic scenarios. Successful GSL experiments were carried out in the same indoor scenario as before, using a measurement time of 2 minutes and a step size of 20 cm.

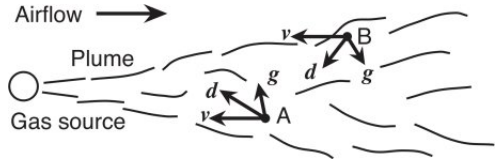


Figure 1.28: Computation of gas source direction **d** from the wind vector **v** and the concentration gradient **g** (from [147]).

3D zig-zag. Russell et al. [148] developed a 3D version of the zig-zag algorithm previously presented by Ishida et al. [22]. A 3D raster scan is initially performed to find the gas plume, which is modeled as having a circular cross-section (Figure 1.29). At each time step, the robot estimates the plume shape by measuring the concentration in three points of the boundary of the plume. The method was tested in a very small indoor scenario (1 x 0.6 m2) with artificial ventilation and succeeded in 5/8 trials. No indication about the exploration time was provided. The assumption of a circular plume with well-defined boundaries might not hold in many real scenarios.

3D casting-2D surge. Eu et al. [152] adapted three bioinspired algorithms (zigzag-surge, spiral-surge and spiral-zigzag-surge) to 3D by varying the tracker's altitude during the plume finding and re-acquisition stages. The plume height is initially determined by

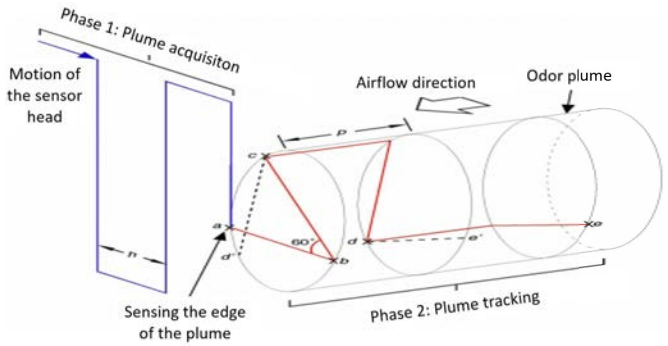


Figure 1.29: Schematic illustration of the 3D zig-zag algorithm (from [148]).

performing an oscillatory 3D movement and, then, a surge (upwind movement) at constant height is performed (Figure 1.30).

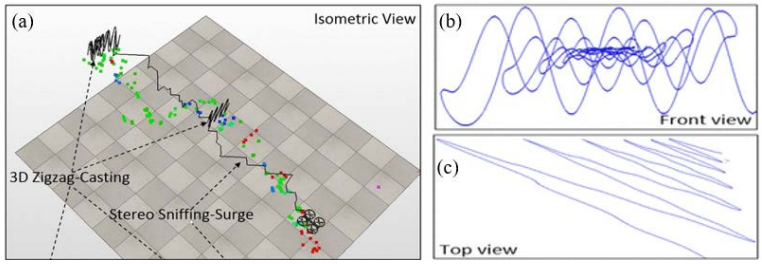


Figure 1.30: 3D zigzag-2D surge algorithm. (a) Isometric view; (b) Front view of the zigzag cast; (c) Top view of the zigzag cast (from [152]).

The authors experimentally validated the algorithms using a quadrotor equipped with four MOX sensors below the propellers. Experimental results in a small indoor arena with strong air flow show that spiral-zigzag-surge exhibits better overall performance as compared to the other two algorithms. The starting position has a strong influence on the time required to locate the source, which varied from 2 to 15 minutes.

Chemo-ortho-klino-kinesis (COKK). Rutkowski et al. [153] developed a 3D algorithm inspired by the upwind and counter-turning movements of moths tracking pheromone plumes. The motion of the UAV is decomposed into the tangential and perpendicular components to the wind vector. In the tangential direction, the UAV moves forward (i.e. upwind) or backward (i.e. downwind) based on binary chemical detections. In the perpendicular plane, the turn rate is controlled by the derivative of concentration (i.e. COKK). A decrease (resp. increase) in concentration increases (resp. decreases) the turn rate, and if the concentration remains constant the turn rate is gradually decreased. The algorithm was tested in a small simulation environment (0.6 x 1.5 m^2) using a plume model and constant uniform wind conditions. The results show that the tracker effectively reaches the

source and wanders around it (Figure 1.31). Validation in a real scenario may be required to confirm the utility of the proposed algorithm.

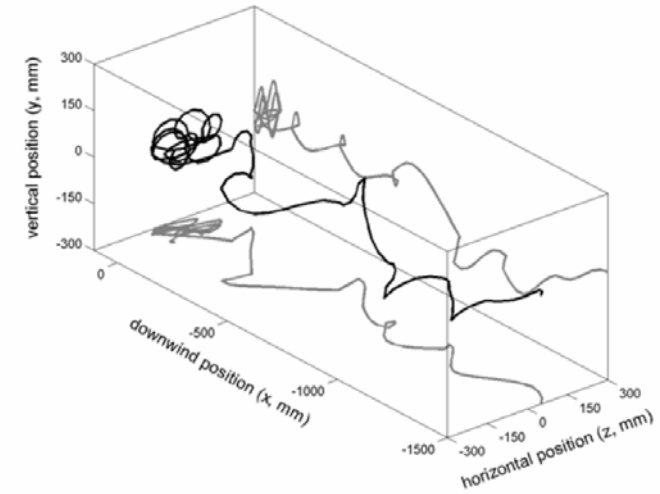


Figure 1.31: Representative flight track of the COKK algorithm. The gas source is at the origin of coordinates (from [153]).

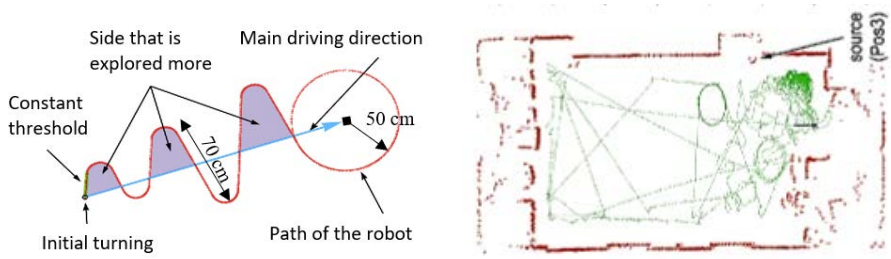
1.2.3 Weak turbulent fluid flow

The approaches described so far rely on the existence of a chemical plume, which is created in practice by forcing a strong airflow in the environment using fans. However, real environments in which the robot should be used are not supposed to be modified, which means that only poor assumptions about the airflow and the underlying functional form of the gas dispersion could be made. The algorithms described in this section rely exclusively on chemical sensors, i.e. they do not use anemometry, and do not make strong assumptions about the gas distribution.

Adapted silkworm moth

Reactive plume tracking algorithms require knowledge of the wind direction, which hinders their application to environments with weak air flow (e.g., indoors). Lilienthal et al. [154] adapted the silkworm moth algorithm [111] to environments with weak airflow by replacing the upwind surge with a straight movement at a random direction. This direction is reset after getting close to an obstacle. After detecting a gas patch, the robot starts a zigzag pattern oriented 65 degrees to the side at which the higher concentration was sensed (Figure 1.32(a))

Experiments carried out in a large indoor environment seem to indicate that the robot effectively wanders around the source location (Figure 1.32(b)), although the authors could not demonstrate statistically that this method was superior to a random walk.

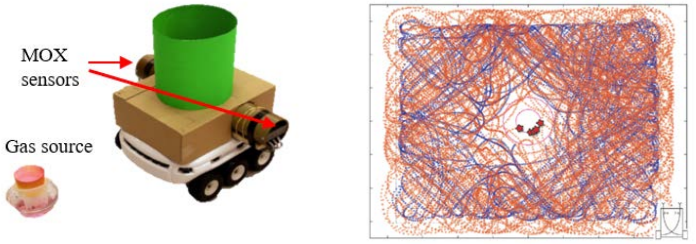


(a) (b)

Figure 1.32: Adapted silkworm moth algorithm. (a) Motion pattern that is executed by the robot in response to an odor detection; (b) Trajectories recorded in real experiments in a 15x5 m^2 indoor environment (from [154]).

Reverse Braitenberg vehicle

A reverse Braitenberg vehicle (Figure 1.33(a)) reactively avoids local concentration maxima and tends to wander around low concentration areas [30]. From visual analysis of the robot trajectories it is easy to determine regions with high concentrations, as the robot simply leaves them out (Figure 1.33(b)). This approach involves long exploration times and might trap the robot in local maxima, however it might be convenient when the released gas could damage the robot or the gas sensors.



(a) (b)

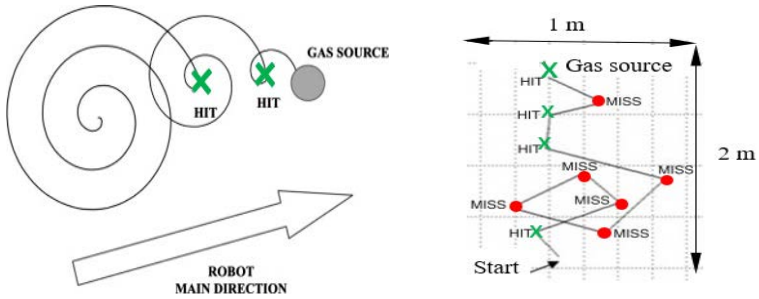
Figure 1.33: (a) Reverse Braitenberg robot and gas source; (b) Path of the robot during 3 hours in a 3.7 x 3 m^2 arena. The gas source is placed in the middle of the room (from [30]).

SPIRAL

Ferri et al. [122] proposed a bioinspired CSL method for weak air flow conditions, called SPIRAL. The algorithm uses a proximity index (PI), defined as a linear combination of the mean concentration (μ) and number of peaks (P) of the sensor signals during a measurement window ΔT

$$PI = K_\mu \mu + K_p P \tag{1.6}$$

where K_μ and K_p are the empirical weights for the mean and the peaks, respectively. The robot follows a spiral trajectory, stopping at predefined points for 30 s to compute the PI. If the current PI exceeds the last PI, a 'hit' is declared and the robot starts a new spiral; otherwise, it continues in the current one (Figure 1.34(a)). This behaviour is supposed to bring the robot closer to the source.



(a) (b)

Figure 1.34: SPIRAL algorithm. (a) Sketch of the navigation strategy (the green crosses represent odor hits); (b) Experimental trial in an indoor room without artificially induced air flow (Adapted from [122]).

SPIRAL was compared to a bioinspired (e. coli bacteria) and a random walk algorithm in a small room with weak airflow (Figure 1.33(b)). The maximum initial distance from the robot to the source was only 1.8 m. The mean time to reach the source (across 15 trials) was 6.5 min for SPIRAL, 7.5 min for the e. coli algorithm and 22.5 min for the random walk. Despite SPIRAL seems to be the most efficient of the three algorithms, the low exploration speed (1 m^2/min) might not be adequate for larger environments.

Instantaneous response

Lilienthal et al. [81] observed that in a weakly ventilated corridor the maximum response of a MOX sensor often corresponds to the approximate location of a gas source, if the sensor readings are acquired in motion (Figure 1.35). Such a correlation was never observed if the concentration measurements are collected with a stop-sense-go strategy. This behaviour, which was previously reported by Atema et al. [155] and confirmed by Farah and Duckett [156], is attributed to the long recovery time of MOX sensors. If a MOX sensor is exposed to two consecutive gas patches, the response to the second stimulus will occur when the sensor has not yet recovered from the first exposure. The overall response to the second patch will be higher than if the sensor was fully recovered from the first patch. Since the local density of gas patches is maximized near the source, the encounter rate with gas patches is higher if the robot is moving towards the source.

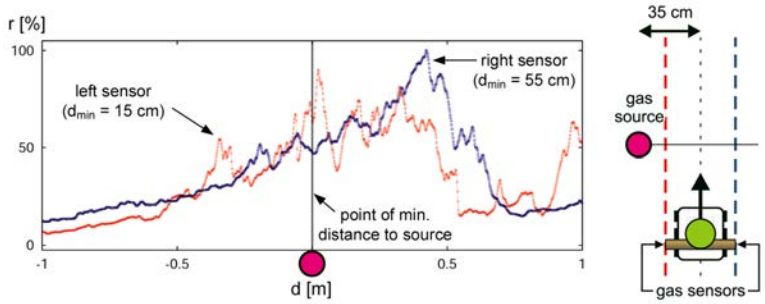


Figure 1.35: Normalized sensor response along the exploration of an unventilated indoor corridor. One MOX sensor is mounted on each side of the robot (from [157]).

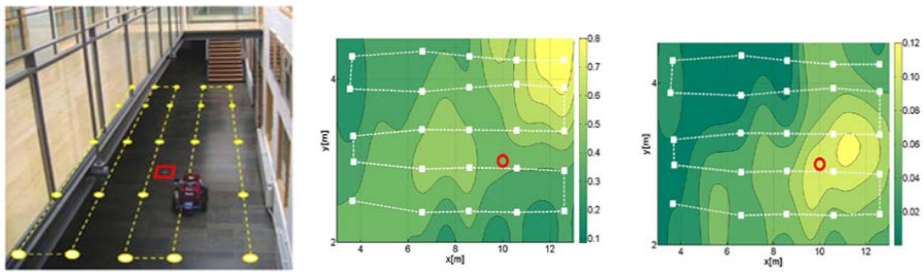
This suggests a straightforward CSL method that consists on driving the robot along a predefined path or in a random walk, collecting concentration measurements, and eventually declaring the source at the point of maximum concentration. This might only be appropriate for small exploration areas or fast robots, provided the source is reachable by the sensing unit. Rossi et al. [47] and Carrozo et al. [49] obtained a good estimate of source location using the instantaneous response in outdoor experiments using a micro drone.

Gas concentration gridmaps

Instantaneous concentration measurements taken at different locations and times can be integrated into a 2D or 3D statistical model (i.e. a map), representing the mean [82] (Figure 1.36(b)), the variance [18] (Figure 1.36(c)) or the maximum [158] of the concentration. It has been shown that, under certain conditions, the cell with the highest value in each of these maps can be used to estimate the source location. For example, the grid cell with maximum value in a mean gas distribution map is a good approximation to the true location of the gas source when the shape of the distribution is roughly circular with a strong central peak [82]. The variance maps usually lead to more accurate source location estimates than the mean maps in weakly ventilated indoor environments [18, 28].

To build a gas distribution map, the robot typically navigates the environment along a rectangular sweeping path with predefined measurement points where it stops to sample the gas concentration for a certain amount of time (typically 30 s) [18, 28, 159–161] (Figure 1.36(a)). Then, measurements are interpolated, for example using a Gaussian kernel to smooth the maps. This method assumes that the environmental conditions and the activity of the gas source are stable within the time frame of the mapping process, which might be not always true due to the long time required to map a large area using a stop-sense-go strategy. The map building process can be speed up by using multiple collaborative robots [29, 113, 162–164] or by adaptive sampling strategies [165].

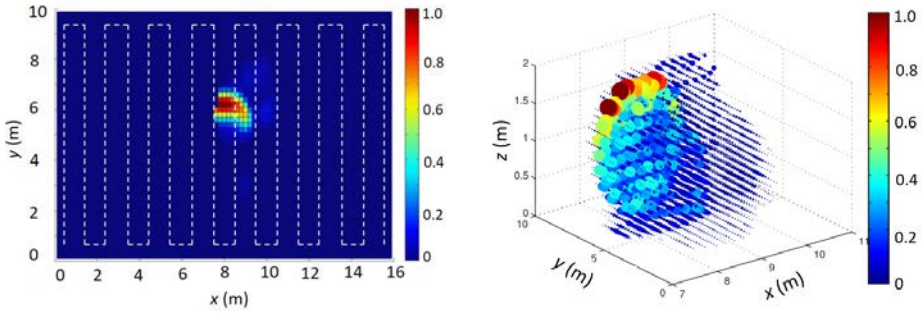
The suitability of gas concentration gridmaps map for GSL has been also demonstrated with aerial robots.mNeumann et al. [165] found that the accuracy in the source location estimated from a 2D gas distribution map might degrade in subsequent trials due to the effect of the rotors on the local gas distribution [165]. Luo et al. [166] built a 3D gas



(a) (b) (c)

Figure 1.36: Gas distribution mapping in an indoor corridor (14 x 2 m^2). (a) Predefined trajectory with sampling stops of 30 s. (b) Predictive mean map; (c) Predictive variance map (Adapted from [28]).

distribution map in a large outdoor environment (10 x 16 m^2) using a micro-UAV (800 g). A predefined flight path consisting of two 2D rectangular sweepings at different heights (0.3 and 1 m) was performed, without stopping at predefined locations for measuring. The cell with maximum value of the resulting mean map coincided approximately with the source location in the x-y plane (Figure 1.37(a)), but not in the z-axis (Figure 1.37(b)). The short exploration time (9 min) for the large area being considered (160 m^2) makes this algorithm appropriate for large exploration areas.



(a) (b)

Figure 1.37: 3D gas concentration gridmap algorithm. (a) Top view (the flight path is depicted as a white dashed line). The gas source is located at $(x,y,z) = (8,6,0)$ m. (b) 3D view (adapted from [166]).

1.2.4 Long range algorithms

To move the robot near the source, previous works considered using long-range sensors, such as CCD cameras [167–172], TDLAS [4, 173] or OGI cameras [19, 174, 175]. Using a CCD camera, odor source candidates can be detected from a long distance provided

that: (i) candidates are identifiable from visual features and (ii) the algorithm has been trained with multiple images of the odor source, taken from different perspectives. The works using TDLAS devices and OGI cameras were focused in demonstrating the ability of the instrument to detect a gas leak when the robot follows a pre-defined navigation path, rather than designing autonomous GSL algorithms. Neumann et al. [56] used a micro-UAV equipped with a TDLAS sensor to remotely detect a CH_4 leak outdoors. The UAV and the TDLAS were remotely controlled by a human operator that aimed the TDLAS towards the known location of the gas source, which was detectable at 25 m.

1.3 Gas sensors used in mobile robots

The key requirements of a gas sensor for a mobile robot application are fast response time, low detection limit, high power efficiency and low weight. The faster the robot can move, the faster the gas sensor should be to avoid spatial averaging over large areas. The weight and power consumption are critical in UAVs, due to the limited payload and available energy. A low detection limit is especially important when the search area is large or when the source is weak. In this section, we describe the gas sensors that have been used in mobile robots, which are just a subset of the available gas sensing technologies (see a comprehensive review by Nanto et al. [176]).

1.3.1 In-situ gas sensors

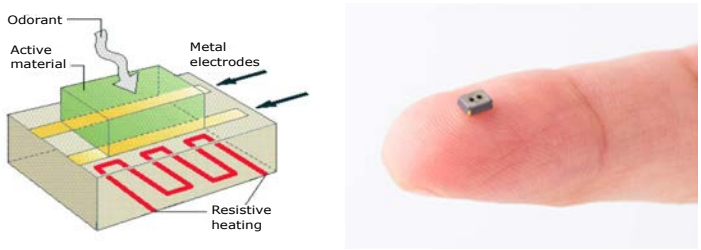
In-situ gas sensors are based on the change of the properties of a material (e.g., conductivity, resonant frequency, etc.) when a gas interacts with it. An electronic circuit inside the sensor (i.e., the transducer) converts the variations of the physical property into an electrical signal. These sensors are usually very simple and cheap, but they are usually not selective (different gases may produce a similar response) and the response can show a non-linear dependency with the concentration of the gas.

Metal oxide semiconductor (MOX)

MOX sensors are based on the property of certain semiconducting metal oxides of changing its electrical resistance upon exposure to gases at elevated working temperatures (typically in the range 150 to 500ºC). The sensing material (typically tin dioxide SnO_2) is deposited over a substrate provided with electrodes, for readout of the electrical resistance. The heater resistor is separated from the sensing layer and the electrodes by an electrical insulating layer. A diagram is shown in Figure 1.38(a).

MOX sensors are highly sensitive to many VOCs (LOD $\approx$ 1 part-per-million), easy to use, extremely cheap ($\approx$ 10 eur) and can last for decades. The first robots were equipped with metal oxide semiconductor (MOX) sensors due to the commercial availability, ease of use and low cost. This technology has been traditionally criticized by the lack of selectivity, cross-interferences, high-power consumption and slow response time. Whereas the low selectivity is still the main limitation of this technology, improvements in power

[15] `http://www.figarosensor.com/products/entry/tgs8100.html`



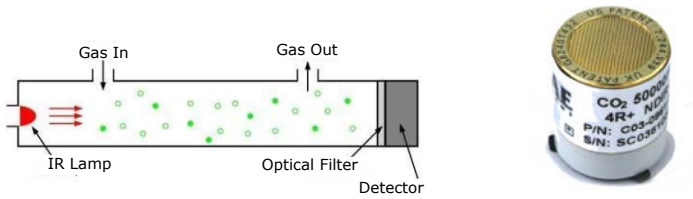
(a) (b)

Figure 1.38: (a) Diagram of the internal structure of a MOX sensor (Copyright 1998, IEEE); (b) Photo of a MEMS MOX sensor[14](Model: TGS8100, Figaro Engineering Inc).

consumption, size and response time are possible nowadays thanks to microelectromechanical systems (MEMS) technology. MEMS sensors (e.g., Figure 1.38(b)) incorporate a miniaturized sensing layer deposited over a micro hotplate, that leads to a smaller size (few mm^2) and increased power efficiency (15 mW). These features make them attractive for mobile robots, especially for platforms with limited payload such as nano-drones.

Nondispersive infrared (NDIR)

The main components of an NDIR sensor are an IR lamp, a sample chamber, a narrowband optical filter and an IR detector (Figure 1.39(a)). It is called nondispersive because the optical filter rejects all light except the wavelengths corresponding to the absorption band of the target gas. The difference between the amount of light emitted by the IR lamp and the light received by the detector is proportional to the concentration of the target gas in the sample chamber (Figure 1.39(b)). Whereas CO_2 can be selectively detected by NDIR sensors due to its unique absorption spectrum at 4.26 μm, CH_4 detection based on CH vibration is cross-sensitive to many hydrocarbons, e.g., ethanol (C_2H_5OH). Watai et al. [44] used NDIR technology on board of a UAV to measure atmospheric CO_2 variations at heights between 650 and 2000 m.

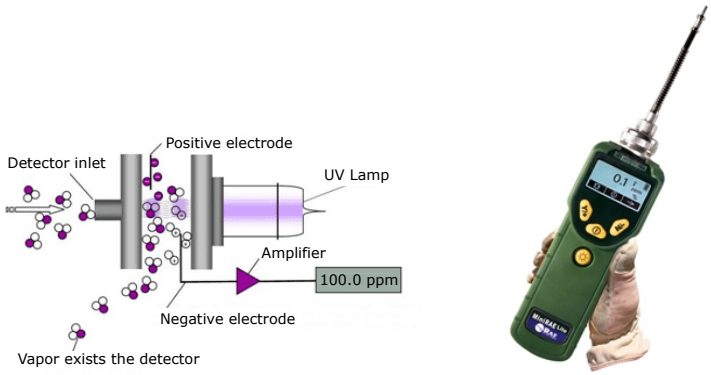


(a) (b)

Figure 1.39: (a) NDIR working principle[16]; (b) NDIR CO_2 sensor[17](Model: C03-0961-000, RAE Systems Inc).

Photo-ionization detector (PID)

A PID uses an ultraviolet lamp (typically 10.6 eV) to ionize the input sample. The electrical current produced by the ions reaching the detector provides an estimate of the gas concentration (Figure 1.40(a)). PIDs are fast (typically 1 s response time), highly sensitive (LOD of parts-per-billion) and are relatively unaffected by environmental factors such as humidity or temperature. Because of this, hand-held instruments (Figure 1.40(b)) are commonly used in industrial sites and military applications for monitoring toxic VOCs. The main problem of the instrument is the relatively high cost and the low specificity, since they respond to every compound with ionization energy below the energy of the lamp. PIDs have been sometimes used in mobile robots due to its fast response time and because they are often factory-calibrated [143, 161, 177].



(a) (b)

Figure 1.40: (a) Working principle of a PID[21]); (b) Photo of a PID[22](Model: miniRAE Lite, RAE Systems).

Electronic nose (e-nose)

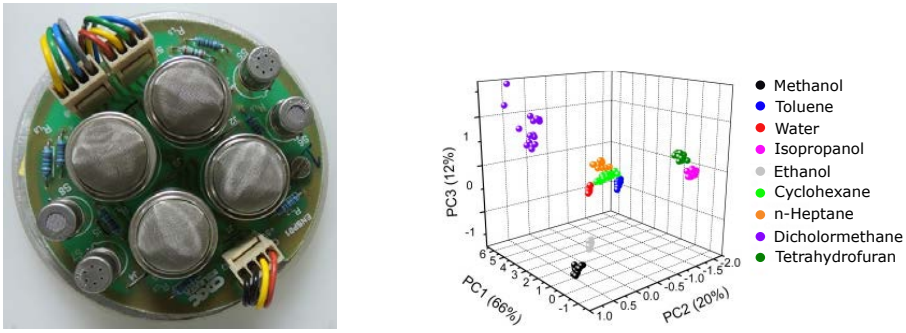
An electronic nose or sensor array is a hardware system that combines multiple partially selective sensors (e.g., MOX, EC, SAW, CPs), with different chemical properties, and an appropriate pattern recognition algorithm [178,179] (Figure 1.41(a)). The selectivity of the e-nose system is much higher than the selectivity of the individual sensors. Conventional statistical methods, such as principal components analysis (PCA) or partial least squares (PLS), and biologically inspired methodologies, such as artificial neural networks (ANN) are commonly used for gas quantification or discrimination [180] (Figure 1.41(b)).

[19]https://www.co2meter.com/blogs/news/6010192-how-does-an-ndir-co2-sensor-work

[20]https://www.gasdetectorshop.com/C03-0961-000-RAE-Carbon-dioxide-CO2-NDIR-Sensor-p/c03-0961-000.htm

[16]https://www.envieq.com/pid-photoionization-detector

[17]https://www.envieq.com/pid-minirae-lite



(a) (b)

Figure 1.41: Electronic-nose. (a) Hardware system composed of 8 MOX sensors[18]; (b) PCA of nine compounds measured with an e-nose, showing the good discrimination capabilities of the system (from [181].

The sensor array can be heterogeneous (i.e. contains different sensor technologies) or homogeneous (i.e. sensors of the same technology are exclusively used). In this latter case, the sensor units are selected from different models or they are operated in a different manner, for example by changing the working temperature in MOX sensors.

Electronic noses have been used in robotic experiments to discriminate among multiple gas sources for localization [126] and mapping [159, 182–188]

1.3.2 Remote optical sensors

Current robots are starting to incorporate remote optical sensors, because they are selective and provide a quick visualization of large areas that would require thousands of measurements using point-like sensors. These instruments are based on the characteristic molecular vibrational frequencies of certain chemical species when they are excited with infrared (IR) light (Figure 1.42). The two main types of remote sensors used in mobile robots are OGI cameras and TDLAS.

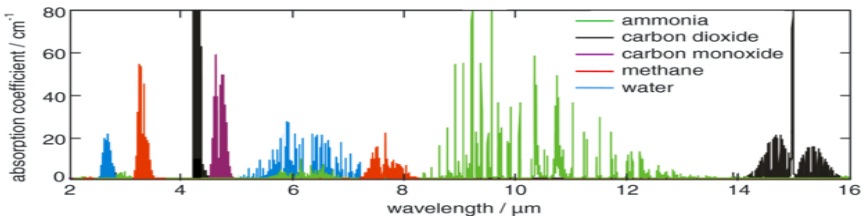


Figure 1.42: Absorption spectra for five gases in the mid-IR region of the spectrum (from [189].

[18] http://pib.nic.in/newsite/PrintRelease.aspx?relid=123091

Optical gas imaging (OGI)

Most gases are transparent to the naked eye (i.e. in the visible spectra), but there may be IR wavelengths where they are opaque due to optical absorption (Figure 1.42). An OGI camera (Figure 1.43(a)) is a thermal (IR) imaging camera that mounts a IR filter that only allows to pass radiation corresponding to the absorption band of the target gas [190]. To the "eyes" of the detector, the gas will block the radiation coming from objects in the background. In the recorded video stream, there will be a thermal contrast between the gas leak and the background, that can be enhanced with image processing algorithms (Figure 1.43(b)).

(a) (b)

Figure 1.43: (a) OGI camera[19](Model: GX320, FLIR Systems Inc.); (b) Visualization of a gas leak (colored in red) using an OGI camera[20].)

Mid wave (MW) and long wave (LW) OGI cameras are tuned for the 3.3 and 10.5 μm wavelength, respectively. A MW camera is most responsive to hydrocarbons (e.g., CH_4), thus it has a wide range of potential uses in the petrochemical industry. LW cameras are responsive to sulphur hexafluoride (SF_6), a gas commonly used as insulator in high voltage equipment in the electrical distribution industry. The 10.5 μm wavelength band is also suitable for the detection of leaks of refrigerant gases such as ammonia (NH_3).

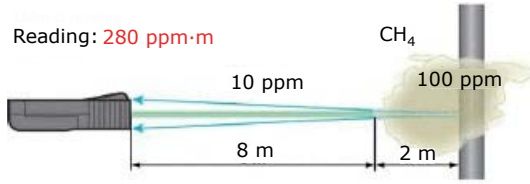
OGI cameras alone do not provide quantitative information about the gas, cannot detect small leaks, are expensive (100k euro) [191] and they can only be used for a single application. For example, a LW camera will not see CO or CO_2 leaks. Manual adjustment of the camera settings, such as thermal tuning or positioning might be necessary to improve the visibility of a gas leak. Operating the camera in automatic mode is possible but it may increase the chance of missing leaks, which restricts its application in autonomous mobile robots. Nonetheless, OGI cameras have been mounted in terrestrial robots (Figure 1.3(c)) and UAVs (Figure 1.7(d)) to remotely detect gas leaks.

Tunable diode laser absorption (TDLAS)

TDLAS is the most common laser-based absorption technique for quantification of gas mixtures [5]. A TDLAS instrument is basically composed of a tunable diode laser light

[21] https://www.flir.com/products/gf346
[22] http://www.seikey.com/page/133_FLIR+Gas+Detector+A6604

source and a photodiode. The laser-center frequency is locked onto to the absorption line of the target gas, a light beam is emitted towards the sample and the intensity of the received radiation is measured (Figure 1.44(a)). The measured intensity can be related to the integral of the concentration of the target gas across the light beam, using the Beer-Lambert law [192].



(a)

(b)

Figure 1.44: (a) Measurement principle of TDLAS. The device reports an integral concentration measurement of 280 ppm · m with the infra-red beams traveling a 10 m path, in which a background concentration level of 10 ppm and a CH_4 patch of 200 ppm are present (10 ppm $\times$ 8 m + 100 ppm $\times$ 2 m = 280 ppm · m) (from [56]). (b) Remote methane leak detector (RMLD), from Hermann Sewerin GmbH[23]

TDLAS is a highly selective and sensitive technique (sub-ppm detection limits), fast (response time < 1 s) and generally insensitive to cross-species interference. Over the past decade, TDLAS has evolved from specialized laboratory equipment to commercial portable devices, such as the remote methane leak detector (RMLD) [6] (Figure 1.44(b)). Recently, several authors have explored the use of TDLAS-based devices for mapping and localization of CH_4 in industrial facilities and landfill sites using terrestrial robots [28, 31, 193] and UAVs [56]. Portable TDLAS detectors are expensive, too heavy for platforms with limited payload and currently available only for detection of CH_4.

1.3.3 Calibration of in-situ gas sensors

The response of an in-situ gas sensor indicates how the properties of the sensing material vary when a gas interacts with it. This variation is normally non-linear with the concentration of the gas and is affected by environmental factors such as humidity or temperature. To convert the raw sensor response to absolute concentration units, a calibration curve

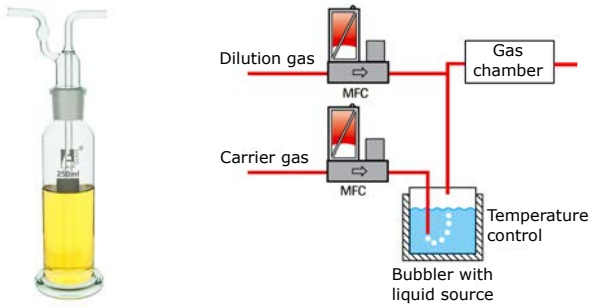
[23] https://www.sewerin.com/cms/en/our-products/detailview/rmld-is.html

must be obtained. A calibration curve models the relationship $y = f(x)$ between the sensor response y and the concentration x of the target gas. To obtain a calibration curve, the sensor is exposed to N known concentrations that span the working range of the device. Sensors that are not linear over the measurement range require a multi-point calibration (i.e. $N > 2$) to elucidate the best non-linear model that fits the data. A calibration system is composed of two stages: gas mixture generation and gas measurement.

Gas generation

Gas concentrations can be generated using static or dynamic systems [194]. In a static system, a known amount of analyte (either in gaseous or liquid phase) is added to a known volume of air in a container such as a plastic bag or a glass bottle. After some time, the mixture inside the container reaches equilibrium and can be delivered to the sensor under test. For a gaseous analyte, the concentration is the ratio between the volume of the analyte and the volume of air. For a liquid analyte, the concentration formula involves the saturated vapor pressure of the analyte and the temperature of the system, which must remain constant for the equation to produce reliable results. Two major disadvantages of static methods are that (1) the volume of the sample is limited by the volume of the container and (2) the waiting time to reach equilibrium is unknown.

A dynamic generation system provides for continuous variation of mixture composition, including humidification. The basic components of dynamic systems are pressurized air supply, mass flow controllers (MFCs) and the necessary tubing. The MFCs normally control three gas streams: a dry carrier gas, a humidified carrier gas and the test gas. The blending process and the resulting composition are regulated by varying the gas flow through the MFCs. The humidification of the gas mixture can be simply performed by the evaporation method using a glass bubbler (Drechsler bottles) (Figure 1.45(a)). Humidification is important in most sensing technologies because the baseline and sensitivity of the sensor are dependent on the humidity content of the sample.



(a) (b)

Figure 1.45: (a) Drechsler bottle[24]; (b) Vapor pressure generation system (adapted from Bronkhorst brochure[25]).

High purity gases such as O_2, CO, N_2 or CO_2 are available in pressurized cylinders. Generation of dynamic gas mixtures from a pressurized cylinder is a very convenient and accurate method (concentration uncertainty of a CO bottle can be as low as $\pm 0.5\%$ [195]). However, gas cylinders are expensive and not available for every test gas. VOCs such as ethanol, acetone, are usually found in liquid phase. The generation of test gas from liquid analytes can be made by alternative methods, such as syringe injection, diffusion, evaporation or permeation [196]. The most frequently used system for liquid analytes is evaporation. An evaporation system is composed of a container for the liquid analyte over which a small flow of a carrier gas is passed (Figure 1.45(a)). The resulting stream saturated with vapor of the liquid analyte can be diluted using a second air stream before being injected into the gas chamber (Figure 1.45(b)). By adjusting the temperature of the liquid and the flow rate of the diluting gas, the generated concentrations can cover a wide calibration range (e.g., 1- 10,000 ppm). The output concentration can be estimated analytically, from the vapor pressure of the liquid, or empirically, using a calibrated reference device. Stable concentrations can only be achieved if pressure, temperature and airflow within the liquid container are actively controlled. Minor fluctuations of these parameters can produce large deviations of the expected concentration. Permeation methods are preferred for preparing low concentrations of liquid analytes, i.e. at ppm to ppb levels, due to the high accuracy of the technique.

Gas measurement

The generated gas mixtures are delivered to the sensors under test, which are normally enclosed inside an air-tight gas chamber, where concentrations, environmental conditions, gas exposure times, airflow and, in general, all relevant constituents and interferences that might affect the sensor response are known and controlled. The conditions within the gas chamber (e.g., humidity, temperature or air flow) must be as similar as possible to the conditions of the target scenario. The gas exposure time must be long enough to capture the steady state values of the sensor response, which are then used to build a calibration model.

A radically different approach, where the sensors are calibrated in an open sampling system (i.e. without gas chamber), was proposed by Monroy et al. [197]. The sensors are placed at 50 cm from an odor blender which emits different concentrations of the target substance every 2 minutes. A PID is used as a reference device to measure the gas concentration reaching the sensors. This method simplifies the calibration process and brings the calibration conditions closer to the test conditions, by delivering the sample through turbulent air flow instead of through a controlled delivery system. The main disadvantage is that the calibration model has less predictive power than when it is obtained in a closed sampling system. The best model reported by the authors yielded average relative prediction errors between 20 and 30 %, depending on the sensor model.

[34] `http://www.betterequipped.co.uk/dreschel-gas-wash-bottle-prd8263p-2296`

[35] Liquid Delivery System with Vapour Control: `https://www.bronkhorst.com/products/vapour-flow`

Bad calibration practices in the robotics literature

Sensor calibration is something not usually done in the field of mobile robot olfaction (MRO) since calibration of gas sensors is time-consuming and requires expensive dedicated equipment. The sensor signals are often uncalibrated [14, 167, 170, 186, 198], scaled to unit range [154, 159, 199, 200] or simply normalized by the baseline response [100, 201]. Scaling and normalization are linear operations that cannot remove non-linearities of the original signals. If GDM experiments are performed with uncalibrated sensors, the resulting gas distribution maps represent the sensor response instead of the absolute concentration. Response maps are not easily interpretable due to the non-linear sensor response and are sensor-dependent; i.e., two response maps obtained with different gas sensors may show notable differences even if the underlying gas distribution is similar. Calibration is also important in GSL algorithms that combine the response of multiple gas sensors, for example to compute the chemical gradient [100, 201]. In the likely case that the different sensors have different sensitivity curves, the resulting gradient will be incorrectly computed.

Some authors assume a linear response within a certain concentration range [14, 202]; however, for most gas sensing technologies (e.g., MOX, gasFETs or thermoelectric sensors) this is only true within a narrow range of concentrations (e.g., 0–10 ppm) often exceeded by the typical concentrations obtained in real experiments. Other authors resort to calibration curves previously reported in the literature and apply it to their sensors [36, 158, 202]. Calibration functions are generally not transferable over long times or across sensor units, due to sensor instability and batch-to-batch variation in sensor characteristics, respectively. For a calibration transfer to be effective, at least some measurements need to be taken [203]. An extreme case is found in the work of Kowadlo et al. [202], where they used the calibration curve (and its parameters) from a paper presented 20 years before. In such a long time, the fabrication process of that sensor model might have changed drastically and therefore the sensor response can be completely different.

Among those works that indicate that calibration was performed, there is often lack of rigor in performing and describing the calibration procedure. In some cases, the calibration procedure is not described at all [118, 146] or it is vaguely described [122, 129, 140, 163, 166, 204, 205]. For example, neither the concentration values nor the way in which the concentrations were generated are described in [122, 204]. Luo et al. [166] reported the calibration concentrations but did not describe how they were generated. On top of that, they represented the gas distribution maps in unknown units in the range $[0, 1]$. Other authors describe the calibration method but do not report the calibration standards [129, 140, 163]. In none of the above works the uncertainty of the concentrations or the instrumental signals was reported.

The sensitivity of both the calibration method and the calibration curve to environmental parameters is often neglected in the literature. For example, calibration by the evaporation method with apparently lack of control over the relevant environmental parameters or without a reference device has been reported [205, 206]. It should be noted that the calibration curve has a domain of validity closely tied to the set of conditions covered by the calibration points. Departure from those conditions casts doubt on the validity of the calibration curve. For example, a calibration curve obtained at 30 %r.h. might not repre-

sent the sensor characteristics at 60 %r.h if the sensor is cross-sensitive to humidity (e.g., MOX sensors). In these cases, it is important to ensure that the calibration temperature and humidity are as similar as possible to those of the target application. Nonetheless, Ferri et al. [122] calibrated six MOX sensors simultaneously without measuring the humidity and temperature within the gas chamber. Due to the high power dissipated by multiple heaters, the conditions within the chamber may significantly vary with respect to the ambient conditions in which the device was later on operated.

1.3.4 Limit of detection (LOD)

The LOD is defined as the "smallest measure that can be detected with reasonable certainty for a given analytical procedure", by the International Union of Pure and Applied Chemistry (IUPAC) [207]. The LOD is a fundamental characterization of a gas sensor that allows to determine if a given reading is product of contact with the target gas or it is due to background noise. The background noise of MOX sensors is produced by the presence of other gases, changes in environmental conditions such as humidity or due to the intrinsic stability of the sensing layer. This creates random fluctuations of the blank sensor response (i.e. response without analyte) that overlap the distribution of responses corresponding to low concentrations of the analyte (Figure 1.46).

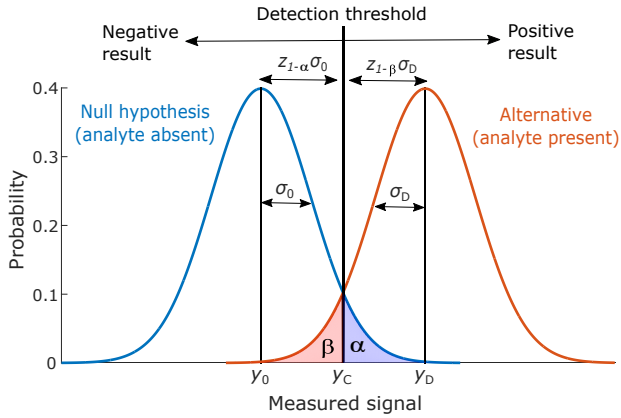


Figure 1.46: Schematic representation of the LOD in the signal domain. The blank response (blue curve) and the response at the LOD (red curve) are assumed to be normal distributions $N(y_0, \sigma_0)$ and $N(y_D, \sigma_D)$, respectively. The rest of the symbols are explained in the text.

The LOD is fundamental in odor robots because it determines the maximum distance at which the odor source can be detected, and is directly related to the threshold required by most GSL algorithms to detect the chemical plume. The estimation of the LOD in in-situ gas sensors is also relevant to applications outside mobile robotic olfaction, such as biomedicine or environmental monitoring. In the biomedical field, MOX sensors have been proposed to measure VOCs in expired breath, such as ammonia (NH_3) and carbon monoxide (CO), which might be related to respiratory and gastrointestinal dysfunc-

tion [208–210]. Trace concentrations of these VOCs higher than 1 ppm might indicate renal failure [209] and asthma [210], respectively. Thus, only sensors with LOD below this limit can be used for such applications.

Univariate LOD

When the sensor response is univariate, IUPAC derives the LOD in terms of statistical hypothesis testing with default values for the probabilities of false positives (α) and false negatives (β) both equal to 0.05. Assuming that measurements corresponding to blank samples and gas concentrations near the LOD are normally distributed with known variance (σ_0 and σ_D, respectively), the decision limit (y_C) and the detection limit (y_D) in the signal domain are given by

$$y_C = y_0 + z_{1-\alpha}\sigma_0 \tag{1.7}$$

$$y_D = y_C + z_{1-\beta}\sigma_D \tag{1.8}$$

where $z_{1-\alpha}$ and $z_{1-\beta}$ are the critical values of the normal distribution for significance level α and β, respectively (usually $\alpha = \beta = 0.05$). In many practical applications the true values σ_0 and σ_D are unknown, so a Student t-distribution is used to approximate the Gaussian curve. The critical values $z_{1-\alpha}$ and $z_{1-\beta}$ are replaced by the one-sided t-critical value $t_{1-\alpha,\nu}$ for the chosen confidence level (α) and degrees of freedom (ν). Similarly, the true standard deviations σ_0 and σ_D are replaced by the corresponding estimates s_0 and s_D, respectively. Because s_D corresponds to the (unknown) variability of the response at y_D (in fact, y_D is what we are trying to estimate), s_D is replaced by s_0 (assuming homoscedasticity). Finally, replacing the gross signal y by the net signal $y - y_0$ and recalling that we chose $\alpha = \beta$, Equations 1.7 and 1.8 can be simplified to

$$y_C = t_{1-\alpha,\nu}\sigma_0 \tag{1.9}$$

$$y_D = 2t_{1-\alpha,\nu}\sigma_0 \tag{1.10}$$

When the number of samples is large, $t_{1-\alpha,\nu} \to 1.65$ and y_C and y_D can be approximated by $1.65\sigma_0$ and $3.3\sigma_0$, respectively. This means that, even for the simplest formulas, at least the variability of the blanks needs to be determined. The LOD in the signal domain can be converted to concentration units by application of the calibration curve. The IUPAC only considers the linear calibration function

$$y = Ax \tag{1.11}$$

where y is the net sensor response, A is the real (unknown) sensitivity and x is the concentration. The LOD in the signal domain (assuming many samples) is found by inserting Equation 1.10 into Equation 1.11 and solving for x

$$x_D = \frac{3.3\sigma_0}{A} \tag{1.12}$$

To generate this simplified LOD formula, it is assumed that the response is linear with concentration, and the residuals of the linear fit are homoscedastic and normally distributed. Direct application of the above formulas when any of this assumption is violated may yield inconsistent LOD values. This is the case of MOX sensors because the response is highly non-linear and heteroscedastic (variance tends to be higher at low concentrations). Thus, reliable LOD estimation in MOX sensors is still an open problem.

Multivariate LOD

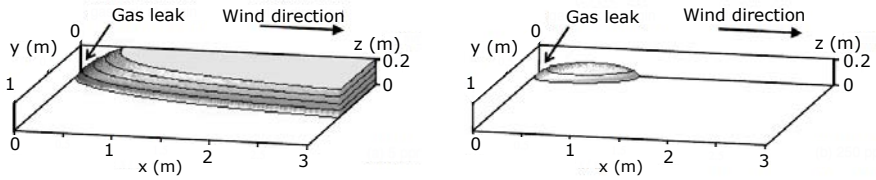
In some cases, the output of a sensing system is multivariate, i.e. each measurement contains several variables. This kind of first-order data arises naturally in analytical instruments such as the mass spectrometer or gas cromatography where a measurement is a spectra instead of a scalar value. However, multivariate data can also result from but can also happen in gas sensor arrays or temperature-modulated MOX sensors. The estimation of the LOD in multivariate calibration is not standardized (see a recent review by Olivieri [211]). The main issue is the lack of agreement between two competing methods: (i) analytically, by propagating the errors of the calibration concentrations and measured signals; (ii) empirically, by carrying out predictions on validation samples. The analytical method assumes that errors are independent and identically distributed (i.i.d.) and that the response is a contribution of additive signal and random noise. These assumptions are unrealistic in temperature-modulated MOX response patterns are highly non-linear and correlated. Furthermore, the analytical method requires that the variance of the errors is estimated a priori (e.g., from replicates), which is costly and sometimes impractical.

In the empirical approach, the prediction uncertainty is estimated through a univariate "surrogate" variable computed on the calibration or (ideally) validation samples. The univariate LOD formula (Equation 1.12) is then applied to the surrogate variable. Examples of surrogate variables are the net analyte signal (NAS) [212], the predicted concentrations [213] or the scores of the first PCA component [214]. Although the NAS approach is supported by the IUPAC [215], there are different approaches to compute the NAS which produce slightly different outcomes [216].

LOD in GSL algorithms

Most GSL algorithms require a detection threshold that is used to differentiate when a given sensor reading is product of contact with the gas plume and when it is due to background noise. Bioinspired algorithms use such a threshold in the plume finding phase and during plume tracking to detect when the plume has been lost. Algorithms based on 'odor hits' or 'odor events' [113, 142, 144] require such a threshold to binarize the sensor response. In the absence of the target gas (e.g., during gas finding), the threshold is exactly the definition of the limit of decision (y_C or x_C) of the sensor. The lower is such value (or equivalently the LOD) the longer the distance at which the source can be detected with a low probability of false alarm (Figure 1.47).

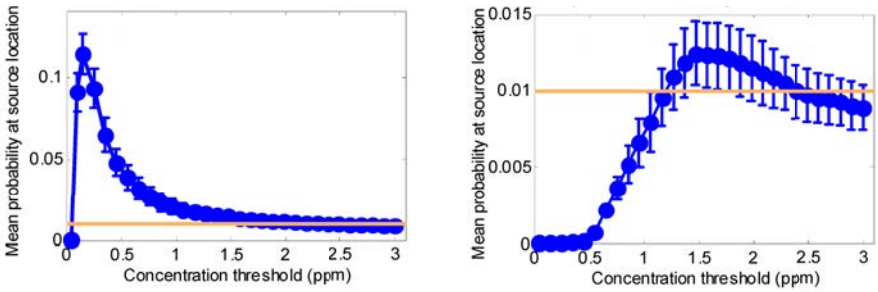
The background concentration is normally low far from the source and increases with proximity to the gas source (especially in indoor environments). Then, the detection threshold estimated at the beginning of the exploration (when the gas concentration was probably



(a) (b)

Figure 1.47: Distance at which a gas leak can be detected for an instrument with LOD of (a) 5 ppm and (b) 250 ppm. The values of 5 ppm and 250 ppm were chosen to be representative of a flame ionization detector (FID) and a pellistor sensor, respectively (adapted from [217]).

negiligible) might be no longer valid for detecting the plume close to the source. Pomareda et al. [143] found that the optimum threshold in Pang and Farrell's algorithm [142], i.e. the threshold that maximizes the mean probability at the true source location, increases with the level of background interference (Figure 1.48).



(a) (b)

Figure 1.48: Dependency of the mean probability at source location with the concentration threshold of Pang and Farrell's algorithm [142]. (a) Low background level; (b) High background level. The orange line shows the default equiprobable value ($\frac{1}{N_c}$) assigned to every cell. (from [143]).

Assuming that the background concentration varies slowly, Pomareda et al. proposed an adaptative background estimation algorithm that automatically subtracts the background from the instantaneous signals. Using the background-corrected signals, x_C can be used as a fixed threshold across the entire exploration area. Algorithms based on the derivative of the signal [121, 145, 218] can also implement a fixed threshold because the derivative is insensitive to background changes. Li et al. [219] and Lochmatter et al. [29] defined an adaptive threshold as a linear combination of the instantaneous reading and the average of the signal since the start of the experiment.

Despite of its importance, the detection threshold is often not estimated properly in the robotics literature. In some articles, neither the value nor how it should be computed is reported [114, 126, 129, 141, 142, 148]. Other works provide the value but do not explain how it was obtained [20, 154, 201, 202, 220]. Among those works that explain the computation,

some of them do not report the value [122] and others ignore the LOD definition and provide alternative definitions. Examples of reported thresholds are: $0.8 \times r_0$ [201], $4\sigma_0$ [113], $5\mu_0$ [144] or $\max(r_0)$ [40], where r_0 is the instantaneous response in clean air, μ_0 the mean blank response and σ_0 the standard deviation of r_0. It is clear that these definitions do not follow the LOD theory, which fixes the decision limit at $\mu_0 + 1.65\sigma_0$.

The rapid decay in chemical concentration with increasing distance from the source combined with incorrectly estimated detection thresholds results in a reduced area where the plume can effectively be detected (if threshold is too high) or a high number of false alarms (if threshold is too low). In outdoor applications, the search zone might involve several km^2 so it becomes very important to set the detection threshold very close to y_C. This has not been an issue in reported GSL experiments because the concentration was unrealistically high and the arena too small, i.e. the robot was too close to the source.

In some cases, the detection threshold is calculated when the instrument is calibrated in laboratory conditions and then it is (incorrectly) assumed that in future and field operation the threshold is still valid. However, in field operation the range of variation of interferences such as temperature and humidity can go beyond the calibration range or the instrument might have drifted. An illustrative example is found in Farrell et al. [40], where an AUV is used for CSL. The threshold of their binary-based algorithm was computed from a set of initial measurements taken in San Diego Bay in the absence of the chemical. They set the threshold as the maximum of the blank response, which represents a lower false positive rate than the 5 % established in the LOD definition and will probably reduce the distance at which the source can be detected. In addition to that, the field tests were conducted 10 months later in a different location (Duck, NC). By changing the scenario and the time, the background composition might have changed because of different water properties or the sensor might have drifted. Because the sensor equipped in the AUV is not mentioned, it is not possible to discuss about the potential cross-sensitivities or the temporal stability of such sensing technology. This example highlights the poor attention that is sometimes paid to the gas sensor system in the CSL literature.

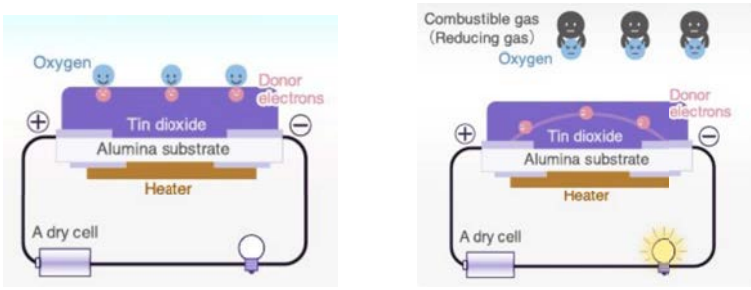
1.4 MOX gas sensors

MOX sensors deserve an independent section because they are the most common gas sensors in robotic applications. Besides of that, they have been used to solve problems in a number of fields: environmental monitoring [221–224], food quality control [225–229], biomedicine [230–232], safety and security [233–237], automotive [238–240], household applications [241, 242], energy [243], industrial processes [244], fire alarm systems [245], agriculture [246], law enforcement [247], air and space [248], water quality monitoring [249] and microbiology [250], among others.

1.4.1 Working principle

The working principle of a n-type MOX sensor is illustrated in Figure 1.52. In clean air (Figure 1.49(a)), free electrons inside the metal oxide are attracted towards air oxygen absorbed on the material surface, forming a potential barrier that prevents current flow

[24] http://www.figaro.co.jp/en/technicalinfo/principle/mos-type.html



(a) (b)

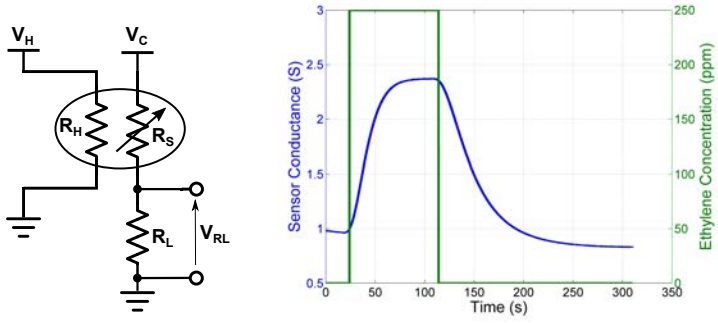
Figure 1.49: Operating principle of a n-type MOX sensor[36]. (a) In clean air; (b) Under gas exposure.

and increases the sensor resistance. In the presence of a reducing gas (Figure 1.49(b)), the density of adsorbed oxygen on the hot sensing surface decreases and the sensor resistance decreases.

1.4.2 Basic measuring circuit

The basic measuring circuit of an analog MOX sensor, depicted in Figure 1.50(a), requires two constant voltage inputs (V_H and V_C) and a load resistor (R_L) connected in series with the sensor output. The heater voltage (V_H) is applied to the heater resistor (R_H) to heat up the sensing element at a specific temperature which is optimal for sensing, whereas (V_C) allows measuring the voltage across R_L, which can be related to the sensor resistance (R_S) by

$$R_S = R_L \times \frac{V_C - V_{RL}}{V_{RL}} \tag{1.13}$$



(a) (b)

Figure 1.50: (a) Voltage divider; (b) Sensor response to a step of ethylene.

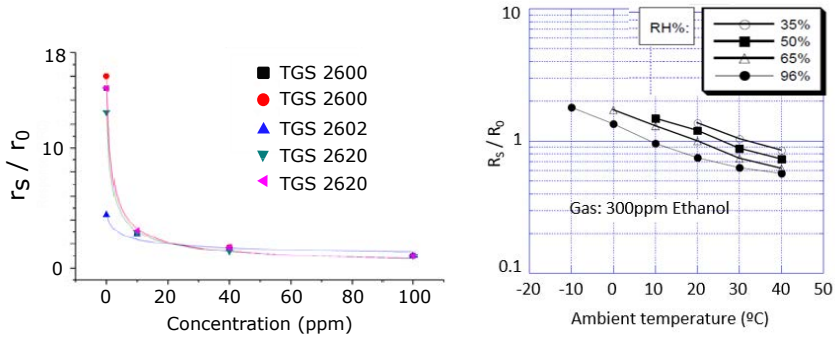
The load resistor might be tuned for the gas concentration range of the target application, and the output of the voltage divider can be amplified and filtered. Because of the huge dynamic range of R_S (it may vary through several decades), resistance-to-time conversion (RTC) methods [251] or amplifiers covering a measurement range of eight orders of magnitude [252] have been proposed to improve the resolution of the measurement.

1.4.3 Non-linear response

The response of a MOX sensor is highly non-linear. It usualy shows high sensitivity at low concentrations (e.g., < 10 ppm) and low sensitivity at mid-high concentrations (Figure 1.51(a)). In the ideal case of a single gas mixture, the Clifford-Tuma model [253] mathematically describes the response of a MOX sensor as a power law

$$r_s = r_{air} \cdot (1 + kc)^{-\beta} \tag{1.14}$$

where r_s ($k\Omega$) is the sensor resistance, r_{air} ($k\Omega$) the sensor resistance in clean air, c (ppm) the concentration of the target gas, k (ppm^{-1}) is a gas-dependent constant and β is the power law slope characteristic of the sensor.



(a) (b)

Figure 1.51: Calibration curve of five MOX sensors to ethylene (Models: TGS 2600, 2602 and 2620, Figaro Engineering Inc). r_0 is sensor resistance at 100 ppm of ethylene (from [254]); (b) Normalized sensor resistance (R_0 is the response in clean air) as a function of temperature and humidity[25].

The parameters k, β and r_{air} are typically found by an individual calibration procedure because they show high variance even for different units of the same sensor model. Once these parameters are determined, the measured sensor resistance can be converted to gas concentration by solving for c in Equation 1.14. The validity of Equation 1.14 was confirmed for CH_4, H_2, CO and H_2O, when only one of these constituents was variable and the rest were inexistent or constant.

[25] http://www.figarosensor.com/products/2620pdf.pdf

1.4.4 Selectivity and cross-sensitivity to environmental factors

In the complex case of a dynamic mixture of gases (i.e. several of them can vary simultaneously), the sensor response follows a power law that involves multiplicative interaction between the gases in the mixture and quadratic terms [253]

$$r_s = r_{air} \cdot \left(1 + k_{CH_4} c_{CH_4} + k_{H_2O} c_{H_2O} + k_{CO} c_{H_2O} c_{CO} + k_{CO} c_{H_2O} c_{CO}^2\right)^{-\beta} \quad (1.15)$$

where c_{CH_4}, c_{H_2O} and c_{CO} are the concentrations (in ppm) of CH_4, H_20 and CO, respectively, and k_{CH_4}, k_{H_2O} and k_{CO} are the corresponding constants (ppm^{-1}). According to Clifford and Tuma [253], competitive and synergistic interactions among gases and environmental factors (e.g., air humidity) could enhance or mask the target gas. Converting sensor resistance measurements into gas concentration is thus non-trivial, as device's response is non-linear with respect to the target gas and interfering substances (which are typically unknown). In other words, if the target gas is CH_4 it is not possible to solve Equation 1.15 for c_{CH_4} without knowing the values of c_{H_2O} and c_{CO}. This illustrates one of the biggest limitations of MOX sensors, the lack of selectivity.

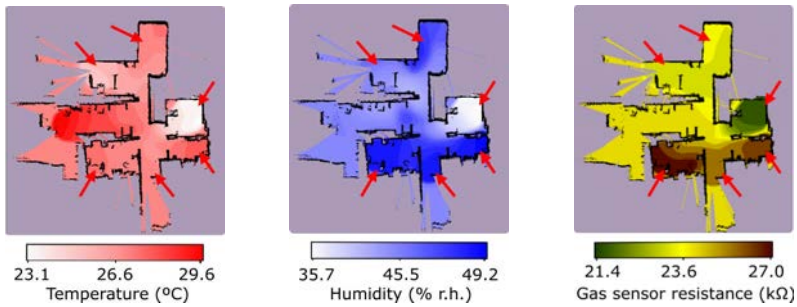
Humidity is probably the environmental factor that most influences the performance of MOX sensors, as water molecules adsorbed on the sensing surface decrease the area available for oxygen chemisorption reactions, decreasing the baseline and sensitivity [255]. In fact, metal oxides are commonly used in the fabrication of humidity sensors [256]. Prolonged exposure to extremely dry or wet atmospheres creates a "memory effect" in the sensor response [257]. Ambient temperature and barometric pressure may also affect the sensor response [258] (Figure 1.51(b)).

The lack of selectivity and cross-sensitivity to environmental factors may hinder gas sensing tasks with a mobile robot. Kamarudin et al. [259] studied the effect of temperature and humidity on gas distribution mapping, in a scenario in which the spatial distribution of temperature and humidity varied up to 6.5 ºC and 13.5 %r.h., respectively (Figure 1.27(a)). Despite no gas source was present, the predictive mean map (Figure 1.27(a)) seems to indicate the presence of a gas source in the east side of the map (lower resistance means higher concentration). The robot might be therefore misled towards areas with high humidity areas or high concentration of background constituents.

The selectivity of a MOX sensor can be improved by (1) technology and material improvements, (2) operating mode or (3) signal processing [260]. The recipe for the active sensor material can be designed to enhance the response to specific gases, for example adding noble metals such as platinum (Pt), gold (Au), palladium (Pd) or silver (Ag) [261]. The temperature of the semiconductor surface also has a strong impact into the selectivity. There is usually an optimum temperature for each target gas, but only limited selectivity can be achieved by working at a constant temperature [262]. In the example of Figure 1.53, the selectivity to acetone with respect to acetic acid is maximized at 250 ºC, but at this temperature there is also high interference to 2-butanone.

Temperature modulation

The simplest way to exploit the temperature-dependent sensor response to discriminate gases is to modulate the sensor temperature by applying a step, sinusoidal or ramp wave-



(a) (b) (c)

Figure 1.52: Effect of temperature and humidity on gas distribution mapping (adapted from [259]). (a) Mean temperature map; (b) Mean humidity map; (c) Predictive mean map (gas). Although no gas source is present, the predictive mean map is clearly influenced by the humidity map.

form to the heater voltage [263, 264] (Figure 1.54(a)). More complex modulations strategies such as pseudorandom binary sequences [265] or self-adapted modulations [266] have been also reported. The multivariate response patterns resulting from temperature modulation are normally fit by a multivariate calibration model, such as Principal Component Analysis (PCA) [267] or Partial Least Squares (PLS) [268] (Figure 1.54(b)). The latter is the de-facto method for regression purposes, as it can deal with high levels of collinearity in the response.

The length of the modulation cycle is a trade-off between discrimination power and effective sampling rate. The longer the heating cycle (up to a limit), the richer will be the sensor response and thus the higher the potential for gas discrimination [263]. However, long cycles reduce the effective sampling rate of the sensor because a measurement point is obtained after the heating cycle is completed. Temperature modulation was proposed for robotic gas sensing tasks by Purnamadjaja et al. [220], although the authors do not describe how they processed the multivariate response patterns. Temperature-modulated MOX sensors have been occasionally used within a sensor array [269, 270].

Signal processing

Signal processing and data fusion algorithms to determine and compensate changes of environmental conditions have been proposed in the literature [233, 271–275]. Sohn et al. [271] demonstrated that the humidity level in clean air samples can be accurately predicted by fitting a PLS model with the response of a MOX sensor array. However, these preliminary experiments did not prove the ability of the proposed model to compensate humidity fluctuations in samples with concentration higher than zero. Hossein-Babaei et al. [272] trained an artificial neural network (ANN) using sensor response patterns recorded at ambient temperature and relative humidity ranging from 23 to 43 °C and 10 to 80 %r.h., respectively. They were able to reduce the average relative prediction error to 3% in independent test samples of methanol in the range 0-2000 ppm (Figure 1.55).

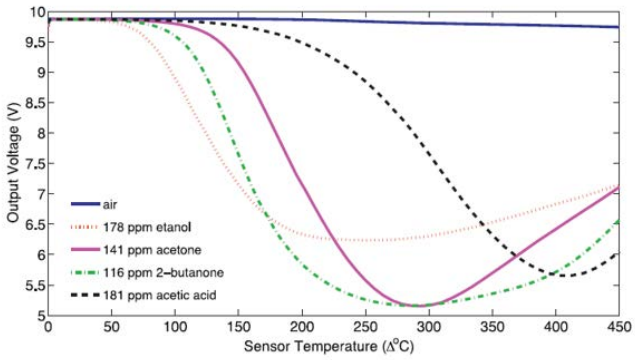


Figure 1.53: Dependence between the response of a MOX sensor and the temperature of the sensing layer. The selectivity to certain compounds can be increased by tuning the operating temperature (from [262]).

Masson et al. [274] fitted a second order polynomial to the sensor response under CO concentrations between 0 and 2.8 ppm, and changes in temperature from 0 to 35 °C. They obtained moderate prediction errors of 0.22 to 0.48 ppm. Piedrahita et al. [273] proposed a linear calibration model which includes independent terms for absolute humidity and temperature. Huerta et al. [275] proposed an energy band model to correct for humidity and temperature fluctuations in clean air conditions. The efficacy of the latter two compensation methods cannot be assessed because the results were not benchmarked against non-compensated data.

Sensor manufacturers on the other hand implement compensations of temperature and humidity via hardware, either using a thermistor-based analog circuit (as described in a technical document for the Figaro TGS 203 [26]) or by digital compensation using an external temperature/humidity sensor [252].

1.4.5 Stability

MOX sensors present short- and long-term instability, which is defined by Ionescu et al. [257] as "a variation of the sensor response when it is exposed to the same analytes under identical conditions". The short-term instability comprises two transient behaviours, named the "burn-in" and "run-in" phases by AMS manufacturer [27], which occur immediately after powering up the device.

Burn-in phase

Right after the first power up, the sensor performs a special "burn-in" behavior, characterized by a drift in the sensor response which might take up to one week (Figure 1.56(a)). It

[26] `http://pdf.dzsc.com/88888/2008617161056961.pdf`

[27] AS-MLV-P2 datasheet: `https://ams.com/documents/20143/36005/AS-MLV-P2_DS000359_1-00.pdf`

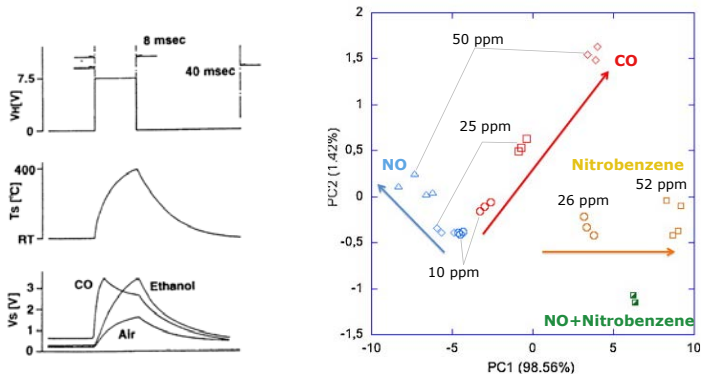


Figure 1.54: (a) Temperature modulation using a step heating waveform: (top) Heater voltage; (middle) Sensor temperature; (bottom) Multivariate response patterns (from [263]); (b) PCA analysis of a temperature-modulated MOX sensor, showing good discriminative power between four compounds. The arrow indicates the direction of increasing concentration (from [266]).

is due to manufacturing issues and can be solved by continuously powering the sensor for several days. Once the burn-in phase is completed this behaviour will not show again.

Run-in phase

Immediately after any power up across the lifetime of the device, the sensor enters into an unstable state characterized by a steady increase in the sensor resistance (Figure 1.56(b)), also referred to as "initial action" by Figaro manufacturer (in the application note of the TGS 2610 sensor [28]). The standard duration of this transient behaviour depends on the sensor model and how long the sensor was unenergized; the longer the sensor was unpowered, the longer the run-in phase takes. This means that a reliable and stable measurement can only be obtained after waiting for that length of time, which represents a practical issue for devices intermittently operated or with limited energy budget. Immediate stable measurements can be obtained if the MOX sensor is continuously powered, however the power consumption in this case might be considered too high for certain battery-operated application such as robots [276] and, especially, micro aerial vehicles.

Even after powering the sensor long enough to get a stable response, replicate measurements under similar conditions can be noisy. An important component of this drift may come from the presence (even at low concentration) of traces of unknown and time-varying compounds. Another part of this apparently random noise may be caused by uncontrolled variations of the sensor temperature, due to (1) the exothermic nature of chemical reactions between the gas and the metal oxide, (2) Joule heating of the sensor during the measurement or (3) changes in the air flow [258]. In the technical information sheet [28], Figaro manufacturer indicates that air flow cools down the sensor surface and thus influences

[28] Application note Figaro TGS 2610: `http://www.figarosensor.com/products/2610app.pdf`

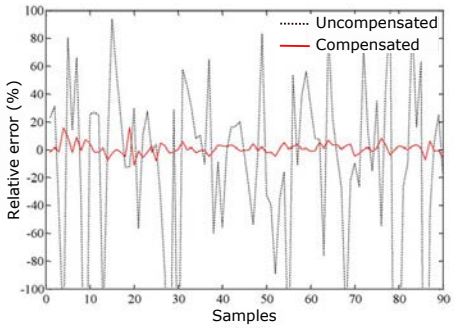
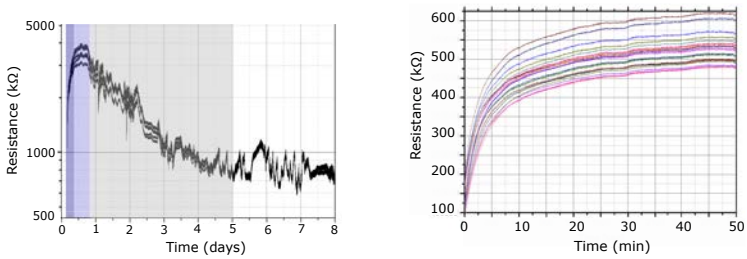


Figure 1.55: Comparison between relative error for humidity-compensated and uncompensated signals (from [272]).



(a) (b)

Figure 1.56: Instability of MOX sensors. (a) Burn-in and (b) run-in phases of several units of the AMS AS-MLV-P2 sensor (adapted from the AS-MLV-P2 datasheet[27]).

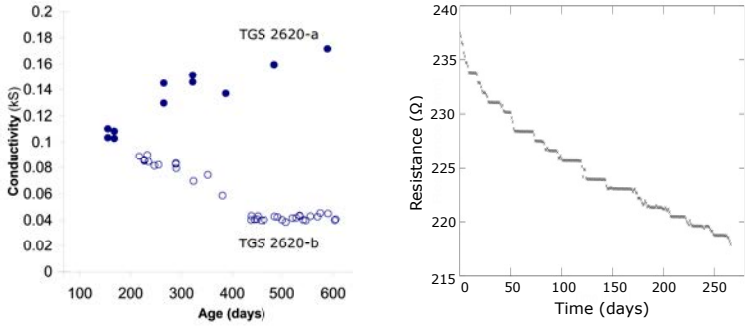
its response. The heater resistor's TCR is also susceptible to changes in environmental conditions. Without appropriate temperature control, deviations in the operating temperature can shift the baseline and sensitivity. Stable operation independent of the surrounding temperature can be achieved by specialized circuitry to drive the heater voltage [29], or by integrating a temperature sensor on the hotplate to actively control its operating temperature [252, 277].

Long-term drift

Long-term systematic drift of the sensor response can be noticeable in a time-scale of several days (Figure 1.57(a)). It may be caused by sensor aging, prolonged exposure to highly humid environments, thermo-mechanical degradation or poisoning, among other factors [218]. Drift in heater resistance over time was observed for a MiCS-5525 sensor [274] (Figure 1.57(b)). Long-term drift compensation can be achieved by modifica-

[29] Alphasense Application Note AAN 601-02: `http://www.alphasense.com/WEB1213/wp-content/uploads/2017/07/AAN-601-02.pdf`

tions in sensor technology and design, changes in the operating mode or signal processing methods [260, 274, 278–283]



(a) (b)

Figure 1.57: (a) Long-term drift of two units of the Figaro TGS 2620 sensor, continuously exposed to the same conditions within a gas chamber (from [284]); (b) Drift in heater resistance of a SGX MiCS 5525 sensor over a period of nearly nine months. Ambient temperature and humidity were both stable (from [274]).

1.4.6 Response time

The response of MOX sensors is characterized by a fast response time (typically 1-3 s) upon exposure to the gas, followed by a slow recovery of the response after gas is removed (typically 20-30 s). A large component of the slow recovery time is due to the housing of the sensor, which acts as a small gas chamber where the gas accumulates before reaching the sensing layer, producing a low-pass filtering of the response [285] (Figure 1.58). MEMS sensors are thus faster than conventional sensors because of the reduced size of the housing. Removing the sensor housing notably improves the response time [98, 285, 286] but the sensor might become less stable and more prone to damage. Martinez et al. [285] showed that a MOX sensor without housing can resolve concentration fluctuations in the seconds scale (1 Hz), which represents an improvement of one order of magnitude with respect to the bandwidth of a sensor with housing (0.1 Hz). According to Martinez et al., the response time of a naked MOX sensor is still one order of magnitude worse than the bandwidth of the moth's antenna (10 Hz).

The response time can be improved either by novel hardware designs [204, 286] or by signal processing [121, 287–290]. Batog et al. [286] developed a custom hardware piece where a small DC fan creates a downwards air flow that bounces on the ground and passes through the perforated housing of a MOX sensor (Figure 1.59(a)). Experimental results using a mobile robot for chemical trail following indicate that this device, which was specifically developed for this task, improves the detection of the boundaries of the trail (Figure 1.59(b)).

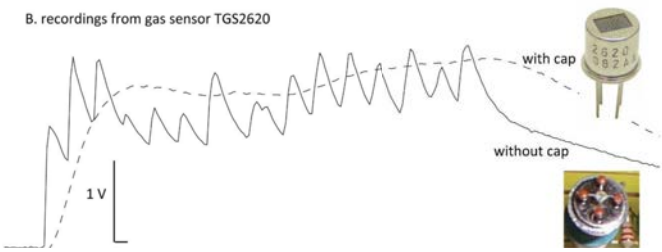
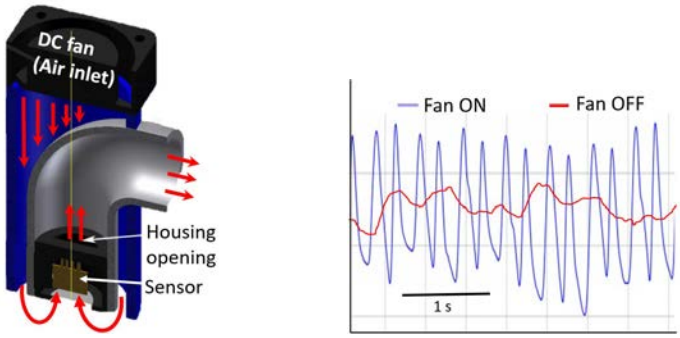


Figure 1.58: Recordings from gas sensor TGS2620 in response to fluctuations in concentration of ethanol, with (dashed line) and without (solid line) the cap (from [285]).

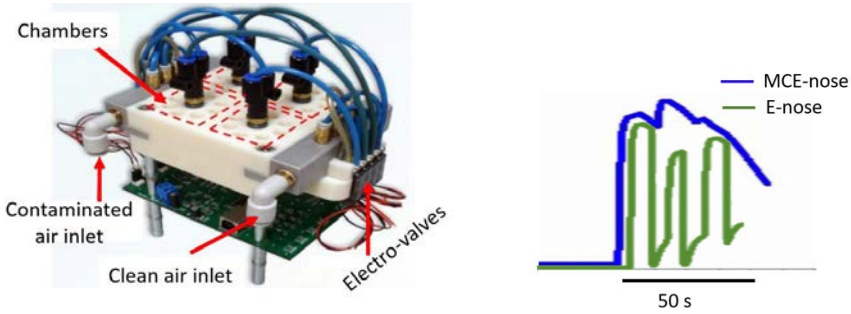


(a) (b)

Figure 1.59: (a) Custom enclosure that integrates a DC fan and a MOX sensor with a perforated housing; (b) Comparison between the sensor response when the fan is turned ON or OFF, during multiple crossings of a chemical trail (adapted from [286]).

Gonzalez-Jimenez et al. [204] designed a multiple chamber e-nose (MCE-nose) in which identical MOX sensors are accommodated in four separate chambers (Figure 1.60(a)). When the system detects that the active sensor is in decay phase, it delegates the sensing task to one of the sensors in a "clean" chamber. The overall output signal is the concatenation of rise phases of clean MOX sensors. The system clearly improves the respone time of a conventional e-nose (Figure 1.60(b)), but the use of multiple chambers with electrovalves increases the cost, size, weight and power consumption, limiting the applicability of the device to platforms with limited payload.

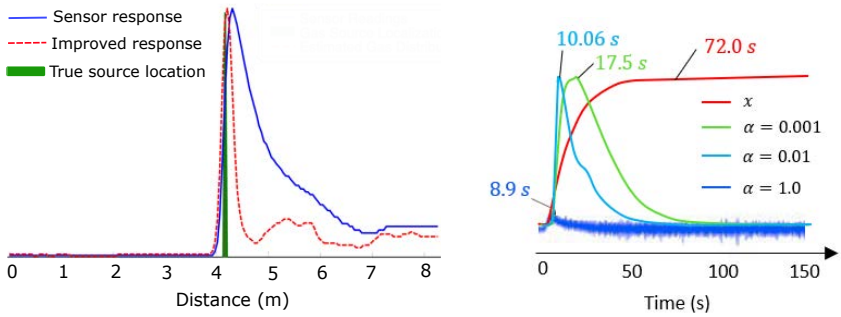
Regarding signal processing approaches, inverse dynamical models have been proposed to retrieve the (fast) excitation signal from the slow sensor response [287, 289, 290]. Marco et al. [289] demonstrated that ANNs and kernel estimation methods were superior than functional expansions. Pardo et al. [290] refined the previous work by including the dynamics of the sampling chamber into the model and concluded that, among different kernels, Wiener kernels lead to smaller errors and faster estimation. Monroy et al. [287] modelled the MOX response using two exponentials with different time constants for rise



(a) (b)

Figure 1.60: Multi chamber e-nose (MCE-nose). (a) Picture of the hardware system; (b) Response of a conventional e-nose (blue line) and the MCE-nose (green line) mounted on a mobile robot, as it passes by three consecutive gas sources separated 2 m from each other (adapted from [204]).

and recovery (Figure 1.61(a)). The practical limitation of this approach is that it requires a complex setup to calibrate the time constants for each substance and sensor.



(a) (b)

Figure 1.61: Improvement of response time by signal processing. (a) Inverse modelling. The raw sensor response as a robot passes by a chemical source is depicted as a blue line, the improved response as a red line and ground truth as a green line (adapted from [287]); (b) Smoothed derivative (adapted from [291]).

Fonollosa et al. [288] used neural networks to compensate the slow response time of a MOX sensor array. Experimental results in a gas chamber with mixtures of CH4, CO and Ethylene yielded relative prediction errors of 4.5-8.5 %. It is unclear the suitability of the proposed approach for open sampling systems (OSS) where the chemicals are transported by turbulence rather than by a controlled flow delivery system.

Other authors have used the derivative of the response for improving the response time. Muezzinoglu et. al [291] first noticed that the height of the peak of the smoothed derivative (Equation 1.16) can anticipate the steady-state value of the response (Figure 1.61(b)).

$$y_t = (1-\alpha)\cdot y_{t-1} + \alpha\cdot(x_t - x_{t-1}) \tag{1.16}$$

where α controls the amount of smoothing. As α increases, the closer is the y_t to the derivative and the earlier the peak appears in time. However, being close to the derivative means also higher noise. Thus, α governs a trade-off between response time and signal to noise ratio (SNR).

Schmuker et al. [121] used the smoothed derivative for extracting short-scale features of turbulent plumes (Figure 1.62). In wind tunnel experiments, the authors found that the frequency of rising edges of y_t (the so-called 'bouts'), after filtering out low-amplitude bouts produced by noise, is strongly correlated to the distance to a gas source: the higher the bout frequency, the closer the sensor to the gas source. Comapared to previous signal processing methods, the advantage of this approach is that no complex calibration procedures are required. Nonetheless, as of now there is no systematic approach to determine the smoothing factor α and the threshold to discard low-amplitude bouts, and both parameters may have a strong impact on the results.

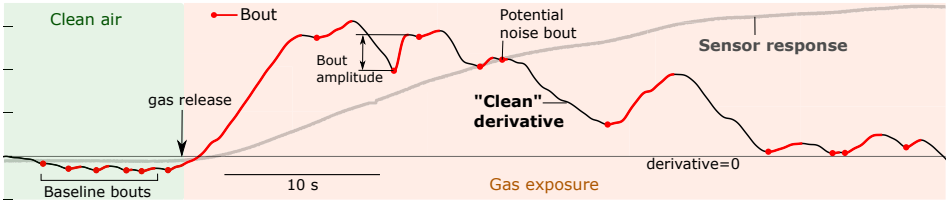


Figure 1.62: Schematic representation of the raw sensor response, its "clean" derivative and the detected bouts when the sensor is exposed to a chemical plume. The figure illustrates that the dynamics of the derivative are much faster than those of the raw response (Source: unpublished results).

1.4.7 Surface mount devices (SMD)

There is a trend towards small SMD housings that integrate the sensing elements with an analog (e.g., SGX MiCS family[30]) or digital interface (e.g., Bosch BME680[31], Sensirion SGP-30[32], AMS CCS811[33]). The digital interface simplifies the integration of sensors into microprocessor applications by accepting a digital power voltage (e.g., 3.3 or 5 V) and providing a digital output signal which can be directly used without further processing from the user. Internal algorithms in digital sensors can provide baseline and humidity compensation (e.g., using the signal from an external humidity sensor) and convert raw signals into calibrated outputs in concentration units [252]. The main disadvantage of using digital sensors is that the sampling frequency is usually limited by the chip, and there is no access to the raw output signal. Some SMD sensors combine two (e.g., MiCS-4514) or three (e.g., MiCS-6814 or Sensirion SGP-30) sensing layers in the same package, each of those tuned

for the detection of specific gases. By fusing information from different sensors they can provide additional outputs such as the indoor air quality (IAQ) index [252].

1.5 Gas flow visualization

Our limited visual experience about how a released gas disperses in the environment (especially indoors) impairs our ability to develop efficient GSL strategies. Much of the research on GSL has relied on the assumption that gases heavier than air, such as ethanol, will flow along ground level [14, 129, 184]. However, simply using the density of the pure substance as the only parameter governing gas dispersion is dangerous, as it is known that, for example, in indoor environments thermal sources or recirculating air currents from forced ventilation systems may have a stronger impact into gas dispersion than the density of the pure gas [292].

Researchers performing GSL and GDM experiments are often interested in determining whether the gas will reach a stable spatial distribution and how quickly will that happen. In most reported experiments, a small amount of time (few minutes) is waited after the gas is released prior to start measuring. This time is normally estimated without knowledge about how the selected gas disperses in the environment. For example, Purnamadjaja et al. [119] assumed that 6 min were sufficient to achieve a stable spatial distribution of ethanol in an indoor environment, whereas Ferri et al. [122] waited for 15 min. Lytridis et al. [293] claimed instead that dispersion of heavy alcohols is a much slower process which may take up to 25 min to be detectable at a distance of 1 m from the point source. If experiments are carried out when the gas distribution is not stable, the results will be dependent on the specific time frame in which the experiments were performed. Lilienthal et al. [294] pointed out that their experimental findings across different trials in an indoor scenario were linked to the time since the source was uncovered, because of the non-stationary nature of indoor gas dispersion.

Despite CFD simulations could in theory shed some light about dispersion of a gas in an environment, measuring the temporal and spatial variations of gas concentrations is the most reliable characterization of gas dispersion, as it does not rely on assumptions nor initial conditions. Such experimental measurements are also necessary for testing predictions from gas dispersion models and CFD simulations. A categorization of the available experimental techniques is possible by distinguishing the operating principle: visualization by tracer particles, optical remote sensing or chemical sensors.

1.5.1 Particle tracer methods

The fluid in which a gas is supposed to be released can be seeded with small tracer particles (e.g., dust, smoke, dye) to make the flow pattern visible. A qualitative analysis of such plumes may be done by a human observer based on video-recorded and digitized

[30] https://www.sgxsensortech.com/sensor-selector/

[31] https://www.bosch-sensortec.com/bst/products/all_products/bme680

[32] https://www.sensirion.com/en/environmental-sensors/gas-sensors/multi-pixel-gas-sensors

[33] https://ams.com/ccs811

images [25, 112, 295, 296]. Quantitive analysis of individual particles can be achieved by particle image velocimetry (PIV) [297], particle tracking velocimetry (PTV) [298] or laser-induced fluorescence (LIF) [24, 299, 300]. The position and velocity of the particles are recorded with high-resolution digital cameras and then shown in the form of a vector plot (Figure 1.63).

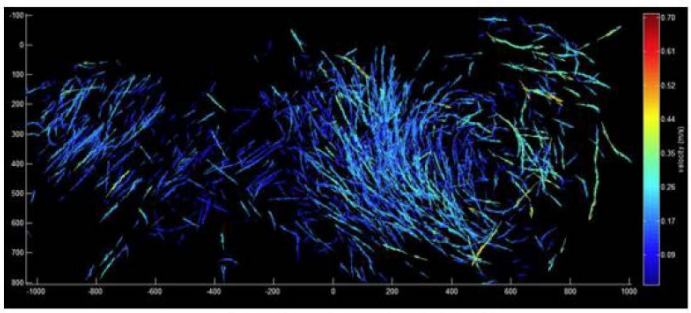


Figure 1.63: Experimental velocity field measured in an aircraft cabin using 3D-PTV (from [301]).

Tracer methods can provide visualizations with very high temporal and spatial resolution (3D), provided that the environment offers sufficiently ample optical access and there are no elements blocking the line of sight. Due to camera resolution and image requirements for accurate particle tracking, only a small fraction of the experimental area can be interrogated. Particle seeding assumes that the motion of the tracer and the target gas are identical, which is not necessarily true. It also requires a complex calibration procedure of the camera system and a depth analysis of the carrier gas in terms of viscosity, surface tension, velocity, etc. The high cost of these techniques together with the demerits summarized above restrict the practical application considerably [302].

1.5.2 Optical remote sensing (ORS)

Optical remote sensing (ORS) combined with computed tomography (CT) is an old technique to visualize 2D gas distributions over room-sized areas. ORS instruments such as the open-path Fourier transform infrared spectrometer (OP-FTIR) [303] or the TDLAS [5] are first used to accurately measure the total number of gas molecules in a beam of radiation. The path-integrated measurements are then converted by means of CT into a representation of the concentration distribution in the sampling plane [304]. This approach has been used to build 2D concentration maps of sulfur hexafluoride (SF_6) or CH_4 (in an indoor [305, 306] and outdoor areas [307]. Apart from the high cost of ORS instruments, mirrors and retroreflectors should be carefully placed over the test environment to avoid unexpected optical effects that would otherwise lead to errors in the measured path-integral concentrations [305].

1.5.3 Hyperspectral imaging

Hyperspectral cameras can be considered a general version of an OGI camera. Instead of acquiring images in a narrow wavelength band, they acquire spectrally resolved images [308]. A hyperspectral image contains hundreds of channels per pixel (a red-green-blue (RGB) image contains three channels per pixel), each corresponding to a narrow wavelength region of the IR spectrum. These extra channels allow for the visualization and discrimination of gas plumes without prior knowledge of the sample. Hyperspectral imaging has been mostly used in outdoor environments for identification of atmospheric pollutants and chemical warfare agents [309, 310] (Figure 1.64).

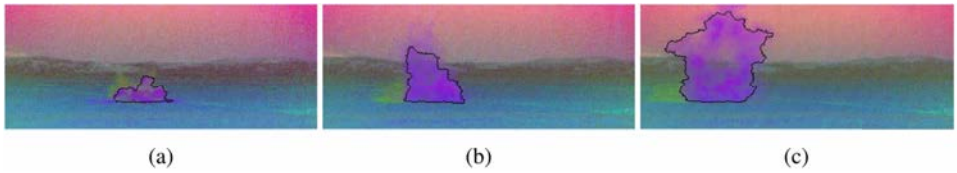


Figure 1.64: Gas plume detection in hyperspectral video (from [308], © 2014 IEEE).

The segmentation of hyperspectral images is still a research topic, due to the interference produced by atmospheric constituents and other objects in the sensor field of view [311]. Large data storage capacities and fast processors are needed for analyzing hyperspectral data. These factors together with the initial cost of the camera greatly increase the cost of the full system.

1.5.4 Chemical sensors

An inexpensive solution for investigating gas dispersion patterns without adding tracers is to use one or more chemical sensors to perform spatially distributed gas concentration measurements. Unlike optical methods, chemical sensors can only measure gas concentration in a very small area around the sensor surface (≈ 0.5 cm^2). This means that either a single sensor is moved to different locations or a grid format of static sensors is used. Metal-oxide semiconductor (MOX) sensors are typically used in both cases, due to their low cost and ease of deployment. Nonetheless, in the case of single-sensor mapping more expensive devices such as PIDs have been ocassionally used [312]

If the gas distribution is expected to be time-invariant, as in controlled environments with constant airflow and gas emission rate, the most straightforward approach to map the gas distribution is to use a single sensor to sequentially measure gas concentration in different locations [23, 100, 112, 119, 137, 148, 158, 312]. The environment is first discretized into a uniform [23, 112, 119, 148, 158] or non-uniform [23, 100, 312, 313] grid and the sensor response is recorded for a certain amount of time (typically 2-8 minutes) in the center of each cell. Then, a map of the gas distribution is created by plotting the average response [23, 100, 112, 137, 148], the maximum response [119, 158, 313] or the odor intermittency [26, 312] in each cell. Interpolation between adjacent cells is often used.

Gas distribution mapping by sequential measurements of a single sensor presents two main inconveniences. First, it has been observed that the movement of the sensor across the

experimental area can affect the air movement, and thus the gas dispersion [206]. Second, building a map of a large area by averaging measurements for several minutes at each location is a slow process. When the gas distribution cannot be assumed to be steady within the time frame of the mapping process (as it is the case in most natural environments), the resulting gas distribution map will show a great temporal averaging. This means that temporary variations in the gas distribution will be mapped as if they were time-constant structures. To reduce the time required to map the 2D gas distribution in a wind tunnel, Vergara et al. [313] performed sequential measurements at 6 downwind distances using a set of 9 sensing units distributed along the crosswind axis. To increase the accuracy in the sensor placement, Haverkamp et al. [312] used a robotic traversing system to move the sensor within the three axis of the environment.

In the general case where the gas distribution is time-varying, the most reliable way to characterize gas dispersion using chemical sensors is to deploy multiple units covering the target area, to perform simultaneous measurements during an extended period of time. With an increasing area, establishing a dense grid of gas sensors would involve an arbitrarily high number of fixed gas sensors, which poses problems such as cost and lack of flexibility. Because of this, this approach has been mostly restricted to 1D [122, 314] and 2D [98, 206, 315–317] measurements in small arenas. Grids of sensors are mostly based on MOX sensors [98, 122, 206, 315, 317] although optical sensors were used in [314, 316].

A dense 2D grid of 72 MOX sensors was presented by Zakaria et al. [206] to visualize the temporal evolution of a chemical plume in a 6 m $\times$ 3 m indoor arena (Figure 1.65). Whereas the plume was clearly visible at the beginning of the experiment, it gradually became less defined as the background concentration increased, becoming hardly distinguishable from the background after 30 minutes of release. This experiment illustrates the importance of characterizing gas dispersion in the environment for reporting and performing GSL experiments, since experiments carried out during the first half hour (when the plume is clearly visible) will show higher success rates than experiments performed later on. Murai et al. [315] deployed a 2D grid of 30 MOX sensors in a 3.5 m $\times$ 3.5 m area within an office room, to study gas dispersion in weak air flow conditions. A fuzzy gas cloud with unpredictable movements was observed, confirming that air flow has a major impact on the gas distribution (at least in the 2D slice observable at ground level).

The three-dimensionality of the environment has been only investigated by few preliminary works [148, 149, 318]. Russell et al. [148] used a mobile robot with a MOX sensor mounted on an extendable vane to map the 3D gas distribution in a small room. Although the gas source was placed near the floor ($h = 0.6$ m) and ethanol is heavier than air, most of the gas concentration was found near the ceiling (Figure 1.66). This supports the idea that dispersion was dominated by convective air currents instead of by the density of the pure substance. Reggente et al. [149] used a mobile robot mounting three MOX sensors at different heights (0.2, 0.4 and 0.6 m) to build a 3D gas distribution map of an indoor environment. The measurements from different sensors were fused using a multivariate Gaussian kernel (so-called 3D-Kernel DM+V). De Vito et al. [318] placed 4 wireless MOX sensor nodes at different heights in a small glass box (0.36 m^3 volume)to build a 3D gas distribution map. The 3D setup in this case was only a proof of concept of the remote logging capability of their propietary wireless modules.

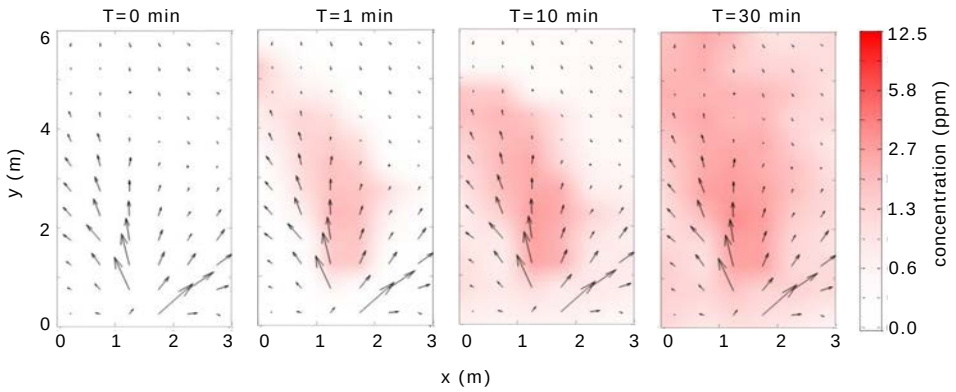


Figure 1.65: Snapshots of gas distribution map in a ventilated indoor arena. The gas source and the fan were placed in (x,y) = (1.5, 0) m. The estimated ethanol concentration is described on the color scale (adapted from [206]).

A practical inconvenience of using a grid of MOX sensors is that the sensors must be calibrated to correct the non-linear response and the variations in sensitivity across different sensor units (even if they are of the same model). The installation of the sensors may slightly perturb flow motion, due to the supports required to hold them at multiple heights and due to the local increase in ambient temperature by the sensor heater. The temporal and spatial resolution of this system cannot be compared to that of optical methods because (1) gas sensors have a slow response time and (2) with an increasing area, establishing a dense grid of gas sensors would involve an arbitrarily high number of fixed gas sensors, which poses problems such as cost and lack of flexibility.

1.6 Summary

Gas source localization by autonomous mobile robots has been a topic of research for almost 30 years and still today remains elusive beyond simple scenarios. Tracking a turbulent gas plume using a robot equipped with gas sensors is a challenging task, mainly due to the complex dynamics of gas dispersal, the limitations of current chemical sensors and the limited mobility of robotic platforms. In this sense, the commercial availability of low-cost UAVs has ignited new research opportunities related to autonomous gas sensing in applications such as environmental monitoring, gas source localization or gas distribution mapping. Whereas large UAVs equipped with complex instrumentation have been proposed for outdoor applications, nano UAVs are better suited for indoor operation. However, these small platforms are subject to constraints related to the reduced flight time and payload, that are not an issue for terrestrial robots or larger drones. For example, algorithms that require long measurement times at fixed positions will quickly drain the batteries of nano UAVs. Similarly, algorithms that rely on heavy instruments such as anemometers or LIDARs are not feasible due to limited payload.

At this time, it is unclear that bioinspired reactive behaviours have better performance than other approaches based on statistical inference from cumulative readings. [28, 29].

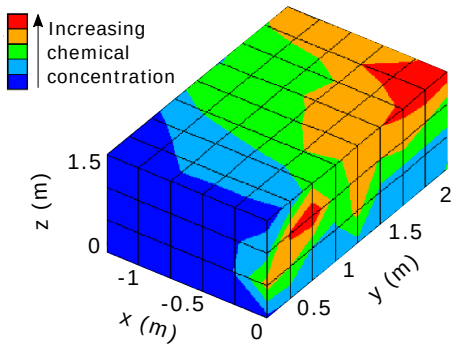


Figure 1.66: 3D gas distribution map of a ventilated indoor arena, obtained by sequential measurements of a MOX sensor. The gas source was an ethanol source placed at $(x,y,z) = (-0.5, 2.5, 0.6)$ m (from [148]).

According to Hernandez et al. [28], current gas sensors are too slow to resolve plume features in the milliseconds scale and mobile robots are not agile enough for performing insect-like reactive movements. Moore and Crimaldi [319] suggest that animals may use olfactory appendages to modify the chemical signal before it contacts the receptor site. On top of that, the reactive behaviors that roboticists have been able to model might be too simple to cope with complex environments. Rapid turns and sudden increases in flight speed upon contact with pheromone molecules are complex behavioral tasks that are not fully understood currently [320]. A practical limitation of bioinspired algorithms is that they require anemometry and real-time obstacle detection, which hinders their use in platforms with limited payload, such as nano-drones.

Engineering-based GSL algorithms are less sensitive to the response time of gas sensors, since instead of reactively tracking the plume they exploit computational resources available to robots, such as memory and processing power. Modeling the gas plume using a mathematical model or moving towards the source aided by a map of odor hits have been both useful GSL methods. However, long measurements are required at each sampling location to obtain statistical representations of the plume or to estimate the density of hits, which significantly increase the exploration time.

A major technical difficulty in designing and testing CSL algorithms is that dispersion of chemicals is difficult to observe since most chemicals produce an invisible plume. This has led to poor assumptions about how gases will disperse in an environment, e.g., that ethanol plumes will disperse at ground level. Lilienthal et al. [14] suggested more than one decade ago that a promising direction for future work is to investigate gas distribution mapping in 3D. The increased degrees of freedom of a UAV requires understanding the dispersion of chemicals in 3D, prior to developing navigation strategies that work in 3D. Visualization of plumes through optical methods is the most accurate method but it requires a complex and expensive setup that limits its practical applicability. Using a grid of chemical sensors is a low-cost approach that can be used as a first approximation.

Changes in environmental factors and trace concentrations of unknown gases hinder the localization a gas source using gas sensors, due to cross-sensitivity. In the presence

of multiple gas sources, a mobile robot equipped with a MOX sensor will struggle to identify the source of interest. To account for that, proper calibration and estimation of the LOD in conditions as similar as possible to the target scenario is necessary. Direct application of the LOD formulas to non-linear sensors is not possible due to the strong statistical assumptions imposed by the IUPAC methodology. The LOD is fundamental at long distances from the source, when the sensors operate near the noise level. In most reported experiments, the robots were placed unrealistically close to the gas source, thus, minimizing the need of such parameter. As pointed out by Hernández et al. [28], new research in the field should consider more general and complex environments.

Chapter 2

Objectives

The main goal of this thesis is to provide a nano-drone with olfaction capability to localize gas emission sources in indoor environments. This application presents several challenges related to the:

- ultra-constrained power envelopes of nano-drones,
- limitations of current gas sensors, e.g. high response time (limits robot speed), low selectivity (increases false positives) and high limit of detection (reduces the distance at which the chemical plume can be detected),
- interaction between propeller-induced turbulence and the gas sensor signals,
- variability of the experimental scenarios, including source characteristics, wind conditions and chemical interferences,
- complex dynamics of turbulent gas plumes.

The first specific objective is therefore to overcome key limitations of MOX sensors for mobile sensing applications, such as the low selectivity, the high power consumption and the high response time. We aim to do that by optimizing the operating mode of the sensor and by signal processing and chemometrics. In this regard, we aim to contribute in the following topics:

1. To develop univariate and multivariate LOD estimation methods for sensors whose response is non-linear, heteroscedastic and (possibly) multivariate (Chapter 3).
2. To develop a low-power heating method to reduce power consumption without degrading the predictive power (Chapter 4).
3. To improve the response time of MOX sensors for rapid tracking of turbulent plumes (Chapter 5).
4. To extract high-frequency features of the sensor response that are most correlated with the distance to a gas source (Chapter 6).

5. To redesign the "bout" detection algorithm from a classical signal processing perspective and optimize the algorithm parameters to increase the robustness against changing wind conditions (Chapter 6).

The second specific objective is to develop a 3D grid of MOX sensors to visualize the gas dispersion patterns (3D) produced by an evaporating chemical source in an indoor scenario, for different configurations of the gas source and air flow profiles (Chapter 7). Within this context, we want to compare the performance of different estimators of source proximity in 3D.

The third specific objective is to integrate a MOX sensor in a commercial nano-drone (Chapter 8). This includes optimizing the sensor location to minimize the effect of rotor-induced turbulences and designing the communication protocol with the base station. The drone should be able to read the MOX signals and send them in real-time to the base station, which would process them and command the drone to the next measurement location.

The final part of the thesis comprises the experimental evaluation of the nano-drone in real-world gas source localization experiments (Chapter 8). We aim at testing the system in a large indoor environment where a single gas source is placed at multiple heights.

Chapter 3

Optimizing the Limit of Detection in MOX Sensors

In this chapter, we propose univariate and multivariate chemometric methods for optimizing the limit of detection (LOD) in temperature-modulated MOX sensors. As we stated in the Introduction section, the LOD is a key figure of merit of any sensing technology but MOX sensors have features incompatible with the standard linear univariate LOD formulas, which assume a univariate linear response with homoscedastic and normally distributed residual errors. Both methodologies were validated using multiple units of the FIS SB-500-12 sensor in a scenario where the goal was to quantify low concentrations of CO (0-20 ppm) and the main source of error was the presence of uncontrolled levels of humidity (20-80 % r.h.).

In the univariate approach (Paper I), we optimize the calibration range and the univariate measurement point (along the multivariate response pattern) such that the response (or its logarithm) can be linearized and the residuals of the linear fit comply with the assumptions of the standard LOD formulas. For that, we fused the output of multiple statistical tests to ensure that a measurement point simultaneously satisfies the assumptions of linearity, normality and homoscedasticity. Compared to the measurement point recommended by the manufacturer, the optimized measurement point reduces the mean LOD from 4.0 to 2.3 ppm and increases the validity of the estimate from 80% to 95%. The optimized calibration models (and consequently the reported LOD values) were stable for at least two weeks if the sensor baseline was corrected using intraday blank samples.

In the multivariate approach (Paper II), we make full use of the data contained in the multivariate response pattern rather than selecting a single measurement point. Since the multivariate response exhibits heteroscedastic and correlated noise, it is not clear which is the best multivariate calibration model for this data. Maximum likelihood principal component regression (MLPCR) is optimal from a statistical point of view as it explicitly includes information on measurement uncertainties in the calibration. On the other hand, partial least squares (PLS) is the de facto calibration model for this kind of data, due to its inherent robustness against high multicollinearity and interference rejection capabilities. We found similar levels of error (0.4 ppm) between PLS and MLPCR, but PLS models

were simpler. We therefore propose a methodology to estimate the LOD in PLS models by applying the well-accepted univariate formulas to the first component of the orthogonalized PLS model. The average LOD value was reduced to 0.77 ppm, which roughly represents a five-fold improvement over the standard operating mode and a three-fold improvement over the optimized univariate models.

Papers I and II are shown below.

PAPER I

Analytica Chimica Acta xxx (2018) 1–13

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

PAPER I

Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models

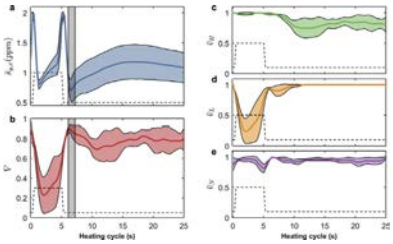
Javier Burgués [a, b, *], Juan Manuel Jiménez-Soto [b], Santiago Marco [a, b]

[a] Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028, Barcelona, Spain
[b] Signal and Information Processing for Sensing Systems, Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028, Barcelona, Spain

HIGHLIGHTS

- A methodology to estimate the LOD in non-linear chemical sensors is proposed.
- A set of statistical tests verify the main LOD assumptions at each detection point.
- The methodology was applied to a scenario of carbon monoxide and variable humidity.
- The optimum univariate detection point differs from manufacturer recommendations.
- The temporal stability of the LOD was explored.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 4 August 2017
Received in revised form
25 January 2018
Accepted 29 January 2018
Available online xxx

Keywords:
Semiconductor gas sensors
Metal-oxide sensors
Limit of detection
Non-linear
Humidity interference
Temperature modulation

ABSTRACT

The limit of detection (LOD) is a key figure of merit in chemical sensing. However, the estimation of this figure of merit is hindered by the non-linear calibration curve characteristic of semiconductor gas sensor technologies such as, metal oxide (MOX), gasFETs or thermoelectric sensors. Additionally, chemical sensors suffer from cross-sensitivities and temporal stability problems. The application of the International Union of Pure and Applied Chemistry (IUPAC) recommendations for univariate LOD estimation in non-linear semiconductor gas sensors is not straightforward due to the strong statistical requirements of the IUPAC methodology (linearity, homoscedasticity, normality). Here, we propose a methodological approach to LOD estimation through linearized calibration models. As an example, the methodology is applied to the detection of low concentrations of carbon monoxide using MOX gas sensors in a scenario where the main source of error is the presence of uncontrolled levels of humidity.

1. Introduction

Many potential applications of gas sensors involve the detection of very low concentrations of the target gases. For instance, in applications related to environmental monitoring the United States Environmental Protection Agency (EPA) specifies four principal pollutants and their corresponding maximum exposure

* Corresponding author. Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028, Barcelona, Spain.
E-mail address: jburgues@ibecbarcelona.eu (J. Burgués).

https://doi.org/10.1016/j.aca.2018.01.062
0003-2670/

concentrations (hourly) in outdoor air: 9 ppm of carbon monoxide (CO), 100 ppb of nitrogen dioxide (NO_2), 70 ppb of ozone (O_3) and 75 ppb of sulfur dioxide (SO_2) [1]. In the case of the commercial automotive market, a sensing range of 5 ppm is required to meet NO_2 regulations [2]. CO detectors integrated in smoking cessation monitors must comply with a detection threshold of 10 ppm of CO in exhaled breath in order to differentiate a smoker from a non-smoker [3].

The figure of merit that characterizes the ability of an instrument to differentiate between the absence of a substance (a blank value) and measurements with low concentration of the target gas is the limit of detection (LOD). The LOD is usually reported in concentration units. The LOD is extremely relevant for the applications described above because only instruments with LODs below the thresholds specified for the target gases could satisfy the applications requirements. The International Union of Pure and Applied Chemistry (IUPAC) defines the LOD for univariate signals in terms of statistical hypothesis tests, considering measurements as random variables [4]. IUPAC simplified formulas are derived from the upper prediction interval of the blank in a linear regression model which transforms the observed responses into concentration values. For binary detection (i.e. analyte present or not present) the estimated net signal or concentration is compared against the critical level (L_C), which only controls the false positives. The LOD (L_D) is, however, a more representative measure of the inherent detection capability of the analytical system because it also considers the false negatives. The distribution of the blanks and the LOD in the signal domain are modelled using a t-distribution with a maximum of 5% of false positives and 5% of false negatives. There are three strong underlying assumptions in the IUPAC LOD formulas based on the use of linear regression models for prediction purposes:

a) **Linearity**: A linear model can be fit to the data only if the relationship between the independent and dependent variables is linear. In other words, the residual errors of a linear fitting should have zero mean.
b) **Homoscedasticity** of the residual errors: If the variance of the residual errors is not constant along the independent variable, the standard error will be biased. A biased standard error will lead to incorrect prediction intervals. Extensions for the heteroscedastic case are described in the IUPAC norm but only for a particular type of heteroscedasticity: variance proportionate to response.
c) **Normality** of the residual errors. For Least Squares estimators to be maximum likelihood estimators of the linear fit the errors should be normal. Additionally, the use of *t*-test for hypothesis testing also requires that the residual errors are normally distributed.

Violation of any of these assumptions may bias the confidence and prediction intervals of the linear regression model. Because the IUPAC LOD is based on the upper prediction interval of the blank, the LOD will be biased too. Even when the assumptions are satisfied, it is important to remark that the LOD estimate has a domain of validity closely tied to the set of conditions covered by the calibration points, including obviously the conditions for the blanks. Departure from those conditions casts doubt on the validity of the LOD estimate (this is true for any analytical instrument).

Instruments based on linear devices such as electrochemical gas sensors usually satisfy these assumptions and therefore the LOD estimates are trustworthy. However, this is not the case for certain chemical sensor technologies. A clear example are semiconductor gas sensors such as, metal oxide semiconductor (MOX) [5], gasFETs [6] or thermoelectric sensors [7]. The non-linear response with respect to concentration will in many cases violate the linearity assumption. In addition, some of these technologies also present cross-sensitivity to environmental interferences (e.g. temperature, humidity or barometric pressure) and lack of stability [8–10]. This might yield non-normal or heteroscedastic residual errors in calibration.

In some cases, the LOD is calculated when the instrument is calibrated in laboratory conditions and then it is (incorrectly) assumed that in future and field operation the LOD will still be the same as in calibration. But, in field operation the range of variation of interferences such as temperature and humidity can go beyond the calibration range or the instrument might have drifted. Since in many chemical sensor applications we expect the system to run without human intervention for extended period of times (months or even years), a relevant question is the stability of the LOD in time. The validity of the LOD estimates in MOX sensors is a critical issue due to the intrinsic drift and cross-sensitivities to environmental conditions and other gases. However, this problem has not been properly addressed in the literature. Some works blindly trust the LOD obtained in calibration [11–14]. Other authors have approached the LOD validation based on the theory of hypothesis testing measuring the probabilities of false positive and false negatives, however, using relatively small samples sizes [15]. The results obtained with this approach are not trustworthy unless there are many measurements and replicates corresponding to concentration standards near the LOD, which is usually unknown.

Probably due to these issues, the LOD is rarely reported in archival papers concerning MOX sensors. A search in the current literature (as of July 2017) reveals that only a handful of articles deal with LOD estimation in MOX sensors [12–14]. Nevertheless, most of these works often ignore the IUPAC assumptions and, for simplicity, assume they hold in their experimental data. Despite the difficulties associated to the estimation of the LOD, the use of MOX sensors in applications that require a certain LOD value has obvious interest. MOX sensors are long-lasting devices which can be cheaply manufactured in a miniaturized form factor while achieving low power consumption (thanks to MEMS technology). The market of MOX sensors has been indeed very active in the last years due to the potential integration into portable devices such as smartphones, wearables or tablets. MOX sensors are being used in diverse applications that range from environmental monitoring [12,16,17], to safety and security [18], food [19] or medical applications [20]. Areas with high levels of pollutants could be easily localized thanks to portable measuring stations based on MOX sensors [21]. In the biomedical field, MOX sensors are currently proposed to monitor breath biomarkers for diagnostic of many diseases, including several types of cancer [22].

The combination of pulsed temperature modulation and univariate CO detection at low temperatures was pioneered in commercial products by Figaro Engineering in the successful product TGS-203 [23]. The marketing of this product with this recommended operational mode started as early as 1980. The sensor manufacturer shows that the best selectivity to CO is achieved when the sensor element temperature is kept under 100 °C. However, at this temperature the sensor becomes very susceptible to water vapor. To eliminate this influence, they propose to use a cycling high/low voltage pulses. They claim that the high heater voltage cleans the sensor surface and removes the influence of water vapor, while the low heater voltage conditions the sensor for measuring CO. This operation mode is still used by several manufacturers like Figaro (TGS3870 A-04), FIS (SB-500-12) or SGX Sensortech (MiCS sensors). Newer sensors using either beads or micromachined substrates are faster from a thermal point of view and they allow to use shorter thermal cycles, but the recommended operation mode remains basically the same. In all cases, the

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

detection of CO remains univariate for simplicity of use. The reduction of humidity influence in this operation mode has been confirmed later [24–26]. It is also true that the selectivity of MOX sensors can be improved by temperature modulation followed by a multivariate prediction model. However, from a user perspective, univariate detection has the advantage of simplicity and probably for this reason is still the recommended operation mode by the manufacturers. In this work, we will be using FIS SB-500-12 sensors. Our measurements are based on univariate features: single detection points at certain locations within the temperature modulation cycle. Although there is potential for improving the LOD by using the full response pattern, it is not the scope of this work to explore the estimation of the LOD from the multivariate response pattern. The problem of multivariate LOD estimation is that there is no well-accepted LOD formula, as it is the case for univariate estimation. There are several multivariate LOD proposals [27] and most of them produce different outputs [28]. Instead of dealing with the multivariate case, the idea behind this work is to optimize the location of the univariate detection point using the standard IUPAC univariate formula for LOD estimation and then applying a set of statistical tests to ensure that the underlying LOD assumptions are satisfied.

The present work, therefore, aims to several purposes: (i) to recall the main ideas underlying the IUPAC theory regarding the univariate limit of detection and their impact in non-linear chemical sensing; (ii) to give guidelines to detect situations in which heteroscedastic, non-normal and biased residuals in calibration hinder the application of the simplified LOD formulas; (iii) to propose a set of statistical tests to verify that the obtained LOD values in calibration are applicable to validation data; (iv) finally to illustrate the problems associated to the blind use of the LOD in future data when using MOX sensors. For that, we proposed a scenario of carbon monoxide (CO) detection under variable humidity conditions, which is representative of various real applications. In the selected scenario, humidity is the main interference and we consider the LOD as a figure of merit to optimize the measurement point for CO sensing reducing the effect of the humidity. We would like to stress that here humidity is a random perturbation and not a systematic influence, and it is not the interest of this work to estimate the humidity, but only to ascertain how humidity variations have an impact on the LOD. Obviously, humidity interference is a major concern for MOX performance. In the case of CO sensing, this has been extensively reviewed by Ref. [29]. Our contribution focuses on the methodological aspects related to the estimation of the LOD in chemical sensors through linearized models. We will present a methodology for evaluating the properties of the sensor data relevant to the LOD formulas.

2. Experimental and methods

To illustrate the challenges described in the introductory section we designed an experimental scenario of CO detection under random humidity conditions using temperature-modulated MOX sensors. The experimental setup was comprised of a test bench where dynamic mixtures of CO in humid synthetic air were generated, an electronic board composed of seven MOX sensors and a temperature/humidity sensor. The data was processed using MATLAB R2009 (The MathWorks) and libraries from the Statistics and Machine Learning Toolbox.

2.1. Sensor board

An electronic board including seven MOX sensors for CO detection (SB-500-12, FIS Inc.), a temperature/humidity sensor (SHT75, Sensirion AG) and the necessary conditioning circuits was assembled in-house. The commercial MOX sensors used in this work are manufactured with large tolerances in baseline (one order of magnitude) and sensitivity (a factor of two) [30]. To obtain statistics of the performance of the sensor family (e.g. mean and variance), several units of the same MOX sensor model were integrated in the sensor board. The Sensirion sensor provided reference humidity and temperature values with tolerance below 1.8% r.h. and 0.5 °C, respectively, every 5 s. According to the manufacturer datasheet, the long-term drift of the SHT75 sensor is less than 0.5% r.h./year and 0.04 °C/year. The humidity values will not be used for sensor data compensation, only to obtain a direct measurement of the humidity values within the sensor chamber.

The heater voltage of the MOX sensors was modulated in the range 0.2–0.9 V in cycles of 25 s, following the manufacturer recommendations (0.9 V for 5s, followed by 0.2 V for 20s) (see Fig. 1). For simplicity, in this work we are not implementing a closed loop control of the heater temperature which has been shown to produce more accurate results [31]. The seven MOX sensors were preheated for one week before starting the experiments to reach stable responses. The manufacturer recommends measuring the sensor response at a single detection point located at the end of low level of the heating cycle (see Fig. 1). The choice of this detection point, that we will call “nominal working point”, is specific to CO detection as the sensors are more sensitive to CO at low temperatures. Nonetheless, we recorded the sensor response during the whole heating cycle to evaluate the performance of the sensor in other measurement points. For that, the MOX read-out circuits consisted of voltage dividers with 1MΩ load resistors. The value of the load resistor allowed proper quantification of the sensor resistance, considering the large dynamic range (20 kΩ–10 MΩ) present at low concentrations of the analyte with the chosen temperature modulation waveform. The output voltage of the sensors was sampled at 3.5 Hz using an Agilent HP34970A/34901A DAQ configured at 15 bits of precision and input impedance greater than 10 GΩ. In this configuration, the errors introduced in the voltage measurements were considered negligible compared to the intrinsic variability of the sensor resistance due to the chemical transduction process.

2.2. Experimental design

Previous experience by the authors on CO sensing suggested that the LOD could be around 5 ppm or even smaller. At this point, it is important to remember that for LOD estimation the IUPAC

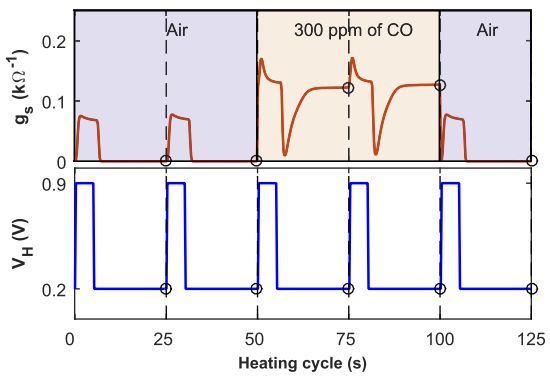


Fig. 1. FIS SB-500-12 operating mode and output signal. The top subplot represents the sensor conductance during an exposure to air (50 s), 300 ppm of carbon monoxide (50 s) and air (25 s). The bottom subplot represents the heater voltage. The black circles indicate the detection points recommended by the manufacturer.

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

recommends that the calibration points should be placed close to the LOD [4]. For this reason, we selected a calibration range of 0–9 ppm. At this point, it is important to remark that 9 ppm is not the upper limit for the detection of CO with MOX sensors, but only a reduced range to follow IUPAC recommendations for LOD estimation. In fact, MOX sensors are routinely used for CO detection in domestic premises with measurement ranges up to 400 ppm [29].

The experimental design was based on the literature recommendations for LOD estimation of using between five and ten calibration standards, eight to ten replicates [32] and equidistant concentration levels [33]. To generate the experimental dataset, five linearly equally spaced calibration standards in the range 0–9 ppm were selected ({0.0, 2.2, 4.4, 6.7 and 8.9} ppm). Ten repetitions were performed at each concentration, with each repetition having a relative humidity value randomly chosen from a uniform distribution between 15 and 70% r.h. (see Fig. 2). Hence, each experiment was composed of 50 conditions (5 concentrations x 10 repetitions) which were performed in random order to reduce the impact of factors that were not explicitly accounted for in the experimental design. The duration of each exposure was 15 minutes. A single calibration experiment lasted approximately 12 hours and was repeated 13 times for repeatability and reproducibility estimation. The dataset was therefore composed of 650 measurements (13 experimental days x 50 measurements/day) which were collected over 15 natural days.

2.3. Generator of dynamic gas mixtures

Gas mixing was performed using mass flow controllers (MFC), which controlled three different gas streams (CO, wet air and dry air). These streams were delivered from high quality pressurized gases in cylinders. The CO bottle contained 1600 ppm of CO diluted in synthetic air with 21 ± 1% O_2. The wet and dry air streams were both delivered from a synthetic air bottle with 99.995% purity and 21 ± 1% O_2. The humidification of the wet stream was based on the saturation method using a glass bubbler (Drechsler bottles) [34]. To reach higher humidity levels, the Drechsler bottles were submerged into warm water at 42 °C. The selected MFCs (EL-FLOW Select, Bronkhorst) had full scale flow rates of 1000 ml_n/min for the dry and wet air streams and 3 ml_n/min for the CO channel. Check valves protected the MFCs against reversed flow in the system. The gas mixtures at different values of concentration and humidity were transferred into a small-sized polytetrafluoroethylene (PTFE) test chamber (250 cm^3 internal volume) containing the sensors under test. Fig. 3 shows a schematic plan of the measuring system.

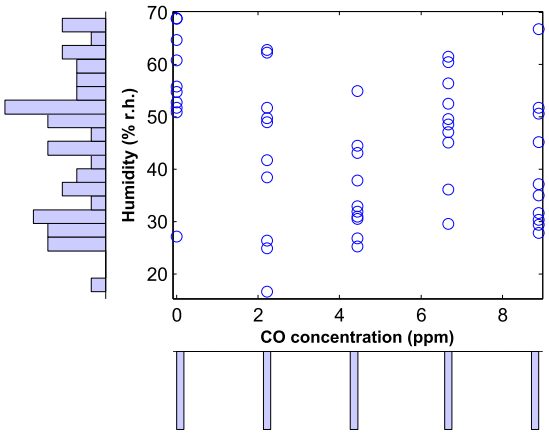


Fig. 2. Scatter plot of the nominal CO concentration (ppm) and the measured humidity (% r.h.) in the first experimental day. The marginal distributions of the nominal CO concentration and the measured humidity are displayed as univariate histograms on the horizontal and vertical axes of the scatter plot, respectively.

At the beginning of each experiment, the gas chamber was cleaned for 15 min using a stream of synthetic air at a flow rate of 240 ml_n/min. After that, the gas mixtures were released at a constant flow rate of 240 ml_n/min for 15 min each. The temperature variations inside the gas chamber, for each experiment, were below 3 °C. The highest relative uncertainty in CO concentration was 5.5% for the first non-zero concentration (2.2 ppm), which corresponds to an absolute uncertainty of 0.12 ppm. A desktop computer running Labview 2015 (National Instruments) commanded the MFCs and integrated the readings from the humidity sensor into a log file.

During each experimental condition, the seven sensor signals, indicative of the gas conditions presented to the sensors, were continuously sampled. In order to evaluate the sensor response to a certain experimental condition, portions of the sensor signal corresponding to a full heating cycle (henceforth called patterns) were extracted from the original signals. To ensure that the extracted patterns correspond to stable CO levels, only the last three cycles of each 15-minute condition were taken into account. These three fragments were averaged to reduce instrumental noise, producing a single multivariate conductance pattern $g_s(t)$, $t \in [0, 25]$ s associated to a stable gas concentration$c$ (ppm). The readings from the humidity sensor were also averaged during the same three heating cycles to provide a reference relative humidity value h (% r.h.).

2.4. Calibration models

As we said earlier in the introduction, MOX sensors present a non-linear response which might be best fit by non-linear models. Nevertheless, within a narrow range of small concentrations the sensor response can be considered quasi-linear and the use of linear models can be justified. This allowed us to use the sound LOD theory already developed for univariate linear calibration models.

In this work, we only considered linear univariate calibration models where the observed response y is predicted from the experimental analyte concentration x by a linear regression model of the form:

$$y = B + Ax + e_y \quad (1)$$

Where B and A are the real (unknown) regression coefficients and e_y is the residual error of y. In Eq. (1) it is assumed that the analyte concentration is error-free because the uncertainty in the preparation of concentrations was negligible (see Section 2.3).

For CO detection, the most widely used method is the Clifford-Tuma model (Eq. (2)) [35]. This empirical model uses a linear relationship between the logarithm of the sensor conductance g_s ($k\Omega^{-1}$) and the logarithm of the analyte concentration c (ppm):

$$\log(g_s) = \log(g_{air}S) + \beta \log(c) = \alpha + \beta \log(c) \quad (2)$$

Where g_{air} ($k\Omega^{-1}$) is the sensor conductance in clean air, S is a gas depending parameter and β is the so-called sensitivity to the gas. Log-log models such as the Clifford-Tuma model provide in general better fitting than normal models but their use for LOD estimation is questionable because the blanks ($c = 0$ ppm) need to be excluded due to the term $\log(c)$. As pointed out in the IUPAC recommendation [4], "*The blank is one of the most crucial quantities in trace analysis, especially in the region of the Detection Limit … Inadequate attention to the magnitude and variability of the overall*

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

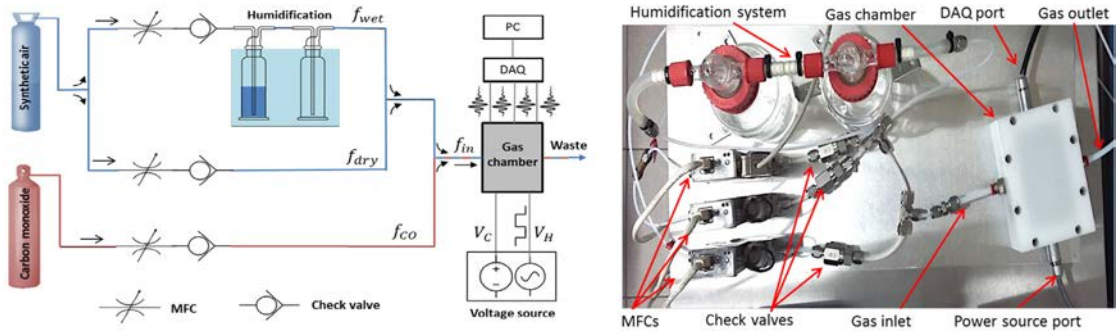


Fig. 3. Experimental test bench for the generation of dynamic gas mixtures and acquisition of the sensor signals. Left: Block diagram. Right: Picture.

blank, may lead to severe underestimation of the detection limit."

To introduce blanks into the model, we propose a log-normal model in which the logarithm is applied only to the sensor conductance (Eq. (3)):

$$\log\left(\frac{g_s}{g_{air}}\right) = \mathrm{B} + \mathrm{Ac} + \varepsilon \tag{3}$$

where B is the intercept, A is the sensitivity and ε is the error term. Dividing g_s by g_{air} corrects the potential changes in sensor baseline. g_{air} was estimated from a blank sample with random humidity measured at the beginning of the day and all the measurements of that day were divided by this value. The use of Eq. (3) for LOD estimation purposes is valid only for a narrow range of small concentrations and for certain working points in which the sensor response (g_s) can be linearized without applying the logarithm to the concentration.

The precision of the calibration function was assessed by the residual standard deviation ($s_{y,x}$) and the estimated standard deviation of the estimated slope ($\sigma_{\widehat{A}}$) and the estimated intercept ($\sigma_{\widehat{B}}$), as recommended by the IUPAC [36]. Because $\sigma_{\widehat{A}}$ and $\sigma_{\widehat{B}}$ only depend on $s_{y,x}$ and the calibration design -which was fixed-, minimizing $s_{y,x}$ also minimizes $\sigma_{\widehat{A}}$ and $\sigma_{\widehat{B}}$. Therefore, we only report $s_{y,x}$ (Eq. (4)).

$$s_{y,x} = \sqrt{\sum_{i=1}^{n}\frac{(y_i - \widehat{y}_i)^2}{n-2}} \tag{4}$$

where n denotes the number of calibration samples, and y_i and $\widehat{y}_i$ are the nominal and fitted value for the calibration sample$i$, respectively.

2.5. Limit of detection (LOD)

In the IUPAC recommendations [37], the LOD is defined as the "smallest measure that can be detected with reasonable certainty for a given analytical procedure". IUPAC derives the LOD in terms of statistical hypothesis testing with default values for the probabilities of false positives (α) and false negatives (β) both equal to 0.05. The fundamental relations (Eqs. (5) and (6)) are formulated based on a generic symbol (L) that represents either net signal (blank-corrected) or analyte concentration:

$$P\left(\widehat{L} > L_C | L = 0\right) = \alpha \tag{5}$$

$$P\left(\widehat{L} < L_C | L = L_D\right) = \beta \tag{6}$$

Where $\widehat{L}$ is an estimate of L, L_C is the threshold used for the detection decision (i.e. analyte present or not) and L_D is the limit of detection or LOD. Assuming that $\widehat{L}$ is normally distributed with known variance (σ), the value of L_C is given by Eq. (7), after solving for L_C in Eq. (5)

$$L_C = z_{1-\alpha}\sigma_0 \tag{7}$$

Where $z_{1-\alpha}$ is the critical value of the normal distribution for significance level α and σ_0 is the true standard deviation of $\widehat{L}$ when $L = 0$. Similarly, L_D can be found by first solving for L_C in Eq. (6):

$$L_C = L_D - z_{1-\beta}\sigma_D \tag{8}$$

Where $z_{1-\beta}$ is the critical value of the normal distribution for significance level β and σ_D is the true standard deviation of $\widehat{L}$ when $L = L_D$. Then, solving for L_D in Eq. (8):

$$L_D = L_C + z_{1-\beta}\sigma_D \tag{9}$$

Inserting Eq. (7) into Eq. (9) yields:

$$L_D = z_{1-\alpha}\sigma_0 + z_{1-\beta}\sigma_D \tag{10}$$

In many practical applications, the true values σ_0 and σ_D in Eq. (10) are unknown, so a Student t-distribution is used to approximate the Gaussian curve. The critical value of the Gaussian distribution is replaced by the one-sided t-critical value ($t_{1-\alpha,\upsilon}$) for the chosen confidence level (α) and degrees of freedom (υ). Similarly, the true standard deviations σ_0 and σ_D are replaced by the corresponding estimates s_0 and s_D, respectively. Because s_D corresponds to the dispersion of $\widehat{L}$ at L_D, which is unknown (in fact, L_D is what we are trying to estimate), s_D is replaced by s_0 (assuming homoscedasticity). Finally, recalling that we chose $\alpha = \beta$, the formula of Eq. (10) is simplified to Eq. (11):

$$L_D = 2t_{1-\alpha,\upsilon}s_0 \tag{11}$$

Eq. (11) is a generic LOD formula that can be used to determine the LOD in the signal domain or in the concentration domain. The LOD is usually reported in concentration units but the measuring instruments report signals instead. Given a set of discrete measurements, there are two ways to compute the LOD: (a) Convert the instrument signals to concentration values and estimate the LOD in

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

the concentration domain using Eq. (11) or (b) compute the LOD in the signal domain using Eq. (11) and then transform the "signal LOD" to a "concentration LOD". Signals and concentrations can be related to each other by the use of a calibration curve. The IUPAC only considers the linear calibration function (c.f. Eq. (1)). In this case, the concentration ($\hat{x}$) corresponding to an observed gross signal (y) is estimated by:

$$\hat{x} = \frac{y - \hat{B}}{\hat{A}} = \frac{\hat{S}}{\hat{A}} \tag{12}$$

Where $\hat{B}$ is the estimated intercept, $\hat{A}$ is the estimated slope of the calibration function and $\hat{S}$ is the net (blank-corrected) response. Application of the fundamental LOD definitions (i.e. Eqs. (5) and (6)) to the concentration domain (i.e. option (a) above) requires normality in the estimator $\hat{x}$, which is a strong assumption considering that $\hat{A}$ will be estimated with uncertainty. Therefore, the IUPAC recommends the computation of the LOD in the signal domain (Eq. (11)) and the posterior transformation to concentration units through Eq. (12).

The only variable that we need to estimate in Eq. (11) is s_0, the standard deviation of the estimated net response $\hat{S}$ when the analyte is absent ($x = 0$). From linear regression [36], we know that the standard deviation of a predicted value y_p at position x_i is:

$$s_{y_p} = s_{y,x}\sqrt{1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_j - \bar{x})^2}} \tag{13}$$

Where $s_{y,x}$ is the standard error of regression, n is the number of samples, x_j are the calibration standards and $\bar{x}$ is the mean calibration concentration. Solving for a blank ($x_i = 0$) yields:

$$s_0 = s_{y,x}\sqrt{1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum (x_j - \bar{x})^2}} = s_{y,x}\eta \tag{14}$$

where $\eta = \sqrt{1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum (x_j - \bar{x})^2}}$ is a design parameter reflecting the calibration design (i.e. relative position of the calibration standards and number of samples) [37].

Substituting Eq. (14) into Eq. (11), yields the LOD in the signal domain (S_D):

$$S_D = 2t_{1-\alpha}s_{y,x}\eta \tag{15}$$

The LOD in the concentration domain is obtained by inserting Eq. (15) into Eq. (12):

$$x_D = \frac{S_D}{\hat{A}} \cdot \left(\frac{K}{I}\right) = \frac{2t_{1-\alpha}s_{y,x}\eta}{\hat{A}} \cdot \left(\frac{K}{I}\right) \tag{16}$$

The correction factor $\left(\frac{K}{I}\right)$ accounts for the uncertainty in the regression slope ($\hat{A}$) [37] and it is defined in Eq. (17):

$$\frac{K}{I} = \frac{1 - r(B,A) \cdot \left(\sigma_{\hat{B}}/\sigma_0\right) \cdot \left[t_{1-\alpha,\nu} \cdot \left(\sigma_{\hat{A}}/A\right)\right]}{1 - \left[t_{1-\alpha,\nu} \cdot \left(\sigma_{\hat{A}}/A\right)\right]^2} \tag{17}$$

Where r(B, A) is the correlation coefficient between the slope and intercept of the calibration line (Eq. (18)), $\sigma_{\hat{B}}/\sigma_0$ is the relative standard deviation of $\hat{B}$ (estimated intercept), σ_0 is the standard deviation of the blanks and $\sigma_{\hat{A}}/A$ is the relative standard deviation of $\hat{A}$.

$$r(B,A) = \bar{x} \Big/ \sqrt{\sum_{i=1}^{n} x_i^2 \Big/ n} \tag{18}$$

The interpretation of the correction factor $\left(\frac{K}{I}\right)$ is as follows. If the slope (A) is estimated with low uncertainty (i.e. $(\sigma_{\hat{A}}/A) \to 0$), the term in square brackets in Eq. (17) goes to zero, and both K and I are equal to 1, yielding $\left(\frac{K}{I}\right) = 1$. In this case, the LOD is not affected by the correction factor. At the other extreme, when the uncertainty in estimating A is very high (as $\sigma_{\hat{A}}/A$ approaches $1/t_{1-\alpha,\nu}$), the denominator (I) goes to zero, the ratio $\left(\frac{K}{I}\right)$ goes to infinity and x_D is unbounded. As it can be seen in Eq. (16), the estimated slope ($\hat{A}$) is the only parameter of the regression line that affects x_D. However, the correction factor includes the effect of the uncertainty of $\hat{B}$ when estimating A. If A and B are estimated individually, the correlation coefficient r (B, A) is zero and the ratio $\left(\frac{K}{I}\right)$ only depends on the uncertainty of $\hat{A}$. When A and B are estimated jointly, they will be negatively correlated and the numerator (K) in Eq. (17) will increase with the uncertainty of $\hat{B}$.

2.5.1. Statistical tests to check the underlying LOD assumptions

In the preceding section, we presented a formula for estimating the LOD in the concentration domain, assuming certain properties of the instrumental signals. In particular, the IUPAC LOD definition [4] assumes that the residual errors are unbiased, normally distributed and homoscedastic. The violation of any of these assumptions may yield inconsistent LOD values. In this section, we present a set of statistical tests to systematically verify the underlying LOD assumptions:

- **Homoscedasticity:** Levene's test [38] was used to test the homogeneity of variance across the residuals of the I concentration levels (J samples each). This test assumes normally distributed data but is less sensitive to departures from normality than other tests for comparison of variances.
- **Normality:** The Saphiro-Wilk test [39] was applied to the residuals of each concentration level to assess normality (I tests with J samples each). This test assumes that the sample size is large enough (30 or more) because for small sample sizes normality tests have little power to reject the null hypothesis of normally distributed data. Nonetheless, it has been proved that Saphiro-Wilk test is the most powerful normality test for small sample size [40]. When the residuals at every concentration level are normally distributed the full set of residuals will be normally distributed too. Based on this, an additional Saphiro-Wilk test was applied to all the residuals (one test with IxJ samples) to increase the power of the test.
- **Linearity:** If the data behaves linearly, the residual errors in each concentration level should be unbiased or, in other words, symmetrically distributed around zero. A two-tailed t-test of zero mean [41] was applied to the residuals of each concentration level and to all residuals to verify the linearity assumption. This test assumes normally distributed data. The power analysis of this test indicated that the minimum detectable changes for the group tests (10 samples) and for the full set (50 samples) were 1.5 and 0.3 ppm respectively, which we considered acceptable values.

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

Table 1
Summary of statistical tests used to verify the LOD assumptions.

Assumption	Test	Null hypothesis	Alternative hypothesis
Homoscedasticity	Levene's test	The multiple data samples have equal variances	At least two of the data samples do not have equal variances
Normality	Saphiro-Wilk test	Sample data is normally distributed	Sample data comes from a non-normal distribution
Linearity	One-sample T-test	The sample data comes from a normal distribution with mean equal to zero	The sample distribution does not have a mean equal to zero

The Holm-Bonferroni correction for multiple comparisons was applied to the family of tests to reduce the false positive rate [42]. Let H_1, H_2, H_3 be the null hypothesis of the homoscedasticity, normality and linearity assumptions (see Table 1), sorted by ascending p-values P_1, P_2, P_3. The Holm-Bonferroni algorithm will first find the minimum index k such that the p-value P_k satisfies the following condition:

$$P_k > \frac{\alpha}{m+1-k} \tag{19}$$

Where m is the number of hypothesis ($m = 3$ in our case) and α is the significance level ($\alpha = 0.05$ in our case). The null hypothesis $H_1, \cdots, H_{k-1}$ are rejected and $H_k, \cdots, H_3$ are not rejected. We define the validity of a hypothesis H, as the Boolean variable v:

$$\begin{aligned} v &= 0, \ \text{if } H \text{ is rejected} \\ v &= 1, \ \text{if } H \text{ is not rejected} \end{aligned} \tag{20}$$

v_H, v_N and v_L represent the validity of the homoscedasticity, normality and linearity assumptions, respectively. The joint validity of the LOD assumptions ($v_{H,N,L}$) is a Boolean variable defined as the AND operation (represented by the logic operator $\wedge$) of the marginal validities:

$$v_{H,N,L} = v_H \wedge v_N \wedge v_L \tag{21}$$

2.5.2. Validity of x_D

The value of x_D only makes sense if it is positive, the calibration function is estimated with reasonable accuracy and the underlying LOD assumptions are satisfied. The constraint in Eq. (22) is used to avoid negative values resulting from Eq. (16) due to the random variable $\widehat{A}$.

$$x_D > 0 \tag{22}$$

The constraints in Eqs. (23) and (24) set an upper limit in the uncertainty of the regression parameters (slope and intercept, respectively).

$$\sigma_{\widehat{A}} \Big/ A \leq 1 \tag{23}$$

$$\sigma_{\widehat{B}} \Big/ \sigma_0 \leq 1 \tag{24}$$

The joint validity of the LOD assumptions (see Section 2.5.1) must be equal to one:

$$v_{H,N,L} = 1 \tag{25}$$

The values of x_D which did not satisfy the above constraints were considered invalid. Combining Eqs. (22)–(25), we obtain the formula for the validity of x_D (V):

$$V = (x_D > 0) \wedge \left(\sigma_{\widehat{A}} \Big/ A \leq 1\right) \wedge \left(\sigma_{\widehat{B}} \Big/ \sigma_0 \leq 1\right) \wedge (v_{H,N,L} = 1) \tag{26}$$

Here, V is a Boolean variable which takes the value '1' if x_D is considered valid and '0' otherwise.

2.6. Optimization of the working point

To optimize the working point of a given sensor, we used the measurements of the first experimental day to fit the model described by Eq. (3) for every working point t_i of the response pattern. Then, from the subset of the models that satisfied Eq. (26), we selected the working point that minimized the error of regression $s_{y,x}$ (Eq. (4)). We denote this optimum point as $t_i^*(s)$, to indicate that there is one optimum point for each sensor s.

To report robust indicators of x_D and V, for a given working point and a given sensor s, the individual estimates $x_D(s,d)$ and $V(s,d)$ obtained in different experimental days d ($d = 1, \ldots, N$) were averaged according to Eqs. (27) and (28):

$$\overline{x}_D(s) = \frac{1}{N}\sum_{d=1}^{N} x_D(s,d) \tag{27}$$

$$\overline{V}(s) = \frac{1}{N}\sum_{d=1}^{N} V(s,d) \tag{28}$$

where $\overline{x}_D(s)$ and $\overline{V}(s)$ represent the average LOD and average validity for sensor s, respectively.

2.7. Validation of the optimum working point

Previously we described the assumptions imposed by the IUPAC on the calibration model to reliably calculate a LOD value. The IUPAC recommendations only concern the calibration process and expect that the unseen future data used in the prediction step have the same properties as the calibration data. While this can be the case in linear stable instruments operating under controlled lab conditions, MOX sensors tend to drift over time and are very sensitive to interferences. Slight changes in experimental conditions during the measurements, influences of temperature, pressure or humidity, can also cause the baseline to drift away from its original level. As a result, data obtained days after calibration might be shifted with respect to the calibration data or show different variance (see Fig. 4). In any of these cases, the LOD value obtained in calibration will be biased and the number of false positives and false negatives can be different from the 5% established by the IUPAC norm.

The calibration models built on the kth day were externally validated using the samples of the consecutive days k+1, k+2, ..., N. The concentration of the external validation samples was predicted using the inverse of Eq. (3) and the root mean squared error in prediction (RMSEP) was calculated:

PAPER I

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

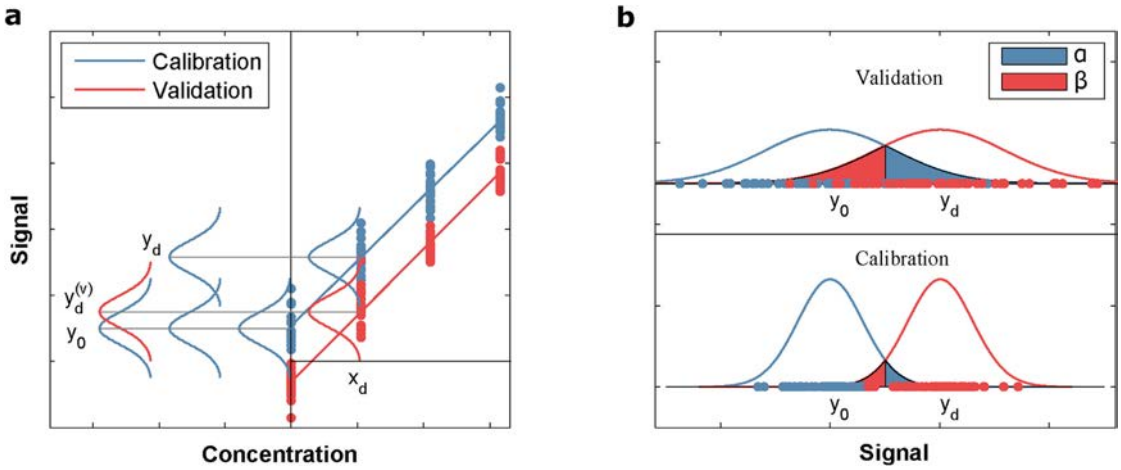


Fig. 4. Simulated examples of degradation of calibration model due to: (a) shifted validation data; (b) increased variance of the validation data. The blue gaussian curve centered at y_0 is the response in blank samples and the red gaussian curve centered at y_d is the response at the LOD level. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$\text{RMSEP} = \sqrt{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2 / n} \quad (29)$$

Where n denotes the size of the validation set, and x_i and $\hat{x}_i$ are the true and predicted concentration for sample i, respectively.

3. Results and discussion

In order to illustrate the proposed methodology in a real data set and face some of the challenges associated to LOD estimation in non-linear sensors, we designed a scenario representative of real world applications: CO detection under random variable humidity conditions. The working principle of MOX sensors is based on the change of the sensor conductance when the sensor is exposed to

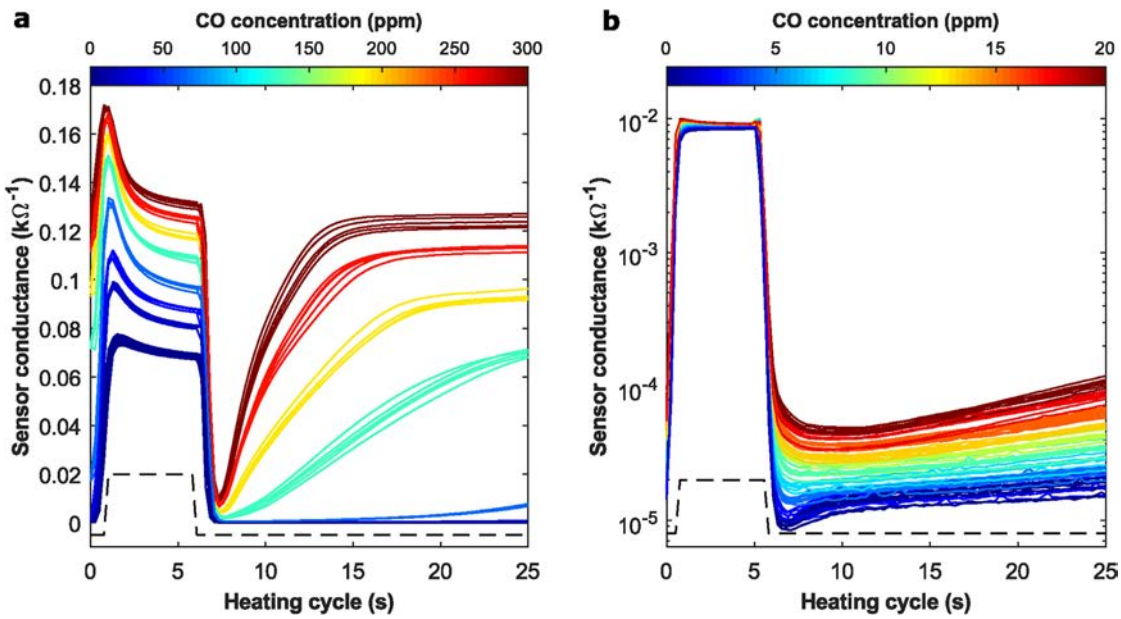


Fig. 5. Multivariate sensor conductance patterns of a FIS SB-500-12 sensor exposed to (a) 0–300 ppm of CO and (b) 0–20 ppm of CO. The colormap indicates the CO concentration (see colorbar at the top of each subplot). The black dashed line represents the heater voltage (a.u.). The logarithm is applied to the sensor conductance in (b) to facilitate the visualization of the low heating temperatures.

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

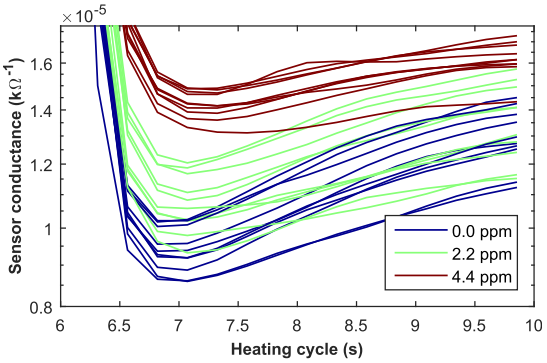


Fig. 6. Zoom-in of the region $t = 6 - 10$ s of the logarithmic sensor conductance patterns shown in Fig. 5b. Only the three lowest concentration levels (0.0, 2.2 and 4.4 ppm) are displayed for visual clarity.

the target gas. The sensitivity to the target gas depends on the temperature (working point) of the sensing surface. When the sensor is exposed to relatively high CO concentrations the sensitivity to CO is maximized at the end of the low heating cycle (Fig. 5a). This is why the nominal detection point recommended by the manufacturer is $t = 25$ s. At lower concentration ranges (Fig. 5b), similar values of the sensitivity are obtained in a range of measurement points in the low level of the heating cycle ($t \in [5, 25]$ s). At the lower end of the concentration range, the separability across concentrations is higher in a narrow range of working points located at the beginning of the low temperatures ($t = 6.5 - 7.5$ s). Because this is barely visible in Fig. 5b and it is a relevant finding, we have created a zoomed-in version of that area only for the three lowest concentrations (see Fig. 6).

At high concentration ranges (e.g. 0–300 ppm), the log conductance – log concentration response curve shows a non-linear behavior, for every working point of the response pattern (Fig. 7a). The response at high working temperatures and at the beginning of the low temperatures ($t < 7.5\ s$) follows a slight quadratic behavior (it cannot be seen in the figure), whereas for most working points in the low heating temperatures ($t > 7.5\ s$) the response might be approximated by a higher order polynomial. Because the IUPAC univariate LOD formulas are based on a linear regression model, no working point satisfied the LOD assumptions in this concentration range. We found that the logarithm of the sensor conductance in a reduced concentration range (0–9 ppm) linearized the response for most working temperatures. In this reduced concentration range, the working point can affect both the slope of the calibration graph (sensitivity) and the distribution of the data in each concentration level. In the example shown in Fig. 7b, the sensor output at the working point $t = 6.3$ s showed more sensitivity and less variability than at the nominal working point ($t = 25$ s). This resulted in a lower LOD estimate ($x_D = 1.82$ versus 4.77 ppm) and smaller residual error of the calibration model ($s_{y,x} =$ 0.55 versus 1.29 ppm). The reliability of the LOD estimate was also higher at $t = 6.3$ s because the LOD assumptions were satisfied, whereas at $t = 25$ s the homoscedasticity assumption was violated. The Levene's test determined that there was more variability at 2.2 ppm than at 4.4 ppm. Predictions with unbiased, normally distributed and homoscedastic residuals are the basis of LOD theory. The LOD values obtained when any of these assumption is violated are not trustworthy. This highlights the importance of assessing the validity of the underlying LOD assumptions before reporting a LOD value.

To find the optimum working point for each sensor, the procedure described in Section 2.6 was applied. The measurements corresponding to one experimental day (d = 11) were excluded from the dataset due to abnormal readings caused by an external factor. The MFC controlling the CO injection channel reported

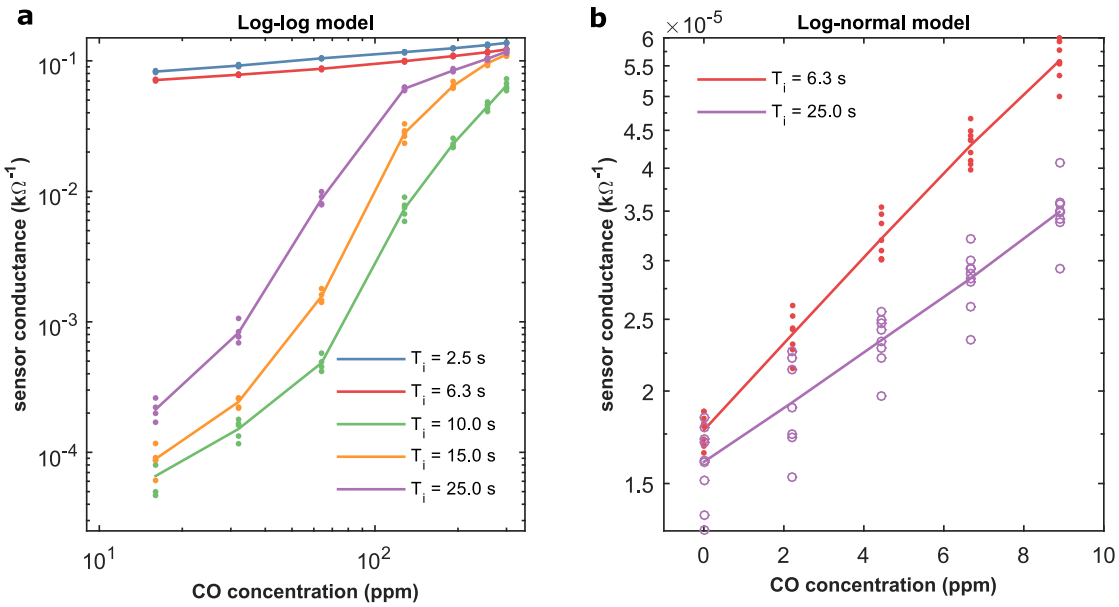


Fig. 7. Dependence between sensor conductance and CO concentration at several detection points of the heating cycle. (a) log-log model in the range 0–300 ppm, (b) log-normal model in the range 0–9 ppm. In both (a) and (b), each point is the average of ten replicate measurements each one having a random humidity value in the range 30–70% r.h.

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

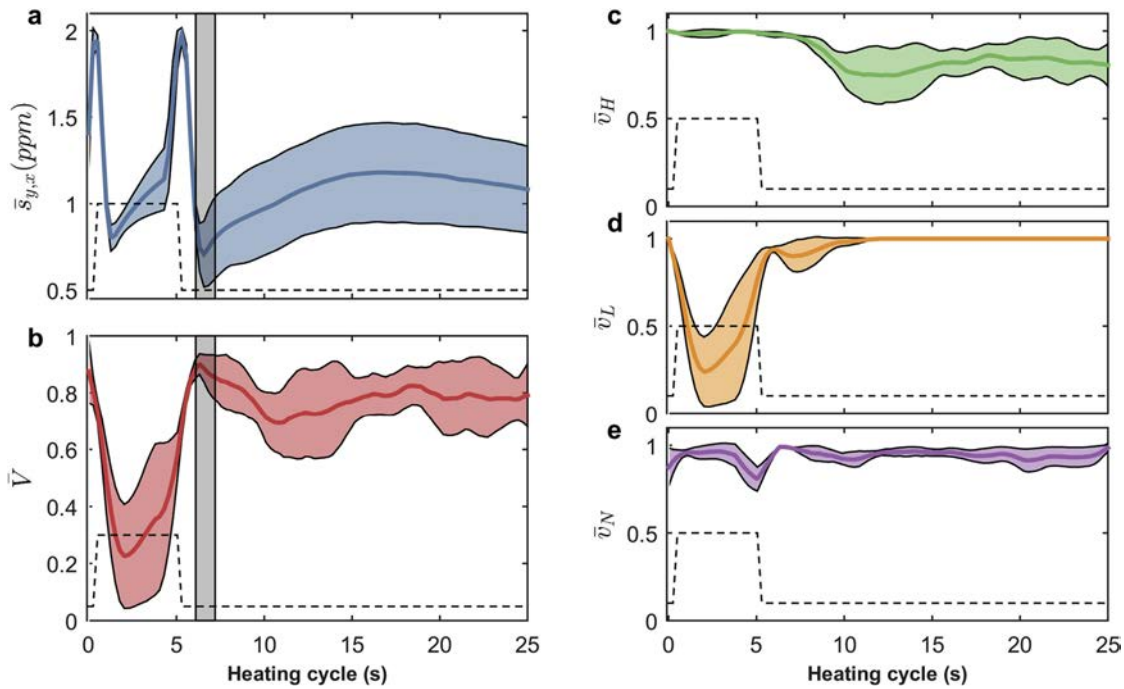


Fig. 8. Main parameters for the optimization of the working point in the log-normal calibration model (Eq. (3)) in the range 0–9 ppm, as a function of the heating cycle. (a) Standard error of regression; (b) validity of x_D; (c–e) validity of the homoscedasticity, linearity and normality assumptions, respectively. The values were first averaged across calibration days and the dispersion of the average values among sensor units is represented as an area plot. The solid line surrounded by a shaded area indicate the mean ± 1 mean absolute deviation (MAD). MAD $= \frac{1}{n}\sum_{i=1}^{n}|x_i - mean(x)|$. The vertical gray shaded area in subplots a-b indicate the optimum working range for the sensor family.

unusually low flow rates, probably due to an accidental manipulation of the manometer that controls the injection of CO into the gas chamber was accidentally manipulated. Similarly, one of the seven MOX sensors (sensor #2) was not considered for the analysis because its response was clearly different from the other six sensors. This sensor, which probably belongs to a different fabrication batch, was acting as an outlier for the statistical analysis. As a result, the following analysis was performed using data extracted from the remaining twelve days and six sensors. Fig. 8 shows a summary of the main parameters used during the optimization process, for the entire sensor family. The values were averaged across experimental days -which can be understood as repetition of the calibration procedure-to produce more robust estimates. The first thing we can see in Fig. 8 is that the temperature transitions ($t = 0$ s and $t = 5$ s) are not good detection areas because s_{yx} was quite high, which means that the calibration model was not very accurate in this area. High heating temperatures ($t \in [0,\ 5]$ s) produced better calibration models (values of s_{yx} were moderate) but the obtained models rarely passed the linearity test. It should be recalled that values of V, v_H, v_N or v_L close to 1 indicate that the given constraint was valid in 100% of the repetitions (i.e. experimental days). The logarithmic transformation was more effective in linearizing the sensor response at low heating temperatures ($t \in [5,\ 25]$ s). The normality assumption was satisfied in both high and low heating temperatures. The data became slightly heteroscedastic in the second half of the heating cycle ($t \in [10,\ 25]$ s). The optimum working points for the six sensors were located at $t = \{6.5,\ 7.6,\ 6.8,\ 6.6,\ 6.1$ and $7.1\}$ s, respectively. The optimum working point was slightly different for each sensor but an optimum range consistent for the sensor family was found at the beginning of the low heating temperatures ($t \in [6.1,\ 7.6]$ s). Within this narrow range, minimum s_{yx} values and maximum values of V were obtained. This variability in optimum working points may be related to device tolerances regarding temperature at the sensing surface. Due to the mini-bead technology of FIS for this sensor series, we can have small tolerances both in thermal dynamics and steady temperature values at low power cycle. These small variations on the surface temperature, then will lead to variations in the chemical dynamics of the reactions taking place and ultimately instability of the optimum working point across devices.

Using the optimum working point of each sensor s, calibration models were computed for every experimental day d. The LOD $x_D(s,d)$ and the validity of the LOD $V(s,d)$ of these models were averaged across experimental days to produce more robust estimates (Eqs. (27) and (28)). A comparison between $\overline{x}_D(s)$ and $\overline{V}(s)$ and the equivalent values obtained at the nominal working point is shown in Fig. 9. Compared to the nominal detection point, the optimized models improved both the mean LOD (2.3 versus 4.0 ppm) and the mean validity (95% versus 80%). Because the optimum LOD value is at the low end of the concentration range tested in this work, to fully confirm the results more tests in a reduced range centered around the LOD (e.g. 0–4 ppm) might be required.

If the LOD value is going to be taken as reference for future operation, it is key to check the stability of this parameter. In other words, it is important to check if the calibration model remains stable. The prediction error may increase due to a shift of the data, a change of sensitivity or an increase in the variance. In any of these

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

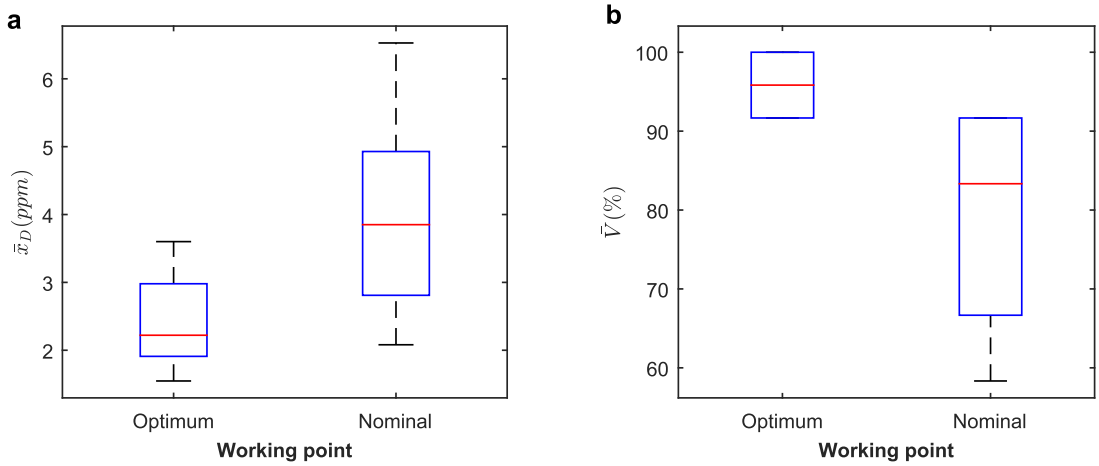


Fig. 9. Comparison between the optimized working point and the nominal working point in terms of (a) LOD and (b) its validity. The values were first averaged across calibration days and the dispersion of the average values among sensor units is represented as a box plot. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers.

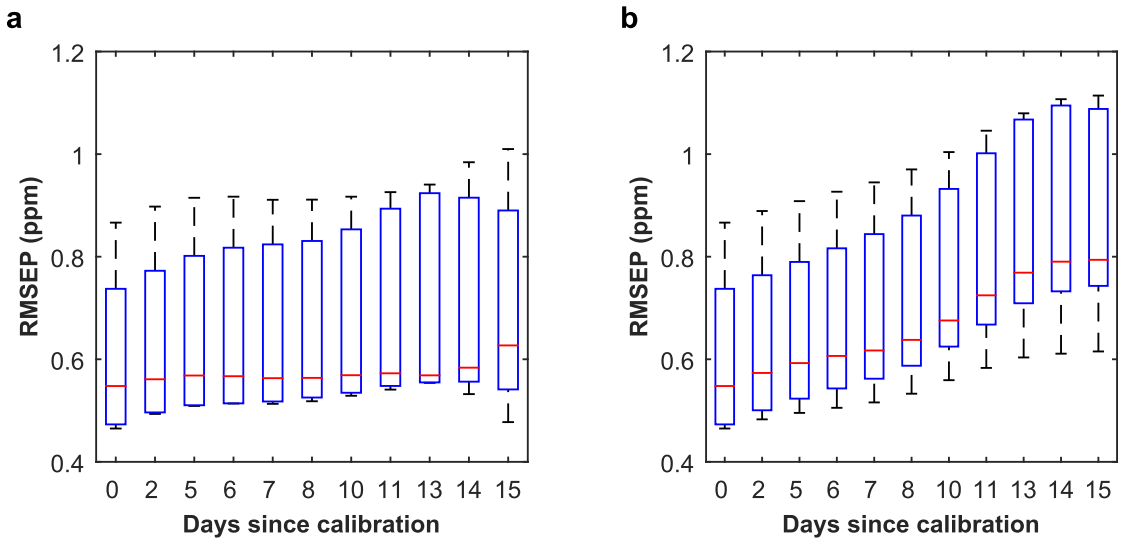


Fig. 10. Temporal evolution of the RMSEP of the optimized calibration models in the range 0–9 ppm. (a) Baseline correction; (b) No baseline correction. The dispersion of the values among sensor units is represented as a box plot. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. The x-axis indicates the elapsed time between the calibration day d_c and the validation day d_v. Each x value combines those groups of days in which $\{d_v - d_c = x\}$. $x = 0$ represents the RMSE in calibration.

situations, the reported LOD will not represent the detection capabilities of the sensor beyond the calibration day. The temporal evolution of the mean RMSEP of the optimized models is presented in Fig. 10. In the left graphic, we can see that the mean RMSEP slightly increased as the elapsed time between calibration and validation increased. In 15 days, the average RMSEP increased 0.09 ppm (0.61 versus 0.70 ppm), which represents a 3.9% of the mean LOD value (2.3 ppm). This means that the LOD value can be trusted in the studied temporal period provided that baseline correction was performed (i.e. dividing g_s by g_{air} in Eq. (3)). This baseline subtraction method is based on the measurements of a blank sample, which might not be easy to obtain in a real scenario. Blank samples could be generated either by exposing the sensor system to a reference clean air or by directly measuring the target atmosphere when the target gas is not present. The first approach increases the complexity of the system by adding a container to store the reference gas and the extra fluidic components. The second method assumes that at certain time the gas sensor will not be exposed to the target gas or at least not to concentrations near the limit of detection. For example, background concentrations of 150 ppb can be sometimes found in “clean” air. Since this residual concentration is one order of magnitude below the estimated limit

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

of detection, the baseline correction method will still work.

If the baseline is not corrected, the stability of the models is degraded (Fig. 10b). We can see that the mean RMSEP monotonically increased as the elapsed time between calibration and validation increased. In 15 days, the average RMSEP increased 0.32 ppm (0.64 versus 0.96 ppm), which represents a 13.9% of the mean LOD value (2.3 ppm). This may put into question the reliability of the LOD in a 2-week period. Further investigation revealed that the sources of error were small shifts in the measurements from one day to another, probably due to the instability of the sensor response at blank measurements, the open-loop control of the sensor temperature and the intentional changes in the levels of interference (humidity). The sensitivity and the variance of the measurements were not significantly affected. These findings greatly encourage baseline correction whenever possible.

4. Conclusions

We have proposed a methodology to overcome the challenges associated to LOD estimation in non-linear sensors, in this case MOX sensors aiming at CO detection in the presence of uncontrolled humidity levels. A main contribution of the paper is the selection of the sensor measurement point that should show small prediction error and simultaneously fulfilled the statistical requirements imposed by the IUPAC for linear univariate LOD estimation. The calibration models with logarithmic transformation at the input in a reduced concentration range (0–9 ppm) satisfied the linearity assumption at low temperatures of the heating cycle. The working points that yielded the most accurate calibration models and highest validity ratios of the LOD assumptions were all found at the beginning of the low level of the heating cycle ($t \in [6.1,\ 7.6]$ s). Average LOD values of 2.3 ppm and average validity ratios of 95% were found in the optimized models. The nominal working point suggested by the manufacturer yielded higher LOD values (4.0 ppm) and satisfied the LOD assumptions only in 80% of the repetitions.

The calibration models were stable in a scenario of varying levels of interferences, if baseline correction was performed. In this case, a maximum RMSEP increase of 0.09 ppm was observed (3.9% of the mean LOD value). This suggests that the LOD obtained in calibration is representative of the detection capabilities of the analytical system at least 2 weeks after calibration. Nonetheless, longer experiments might be required to confirm these results. If the changes in baseline were not corrected, the average RMSEP increased 0.32 ppm in 15 days, which represents a 13.9% of the mean LOD value. This might question the reliability of the LOD estimates in the studied time period.

This study demonstrated that linearized MOX sensors calibrated with optimized univariate features and periodic baseline adjustments can be used for gas sensing applications requiring a detection threshold. However, we should distinguish between two scenarios: (i) the LOD required by the application is much higher than the estimated LOD of a given sensor; (ii) the LOD required by the application is comparable to the sensor's LOD. Whereas in the first case the sensor will definitely meet the application requirements, in the second case the use of MOX sensors is questionable. It should be noted that MOX sensors are characterized by their intrinsic instability and, therefore, taking the LOD as a hard threshold in the second scenario might result in more false positives or false negatives than it was expected. This advice is especially important in critical applications in which a false positive or a false negative could result in legal consequences. Due to the high tolerance in sensitivity between MOX sensor units, the values reported in this work are specific to the sensor units used in this work. If the same experiments were repeated using different units the results might slightly change.

The main limitation of the current proposal is that the LOD must lie within the range of linear behavior. The current approach will fail if the LOD is larger than this limit. Additionally, it must be remarked that the optimum working temperature and the reported LOD values are specific to the scenario of carbon monoxide as target gas and humidity as the main interference. If the sensors were exposed to a complex gas mixture, the LOD values will probably increase. The methodology proposed in this paper could be applied to find a better working temperature for the given gas mixture.

Acknowledgements

This work was partially funded by the Spanish MINECO program, under grants TEC2014-59229-R (SIGVOL), PCIN-2013-195 (SENSIBLE) and BES-2015-071698 (SEVERO-OCHOA). The Signal and Information Processing for Sensor Systems group is a consolidated Grup de Recerca de la Generalitat de Catalunya and has support from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2014-SGR-1445). This work has received support from the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya and the European Social Fund. Additional financial support has been provided by the Institut de Bioenginyeria de Catalunya (IBEC). IBEC is a member of the CERCA Programme/Generalitat de Catalunya.

References

[1] EPA, Criteria air pollutants, Am. Child. Environ. (2015) 1–22. http://www.epa.gov/criteria-air-pollutants.
[2] P.K. Sekhar, E.L. Brosha, R. Mukundan, W. Li, M.A. Nelson, P. Palanisamy, F.H. Garzon, Application of commercial automotive sensor manufacturing methods for NOx/NH3 mixed potential sensors for on-board emissions control, Sensor. Actuator. B Chem. 144 (2010) 112–119, https://doi.org/10.1016/j.snb.2009.10.045.
[3] E. Beard, R. West, Pilot study of the use of personal carbon monoxide monitoring to achieve radical smoking reduction, J. Smok. Cessat. 7 (2012) 12–17.
[4] L.A. Currie, Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995), Pure Appl. Chem. 67 (1995) 1699–1723.
[5] A. Burresi, A. Fort, S. Rocchi, B. Serrano, N. Ulivieri, V. Vignoli, Dynamic CO recognition in presence of interfering gases by using one MOX sensor and a selected temperature profile, Sensor. Actuator. B Chem. 106 (2005) 40–43.
[6] C. Bur, M. Bastuck, D. Puglisi, A. Schuetze, A.L. Spetz, M. Andersson, Discrimination and quantification of volatile organic compounds in the ppb-range with gas sensitive SiC-FETs using multivariate statistics, Sensor. Actuator. B Chem. 214 (2015) 225–233.
[7] S. Udina, M. Carmona, G. Carles, J. Santander, L. Fonseca, S. Marco, A micromachined thermoelectric sensor for natural gas analysis: thermal model and experimental results, Sensor. Actuator. B Chem. 134 (2008) 551–558.
[8] J.W. Gardner, P.K. Guha, F. Udrea, J.A. Covington, CMOS interfacing for integrated gas Sensors : a review, IEEE Sens. J. 10 (2010), https://doi.org/10.1109/JSEN.2010.2046409.
[9] A.C. Romain, J. Nicolas, Long term stability of metal oxide-based gas sensors for e-nose environmental applications: an overview, Sensor. Actuator. B Chem. 146 (2010) 502–506, https://doi.org/10.1016/j.snb.2009.12.027.
[10] M. Padilla, J. Fonollosa, S. Marco, Improving the Robustness of Odour Sensing Systems by Multivariate Signal Processing, 2013. Hum. Olfactory Displays Interfaces Odor Sens. Present. Inf. Sci. Ref. IGI-Global.
[11] L. Valentini, I. Armentano, J.M. Kenny, C. Cantalini, L. Lozzi, S. Santucci, Sensors for sub-ppm NO 2 gas detection based on carbon nanotube thin films, Appl. Phys. Lett. 82 (2003) 961–963.
[12] J. Nicolas, A.-C. Romain, Establishing the limit of detection and the resolution limits of odorous sources in the environment for an array of metal oxide gas sensors, Sensor. Actuator. B Chem. 99 (2004) 384–392, https://doi.org/10.1016/j.snb.2003.11.036.
[13] J. Fonollosa, A. Vergara, R. Huerta, S. Marco, Estimation of the limit of detection using information theory measures, Anal. Chim. Acta 810 (2013) 2013–2015, https://doi.org/10.1016/j.aca.2013.10.030.
[14] A.M.Z.B. Erna, S.T.T. Rowell, W.I.E.S.C. Ynkar, Comparison of metal oxide-based electronic nose and mass spectrometry-based electronic nose for the prediction of red wine spoilage, J. Agric. Food Chem. 56 (2008) 3238–3244.
[15] R. Boqué, M.S. Larrechi, F.X. Rius, Multivariate detection limits with fixed probabilities of error, Chemometr. Intell. Lab. Syst. 45 (1999) 397–408, https://doi.org/10.1016/S0169-7439(98)00195-6.

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

[16] C. Borrego, A.M. Costa, J. Ginja, M. Amorim, M. Coutinho, K. Karatzas, T. Sioumis, N. Katsifarakis, K. Konstantinidis, S. De Vito, E. Esposito, P. Smith, N. André, P. Gérard, L.A. Francis, N. Castell, P. Schneider, M. Viana, M.C. Minguillón, W. Reimringer, R.P. Otjes, O. v Sicard, R. Pohle, B. Elen, D. Suriano, V. Pfister, M. Prato, S. Dipinto, M. Penza, Assessment of air quality microsensors versus reference methods: the EuNetAir joint exercise, Atmos. Environ. 147 (2016), https://doi.org/10.1016/j.atmosenv.2016.09.050.

[17] M. Kuske, M. Padilla, A.C. Romain, J. Nicolas, R. Rubio, S. Marco, Detection of diverse mould species growing on building materials by gas sensor arrays and pattern recognition, Sensor. Actuator. B Chem. 119 (2006) 33–40, https://doi.org/10.1016/j.snb.2005.02.059.

[18] T. Hübert, L. Boon-Brett, V. Palmisano, M.A. Bader, Developments in gas sensor technology for hydrogen safety, Int. J. Hydrogen Energy 39 (2014) 20474–20483.

[19] A. Loutfi, S. Coradeschi, G.K. Mani, P. Shankar, J.B.B. Rayappan, Electronic noses for food quality: a review, J. Food Eng. 144 (2015) 103–111, https://doi.org/10.1016/j.jfoodeng.2014.07.019.

[20] M. Righettoni, A. Amann, S.E. Pratsinis, Breath analysis by nanostructured metal oxides as chemo-resistive gas sensors, Mater. Today 18 (2015) 163–171.

[21] G. Neri, First fifty years of chemoresistive gas sensors, Chemosensors 3 (2015) 1–20.

[22] C. Di Natale, R. Paolesse, E. Martinelli, R. Capuano, Solid-state gas sensors for breath analysis: a review, Anal. Chim. Acta 824 (2014) 1–17.

[23] F.U.S.A. Inc, TGS203-Carbon Monoxide Sensor Specification & Technical Information for Carbon Monoxide Sensors, 1999 pdf.dzsc.com/88888/2008617161056961.pdf.

[24] P. Van Geloven, M. Honore, J. Roggen, S. Leppavuori, T. Rantala, The influence of relative humidity on the response of tin oxide gas sensors to carbon monoxide, Sensor. Actuator. B Chem. 4 (1991) 185–188.

[25] G. Korotcenkov, B.K. Cho, Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey), Sensor. Actuator. B Chem. 156 (2011) 527–538, https://doi.org/10.1016/j.snb.2011.02.024.

[26] G. Korotcenkov, B.K. Cho, Engineering approaches to improvement of conductometric gas sensor parameters. Part 2: decrease of dissipated (consumable) power and improvement stability and reliability, Sensor. Actuator. B Chem. 198 (2014) 316–341.

[27] A.C. Olivieri, N.M. Faber, J. Ferré, R. Boqué, J.H. Kalivas, H. Mark, Uncertainty estimation and figures of merit for multivariate calibration (IUPAC Technical Report), Pure Appl. Chem. 78 (2006) 633–661, https://doi.org/10.1351/pac200678030633.

[28] A.M. de la Peña, A. Espinosa-Mansilla, M.I.A. Valenzuela, H.C. Goicoechea, A.C. Olivieri, Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxycillin and clavulanic acid by stopped-flow kinetic analysis, Anal. Chim. Acta 463 (2002) 75–88.

[29] N. Barsan, U. Weimar, Understanding the fundamental principles of metal oxide based gas sensors; the example of CO sensing with SnO2 sensors in the presence of humidity, J. Phys. Condens. Matter 15 (2003) R813.

[30] F.I.S. Inc, FIS GAS SENSOR SB-500-12, 2017. http://www.fisinc.co.jp/en/common/pdf/SB50012E%7B_%7DP.pdf.

[31] T. Conrad, P. Hiry, A. Schutze, PuMaH-a temperature control and resistance read-out system for microstructured gas sensors based on PWM signals, in: Sensors Conference, 2005 IEEE, 2005, p. 8.

[32] I. Lavagnini, F. Magno, A statistical overview on univariate calibration, inverse regression, and detection limits: application to gas chromatography/mass spectrometry technique, Mass Spectrom. Rev. 26 (2007) 1–18.

[33] J. Mocák, I. Janiga, E. Rábarov Á, Evaluation of IUPAC limit of detection and iso minimum detectable value electrochemical determination of lead, Nov. Biotechnol 9 (2009) 91–100.

[34] G. Nelson, Gas mixtures: Preparation and Control, CRC Press, 1992.

[35] P.K. Clifford, D.T. Tuma, Characteristics of semiconductor gas sensors II. Transient response to temperature change, Sensor. Actuator. 3 (1982) 255–281.

[36] K. Danzer, L.A. Currie, Guidelines for calibration in analytical chemistry. Part I. Fundamentals and single component calibration (IUPAC Recommendations 1998), Pure Appl. Chem. 70 (1998) 993–1014.

[37] L.A. Currie, Detection: International update, and some emerging di-lemmas involving calibration, the blank, and multiple detection decisions, Chemometr. Intell. Lab. Syst. 37 (1997) 151–181.

[38] H. Levene, Contributions to Probability and Statistics, Essays Honor Harold Hotelling, Stanford University Press, 1960, pp. 278–292.

[39] S.S. Shapiro, M.B. Wilk, An analysis of variance test for normality (complete samples), Biometrika 52 (1965) 591–611.

[40] N.M. Razali, Y.B. Wah, others, Power comparisons of shapiro-wilk, Kolmogorov-smirnov, lilliefors and anderson-darling tests, J. Stat. Model. Anal 2 (2011) 21–33.

[41] G.E.P. Box, W.G. Hunter, J.S. Hunter, Statistics for Experimenters: an Introduction to Design, Data Analysis, and Model Building, JSTOR, 1978.

[42] S. Holm, A simple sequentially rejective multiple test procedure, Scand. J. Stat. (1979) 65–70.

PAPER I

Please cite this article in press as: J. Burgués, et al., Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.01.062

Analytica Chimica Acta 1019 (2018) 49–64

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors

Javier Burgués [a, b, *], Santiago Marco [a, b]

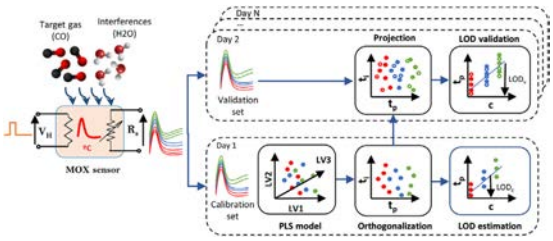
[a] Department of Electronic and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028, Barcelona, Spain
[b] Signal and Information Processing for Sensing Systems, Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028, Barcelona, Spain

PAPER II

HIGHLIGHTS

- A methodology to estimate the LOD in orthogonalized PLS models is proposed.
- It reveals the role that interferences (chemical noise) play in the model structure.
- The methodology allows the estimation of the LOD in temperature modulated metal oxide sensors.
- As an illustration, we estimated the LOD of carbon monoxide in the presence of uncontrolled humidity.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 28 August 2017
Received in revised form
2 March 2018
Accepted 5 March 2018
Available online 20 March 2018

Keywords:
Metal oxide sensors
Partial least squares
Orthogonal projection to latent structures
Maximum likelihood principal component regression
Limit of detection
Temperature modulation

ABSTRACT

Metal oxide semiconductor (MOX) sensors are usually temperature-modulated and calibrated with multivariate models such as partial least squares (PLS) to increase the inherent low selectivity of this technology. The multivariate sensor response patterns exhibit heteroscedastic and correlated noise, which suggests that maximum likelihood methods should outperform PLS. One contribution of this paper is the comparison between PLS and maximum likelihood principal components regression (MLPCR) in MOX sensors. PLS is often criticized by the lack of interpretability when the model complexity increases beyond the chemical rank of the problem. This happens in MOX sensors due to cross-sensitivities to interferences, such as temperature or humidity and non-linearity. Additionally, the estimation of fundamental figures of merit, such as the limit of detection (LOD), is still not standardized in multivariate models. Orthogonalization methods, such as orthogonal projection to latent structures (O-PLS), have been successfully applied in other fields to reduce the complexity of PLS models. In this work, we propose a LOD estimation method based on applying the well-accepted univariate LOD formulas to the scores of the first component of an orthogonal PLS model. The resulting LOD is compared to the multivariate LOD range derived from error-propagation. The methodology is applied to data extracted from temperature-modulated MOX sensors (FIS SB-500-12 and Figaro TGS 3870-A04), aiming at the detection of low concentrations of carbon monoxide in the presence of uncontrolled humidity (chemical noise). We found that PLS models were simpler and more accurate than MLPCR models. Average LOD values of 0.79 ppm (FIS) and 1.06 ppm (Figaro) were found using the approach described in this paper. These values were contained within the LOD ranges obtained with the error-propagation approach. The mean LOD increased to 1.13 ppm (FIS) and 1.59 ppm (Figaro) when considering validation samples collected two weeks after calibration, which represents a 43% and 46% degradation,

* Corresponding author. Signal and Information Processing for Sensing Systems, Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona Spain.
E-mail address: jburgues@ibecbarcelona.eu (J. Burgués).

https://doi.org/10.1016/j.aca.2018.03.005

respectively. The orthogonal score-plot was a very convenient tool to visualize MOX sensor data and to validate the LOD estimates.

1. Introduction

Gas sensor technology is evolving towards miniaturized devices with low power consumption that can be integrated in smartphones and wearable devices [1–6] .With millions of units sold worldwide each year, metal oxide semiconductor (MOX) is the most commercially successful gas sensor technology. MOX sensors have been used to solve problems in a number of fields: energy [7], food [8,9], automotive [10–12], environmental monitoring [13–16], safety and security [17–21] or biomedicine [22,23]. MOX sensors offer high sensitivity to many volatile organic compounds (VOCs), low fabrication cost, miniaturization potential thanks to microelectromechanical systems (MEMS) technology and are very robust devices which can last for decades. However, after more than 50 years in the market, MOX sensors still suffer from very important limitations, namely poor stability [24,25], low selectivity [26], slow recovery time [27,28] and large tolerances in specifications for identical devices [29–31]. Among those, the lack of selectivity and stability (temporal drift) might be considered the most important ones.

The transduction mechanism of MOX sensors consists on measuring resistance changes resulting from chemical interactions between the gas species and the metal oxide surface. Clifford-Tuma [32] found that the response of a S_nO_2 based Taguchi-type sensor in the presence of CO, H_2O and CH_4 follows a power law which contains products between several gas concentrations and quadratic terms. This non-linear response is due to competitive and synergistic interactions among gases that make the effect of the target gas can be either masked or enhanced by the presence of other gases (i.e. "matrix-effect"). The sensor response is also cross-sensitive to environmental factors, such as temperature, humidity or pressure. The temperature of the sensing surface affects the rate at which the chemical reactions occur. There is usually an optimum working temperature (in terms of sensitivity) for each target gas but only limited selectivity can be achieved by working at a constant temperature. Selectivity to a given analyte can be increased by modulating the operating temperature with a periodic heating power waveform [33]. The resulting multivariate response patterns capture the selectivity-temperature dependence characteristic of the target gas plus the effect of the interferences. The variables of the response pattern are highly correlated because the shape of the heating waveform smoothly varies the sensor temperature (except for the temperature transitions).

Using a set of calibration samples with known concentration of the target gas and varying contributions of potential interferences, a multivariate calibration model is built to predict the concentration of the target gas in future samples. Temperature-modulated MOX sensors are usually calibrated with partial least squares (PLS) [34] models, which can deal with the highly collinear response patterns. The good performance of PLS from a theoretical point of view has been discussed by Helland [35,36]. PLS works optimally when the measurement errors are independent and identically distributed (iid). Because temperature-modulated MOX sensors usually present heteroscedastic and correlated measurement errors (i.e. non-iid), maximum likelihood methods, such as maximum likelihood principal components regression (MLPCR) [37], should in principle outperform PLS. MLPCR takes into account the measurement error structure for building the multivariate model. However, in a comparison between MLPCR and PLS, M.S. Reis et al. [38] have found that PLS seems to be more effective than MLPCR to reject heteroscedasticity. In fact, PLS prediction errors are close to those found by MLPCR in most scenarios. Reis attributes the good results of PLS to the effective way of this method to find a low dimensional predictive space, onto which the regressors are projected prior to being used for predicting the response. This projection is an effective filter that minimizes the effect of noise in X, regardless of the structure of the noise. The authors of this study also claim that the use of orthogonal projections benefit PLS for heteroscedastic data, when compared to maximum likelihood projections used by MLPCR. However, there are no references in the available literature regarding the suitability of MLPCR for the calibration of temperature-modulated MOX sensors. One contribution of this paper is the comparison between PLS and MLPCR in modelling temperature-modulated MOX sensor data, in a concentration range near the limit of detection (LOD).

After the calibration model is built, it is common to report figures of merit, such as sensitivity, selectivity, precision or LOD, that characterize the chemical measurement process. The LOD can be considered the most relevant figure of merit for a myriad of applications requiring the detection of low concentrations of the target gases [10,11,13,39]. Despite of its relevancy, the estimation of the LOD in multivariate calibration is not standardized (see a recent review in [40]). The main issue is the lack of agreement between two alternative methods to estimate the LOD. The LOD is a consequence of uncertainty in analyte prediction at the blank level. As it is pointed out by Olivieri in [41], this uncertainty in prediction can be obtained either (i) analytically, by propagating the errors of the calibration concentrations and measured signals or (ii) empirically, by carrying out predictions on validation samples.

In the error-propagation approach, the uncertainties of the measured signals (both calibration and test samples) and the calibration concentrations are propagated through the model to estimate the variance in the predicted concentration of an unknown sample. Then, a sample-specific LOD can be computed from the prediction interval of blank samples. The most complete and general formula for the multivariate prediction error was proposed by Faber and Kowalski [42]. It incorporates the above sources of uncertainty plus the potential bias introduced by lack of model adequacy to the test sample. When a preprocessing method such as, orthogonal signal correction (OSC) or net analyte signal (NAS) is applied, Olivieri [41] shows that a correction needs to be applied to this formula to account for the uncertainty introduced by the filtering matrix. The uncertainty of the measured signals and the model parameters (due to variance in the calibration data) take the form of covariance matrices [43], which introduce complexity into the formulas. To simplify the covariance matrices to scalar values, it is assumed that the errors are iid. However, this assumption is unrealistic in MOX response patterns because the variables are highly correlated, and the errors of the variables close to the temperature transitions tend to be higher than at smoother temperature areas. Moreover, the formula assumes that real data is a contribution of additive signal and random noise and, as we previously discussed, MOX sensor response is highly non-linear. A practical limitation of using this formula is that the variance of the

errors must be estimated a priori (e.g. from replicates), which is costly and, sometimes, impractical. The Faber and Kowalski formula [42] was tested in Montecarlo simulations [44], yielding unexpected results. Nonetheless, Allegrini et al. [45] developed a LOD definition based on the formulations of Faber and Kowalski [42], particularized for PLS models. The result is a range of LOD values: from the lowest (i.e. *LODmin*), corresponding to the lowest blank leverage, to the largest (i.e. *LODmax*), corresponding to the highest blank leverage. The LOD range clearly reflects that the composition of the sample in multivariate calibration affects the LOD estimate. However, the detection decision is slightly complex when the predicted concentration c of the test sample falls within the LOD range (i.e. $LODmin < c < LODmax$). In this case, it is unclear whether the decision to be made should be "detected" or "not detected". Allegrini et al. [45] propose to recompute a specific LOD for the test sample, approximating the leverage of the test sample by that of a blank. Then, the predicted concentration is compared against this single LOD value.

In the empirical approach, the uncertainty in prediction is estimated through a univariate "surrogate" variable computed on the calibration or (ideally) validation samples. The use of cross-validation for error estimation is a well-accepted practice and has been recommended among others by Efron and Gong [57]. The well-accepted univariate LOD formula [46] is then applied to the pseudounivariate calibration graph that relates the univariate surrogate variable to the analyte concentration. Examples of surrogate variables are the net analyte signal (NAS) [47], the predicted concentrations [48] and the scores of the first component in principal component analysis (PCA) [49,50]. Among these, the approach based on the NAS is supported by the IUPAC [51]. The NAS of a certain sample is (a vector) defined as the part of the gross signal which can be uniquely attributed to the target analyte. The norm of the NAS (a scalar) can be used as the surrogate variable. However, there are different approaches to compute the NAS [47,52–55] which produce slightly different outcomes [56]. This is an issue because the resulting LOD estimates might differ depending on which method was used to estimate the NAS. The use of the predicted concentrations as the pseudounivariate variable resorts to the mathematical proof that the LOD is invariant for linear transformations of the response variable [48]. The predictions given by the model are obtained by multiplying the measured signals by the regression vector, which is a linear operation. However, this ignores the uncertainty in the estimated model parameters resulting from the use of noisy calibration data. Finally, computing the LOD directly from the scores of the model [49,50] is a promising approach because there is a unique way to compute the scores and they reveal the underlying structure of the model. However, this approach is limited to those cases in which the input data can be sufficiently well represented by a single component. In temperature-modulated MOX sensors, usually more than one PLS component is necessary to properly fit the calibration data. Due to the non-linearities and noise in the sensor conductance patterns that are not relevant for predicting the concentration of the target gas, several PLS components are needed to model one hyperplane or "latent direction" that relates the predictors to the predictand. Thus, the percentage of variance captured by the model is distributed among several latent variables and the LOD formulas based on the first component cannot be used.

If the error of the instrumental signals is known a priori, the LOD formula based on error-propagation should in principle provide more accurate results than any empirical approach. However, it is not usually the case that measurement errors (including all potential sources of error) are known a priori. In this case, replicate measurements must be used to compute the error covariance matrix, prior to applying the error-propagation formula. In multivariate calibration, these "replicate" measurements must contain the potential interferences that the analytical system might find during operation. The empirical approach does not strictly require replicate measurements but the obtained LOD is an average value instead of a sample-specific LOD. In this case, it is unclear whether the resulting LOD truly represents the degree of uncertainty of prediction at the blank level. Nevertheless, if replicate measurements are available (including blank replicates), then the empirical approach can estimate the uncertainty at the blank level by computing the errors of the model (i.e. the variance of the surrogate variable) in the blank replicates. One research question that this paper will address is how different is the LOD computed analytically or computed empirically, when replicate measurements are available.

As the number of latent factors increases, not only the computation of the LOD becomes difficult but the interpretability of the model decreases [58]. A simpler model is advantageous to understand the underlying structure of the model by, for example, answering some of the following questions: How is the model induced from the data? Which variables are relevant for prediction? Which variables/observations are more affected by the interferences? Which observations can be considered outliers? Is there any time-based trend in the observations? To address the interpretability issues of PLS models mentioned above, Trygg et. al. [59] proposed a new method called orthogonal projections to latent structures (O-PLS). In the O-PLS algorithm, the structured noise in the input signals is removed by an integral OSC filter, prior to PLS modelling. This yields a more simpler and relevant model with the same predictive power as the original PLS model. An extension of O-PLS, called O2-PLS, was presented by the same authors [60]. The advantage of O2-PLS over its predecessor is that it can estimate the pure constituent profiles, in the case of multiple predictands. However, for a single predictand case O-PLS and O2-PLS are identical. Ergon [61] proved that the same results of O-PLS could be obtained by post-processing the PLS model using a similarity transformation (ST). The resulting algorithm was named PLS + ST. O-PLS has been extensively applied in a wide diversity of fields such as spectroscopy [62], metabolomics [63], bioinformatics [64] and process monitoring [65]. Surprisingly, 15 years after its invention (as of July 2017) we have not found any work applying O-PLS, PLS + ST or similar methods to chemical sensor data. An alternative method to simplify a PLS model into a two component model was proposed by Ergon [66]. The idea behind 2PLS (also referred as bi-orthogonal PLS in [67]) is to keep the first PLS latent variable (as it is considered by some authors the best estimate of the spectral profile of the target gas) and compress the remaining latent variables (which in theory model the noise and interferences) into a new variable, allowing a 2D representation of the model. The usage of 2PLS in the literature is very scarce and has been exclusively used in process monitoring [68].

We consider that O-PLS (equivalently PLS + ST) provide a more convenient representation than 2PLS because the first component of the simplified model is in the direction of the regression vector (instead of the first latent variable, as in 2PLS). The importance of the regression vector is that it models the part of the signal that is useful for prediction. Thus, the projection of the scores onto the first component describes the variations of the predictors (x) related exclusively to the predictand (y). We will call the first component the predictive component. Taking advantage of this property, we propose a new method to estimate the LOD by applying the well-accepted univariate formulas [46] to the predictive component of the orthogonalized model. We consider this is an improvement over existing methods based on the first PCA component [49,50] because our approach ensures that the first component condenses all the variance related to the analyte

concentration. This approach yields the same results as the IUPAC accepted multivariate formulas based on the NAS [51], when the NAS is computed through the regression vector [53]. It is also equivalent to the method proposed by Ortiz et. al [48] in which the surrogate variable are the concentrations predicted by the model. Compared to these methods, the main advantage of computing the LOD through the orthogonal model is that the role of the interferences and the target gas can be visualized, which is beneficial for interpreting the LOD.

An underlying assumption when building a calibration model is that the characteristics of the prediction samples are the same as for the calibration samples. This is usually the case in analytical instruments, such as spectrometers, which are highly selective and usually calibrated and operated in laboratory conditions. However, MOX sensors are calibrated in laboratory conditions and then are operated without supervision outside the lab: in domestic scenarios (e.g. CO toxic alarm), mobile platforms (e.g. smartphone, robot) or in environmental samples (e.g. monitoring traces of contaminants), among others. The greater variability of the environmental factors in such scenarios, the more complex composition of the gas mixtures or even device shift will introduce bias and variability in the predictions. Subsequently, the LOD obtained in calibration might not be representative of the model performance in the target scenario. In this work, we will show that the validation of the calibration model using future samples is essential to quantify the degradation of the model and report more realistic LOD estimates.

2. Materials and methods

To illustrate that orthogonal PLS (computed either by pre- or post-processing) is a very convenient tool to treat chemical sensor data, we designed a scenario of CO detection in the presence of random humidity, using MOX sensors with different degree of cross-sensitivity to environmental interferences. In this section, we describe the dataset used in the experiments and how this dataset was used to build MLPCR and orthogonalized PLS models. Finally, we explain the estimation and validation of the LOD from the orthogonal model. Figure 1 shows a block diagram of the full process.

The analysis of the data was done using MATLAB R2009a (The Mathworks) and functions *pls* and *orthogonalizepls* from the PLS_Toolbox_8.0.2 (Eigenvector Research).

2.1. Dataset

For the experiments presented in this work we used the dataset described in detail in [69]. Therefore, only a brief summary is described in this section. Seven units of two commercial MOX sensors (SB-500-12 and TGS 3870-A04, provided by FIS and Figaro, respectively) were exposed to dynamic mixtures of CO and humid synthetic air. The heater voltage was modulated following the manufacturer recommendations (0.9 V for 5s, followed by 0.2 V for 20s (FIS) or 15s (FIGARO)). The resulting multivariate sensor conductance $\boldsymbol{g}_s$ was measured using a voltage divider with a load resistor of 1 MΩ. The sensors were pre-heated for one week before starting the experiments.

Dynamic mixtures of CO and humid synthetic air were delivered from high purity gases in cylinders to a gas chamber, by means of a piping system and mass flow controllers (MFCs). Each experiment consisted on 100 measurements: 10 experimental concentrations uniformly distributed in the range 0–20 ppm and 10 replicates per concentration (see Fig. 2). Each replicate had a relative humidity randomly chosen from a uniform distribution between 15 and 75% r.h. A temperature/humidity sensor (SHT75, from Sensirion) was used to measure the humidity. The grid of concentration versus humidity was designed irregularly as we considered it was more representative of a real scenario. The samples were presented to the sensors in random order, using an exposure time of 15 minutes per sample. This exposure time ensured that the sensors reached stable responses and it was determined by the size of the gas chamber and the gas flow rates used. A single experiment lasted 25 hours (100 samples x 15 minutes/sample) and was replicated on 13 working days spanning a natural period of 17 days. The uncertainty in the concentrations generated by our system was derived through error propagation, taking into account all known sources of errors (bottle uncertainty, mass flow control reading errors, etc.). A maximum uncertainty of 0.12 ppm was associated to the lowest concentration (2.2 ppm), because the generation of this concentration requires the MFC which dilutes the CO in air to operate at the maximum

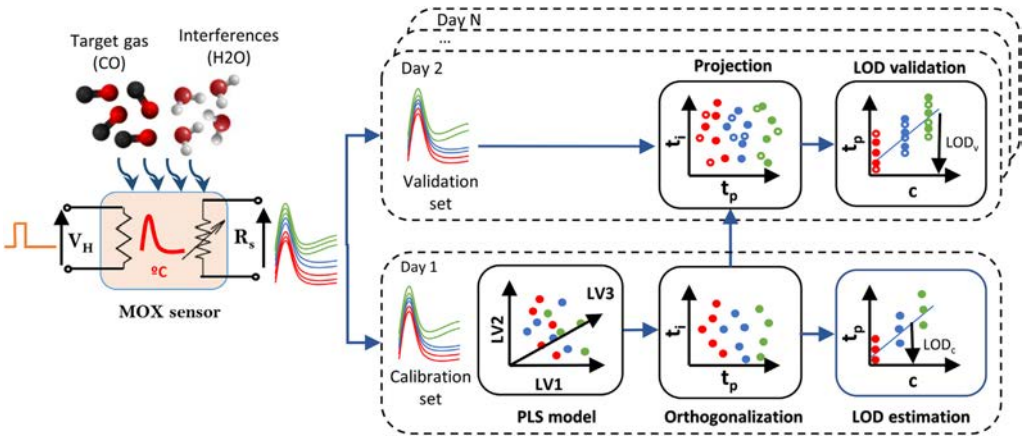


Figure 1. Block diagram of the LOD estimation/validation method described in this paper. The MOX sensor is exposed to dynamic gas mixtures of CO and Humidity. The sensor temperature is modulated using a square heating waveform (V_H). Measuring the sensor resistance (R_s) during a heating cycle produces patterns characteristic of the target gas and the interferences. A PLS model is built with the signals obtained in Day 1. This model is then orthogonalized, and the LOD is estimated from the scores of the first orthogonal component. The LOD estimate is validated by projecting signals measured in the following days into the orthogonal PLS model.

PAPER II

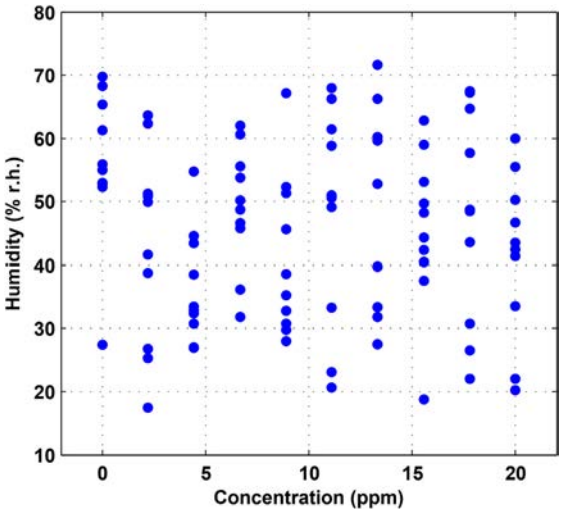


Figure 2. Combinations of CO concentration (nominal value) and relative humidity (measured using a Sensirion SHT75) that were presented to the sensors in each experimental day.

flow rate (the error of the MFC increases with the flow rate).

In each experimental day, the sensor conductance signals $\boldsymbol{g}_s$ were baseline-corrected and logarithmically transformed. The baseline correction consists on dividing the sensor conductance $\boldsymbol{g}_s$ ($k\Omega^{-1}$) by the sensor conductance in air $\boldsymbol{g}_0$ ($k\Omega^{-1}$), which was measured at the beginning of each experimental day at a random humidity and temperature. The logarithmic transformation of the response in this sensor model linearizes the response at low concentrations, which produces simpler and more accurate models [69]. The data matrix $\mathbf{X}$ ($n \times m$) consists of $n = 100$ observations (samples) of the baseline-corrected logarithmic sensor conductance patterns, sampled at $m = 100$ uniformly distributed points (variables) along the heating cycle. It should be noted that the replicates of each sample were included in the matrix $\mathbf{X}$. The vectors $\mathbf{y}$ ($n \times 1$) and $\mathbf{h}$ ($n \times 1$) contain the true CO concentration (ppm) and the humidity content (% r.h.) of the 100 samples, respectively.

2.2. PLS models

Given a matrix $\mathbf{X}$ ($n \times m$) of n samples and m predictors or variables and their corresponding concentration $\mathbf{y}$ ($n \times 1$), the PLS algorithm [70] can be used to solve the multiple linear regression problem in the inverse formulation (Eq. (1)), for the case where there are more variables than samples ($m > n$).

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \tag{1}$$

Here, $\mathbf{b}$ ($m \times 1$) is the vector of regression coefficients and $\mathbf{e}$ is the residual error. PLS solves the inverse calibration problem by finding a reduced number of underlying latent factors, also called latent variables (LVs), that account for as much as possible variance in $\mathbf{X}$ while also capturing the correlation to $\mathbf{y}$. The advantage of inverse calibration is that, because $\mathbf{y}$ is modelled as a function of $\mathbf{X}$, no explicit knowledge of the potential interfering species is required (although they must be sufficiently represented in the calibration set). The latent space is defined by a set of A basis vectors or *loadings* $\mathbf{P}$ ($m \times A$) and $\mathbf{q}$ ($1 \times A$), which give information about the weights for each predictor in $\mathbf{X}$ when calculating each latent variable. The coordinates of each sample in the latent space are the scores $\mathbf{T}$ ($n \times A$). The score matrix is able to summarize $\mathbf{X}$ and predict $\mathbf{y}$ with small errors in $\mathbf{E}$ ($n \times m$) and $\mathbf{F}$ ($n \times 1$):

$$\mathbf{X} = \mathbf{TP}^{\mathrm{T}} + \mathbf{E} \tag{2}$$

$$\mathbf{y} = \mathbf{Tq}^{\mathrm{T}} + \mathbf{F} \tag{3}$$

$\mathbf{X}$ was mean-centered and standardized to unit variance prior to PLS modelling. The vector of responses $\mathbf{y}$ was mean-centered. PLS models were built using the samples corresponding to the first experimental day, which we denote as the calibration set. The complexity of the models was optimized by performing a 5-fold cross-validation (CV). The 5-fold CV involves randomly splitting the data into five groups and fitting the model using four folds (i.e. using 80% of the samples). The remaining fold (i.e. 20% of the samples) is hold out as the test set to compute the root mean squared error (*RMSE*) (Eq. (4)). The folds are made such that the 80/20% split is maintained in each concentration level.

$$RMSE = \sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \Big/ n} \tag{4}$$

where n denotes the size of the test set, and y_i and $\hat{y}_i$ are the true and predicted value for test sample i, respectively. This process is repeated five times so that each fold is used exactly once as the test set. The optimum number of latent variables (LVs) is found by inspection of the average *RMSE* of the five folds versus number of LV curve. The final models were built using the selected number of LVs and the full calibration set.

2.2.1. Orthogonalization of the PLS models (PLS + ST)

The PLS models were post-processed with an orthogonalization step (PLS + ST). This produces simpler models with orthogonalized loadings and scores, which condense all the $\mathbf{y}$ -variance in the first weight and loading. The first loading of the orthogonalized model is in the same direction of the regression vector for data in which the structured noise has been filtered out. The predictions of the new model are identical to the non-orthogonalized model but the loadings and weights have been rotated. The second loading of the orthogonalized model captures the orthogonal variation to $\mathbf{y}$, which is related to the interferences. There are two main approaches for PLS orthogonalization [59,71], which produce similar results. In [71], a similarity transformation matrix $\mathbf{M}$ (Eq. (5)) is used to remove the orthogonal variation of $\mathbf{y}$:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ q_1^{-1}\mathbf{q}_{2:A} & \mathbf{I} \end{bmatrix} \tag{5}$$

where 0 is a matrix of zeros ,$\mathbf{I}$ is the identity matrix and $\mathbf{q}$ is the loading matrix. Using $\mathbf{M}$, Eq. (3) is transformed into Eq. (6):

$$\mathbf{y} = [\mathbf{t}_1 \quad \mathbf{T}_{2:A}]\mathbf{M}\mathbf{M}^{-1}\begin{bmatrix} q_1 \\ \mathbf{q}_{2:A} \end{bmatrix} + \mathbf{F} = \mathbf{t}_1^* q_1 + \mathbf{F} \tag{6}$$

where

$$\mathbf{t}_1^* = \mathbf{t}_1 + q_1^{-1}\mathbf{T}_{2:A}\mathbf{q}_{2:A} \tag{7}$$

$\mathbf{t}_1^*$ is a rotation of the original $\mathbf{t}_1$ scores. From Eq. (6), it can be seen that $\mathbf{t}_1^*$ has the same direction as $\mathbf{y}$.

Similarly, Eq. (2) is transformed into Eq. (8):

PAPER II

$$\mathbf{X} = [\mathbf{t}_1 \quad \mathbf{T}_{2:A}]\mathbf{M}\mathbf{M}^{-1}\begin{bmatrix}\mathbf{p}_1^{\mathbf{T}} \\ \mathbf{P}_{2:\boldsymbol{A}}^{\mathbf{T}}\end{bmatrix} + \mathbf{E} = \mathbf{t}_1^*\mathbf{p}_1^{\mathbf{T}} + \mathbf{T}_{2:A}\mathbf{P}_{2:\boldsymbol{A}}^* + \mathbf{E} \tag{8}$$

where

$$\mathbf{P}_{2:\boldsymbol{A}}^* = \mathbf{P}_{2:A} - q_1^{-1}\mathbf{p}_1\mathbf{q}_{2:\boldsymbol{A}}^{\mathbf{T}} \tag{9}$$

is the rotated loading matrix for the components orthogonal to **y**.

2.3. MLPCR models

Maximum likelihood principal component regression (MLPCR) [37] is a supervised regression technique that performs a principal component decomposition (PCA) of **X** (i.e. Eq. (2)), considering the structure of the error matrix **E**. The final models are built by ordinary least squares (OLS) regression of the scores **T** and the nominal analyte concentrations **y**. In practice, the error matrix **E** is unknown and must be estimated from a finite number of samples. In the general case, the error variance among all measurement errors can be described by the full covariance matrix **Ω** ($mn \times mn$). However, the MLPCR algorithm [72] cannot handle full covariance matrices because of computational and storage cost. Therefore, several assumptions are required to simplify the problem. First, it is assumed that correlation only exists within each row of the data matrix **X** (i.e. correlation among variables), but not within the columns (i.e. samples). In the case of error covariance in the rows, the error variance **Σ** can be described in terms of n row error covariance matrices $\boldsymbol{\Sigma}_i$ of size $m \times m$:

$$\boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\Sigma}_1 & & \\ & \ddots & \\ & & \boldsymbol{\Sigma}_n\end{bmatrix} \tag{10}$$

Each row error covariance matrix $\boldsymbol{\Sigma}_i$ ($i = 1, \ldots, n$) is estimated in practice from q replicates of sample i:

$$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{(q-1)}\sum_{k=1}^{q}\left(x_{i,k} - \overline{x}_i\right)^{\mathrm{T}} \cdot \left(x_{i,k} - \overline{x}_i\right) \tag{11}$$

where $\widehat{\boldsymbol{\Sigma}}_i$ is the estimated row error covariance matrix corresponding to the i th sample, $x_{i,k}$ is the kth replicate of sample i and $\overline{x}_i$ is the average of the q replicates.

The second assumption of MLPCR is that the row error covariance matrix is similar for each sample. In this manner, a single row error covariance matrix of size $m \times m$, pooled across the different samples, describes the error variance:

$$\widehat{\boldsymbol{\Sigma}}_{pooled} = \frac{1}{n}\sum_{i=1}^{n}\widehat{\boldsymbol{\Sigma}}_i \tag{12}$$

$\widehat{\boldsymbol{\Sigma}}_{pooled}$ is called "equal row error covariance matrix" and is the default error matrix in the MLPCR algorithm.

2.4. Limit of detection (LOD)

The LOD is defined by the International Union of Pure and Applied Chemistry (IUPAC) as the minimum concentration that can be reliably distinguished from blank samples, with stated probabilities of error [46]. In the IUPAC recommendations, the LOD is formulated in terms of hypothesis testing, interpreting the distribution of the LOD and the blanks as gaussian random variables:

$$P\left(\widehat{L} > L_C | L = 0\right) = \alpha \tag{13}$$

$$P\left(\widehat{L} < L_C | L = L_D\right) = \beta \tag{14}$$

Here L is a generic symbol that represents either net signal (blank-corrected) or analyte concentration, $\widehat{L}$ is an estimate of L, L_C is the threshold used for the detection decision (i.e. analyte present or not) and L_D is the limit of detection or LOD. α and β are the probabilities of false positives and false negatives, respectively. Solving the above equations for L_D yields:

$$L_D = z_{1-\alpha}\sigma_0 + z_{1-\beta}\sigma_D \tag{15}$$

where $z_{1-\alpha}$ and $z_{1-\beta}$ are the critical values of the normal distribution for significance level α and β, respectively, while σ_0 and σ_D are the true standard deviation of $\widehat{L}$ when $L = 0$ and $L = L_D$, respectively. In practice, $\alpha = \beta$ and the critical value of the Gaussian distribution is replaced by the one sided t-critical value ($t_{1-\alpha,\upsilon}$) for the chosen confidence level (α) and degrees of freedom (υ). Similarly, the true standard deviations σ_0 and σ_D are replaced by the corresponding estimates s_0 and s_D, respectively. Assuming homoscedasticity, s_D is replaced by s_0 and a simplified formula for the LOD can be obtained (Eq. (16)):

$$L_D = 2t_{1-\alpha,\upsilon}s_0 \tag{16}$$

The only variable that needs to be estimated in Eq. (16) is s_0, the standard deviation of the estimated net signal or predicted concentration of a blank sample. The two competing methods to estimate the LOD, which are explained in detail in the following sections, differ in how they compute s_0.

2.4.1. Pseudo-univariate LOD

In the pseudounivariate LOD approach, s_0 is computed from a "surrogate" univariate variable extracted from the multivariate model. Examples of surrogate variables are the net analyte signal, the predicted concentration in a set of calibration samples, or the first score of an orthogonalized PLS model, as in the present proposal. To compute the LOD, the surrogate variable y is related to the analyte concentration x by a linear regression model (Eq. (17)):

$$y = \widehat{B} + \widehat{A}x \tag{17}$$

where $\widehat{B}$ and $\widehat{A}$ are estimates of the intercept and slope of the calibration curve, respectively.

It should be noted that Eq. (17) represents a direct calibration model and, thus, the meaning of y and x is interchanged with respect to inverse models (e.g. PLS). From linear regression [73], we know that the standard deviation of the predicted value of a blank sample is:

$$s_0 = s_{y,x}\sqrt{1 + \frac{1}{n} + \frac{\overline{x}^2}{\sum(x_j - \overline{x})^2}} = s_{y,x}\sqrt{1 + h_0} \tag{18}$$

where $s_{y,x}$ is the standard error of regression, n is the number of calibration samples, x_j are the concentration of the calibration samples, $\overline{x}$ is the mean calibration concentration and h_0 is the leverage of a blank sample (Eq. (19)).

$$h_0 = \frac{1}{n} + \frac{\bar{x}^2}{\sum (x_j - \bar{x})^2} \tag{19}$$

We can see that h_0 penalizes calibration designs with low number of samples (h_0 grows when n decreases) or high calibration ranges (h_0 grows when the mean calibration concentration $\bar{x}$ increases). Substituting Eq. (18) into Eq. (16), yields the LOD in the signal domain (y_D):

$$y_D = 2t_{1-\alpha,\nu} s_{y,x} \sqrt{1 + h_0} \tag{20}$$

The LOD is usually reported in concentration units. The concentration ($\hat{x}$) corresponding to an observed gross signal (y) is estimated by:

$$\hat{x} = \frac{y - \hat{B}}{\hat{A}} = \frac{\hat{y}}{\hat{A}} \tag{21}$$

where $\hat{y}$ is the net (blank-corrected) signal. The LOD in the concentration domain is obtained by inserting Eq. (20) into Eq. (21):

$$x_D = \frac{y_D}{\hat{A}} \cdot \left(\frac{K}{I}\right) = 2t_{1-\alpha} \cdot \frac{s_{y,x}}{\hat{A}} \cdot \sqrt{1 + h_0} \cdot \left(\frac{K}{I}\right) \tag{22}$$

The correction factor $\left(\frac{K}{I}\right)$ accounts for the uncertainty in the estimated regression parameters ($\hat{A}$, $\hat{B}$) [74] and it is defined in Eq. (23):

$$\frac{K}{I} = \frac{1 - r(B,A) \cdot (\sigma_{\hat{B}}/\sigma_0) \cdot \left[t_{1-\alpha,\nu} \cdot (\sigma_{\hat{A}}/A)\right]}{1 - \left[t_{1-\alpha,\nu} \cdot (\sigma_{\hat{A}}/A)\right]^2} \tag{23}$$

where $r(B,A)$ is the correlation coefficient between the slope and intercept of the calibration line (Eq. (17)), $\sigma_{\hat{B}}/\sigma_0$ is the relative standard deviation of $\hat{B}$ (estimated intercept), σ_0 is the standard deviation of the blanks and $\sigma_{\hat{A}}/A$ is the relative standard deviation of $\hat{A}$.

$$r(B,A) = \bar{x} \Big/ \sqrt{\sum_{i=1}^{n} x_i^2 \Big/ n} \tag{24}$$

The interpretation of the correction factor $\left(\frac{K}{I}\right)$ is as follows. If the slope (A) is estimated with low uncertainty (i.e. $(\sigma_{\hat{A}}/A) \to 0$), the term in square brackets in Eq. (16) goes to zero, and both K and I are equal to 1, yielding $\left(\frac{K}{I}\right) = 1$. In this case, the LOD is not affected by the correction factor. At the other extreme, when the uncertainty in estimating A is very high (as $\sigma_{\hat{A}}/A$ approaches $1 / t_{1-\alpha,\nu}$), the denominator (I) goes to zero, the ratio $\left(\frac{K}{I}\right)$ goes to infinity and x_D is unbounded. The correction factor includes the effect of the uncertainty of $\hat{B}$ when estimating A. If A and B are estimated individually, the correlation coefficient $r(B, A)$ is zero and the ratio $\left(\frac{K}{I}\right)$ only depends on the uncertainty of $\hat{A}$. When A and B are estimated jointly, they will be negatively correlated and the numerator (K) in Eq. (16) will increase with the uncertainty of $\hat{B}$.

2.4.2. Multivariate LOD or error-propagation approach

In the error-propagation approach, the uncertainties of the calibration concentrations and the measured signals (both calibration and test samples) are propagated through the model to estimate the variance $\sigma_{\hat{y}}^2$ in the predicted concentration $\hat{y}$ of an unknown sample (see equation (61) in Reference [42]):

$$\sigma_{\hat{y}}^2 = h_u \cdot \left(\sigma_{\Delta \boldsymbol{y}}^2 + \mathbf{b}^2 \sigma_{\Delta \mathbf{X}}^2\right) + \mathbf{b}^2 \sigma_{\Delta \mathbf{X}}^2 \tag{25}$$

where h_u is the unknown sample leverage for mean-centered data (Eq. (26)), $\sigma_{\Delta \boldsymbol{y}}^2$ is the variance of the calibration concentrations, $\|\mathbf{b}\|$ is the Euclidean norm of the PLS regression vector and $\sigma_{\Delta \mathbf{X}}^2$ is the variance of the instrumental signals of the unknown and calibration samples.

$$h_u = \frac{1}{n} + \mathbf{t_u^T} \cdot \left(\mathbf{T^T T}\right)^{-1} \mathbf{t_u} \tag{26}$$

Here, n is the number of calibration samples and $\mathbf{t_u}$ ($1 \times A$) and $\mathbf{T}$ ($n \times A$) are the PLS scores of the unknown sample and the calibration set, respectively. The expression of Eq. (25) assumes that the measurement errors are iid with zero mean and constant variance, which is not usually the case in chemical sensing. To accommodate other error structures, Allegrini et. al [43] derived a general formula for the prediction error (see equation (11) and Table 1 (Case 2) in Reference [43]) :

$$\sigma_{\hat{y}}^2 = h_u \cdot \left(\sigma_{\Delta \boldsymbol{y}}^2 + \mathbf{b^T} \boldsymbol{\Sigma}_x \mathbf{b}\right) + \mathbf{b^T} \boldsymbol{\Sigma}_X \mathbf{b} \tag{27}$$

where $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_X$ are the error covariance matrices of the instrumental signals of the unknown and calibration samples, respectively. The estimated standard deviation of the predicted analyte concentration of a blank sample can be obtained by particularizing Eq. (27) for $y = 0$ and taking the square root in both sides of the equation:

$$s_0 = \sqrt{h_u \cdot \left(\sigma_{\Delta \boldsymbol{y}}^2 + \mathbf{b^T} \boldsymbol{\Sigma}_X \mathbf{b}\right) + \mathbf{b^T} \boldsymbol{\Sigma}_x \mathbf{b}} \tag{28}$$

Inserting s_0 into Eq. (16) would give the LOD in the concentration domain. However, in contrast to univariate calibration, the fact that there is no single blank sample in PLS calibration will result in multiple values of h_0, one per each blank sample [45]. The result is a range of LOD values: from the lowest (Eq. (27)), corresponding to the lowest blank leverage $h_{0\min}$, to the largest (Eq. (28)), corresponding to the highest blank leverage $h_{0\max}$:

$$x_{D\min} = 2t_{1-\alpha,\nu} \sqrt{\mathbf{b^T} \boldsymbol{\Sigma}_x \mathbf{b} \cdot (1 + h_{0\min}) + h_{0\min} \sigma_{\Delta \boldsymbol{y}}^2} \tag{29}$$

$$x_{D\max} = 2t_{1-\alpha,\nu} \sqrt{\mathbf{b^T} \boldsymbol{\Sigma}_x \mathbf{b} \cdot (1 + h_{0\max}) + h_{0\max} s_{y_{cal}}^2} \tag{30}$$

The details on how to compute $h_{0\min}$ and $h_{0\max}$ are given in [45]. The relationship between the multivariate LOD range (Eqs. (29-30)) and the pseudounivariate LOD (Eq. (22)) can be clarified by noticing that $h_{0\min}$ in Eq. (29) is equal to the leverage of the blank sample h_0 in Eq. (19) [45]. Neglecting the uncertainty in the calibration concentrations in Eq. (29) (i.e. $\sigma_{\Delta \boldsymbol{y}}^2 = 0$) and rearranging the terms inside the square root yields:

$$x_{D\min} = 2t_{1-\alpha,\nu} \sqrt{\mathbf{b^T} \boldsymbol{\Sigma}_x \mathbf{b}} \cdot \sqrt{(1 + h_0)} \tag{31}$$

Comparing Eq. (31) to Eq. (22), we can see that the terms

PAPER II

Table 1
Comparison between x_D, x_{Dmin}, x_{Dmax} and the main parameters related to their computation, for six FIS SB-500-12 sensor units. Sensor S2 was excluded from the analysis due to abnormal behavior. The meaning of the columns is explained in the text. The term $2t_{1-\alpha,v}$ was 3.35, $h_0 = 0.06$, and $\sigma_{\Delta y} = 0.1$ ppm, for all sensor units.

Sensor	PLS LV	$[x_{Dmin}\ ,\ x_{Dmax}]$ (ppm)	x_D (ppm)	$\sqrt{\mathbf{b}^T\Sigma_x\mathbf{b}}$ (ppm)	$[h_{0min}\ ,\ h_{0max}]$	$\frac{s_{y,x}}{\hat{A}}$ (ppm)	$\frac{K}{I}$
S1	7	[0.72, 0.81]	0.73	0.21	[0.06, 0.34]	0.21	0.997
S3	5	[0.67, 0.75]	0.73	0.20	[0.06, 0.31]	0.21	0.997
S4	5	[0.67, 0.73]	0.71	0.20	[0.06, 0.23]	0.21	0.997
S5	7	[0.78, 0.93]	0.91	0.23	[0.06, 0.47]	0.26	0.997
S6	7	[0.85, 0.98]	0.92	0.25	[0.06, 0.38]	0.27	0.995
S7	5	[0.78, 0.90]	0.86	0.23	[0.06, 0.38]	0.25	0.996

$\sqrt{\mathbf{b}^T\Sigma_x\mathbf{b}}$ and $\frac{s_{y,x}}{\hat{A}}$ are equivalent when the correction factor $\left(\frac{K}{I}\right)$ approaches unity. Both terms represent the propagation of the uncertainty in the input signals through the model. In the first case, the propagation occurs directly from the measurement error Σ_x. The coefficients of the regression vector $\mathbf{b}$ are used to weight the variables in Σ_x, according to their importance for prediction. The resulting value is already in concentration units. In the second case, the input signals are first projected onto the regression vector to obtain the surrogate variable. The uncertainty of the surrogate variable is captured by $s_{y,x}$. The slope $\hat{A}$ is used to convert $s_{y,x}$ to concentration units.

3. Results and discussion

The dataset described in this section contains samples measured in 13 experimental days. The problem considered here is to calibrate two MOX sensor families with the objective of predicting the CO concentration and rejecting the humidity interference. First, the sensor conductance patterns and the measurement noise are analyzed. Then, PLS and MLPCR models are built using samples from the first day. The RMSECV and the optimum number of components of both models are compared. Then, PLS models are orthogonalized to simplify the visualization and to calculate the LOD using the pseudo-univariate approach. The resulting values are compared against the multivariate LOD interval. The samples from days 2–13 are used to validate the model and quantify its degradation in terms of LOD. First, a comprehensive analysis of the FIS device is presented and, then, the main results of the Figaro unit will be shown.

3.1. FIS SB-500-12

The response of one sensor unit to varying levels of carbon monoxide (CO) under nearly constant humidity (52 %r.h.) is shown in Fig. 3a. The x-axis represents a single heating cycle in which the sensor is driven at high temperatures for 5 s and then it is cooled down for 20 s. The logarithm is applied to the y-axis to facilitate the visualization of the response at low heating temperatures. Fig. 3b shows the sensor response to varying levels of humidity, for three concentration levels. As it can be seen, the sensor response was cross-sensitive to humidity only in certain areas of the heating pattern (e.g. $t > 10\ s$) and for concentration levels above 15.6 ppm. This indicates a non-linear effect between humidity and CO on the sensor response. From this plot, we can see that the signals corresponding to the blanks are clearly distinguishable from the signals at 11 ppm, at least in the low temperature part of the heating cycle. Therefore, we can intuitively think that the LOD should be lower than 11 ppm. To get an initial estimate of the LOD and as a first step to understand how the PLS algorithm can model these sensor conductance patterns, PLS models were built using samples from the first experimental day. Considering the multiple sensor units, five to seven components were optimum based on the cross-validation results. A scatter plot of the scores (i.e. *score-plot)* of the first two LVs ($\mathbf{t}_1$ vs. $\mathbf{t}_2$) of one calibration model is shown in Fig. 4a. Since the first two LVs already captured 96% of the variance, the scores of the following LVs were not represented. The first thing we can notice in this plot is that there is a non-linear dependence between $\mathbf{t}_1$ and $\mathbf{t}_2$, although the difference in variance between the two axes should be taken into consideration. Second, the projection of the estimated regression vector $\hat{\mathbf{b}}$ (black solid line) into

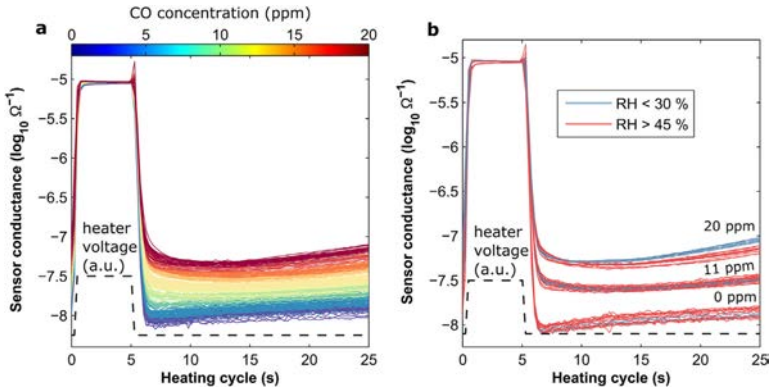


Figure 3. Conductance of a temperature-modulated SB-500-12 sensor (unit S7) during one heating cycle. Data was extracted from the first week of measurements. (panel **a**) Several measurements corresponding to a constant humidity value of 52 ± 3% r.h are overlapped. The color indicates the CO concentration. (panel **b**) Effect of humidity for several concentrations. (0, 11 and 20 ppm). The color indicates the humidity content of the sample. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

PAPER II

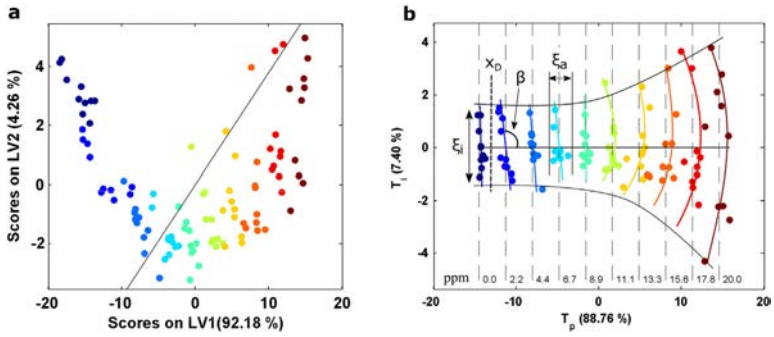


Figure 4. Scores from (panel **a**) PLS and (panel **b**) orthogonalized PLS of a SB-500-12 sensor (unit S7) in the calibration set, demonstrating the better interpretability of the orthogonal scores. The projection of the estimated regression vector $\hat{\mathbf{b}}$ (Eq. (1)) is depicted as a solid black line in panel **a**. The number in brackets indicate the percentage of variance captured by that latent variable. The color code represents the nominal CO concentration (ppm), also labeled at the bottom of the vertical black dashed lines which represent the concentration in the score space (solving for T in Eq. (3)). The symbols in panel **b** are defined in the text. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

this subspace is confusing because of two reasons: (i) the direction of $\hat{\mathbf{b}}$ is not aligned to the analyte-capturing direction and (ii) the interference direction, which is the direction of the intra-class scattering, is not orthogonal to $\hat{\mathbf{b}}$. From this plot, we may (incorrectly) deduce that the PLS model was unable to reject the humidity interference and we cannot precisely assess the LOD.

An improved representation is shown in Fig. 4b, after the model was orthogonalized. In this plot, the *x*-axis captures the **y** -variation and the *y*-axis captures the **y** -orthogonal variation, which is related to the cross-sensitivity of the sensor to the interferences. The $\mathbf{t}_2$ scores were mean-centered at each concentration level to enhance the interpretability of the model. The ability of the model (and obviously of the sensor) to distinguish between two consecutive concentration levels is clearly identified in this plot. The orthogonal score-plot could be seen as a deformation of the sample space (c.f. Fig. 2) due to the combined effect of the sensor, the calibration model and the noise. The smaller the sensor cross-sensitivity to the noise the less scattered will be the points in the vertical direction (ξ_i). The intra-class scattering in the *x*-axis (ξ_a) measures the ability of the PLS model to reject the interference. The smaller the value of ξ_a the higher the selectivity of the model. If the scores of the same concentration level can be approximated by a straight line, the angle β between this line and the horizontal axis also indicates the selectivity of the model for that concentration. $\beta = 90^o$ would mean perfect rejection of the interference and $\beta = 0^o$ means no rejection at all. In this example, we can see that the vertical scattering (ξ_i) increases with concentration, which is coherent with the fact that the cross-sensitivity to humidity of the sensor conductance patterns was higher for 20 ppm than for 2 ppm (c.f. Figure 3b). We can visually check that the first non-zero concentration (2.2 ppm) is clearly separated from the blanks (0 ppm), so the LOD must be located between these two levels. From this picture, we can understand that the LOD is affected by the interferences because the first concentration level is tilted with an angle β. In an interference-free scenario ($\beta = 0^o$), the lesser degree of overlapping between the blanks and the low concentrations would produce a smaller LOD.

From this initial analysis we visually estimated the LOD in the region between 0 and 2.2 ppm. To estimate the LOD, it is recommended to use calibration samples with concentration close to the LOD during the model building step [46]. Therefore, a reduced range of 0–9 ppm was used for the rest of the analysis. The auto-scaled and baseline-corrected logarithmic sensor conductance patterns, in this reduced range, are shown in Fig. 5. The sensor response at low temperatures ($t > 5\ s$) show more sensitivity to higher concentration levels (see how the distance between the red and orange traces is higher than the distance between the dark and light blue signals). On the other hand, the sensor response at high temperatures ($t < 5\ s$) seem more sensitive to the lower concentrations. The variance of the signals at 2.2 ppm was higher than at the other concentration levels (heteroscedasticity), probably due to the uncertainty in the generation of the calibration concentrations (see Section 2.1). However, the greatest contribution to heteroscedasticity occurs along the response pattern, rather than along the samples. The measurement error during the temperature transitions is one order of magnitude higher than in other areas of the pattern. This can be clearly seen in the pooled row error covariance matrix (Fig. 6). Looking at the noise structure, we can

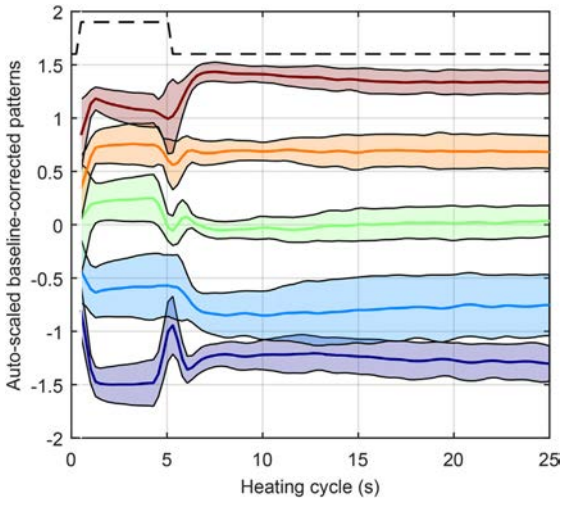


Figure 5. Auto-scaled and baseline-corrected logarithmic sensor conductance patterns of a SB-500-12 sensor (unit S7). The solid line surrounded by a shaded area indicate the dispersion across replicates as the mean ± MAD. MAD $= \frac{1}{n}\sum_{i=1}^{n}|x_i - mean(x)|$. The color indicates the CO concentration: 0 ppm (dark blue), 2.2 ppm (light blue), 4.4 ppm (green), 6.7 ppm (orange) and 8.9 ppm (red). The black dashed line represents the heater voltage. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

PAPER II

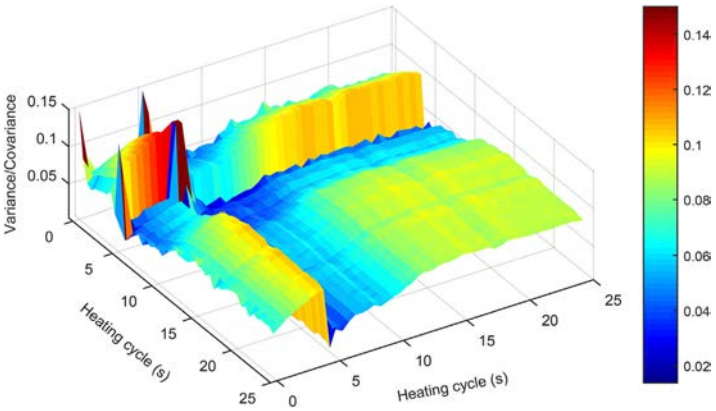


Figure 6. Equal row error covariance matrix of a SB-500-12 sensor (unit S7), obtained by pooling the row error covariance matrices of each sample (Eq. (12)). The row error covariance matrix of each sample was calculated from replicates of the auto-scaled and baseline-corrected logarithmic sensor signals (Eq. (11)). The color indicates the variance/covariance (also indicated by the z-axis). The variance was saturated at 0.15 for visualization purposes (only the temperature transitions exceeded this value). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

see that the highest measurement noise was found in the temperature transitions ($t = 0$ and $t = 5\,s$), probably because the sampling frequency of the acquisition system was not fast enough to accurately capture the rapid changes in sensor conductance during the temperature changes. The measurements corresponding to high temperature of the sensing layer ($t < 5\,s$) showed significant error compared to those at the beginning of the low temperatures ($t \in [6,\ 12]\,s$), where the lowest measurement error was found. A moderate stable noise level was found at the end of the low temperatures ($t > 15\,s$). We computed the correlation coefficient among variables of the response pattern and values above 0.8 were found (except in the temperature transitions).

In the presence of heteroscedastic and correlated measurement errors, MLCPR should produce models with lower number of components and lower prediction error than PLS. However, the empirical studies of Reis [38] confirm that this is not always the case: noise level, correlation level and heteroscedasticity play in role in determining the best linear regression method. Consequently, it seems that an empirical investigation in the particular case of temperature modulated MOX sensors is needed. A comparison of the cross-validation error between PLS and MLPCR is presented in Fig. 7. In both cases, the calibration models were built using 50 samples in the range 0–9 ppm (5 concentrations x 10 replicates) collected in the first experimental day. For the same level of error (i.e. 0.4 ppm), PLS required less components than MLPCR (5 vs 10 LVs). This behavior –which was originally found between PCR and PLS [75]– is related to the fact that PLS latent variables consider the correlation with the Y target vector, providing parsimonious predicting spaces. This behavior has been

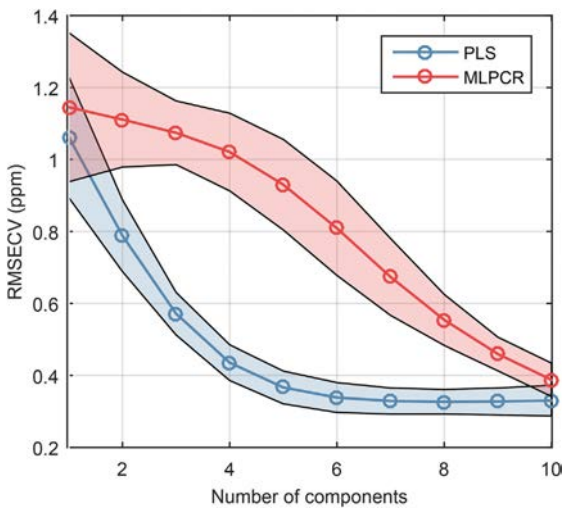


Figure 7. Comparison between PLS and MLPCR in cross-validation error, for six units of the SB-500-12. The solid line surrounded by a shaded area indicate the dispersion of the RMSECV (ppm) across sensor units as the mean ± MAD. MAD = $\frac{1}{n}\sum_{i=1}^{n}|x_i - mean(x)|$

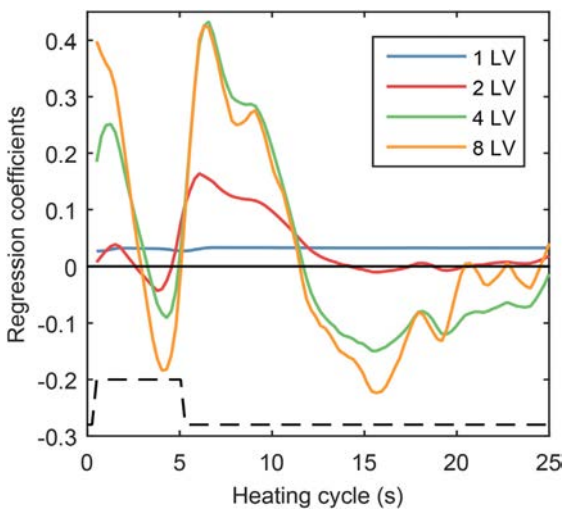


Figure 8. Evolution of the regression vector for different number of latent variables, for a SB-500-12 sensor (unit S7).

confirmed for MLPCR by Reis [38]. The authors attribute the good results of PLS to the effective way of this method to find a low dimensional predictive space, onto which the regressors are projected prior to being used for predicting the response. These results do not agree with the MLPCR theory [37,72], which states that, in the case of non-iid noise, MLPCR should perform better than PLS. On the light of the analysis of Reis, we may assume that PLS is able to reject heteroscedastic noise, when this noise is not correlated with the Y vector. This can be seen by looking at the changes in the shape of the PLS regression vector as a function of the number of latent variables (Fig. 8). When only 1 LV is used, the regression vector equally weights the response pattern, regardless of the error of each variable. When two latent variables are used, suddenly the initial part of the low temperature section is heavily weighted while the rest is rejected. This part of the sensor response exhibits low measurement errors (see Fig. 6) and is less correlated with the rest of the pattern. PLS weights more this area by virtue of the higher correlation with Y, producing a more effective filtering of the noise in X.

Based on these results, we only considered PLS models for the rest of the analysis. New PLS models were built using all samples from the first experimental day in the range 0–9 ppm and the optimum number of LVs. After orthogonalization of the resulting models, x_D was estimated using the first orthogonal component and then compared against the multivariate LOD interval (See Table 1). Values of x_D ranging from 0.71 to 0.92 ppm, with an average value of 0.79 ppm, were found. The multivariate LOD ranges were centered at a mean value of 0.80 ppm and they were relatively narrow (mean width of 0.10 ppm). For the six sensors analyzed, x_D was contained inside [x_{Dmin} , x_{Dmax}]. From this table, we can conclude that both LOD methods provide equivalent results if the model complexity is well estimated. When the model complexity is underestimated or overestimated, x_D was either slightly higher than x_{Dmax} or slightly lower than x_{Dmin}, respectively (Fig. 9). Both LOD methods show a monotonically decreasing relationship with the number of latent variables, like a typical RMSEC vs LV curve. In the simplest model (1 LV), x_D (3.02 ppm) was 4.9% higher than x_{Dmax} (2.85 ppm). As the model complexity reaches its optimum value (5 LV), x_D (0.91 ppm) falls within the LOD range ([0.78, 0.93] ppm). x_D remained within [x_{Dmin}, x_{Dmax}] beyond the optimum number of LVs, up to a point in which it exits the LOD range from the lower end (not shown in the Figure). The width of the LOD range (i.e. $x_{Dmax} - x_{Dmin}$) increased as more latent variables were added to the model (from 0.03 to 0.15 ppm in the first ten LV), due to the term h_{0max} in the computation of x_{Dmax} (Eq. (30)). h_{0max} represents the maximum distance of a given sample to the calibration center in the score space. Because the distance between two given points in a n-dimensional space increases with the dimensionality of the space, h_{0max} forces x_{Dmax} to separate from x_{Dmin} as the dimensionality of the PLS model increases.

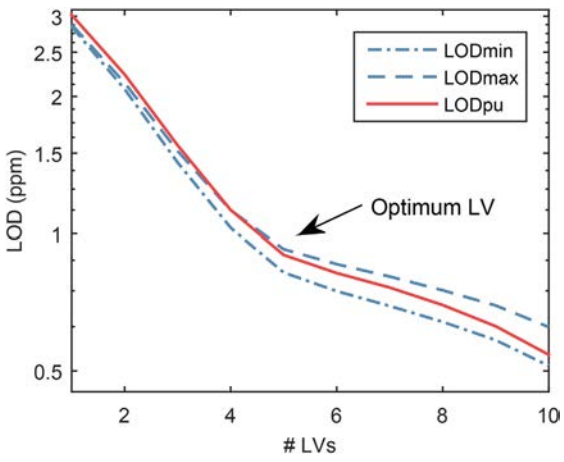


Figure 9. Limit of detection as a function of the number of latent variables, for a SB-500-12 sensor (unit S7). The optimum number of LV is indicated with an ellipse.

Fig. 10 shows the orthogonal scoreplot of the calibration samples for different model complexities, together with the corresponding location of x_D, x_{Dmin} and x_{Dmax}. In the left column (panels **a**, **c** and **e**), moving from top to bottom, we can appreciate the improvement in fitting of the data as the model complexity increases. The scores corresponding to the different concentrations become less scattered in the x-axis and less biased with respect the nominal concentration (black dashed lines), which results in a lower prediction error. Looking now at the right column (panels **b**, **d** and **f**), we can see that x_D , x_{Dmin} and x_{Dmax} shift towards the blanks as more latent variables are added to the model. When the model is too simple (panel **a**), the scores corresponding to different concentrations are partially overlapped, especially at low concentrations. In this case, neither x_D (2.4 ppm) nor the LOD range ([2.1, 2.2] ppm) satisfy the 5% false positive and false negative rates established by the IUPAC, because the t-student distributions of the blanks and the LOD overlap more than 5% (panel **b**). The underestimation of x_D, x_{Dmin} and x_{Dmax} is a side-effect of assuming homoscedasticity and using the pooled standard deviation $s_{y,x}$ in the computation of x_D (Eq. (22)) and the pooled covariance matrix $\widehat{\Sigma}_{pooled}$ in the formulas of x_{Dmin} and x_{Dmax} (Eqs. (29-30)). We can see that the samples of 2.2 ppm show more variability in the x-axis than the rest of concentration levels. When the model complexity is adequate (panels **c** and **d**) or the models are overfitted (panels **e** and **f**), the PLS residuals are homoscedastic and the approximate distribution of the samples at the LOD level can satisfy the 5% overlapping with the blanks, for both LOD estimation methods.

The LOD is estimated using a set of calibration samples, but the sensors will be typically used to predict future samples not seen by the model during the calibration process. Here, a relevant question is whether the calibration LOD will satisfy the 5% false positives and false negatives in the prediction of new samples. To respond this question, we projected validation samples -taken days after calibration- into the calibration score plot (Fig. 11). The plot reveals that the calibration data did not represent all the variability contained in the validation data. The reasons for the increased variability could be related to the overfitting of the model to the calibration data, the cross-sensitivity to humidity and temperature -which varied slightly among different experimental days-, the open-loop control of the sensor temperature or the intrinsic instability of the sensor. The intersection between the blanks and the samples centered at the LOD contains 20 blanks (out of 100) and 20 samples at the LOD (out of 100). This represents an approximate number of false positives and false negatives of 20%, which clearly indicates that, for future operation of the sensor, the LOD was underestimated. To quantify the underestimation factor, we recomputed the LOD using both calibration and validation samples, as a function of the elapsed time between calibration and validation (Fig. 12). The mean LOD (considering multiple sensor units) smoothly increased from 0.79 ppm (calibration) to 1.1 ppm (two weeks after calibration), which represents a 43% degradation. This highlights that the accuracy of the LOD value in MOX sensors is not as critical as in more precise and stable instruments, due to the expected variability of future data. Reporting a LOD with a safety margin would be a sensible choice in this case, so we could say that in our chemical

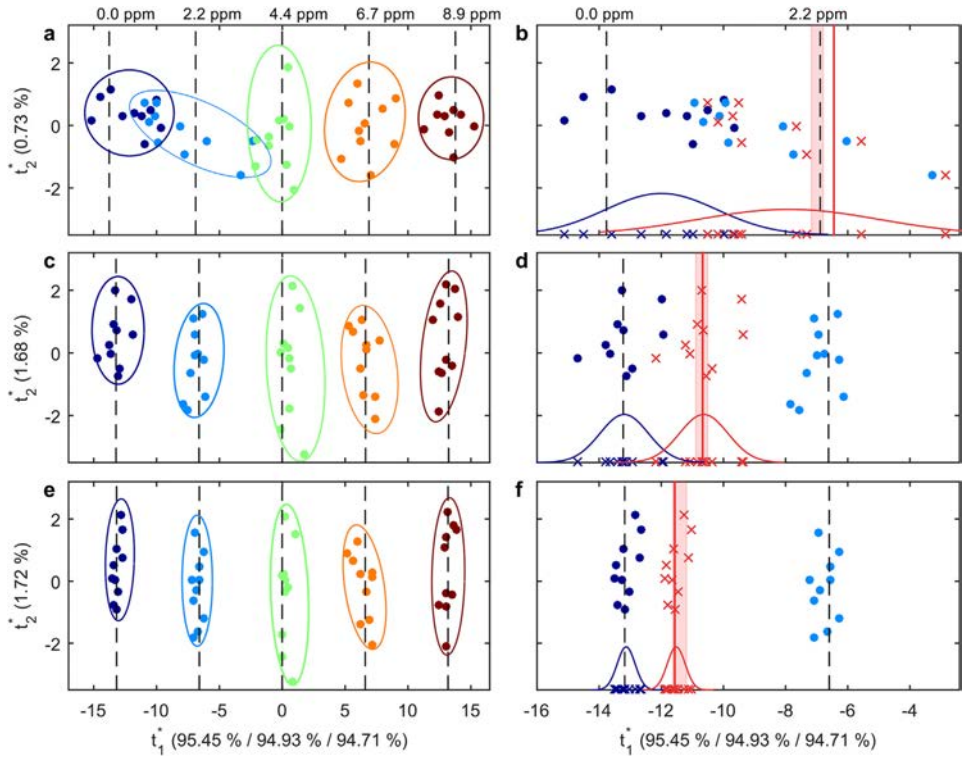


Figure 10. Calibration score plot of the first two orthogonal PLS components of sensor SB-500-12 (unit S7), using 2 LVs (panels **a**,**b**), 4 LVs (panels **c**,**d**) and 10 LVs (panels **e**,**f**). The color code represents the nominal CO concentration (ppm), also labeled at the top of vertical black dashed lines which represent the concentration in the score space (solving for T in Eq. (3)). Panels **b**, **d** and **f** show the blanks and the first concentration level, together with x_D (vertical solid red line) and $[x_{D\min}, x_{D\max}]$ (shaded red area). The red crosses simulate the distribution of the scores at x_D, by replication of the score distribution corresponding to the closest concentration level (i.e. 2.2 ppm in panel b, and 0.0 ppm in panels **d** and **f**). The t-student curves below the blanks and x_D fit the distribution of the scores of the first orthogonal PLS component ($\boldsymbol{t}_1^*$). The number in brackets in the axis label indicates the percentage of variance captured by that latent variable. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

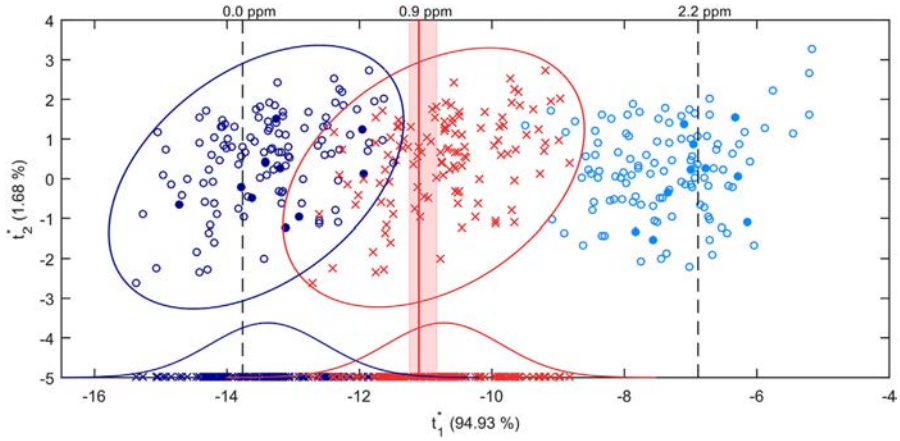


Figure 11. Projection of data from all experimental days into the calibration score-plot of a SB-500-12 sensor (unit S7), using 4 LV. Only the blanks and the first concentration level are shown for visual clarity. The color code represents the nominal CO concentration (ppm), also labeled at the top of the vertical black dashed lines which represent the nominal concentration in the score space (solving for T in Eq. (3)). The solid circles represent the calibration data (Day 1) whereas the open circles represent the projection of the validation samples (Days 2–13). The red crosses simulate the distribution of the scores at x_D, by replication of the scores corresponding to the blanks. The t-student curves below the blanks and x_D fit the distribution of the scores of the first orthogonal PLS component ($\boldsymbol{t}_1^*$). The number in brackets in the axis label indicates the percentage of variance captured by that latent variable. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

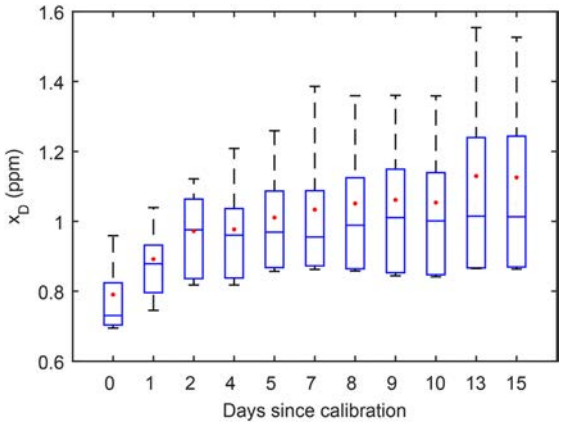


Figure 12. Boxplot of the limit of detection (x_D) of six SB-500-12 sensors, as a function of the days elapsed since calibration. In each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers.

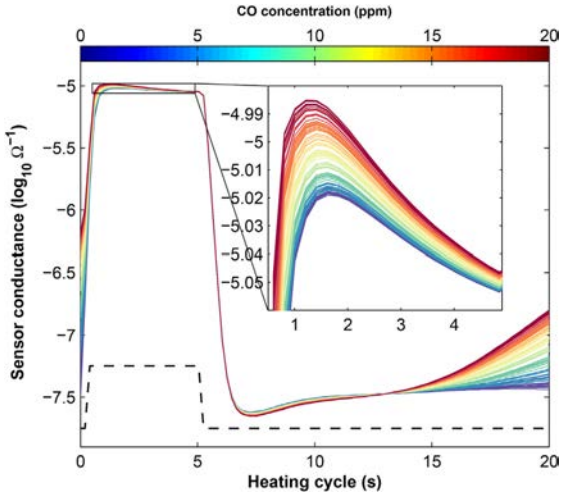


Figure 14. Reconstructed patterns of a TGS 3870-A04 sensor (unit S2), using the first orthogonal PLS component (Eq. (1)). The color code represents the CO concentration (ppm). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

measurement process the sensors can detect concentrations above 2 ppm. In the presence of humidity as the main chemical interference, this sensor model can satisfy the LOD specifications of a carbon monoxide (CO) toxic alarm, where the permissible exposure limit is 35 ppm in 8 h, as established by the Occupational Safety and Health Administration (OSHA) [76]. On the other hand, the use for a smoke cessation monitor application, which requires a detection threshold of few ppm [77], needs to be evaluated considering all possible interferences in a CO concentration range closer to the LOD.

3.2. Figaro TGS 3870-A04

The baseline-corrected logarithmic sensor conductance patterns of the Figaro TGS 3870-A04 (Fig. 13) reveal a high cross-sensitivity to humidity for all the CO concentrations that we tested (see panel **b** for 2.2 ppm). Changing the humidity content from 15% to 65% r.h. produces an effect in the sensor response at least comparable to a change in CO concentration from 0 to 20 ppm (c.f. panel **a**). PLS models with 6–7 components were built using the samples collected in the first experimental day, in the reduced concentration range 0–9 ppm. The models were then orthogonalized to facilitate their interpretation and to estimate the LOD. For example, interference-free signals reconstructed using the first component of the orthogonal model (Fig. 14) clarified which variables of the response pattern are more sensitive to CO. The most important parts of the signal for prediction purposes are the beginning of the high temperatures and the end of the low heater temperatures. This information could not be extracted from the original signals (c.f. Fig. 13a). Regarding LOD estimation, values of x_D ranging from 0.83 to 1.5 ppm, with an average value of 1.1 ppm,

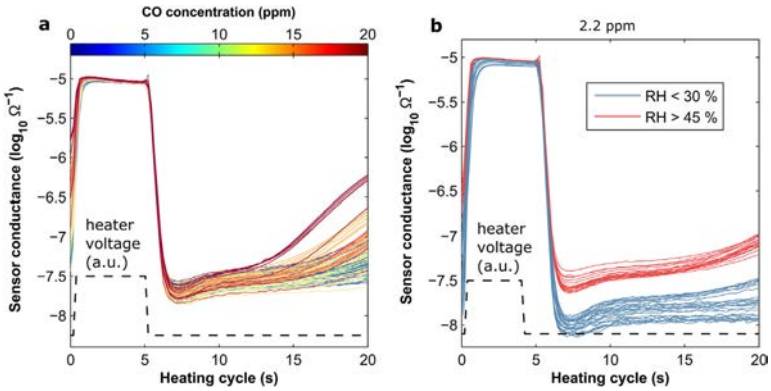


Figure 13. Conductance of a temperature-modulated TGS 3870-A04 sensor (unit S2) during one heating cycle. Data was extracted from the first week of measurements. (panel **a**) Several measurements corresponding to a constant humidity value of 52 ± 3% r.h are overlapped. The color indicates the CO concentration. (panel **b**) Effect of humidity at 2.2 ppm of CO. The color indicates the humidity content of the sample. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

PAPER II

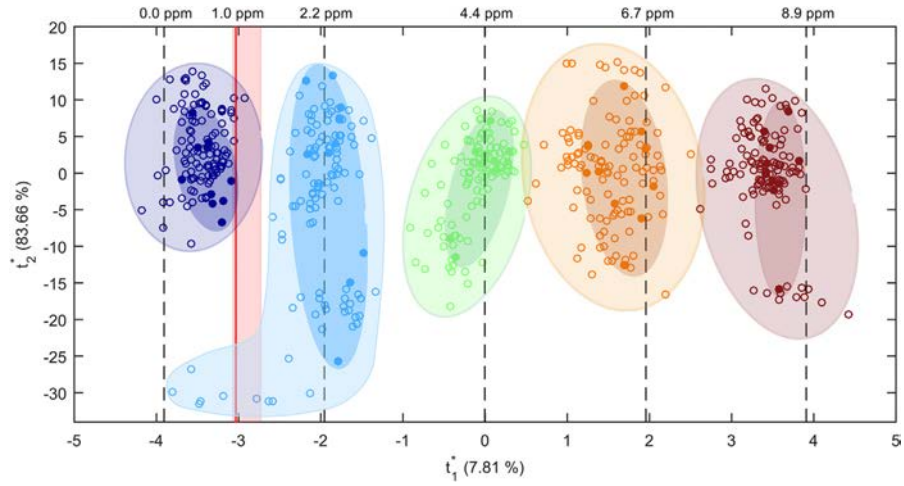


Figure 15. Projection of data from all experimental days into the calibration score-plot of a TGS 3870-A04 sensor (unit S2), using 6 LV. The solid circles represent the calibration data (Day 1) whereas the open circles represent the projection of the validation samples (Days 2–13). The number in brackets in the axis label indicates the percentage of variance captured by that latent variable. The color code represents the nominal CO concentration (ppm), also labeled at the top of vertical black dashed lines which represent the concentration in the score space (solving for T in Eq. (3)). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

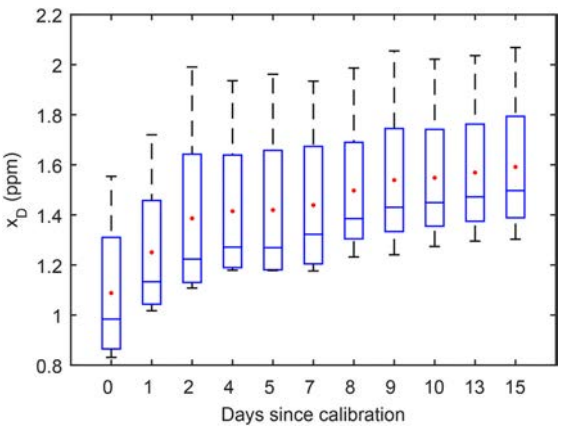


Figure 16. Boxplot of the limit of detection (x_D) of seven TGS 3870-A04 sensors, as a function of the days elapsed since calibration. In each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers.

were found. The multivariate LOD ranges were centered at a mean value of 1.2 ppm and a mean width of 0.32 ppm. For the seven sensors units analyzed, x_D was contained inside the LOD range, similarly to the FIS devices. Fig. 15 shows the score plot of one orthogonalized PLS calibration model, together with the projection of the validation data. One interesting comment about the drift and its impact into the LOD reliability should be remarked. The 2.2 ppm validation samples (cyan open points) with the lowest humidity content (15%) drifted towards the blanks, probably due to the high cross-sensitivity to humidity of this sensor model. Because the blanks and these samples get closer to each other in the x-axis, the model degrades its discrimination power between these two concentration levels. Consequently, the LOD (which was calculated in Day 1 when the scores of both samples were far away) might be no longer valid in Day 13 when the scores are much closer. The orthogonal scoreplot was a useful visualization tool to detect this unexpected situation. The average LOD (over seven sensor units) increased from 1.1 (calibration) to 1.6 ppm (two weeks after calibration), which represents a 46% degradation (Fig. 16).

4. Conclusions

MOX sensors have many challenges to be solved regarding their selectivity and stability. PLS is the de facto multivariate model to calibrate temperature-modulated MOX sensors in the presence of interferences. One research question of this work was if MLPCR would outperform PLS in the analysis of MOX sensor signals with non-iid measurement errors. We found that, for the same level of error, PLS required less components than MLPCR. Unfortunately, the PLS models were still highly complex (5–7 LVs) to properly interpret the models and to estimate common figures of merit, such as the LOD. In the present study, we propose a method to estimate the LOD using the scores of the first component of an orthogonalized PLS model. To illustrate the proposed methodology, we used seven units of two commercial MOX sensors, with different degree of cross-sensitivity to humidity. The sensors tested in this article are just examples of many sensors with similar problems. Post-processing the PLS models with an orthogonalization step resulted in a reduction of the model complexity from 5 to 7 latent variables to two-component models. These simpler models revealed patterns of the data which were hidden in the original PLS representation and yielded average LOD values (estimated using the proposed approach) of 0.79 ppm (FIS) and 1.1 (Figaro). Through the orthogonal score plot, we showed that these LOD values were contained within the multivariate LOD range estimated by error-propagation, provided the model complexity was adequate. When the model complexity was underestimated, the proposed approach yielded slightly higher LOD values than the upper limit of the multivariate LOD range. The average LOD increased to 1.1 ppm (FIS) and 1.6 ppm (Figaro), when validation samples obtained two weeks after calibration were included in the analysis. This

degradation of 43% (FIS) and 46% (Figaro) in the LOD value suggests that, if the interfering factors are not completely controlled, caution should be taken when using MOX sensors for applications requiring LOD values similar to the metrological LOD of the sensors. Adding a safety margin of 1 ppm to the original LOD estimates was recommended in our experiments to guarantee the LOD estimates in a two-week period. This means that the sensors tested in this experimental setup could be used to detect concentrations of CO of around 2 ppm in environments in which the main interfering source is the changes in humidity. This might need to be confirmed with more tests in a concentration range closer to the LOD. In the future work we will include other interfering gases in the sample composition and explore the performance of the sensors in longer timer periods (months).

Acknowledgements

This work was partially funded by the Spanish MINECO program, under grants TEC2014-59229-R (SIGVOL), PCIN-2013-195 (SENSIBLE) and BES-2015-071698 (SEVERO-OCHOA). The Signal and Information Processing for Sensor Systems group is a consolidated Grup de Recerca de la Generalitat de Catalunya and has support from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2014-SGR-1445). This work has received support from the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya and the European Social Fund (ESF). IBEC is a member of the CERCA Programme/Generalitat de Catalunya.

References

[1] M. Utriainen, E. Kärpänoja, H. Paakkanen, Combining miniaturized ion mobility spectrometer and metal oxide gas sensor for the fast detection of toxic chemical vapors, Sensors Actuators B Chem. 93 (2003) 17–24, https://doi.org/10.1016/S0925-4005(03)00337-X.
[2] A. Modi, N. Koratkar, E. Lass, B. Wei, P.M. Ajayan, Miniaturized gas ionization sensors using carbon nanotubes, Nature 424 (2003) 171–174.
[3] H. Li, X. Mu, Y. Yang, A.J. Mason, Low power multimode electrochemical gas sensor array system for wearable health and safety monitoring, IEEE Sens. J. 14 (2014) 3391–3399.
[4] A.J. Bandodkar, J. Wang, Non-invasive wearable electrochemical sensors: a review, Trends Biotechnol 32 (2014) 363–371.
[5] R. Piedrahita, Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, R.P. Dick, Q. Lv, M. Hannigan, others, the next generation of low-cost personal air quality sensors for quantitative exposure monitoring, Atmos. Meas. Tech 7 (2014) 3325.
[6] S. Zampolli, I. Elmi, E. Cozzani, G.C. Cardinali, A. Scorzoni, M. Cicioni, S. Marco, F. Palacio, J.M. Gómez-Cama, I. Sayhan, others, Ultra-low-power components for an RFID Tag with physical and chemical sensors, Microsyst. Technol 14 (2008) 581–588.
[7] G. Adam, S. Lemaigre, X. Goux, P. Delfosse, A.-C. Romain, Upscaling of an electronic nose for completely stirred tank reactor stability monitoring from pilot-scale to real-scale agricultural co-digestion biogas plant, Bioresour. Technol 178 (2015) 285–296.
[8] A. Loutfi, S. Coradeschi, G.K. Mani, P. Shankar, J.B.B. Rayappan, Electronic noses for food quality: a review, J. Food Eng 144 (2015) 103–111, https://doi.org/10.1016/j.jfoodeng.2014.07.019.
[9] A. Perera, A. Pardo, D. Barrettino, A. Hierlermann, S. Marco, Evaluation of fish spoilage by means of a single metal oxide sensor under temperature modulation, Sensors Actuators B Chem. 146 (2010) 477–482.
[10] P.K. Sekhar, E.L. Brosha, R. Mukundan, W. Li, M.A. Nelson, P. Palanisamy, F.H. Garzon, Application of commercial automotive sensor manufacturing methods for NOx/NH3 mixed potential sensors for on-board emissions control, Sensors Actuators, B Chem. 144 (2010) 112–119, https://doi.org/10.1016/j.snb.2009.10.045.
[11] L. Boon-Brett, J. Bousek, P. Moretto, Reliability of commercially available hydrogen sensors for detection of hydrogen at critical concentrations: part II–selected sensor test results, Int. J. Hydrogen Energy 34 (2009) 562–571.
[12] Y.-S. Kim, I.-S. Hwang, S.-J. Kim, C.-Y. Lee, J.-H. Lee, CuO nanowire gas sensors for air quality control in automotive cabin, Sensors Actuators B Chem. 135 (2008) 298–303.
[13] J. Nicolas, A.-C. Romain, Establishing the limit of detection and the resolution limits of odorous sources in the environment for an array of metal oxide gas sensors, Sensors Actuators B Chem. 99 (2004) 384–392, https://doi.org/10.1016/j.snb.2003.11.036.
[14] C. Borrego, A.M. Costa, J. Ginja, M. Amorim, M. Coutinho, K. Karatzas, T. Sioumis, et al., Assessment of air quality microsensors versus reference methods: the EuNetAir joint exercise, Atmos. Environ 147 (2016), https://doi.org/10.1016/j.atmosenv.2016.09.050.
[15] M. Kuske, M. Padilla, A.C. Romain, J. Nicolas, R. Rubio, S. Marco, Detection of diverse mould species growing on building materials by gas sensor arrays and pattern recognition, Sensors Actuators, B Chem. 119 (2006) 33–40, https://doi.org/10.1016/j.snb.2005.02.059.
[16] G.F. Fine, L.M. Cavanagh, A. Afonja, R. Binions, Metal oxide semi-conductor gas sensors in environmental monitoring, Sensors 10 (2010) 5469–5502.
[17] T. Hübert, L. Boon-Brett, V. Palmisano, M.A. Bader, Developments in gas sensor technology for hydrogen safety, Int. J. Hydrogen Energy 39 (2014) 20474–20483.
[18] R. Kwor, Carbon Monoxide Detectors, Carbon Monoxide Toxicity, CRC, Boca Raton, FL, 2000, pp. 61–82.
[19] A.A. Tomchenko, G.P. Harmer, B.T. Marquis, Detection of chemical warfare agents using nanostructured metal oxide sensors, Sensors Actuators B Chem. 108 (2005) 41–55.
[20] T. Šundić, S. Marco, A. Perera, A. Pardo, S. Hahn, N. Bârsan, U. Weimar, Fuzzy inference system for sensor array calibration: prediction of CO and CH 4 levels in variable humidity conditions, Chemom. Intell. Lab. Syst 64 (2002) 103–122.
[21] K. Brudzewski, S. Osowski, W. Pawlowski, Metal oxide sensor arrays for detection of explosives at sub-parts-per million concentration levels by the differential electronic nose, Sensors Actuators B Chem. 161 (2012) 528–533.
[22] M. Righettoni, A. Amann, S.E. Pratsinis, Breath analysis by nanostructured metal oxides as chemo-resistive gas sensors, Mater. Today 18 (2015) 163–171.
[23] A.D. Wilson, M. Baietto, Advances in electronic-nose technologies developed for biomedical applications, Sensors 11 (2011) 1105–1176.
[24] G. Korotcenkov, B.K. Cho, Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey), Sensors Actuators B Chem. 156 (2011) 527–538, https://doi.org/10.1016/j.snb.2011.02.024.
[25] A.C. Romain, J. Nicolas, Long term stability of metal oxide-based gas sensors for e-nose environmental applications: an overview, Sensors Actuators B Chem. 146 (2010) 502–506, https://doi.org/10.1016/j.snb.2009.12.027.
[26] G. Korotcenkov, B.K. Cho, Engineering approaches for the improvement of conductometric gas sensor parameters: Part 1. Improvement of sensor sensitivity and selectivity (short survey), Sensors Actuators B Chem. 188 (2013) 709–728.
[27] J.G. Monroy, J. González-Jiménez, J.L. Blanco, Overcoming the slow recovery of mox gas sensors through a system modeling approach, Sensors 12 (2012) 13664–13680.
[28] J. Fonollosa, S. Sheik, R. Huerta, S. Marco, Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring, Sensors Actuators B Chem. 215 (2015) 618–629.
[29] F.I.S. Inc., Fis Gas Sensor, SB-500-12, 2017. Online, http://www.fisinc.co.jp/en/common/pdf/SB50012E_P.pdf. (Accessed 14 August 2017).
[30] Figaro U.S.A. Inc., TGS 3870-A04-for the Detection of Carbon Monoxide, 2013. www.figarosensor.com/products/3870pdf.pdf. (Accessed 14 August 2017).
[31] O. Tomic, T. Eklöv, K. Kvaal, J.-E. Haugen, Recalibration of a gas-sensor array system related to sensor replacement, Anal. Chim. Acta 512 (2004) 199–206.
[32] P.K. Clifford, D.T. Tuma, Characteristics of Semiconductor Gas Sensors I. Steady State Gas Response 3, 1983, pp. 233–254.
[33] A.P. Lee, B.J. Reedy, Temperature modulation in semiconductor gas sensing, Sensors Actuators B Chem. 60 (1999) 35–42.
[34] S. Wold, A. Ruhe, H. Wold, W.J. Dunn III, The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses, SIAM J. Sci. Stat. Comput. 5 (1984) 735–743.
[35] I.S. Helland, Some theoretical aspects of partial least squares regression, Chemom. Intell. Lab. Syst (2001) 97–107, https://doi.org/10.1016/S0169-7439(01)00154-X.
[36] I.S. Helland, On the structure of partial least squares regression, Commun. Stat. - Simul. Comput. 17 (1988) 581–607, https://doi.org/10.1080/03610918808812681.
[37] P.D. Wentzell, D.T. Andrews, B.R. Kowalski, Maximum likelihood multivariate calibration, Anal. Chem. 69 (1997) 2299–2311.
[38] M.S. Reis, P.M. Saraiva, A comparative study of linear regression methods in noisy environments, Journal of chemometrics (2005) 526–536, https://doi.org/10.1002/cem.897.
[39] P.T. Hernández, A.J.T. Naik, E.J. Newton, S.M.V. Hailes, I.P. Parkin, Assessing the potential of metal oxide semiconducting gas sensors for illicit drug detection markers, J. Mater. Chem. A 2 (2014) 8952–8960.
[40] A.C. Olivieri, Analytical figures of merit: from univariate to multiway calibration, Chem. Rev. 114 (2014) 5358–5378.
[41] A.C. Olivieri, A simple approach to uncertainty propagation in preprocessed multivariate calibration, J. Chemom 16 (2002) 207–217, https://doi.org/10.1002/cem.716.
[42] K. Faber, B.R. Kowalski, Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares, J. Chemom 11 (1997) 181–238, https://doi.org/10.1002/(SICI)1099-128X(199705)11:3<181::AID-CEM459>3.0.CO;2–7.
[43] F. Allegrini, P.D. Wentzell, A.C. Olivieri, Generalized error-dependent prediction uncertainty in multivariate calibration, Anal. Chim. Acta 903 (2016)

PAPER II

51–60.
[44] M. Høy, K. Steen, H. Martens, Review of partial least squares regression prediction error in Unscrambler, Chemom. Intell. Lab. Syst 44 (1998) 123–133.
[45] F. Allegrini, A.C. Olivieri, IUPAC-Consistent Approach to the Limit of Detection in Partial Least-Squares Calibration, 2014.
[46] L. Currie, IUPAC, Commission or Analytical Nomenclature, Recommendations in evaluatior of analytical methods including detection and quantification capabilities, Pure Appl. Chem. 67 (1995) 1699–1723.
[47] A. Lorber, P.O. Box, K. Faber, B.R. Kowalski, Net Analyte Signal Calculation in Multivariate Calibration 69, 1997, pp. 1620–1626.
[48] M. Ortiz, L. Sarabia, A. Herrero, M. Sánchez, M. Sanz, M. Rueda, D. Giménez, M. Meléndez, Capability of detection of an analytical method evaluating false positive and false negative (ISO 11843) with partial least squares, Chemom. Intell. Lab. Syst 69 (2003) 21–33, https://doi.org/10.1016/S0169-7439(03)00110-2.
[49] A. Singh, Multivariate decision and detection limits, Anal. Chim. Acta 277 (1993) 205–214.
[50] M.F. Delaney, Multivariate detection limits for selected ion monitoring gas chromatography—mass spectrometry, Chemom. Intell. Lab. Syst 3 (1988) 45–51.
[51] A.C. Olivieri, N.M. Faber, J. Ferré, R. Boqué, J.H. Kalivas, H. Mark, Uncertainty estimation and figures of merit for multivariate calibration (IUPAC Technical Report), Pure Appl. Chem. 78 (2006) 633–661, https://doi.org/10.1351/pac200678030633.
[52] J. Ferré, N.K.M. Faber, Net analyte signal calculation for multivariate calibration, Chemom. Intell. Lab. Syst 69 (2003) 123–136.
[53] R. Bro, C.M. Andersen, Theory of net analyte signal vectors in inverse regression, J. Chemom 17 (2003) 646–652.
[54] L. Xu, V. Bulatov, V.V. Gridin, I. Schechter, Absolute analysis of particulate materials by laser-induced breakdown spectroscopy, Anal. Chem. 69 (1997) 2103–2108.
[55] H.C. Goicoechea, A.C. Olivieri, Enhanced synchronous spectrofluorometric determination of tetracycline in blood serum by chemometric analysis. Comparison of partial least-squares and hybrid linear analysis calibrations, Anal. Chem. 71 (1999) 4361–4368.
[56] A.M. de la Peña, A. Espinosa-Mansilla, M.I.A. Valenzuela, H.C. Goicoechea, A.C. Olivieri, Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxycillin and clavulanic acid by stopped-flow kinetic analysis, Anal. Chim. Acta 463 (2002) 75–88.
[57] B. Efron, G. Gong, A leisurely look at the bootstrap, the jackknife, and cross-validation, Am. Stat 37 (1983) 36–48.
[58] S. Wold, J. Trygg, A. Berglund, H. Antti, Some recent developments in PLS modeling, Chemom. Intell. Lab. Syst 58 (2001) 131–150.
[59] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), J. Chemom 16 (2002) 119–128.
[60] J. Trygg, O2-PLS for qualitative and quantitative analysis in multivariate calibration, J. Chemom 16 (2002) 283–293.
[61] R. Ergon, PLS post processing by similarity transformation (PLS+ ST): a simple alternative to OPLS Theoretical properties and proofs, (n.d.).
[62] J. Gabrielsson, H. Jonsson, C. Airiau, B. Schmidt, R. Escott, J. Trygg, OPLS methodology for analysis of pre-processing effects on spectroscopic data, Chemom. Intell. Lab. Syst 84 (2006) 153–158.
[63] S.J. Bruce, P. Jonsson, H. Antti, O. Cloarec, J. Trygg, S.L. Marklund, T. Moritz, Evaluation of a protocol for metabolic profiling studies on human blood plasma by combined ultra-performance liquid chromatography/mass spectrometry: from extraction to data analysis, Anal. Biochem 372 (2008) 237–249.
[64] M. Bylesjö, D. Eriksson, A. Sjödin, S. Jansson, T. Moritz, J. Trygg, Orthogonal projections to latent structures as a strategy for microarray data normalization, BMC Bioinformatics 8 (2007) 207.
[65] N. Souihi, A. Lindegren, L. Eriksson, J. Trygg, OPLS in batch monitoring—Opens up new opportunities, Anal. Chim. Acta 857 (2015) 28–38.
[66] R. Ergon, Compression into two-component PLS factorizations, J. Chemom 17 (2003) 303–312.
[67] T.A. Lestander, C. Rhén, Multivariate NIR spectroscopy models for moisture, ash and calorific content in biofuels using bi-orthogonal partial least squares regression, Analyst 130 (2005) 1182–1189.
[68] R. Ergon, Informative score-loading-contribution plots for multi-response process monitoring, Chemom. Intell. Lab. Syst 95 (2009) 31–34.
[69] J. Burgués, J.M. Jimenez-Soto, S. Marco, Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Anal. Chim. Acta 1013 (2018) 13–25.
[70] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst 58 (2001) 109–130, https://doi.org/10.1016/S0169-7439(01)00155-1.
[71] R. Ergon, PLS post-processing by similarity transformation (PLS+ ST): a simple alternative to OPLS, J. Chemom 19 (2005) 1–4.
[72] P.D. Wentzell, M.T. Lohnes, Maximum likelihood principal component analysis with correlated measurement errors: theoretical and practical considerations, Chemom. Intell. Lab. Syst 45 (1999) 65–85.
[73] K. Danzer, L.A. Currie, Guidelines for calibration in analytical chemistry. Part I. Fundamentals and single component calibration (IUPAC Recommendations 1998), Pure Appl. Chem 70 (1998) 993–1014.
[74] L.A. Currie, Detection: International update, and some emerging di-lemmas involving calibration, the blank, and multiple detection decisions, Chemom. Intell. Lab. Syst 37 (1997) 151–181.
[75] I.E. Frank, J.H. Friedman, A statistical of some chemometrics view regression tools, Technometrics 35 (1993) 109–135, https://doi.org/10.2307/1269656.
[76] O. Safety, H.A. OSHA, Carbon monoxide in workplace atmospheres (Direct-reading monitor). https://www.osha.gov/dts/sltc/methods/inorganic/id209/id209.html, 1993.
[77] S.E. Meredith, A. Robinson, P. Erb, C.A. Spieler, N. Klugman, P. Dutta, J. Dallery, A mobile-phone-based breath carbon monoxide meter to detect cigarette smoking, Nicotine Tob. Res. 16 (2014) 766–773, https://doi.org/10.1093/ntr/ntt275.

PAPER II

Chapter 4

Low Power Operation of MOX Sensors

In this chapter, we propose a low-power operating mode to minimize the energy consumption of MOX sensors, which might be considered too high for certain battery-operated applications, such as nano-drones or smartphones. The proposed low-power mode is based on 10% duty-cycling operation of 10-min periods, that represents a tradeoff between power consumption and stability. We benchmarked this mode against the continuous power mode (highest stability) and on-demand mode (lowest power consumption), during two weeks of periodic measurements using multiple units of the FIS SB-500-12 sensor operating under temperature modulation. The proposed duty-cycling powering scheme reduced the prediction error 2.5 times with respect to on-demand operation (2.2 versus 0.9 ppm) and saved up to 90% energy as compared to the continuous operating mode.

Paper III is shown below.

sensors

MDPI

Article

Low Power Operation of Temperature-Modulated Metal Oxide Semiconductor Gas Sensors

Javier Burgués [1,2,*] and Santiago Marco [1,2]

[1] Department of Electronic and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028 Barcelona, Spain; smarco@ibecbarcelona.eu

[2] Signal and Information Processing for Sensing Systems, Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain

* Correspondence: jburgues@ibecbarcelona.eu; Tel.: +34-934-029-070

Received: 27 December 2017; Accepted: 23 January 2018; Published: 25 January 2018

Abstract: Mobile applications based on gas sensing present new opportunities for low-cost air quality monitoring, safety, and healthcare. Metal oxide semiconductor (MOX) gas sensors represent the most prominent technology for integration into portable devices, such as smartphones and wearables. Traditionally, MOX sensors have been continuously powered to increase the stability of the sensing layer. However, continuous power is not feasible in many battery-operated applications due to power consumption limitations or the intended intermittent device operation. This work benchmarks two low-power, duty-cycling, and on-demand modes against the continuous power one. The duty-cycling mode periodically turns the sensors on and off and represents a trade-off between power consumption and stability. On-demand operation achieves the lowest power consumption by powering the sensors only while taking a measurement. Twelve thermally modulated SB-500-12 (FIS Inc. Jacksonville, FL, USA) sensors were exposed to low concentrations of carbon monoxide (0–9 ppm) with environmental conditions, such as ambient humidity (15–75% relative humidity) and temperature (21–27 °C), varying within the indicated ranges. Partial Least Squares (PLS) models were built using calibration data, and the prediction error in external validation samples was evaluated during the two weeks following calibration. We found that on-demand operation produced a deformation of the sensor conductance patterns, which led to an increase in the prediction error by almost a factor of 5 as compared to continuous operation (2.2 versus 0.45 ppm). Applying a 10% duty-cycling operation of 10-min periods reduced this prediction error to a factor of 2 (0.9 versus 0.45 ppm). The proposed duty-cycling powering scheme saved up to 90% energy as compared to the continuous operating mode. This low-power mode may be advantageous for applications that do not require continuous and periodic measurements, and which can tolerate slightly higher prediction errors.

Keywords: smartphone; metal-oxide semiconductor; gas sensor; low power; temperature-modulation; interferences

PAPER III

1. Introduction

Market forecasts [1] indicate that the number of portable air quality devices will increase with respect to fixed air quality stations over the next few years. The high market penetration of smartphones [2] and wearables has opened a market for mobile applications based on gas sensing. Adding a chemical analysis capability to smartphones targets the in situ detection of Volatile Organic Compounds (VOCs). Very low levels of VOCs have been found to contribute to the "sick building syndrome", which degrades workers' health and decreases productivity [3]. At higher concentrations, both short- and long-term exposure to VOCs are known to create health risks [4]. The World Health Organization (WHO) has identified nine harmful compounds, including carbon monoxide (CO), benzene, nitrogen dioxide (NO_2), formaldehyde, and naphthalene, which might be present in indoor

air [5]. In outdoor air, the United States Environmental Protection Agency (EPA) specifies four principal pollutants and their corresponding maximum exposure concentrations (hourly) [6]: 9 ppm of CO, 100 ppb of NO_2, 0.07 ppm of ozone, and 75 ppb of sulphur dioxide. If gas sensing technology is implemented into smartphones, users could assess the air quality wherever they are, perhaps allowing them to avoid unhealthy environments or, at least, indicating to them the need for ventilation. Using the built-in localization and connectivity capabilities of smartphones and wearables, any measured data could simultaneously be geo-tagged and shared to the Internet of Things (IOT) in order to build fine-grain air pollution maps.

A gas-sensitive mobile device has also a direct application to the m-Health sector (defined as medical health practice supported by mobile devices [7]). Such a device can be used to measure the concentration of VOCs present in expired breath, which might be related to respiratory and gastrointestinal dysfunction [8–12]. For example, CO and nitric oxide are biomarkers for asthma [12], chronic obstructive pulmonary disease [8,9], and lung cancer [11], and high levels of ammonia in expired breath are indicative of renal failure [10]. Gas sensors are less accurate than the analytical instruments usually deployed in laboratory settings or hospitals, but may be advantageous in terms of early detection of diseases, reducing healthcare costs, and promoting healthcare equity. Gas-sensitive mobile devices can also be used in healthcare applications not related to diseases. For example, tracking acetone levels in expired breath can be used to monitor sleep quality [13] and the fat-burning process [14]. Expired-air carbon monoxide (eCO) concentrations above 5 ppm might be related to smoking habits. Providing smokers with a personal monitor for measuring such concentrations was found to be a feasible method for reducing the intake of smoke [15]. Commercial eCO monitors, such as the Smokerlyzer (Bedfont Scientific Ltd. Harrietsham, UK), are already sold as plug-ins for mobile devices. Breathalyzers used to measure the blood alcohol content from breath samples are also available as plug-ins for mobile devices [16]. The connection between gas sensors and smartphones is becoming steadily more frequent. The full integration of the gas sensor into the smartphone or wearable would drastically increase the penetration of current gas-sensing applications and the appearance of new ones.

The key specifications of a gas sensor with regard to mobile integration are the cost, size, durability, stability, and power consumption. Among the different gas-sensing technologies, metal oxide semiconductor (MOX) sensors represent the most prominent technology for integration into portable devices. They have the advantages of high sensitivity to a large variety of gases, miniaturization potential, low cost, and long-life. They represent the most commercially successful type of gas sensor, and have been used in applications ranging from environmental monitoring [17–20], energy [21], food [22,23], automotive [24,25], and safety and security [26–30] to biomedicine [31,32]. The main limitations of this technology are their inherent lack of selectivity, high power consumption, and temporal drift [18,33,34].

The main source of power consumption in MOX sensors is the resistor that heats up the sensing surface to promote efficient redox reactions between the adsorbed molecules of the target gas and the metal oxide [35]. The evolution of fabrication technologies has permitted the miniaturization of the sensing chip and heater resistor, leading to smaller and more power-efficient sensors (see Figure 1). Thanks to micro-machined technology (MEMS), MOX sensors can be fabricated nowadays with a very compact form factor and power consumption of mere tens of mW [36].

As opposed to other gas-sensing technologies (e.g., electrochemical sensors) in which power can be issued on-demand (i.e., to take a measurement), MOX sensors require a continuous energy supply. Immediately after a MOX sensor is switched on, the sensor enters into an unstable state characterized by a steady increase in the sensor resistance, known as the "run-in phase" in [37] or "initial action" in [38]. In a separate study we made, the standard duration of this transient behaviour was found to depend heavily on the sensor model and how long the sensor was unenergized (Figure 2). In general, the longer the sensor is unpowered, the longer the run-in phase takes. In some extreme cases (e.g., sensor TGS 2602), the run-in phase can take more than two hours if the sensor has been

unpowered for more than one week. This means that a reliable and stable measurement can only be obtained after waiting for that length of time, which represents a practical issue for intermittently operated devices, such as portable gas detectors. Because of this, MOX sensors have traditionally been powered continuously rather than on-demand.

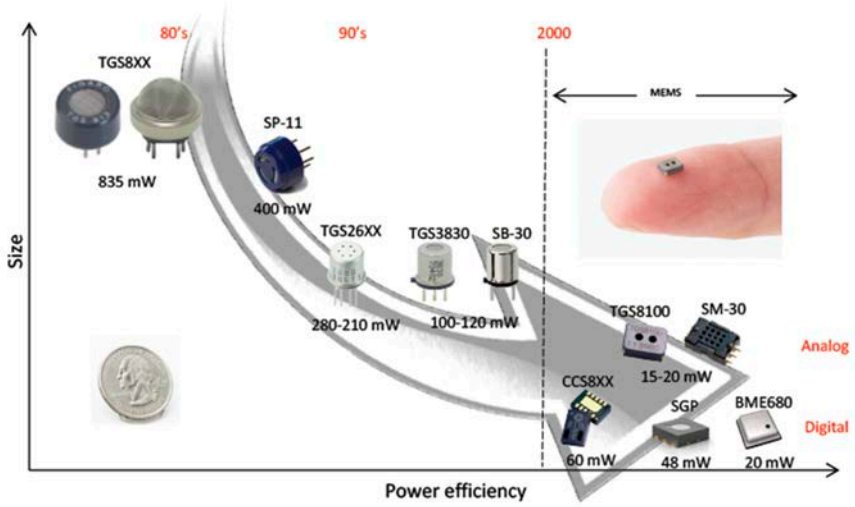


Figure 1. Evolution of metal oxide semiconductor (MOX) technology in terms of size and power efficiency. The background arrow indicates the temporal evolution.

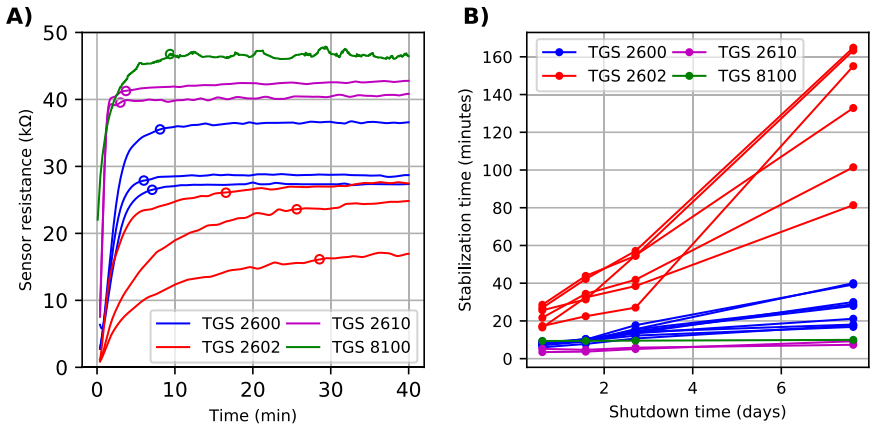


Figure 2. Run-in phase of several isothermally operated MOX sensors (TGS models, Figaro Engineering Inc., Arlington Heights, IL, USA). (**A**) Initial action of the sensor response after a 12-h shutdown. The open circles on top of each trace indicate the stabilization time, defined as the time when the response reaches 95% of its final value. (**B**) Stabilization time as a function of the shutdown period. The sensors were continuously powered for several days before each shutdown.

The power consumption under continuous power supply might be considered too high for certain battery-operated applications. As an example, consider adding a Figaro TGS8100 sensor (one of the most power efficient sensors in the market at 15 mW) to a Samsung Galaxy S7 smartphone, which consumes 462 mW under non-intensive use (see Appendix A for the details). The power

consumption of the MOX sensor represents 3.14% of the total smartphone power usage or, in other words, a reduction in battery lifetime of 45 min. Considering that more than one MOX sensor might be needed to increase the selectivity of the system [39], the situation worsens. For example, the recently released SGP sensor (Sensirion AG) integrates multiple MOX sensors into the same chip, increasing the power consumption to 48 mW. This would represent 10% of the smartphone's power consumption or a reduction in battery lifetime of 2 h and 15 min in our example. Evidently, the power consumption needs to be further reduced.

Many solutions to the problem of power consumption in MOX sensors have been proposed using the technique of duty-cycling. Duty-cycling saves energy by periodically switching on and off the sensor power and, at the same time, it achieves stability because long off periods are split into smaller "chunks". The amount of time the sensor is powered with respect to the duration of the cycle is known as the duty cycle and this is directly proportional to the average power consumption. The duty-cycle period (or duration) can be determined by the frequency of measurements required by the application [40]. Besides this, one needs to take into account that the sensor's stability depends on the duration of the active and inactive (or sleep) phases, which can be calculated from the duty cycle and the cycle period. The minimum duration of the active phase must be longer than the thermal time constant of the heater resistor. In other words, the heater resistor needs to be powered longer than its thermal time constant to ensure that the sensing surface will reach the optimum working temperature. Ceramic sensors have a relatively long thermal time constant, of the order of 3 s. On the other hand, the thermal response time of MEMS micro-hotplates is some 150 times faster, being of the order of 20 milliseconds [41]. The duration of the sleep phase represents a trade-off between power consumption and stability. Long inactive periods reduce the power consumption, but increase the duration of the transient behaviour that occurs after power-up (cf. Figure 2). Thus, the original duty cycle and cycle duration might need to be readjusted based on these criteria.

Previous works have explored how various combinations of duty cycle and cycle period affect the stability of the sensor response and the response time under isothermal operation [40,42,43]. Sayhan et al. [40] found that short heating pulses (below 6 s) did not work well for their proprietary ceramic sensors, but were appropriate for MEMS sensors. Recently, Macías et al. [42] proposed a 0.8% duty cycle in 10-min cycles for the detection of ethanol and cigarette smoke using ceramic Figaro TGS 26XX sensors. This is equivalent to an active phase of 5 s, which is consistent with the results found in [40]. Jelicic et al. [43] explored 20–60% duty cycles and 0.5–1.5 s periods in a MiCS-5525 sensor. They found that a duty cycle of 30% and a period of 0.5 s were optimal. In this case, the duration of the heating phase was 150 ms which is several times the thermal time constant of the sensor (20 ms). Increasing the duty cycle period from 0.5 to 1.5 s increased the limit of detection or minimum detectable concentration.

In a duty-cycling operation, the sensor response is typically measured at the end of the active phase when the response is more stable. Nonetheless, several authors have explored the performance of other features of the sensor response which might lead to lower measurement times [42–45]. Jelicic et al. [43] found an optimum measurement time of only 65 ms using features of the transient response. Rossi et al. [44] extracted features from the frequency spectrum at 20 Hz in the first 512 ms of the transient behaviour.

By definition, duty-cycling means that a given on-off pattern is repeated continuously. However, some authors have explored the response of the sensor to "bursts" of heating cycles after an inactive period [46,47]. Oletic et al. [46] explored the stability of the MiCS-5525 sensor's resistance at the beginning (Rmin) and end (Rend) of the heating pulse after a 30 s pause. They found that Rmin was sensitive to the inactive length of the cycle, so higher duty cycles and lower periods were preferred. On the other hand, the stability of Rend depended on the energy delivered to the sensor, regardless of the duty cycle, period, or number of pulses in the "burst". Bicelli [47] found that a burst of four pulses with high duty cycle and short period were the optimum in a Figaro TGS2442 ceramic sensor.

A common factor of the previous works is that measurements were assumed to be taken periodically. While this might be the case in certain applications, in other situations measurements are taken asynchronously. This means that a measurement can be requested at any moment even after the sensor has been switched off for a long period of time (days, weeks, or months). For example, a MOX-based breathalyser used to check the blood alcohol content in a driver's blood might be turned on only prior to driving. Previous works also do not take possible chemical interferences into account. As Barzan et al. [48] point out, overoptimistic testing conditions in metal-oxide gas sensor research contribute to the excellent performances published. They specifically referred to the absence of changing background conditions (humidity, temperature). In gas-sensing applications in which the user can take a measurement anywhere, and at any time, it is realistic to assume that interferences such as humidity, temperature, and analytes other than the target gas will be present. Therefore, the isothermal operation of a single sensor is not reliable as MOX sensors are intrinsically not selective. Modulating the heater temperature has been shown to be an effective method for increasing selectivity [49]. Among the explored works in low-power modes for MOX sensors, only Vergara et al. [50] used temperature modulation and on-demand powering. They studied the effect of the heating and measurement time in the prediction of gases relevant to fruit quality using Figaro TGS 26XX sensors. The sensors were powered using multi-sinusoidal waveforms of length between 5 and 312 s and both the duty cycle and period were fixed. The measurements were taken periodically twice per day in a time period of 1 month. A measurement time of 39 s was found to be optimal for quantification purposes.

Following on the work of Vergara et al. [50], the current paper proposes a scenario of on-demand measurements of low concentrations of carbon monoxide with background interferences of humidity and temperature. This scenario is motivated by a myriad of applications requiring the detection of CO under variable humidity and temperature as indicated at the beginning of this section. We compare the performance of three power management strategies applied to temperature-modulated FIS SB-500-12 sensors: continuously powered (highest stability), on-demand power (lowest power consumption), and duty-cycling (a power/stability trade-off).

2. Materials and Methods

2.1. Experimental Design

The experiment timeline is shown in Figure 3. It consisted of a calibration phase, which was performed on Day 1, followed by validation measurements taken during the next two weeks. For calibration purposes, the sensors were exposed to five concentrations of CO: 0, 2.25, 4.5, 6.75, and 9 ppm. Ten measurements per day were done for each concentration, with a humidity value randomly chosen from a uniform distribution in the range 15–75% relative humidity (r.h.) (Figure 3a). We believe that this range is representative of various real scenarios, such as indoor air quality or biomedical applications based on expired breath measurements. The temperature was monitored during the two experimental weeks and it was in the range 21–27 °C. The intraday variation was below 2 °C. Dynamic gas mixtures of CO and humid synthetic air were generated using three mass flow controllers (EL-FLOW Select, Bronkhorst High-Tech B.V., Ruurlo, The Netherlands). The resulting 50 gas mixtures (5 concentrations × 10 repetitions) were randomly introduced into a small gas chamber (250 cm^3 internal vol.) at a constant flow rate of 500 ml_n/min for 15 min. The chamber was cleaned at the beginning of the experiment by flooding with synthetic air for 30 min at 500 ml_n/min.

The validation experiments simulated a use case in which the measurements are requested only during the day and at random times, there being no measurements at night (the sensing device might be switched off at night) and no measurements took place on certain days (e.g., days 3 and 4). Validation measurements were performed in a similar manner to calibration, but there were two minor differences with respect to calibration: (i) only three concentration standards (0, 4.5, and 9 ppm) were used and (ii) the gas mixtures were not consecutively introduced into the gas chamber. The 30 measurements

corresponding to validation (3 concentrations × 10 repetitions) were distributed within the first 16 h of the day. This yields an average elapsed time between measurements of 30 min. No measurements took place during the last 8 h of the day.

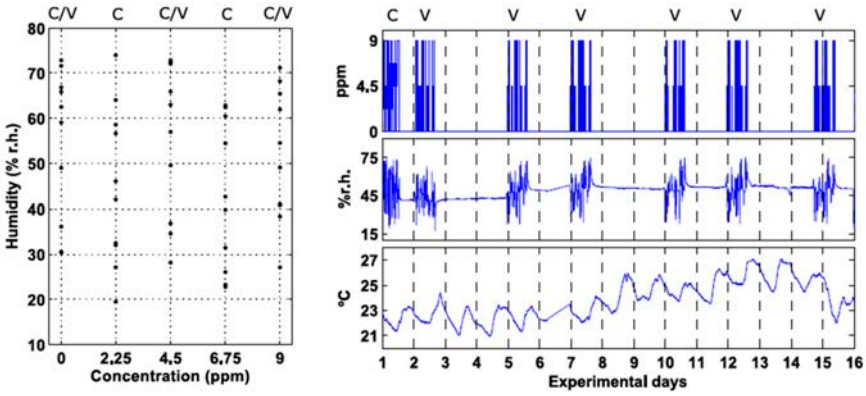


Figure 3. Experimental design. (**a**) Carbon monoxide and humidity values presented to the sensors in calibration and validation. "C" stands for Calibration and "V" for Validation. "C/V" means that the condition was used for both calibration and validation; (**b**) Validation plan. The subplots represent (from top to bottom) the CO concentration (ppm), the measured humidity (% relative humidity (r.h.)), and the measured temperature (°C).

2.2. Sensor Board

A board including twelve MOX sensors of the same type (SB-500-12, FIS Inc., Jacksonville, FL, USA), a temperature/humidity sensor (SHT75, Sensirion AG, Stäfa, Switzerland), and the corresponding read-out electronics was built. The FIS SB-500-12 sensors were powered using the temperature modulation waveform suggested by the manufacturer (0.9 V for 5 s, followed by 0.2 V for 20 s, in cycles of 25 s) [51]. The high heater voltage at the beginning of the cycle cleans the sensor surface and removes the water vapour influence, while the low heater voltage conditions the sensor for measuring CO. The twelve MOX sensors were divided into three groups according to the target operating mode (see Figure 4). Six sensors were operated in continuous mode (i.e., always powered), three sensors were powered on-demand (i.e., powered just before and during the measurement), and the last three sensors were operated in a duty-cycling mode with a 10% duty cycle in periods of 10 min. When any of the sensors was powered, the heating waveform suggested by the manufacturer was continuously applied. For example, a sensor operated on-demand was normally shut down and turned on only for taking a measurement. During the measurement, the heating waveform was continuously applied. For simplicity, we did not implement a closed loop control of the heater temperature although this has been shown to produce more accurate results [52]. The redundancy of sensors included in each power mode accounted for the large tolerance in baseline (one order of magnitude) and sensitivity (a factor of two) between devices [51]. By analysing the results of several units of the same model operating under a specific power mode, statistics on the performance of that mode were obtained. In duty-cycling mode, the temperature modulation waveform was applied only within the active part of the duty cycle. The duty cycle was selected so that the sensors were powered with a small number of heating cycles, as we believed this should greatly increase the stability of the sensor compared to on-demand operation. A time-ON of 60 s allowed for two full heating cycles of 25 s plus the first 10 s of a third heating cycle. Within each heating cycle, the first 5 s at 0.9 V is where the sensing surface reaches the maximum temperature and therefore is where the sensor stability increases the most. With the chosen duty cycle, the sensor was exposed to this high temperature step three times

per cycle. After fixing the duty cycle at 60 s, a cycle period of 10 min was selected to obtain a reduction in power consumption of one order of magnitude. It was beyond the scope of this work to study the effect of other duty cycles. The MOX read-out circuits consisted of voltage dividers with 1 MΩ load resistors. The value of the load resistor allowed for a proper quantification of the sensor resistance, considering the large dynamic range (20 kΩ to 10 MΩ) present at low concentrations of the analyte with the chosen temperature modulation waveform. The output voltage of the sensors was sampled at approximately 3.5 Hz using an Agilent HP34970A/34901A DAQ configured at 15 bits of precision and an input impedance greater than 10 GΩ. In this configuration, the errors introduced into the voltage measurements were considered negligible compared to the intrinsic variability of the sensor resistance due to the chemical transduction process. The sensor voltage sampled at 3.5 Hz during a full heating cycle of 25 s is a multivariate signal $V_s(t)$, $t \in [0,\ 25]$ s, containing 88 variables evenly spaced across the 25 s signal. $V_s(t)$ was then converted to sensor conductance g_s ($k\Omega^{-1}$) using Equation (1):

$$g_s(t) = \frac{V_c - V_s(t)}{V_s(t)R_L} \tag{1}$$

where V_c is the voltage (V) of the voltage divider and R_L is the load resistor (kΩ). The multivariate sensor conductance patterns $g_s(t)$ corresponding to three consecutive heating cycles were averaged to reduce instrumental noise. This is what we call "a measurement". When a measurement was requested, the discontinuously operated sensors were warmed up for 75 s (i.e., three full heating cycles).

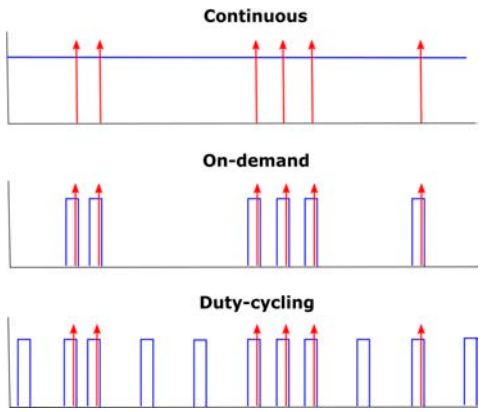


Figure 4. Schematic illustration of the three power management strategies used in this work. The blue solid lines represent the heater voltage and the red arrows indicate when the measurements are taken.

The Sensirion sensor provided reference humidity and temperature values with tolerance below 1.8 % r.h. and 0.5 °C, respectively, every five seconds. According to the manufacturer datasheet [53], the long-term drift of the SHT75 sensor is less than 0.5 % r.h./year and 0.04 °C/year. The readings from the humidity sensor were averaged during the last heating period of each experimental condition to provide a reference relative humidity value h (% r.h.).

2.3. Calibration Models

Multivariate analysis was done in Matlab R2009b (The Mathworks, Inc., Natick, MA, USA) using PLS Toolbox 8.0 (Eigenvector Research Inc., Wenatchee, WA, USA). Partial Least Squares (PLS) [54] calibration models were built individually for each sensor using the measurements collected on the first experimental day. The sensor conductance patterns **G** were normalized by the sensor conductance in air $\mathbf{G}_0$, which was estimated from a blank sample with random humidity measured at the beginning of each experimental day. The design matrix $\mathbf{X} = \mathbf{G}/\mathbf{G}_0$ was mean-centered and auto-scaled to unit

variance and the vector of responses **Y** was mean-centered prior to PLS modelling. The number of latent variables (LV) was optimized via 100 bootstrapping iterations on the calibration samples (see Section 2.1). In each bootstrap iteration, a PLS model was built using 50 samples randomly selected with replacement from the calibration set. The performance of the model in the set of test samples excluded from the selection process was assessed through the root mean squared error (RMSE):

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} \quad (2)$$

where n denotes the size of the test set, and y_i and $\hat{y}_i$ are the true and predicted values for the test sample i. The optimum number of LV was found by inspection of the graph of mean RMSE (across the bootstrap iterations) versus the number of LV.

The stability of the calibration models was evaluated through the root mean squared error in prediction (RMSEP) using external validation samples collected in the two weeks following calibration. The RMSEP was computed using Equation (2), but in this case n denotes the size of the external validation set and y_i and $\hat{y}_i$ are the true and predicted values for the external validation sample i.

PLS models usually require several latent variables to account for the effect that chemical interferences produce in the sensor response patterns. When the model complexity increases beyond the chemical rank of the problem, PLS models are not easily interpretable [55]. A simpler model is advantageous to understand the underlying structure. To address this problem, the PLS models were post-processed with an orthogonalization step [56] that reduced the effective number of latent variables to two [57]. This produces simpler models with orthogonalized loadings and scores, which condense all the variance related to the target analyte into the first weight and loading. The first loading is in the same direction as the regression vector for data in which the structured noise, defined as the systematic variation of **X** not linearly correlated with **Y**, has been filtered out. The second loading of the orthogonalized model captures the variation orthogonal to Y, which is mainly related to the interferences and, to a lesser degree, other noise sources. The orthogonalized model yields the same predictions as the non-orthogonalized model.

3. Results and Discussion

3.1. Drift in Sensor Conductance Patterns

The response of the sensors to the experimental conditions described in Figure 3 was recorded for two weeks. Depending on the operating mode of each group of sensors, different trends in the sensor conductance patterns were observed (Figure 5). The patterns of the continuously operated sensors showed no relevant variations between the calibration and validation samples. The duty-cycling operation introduced a slight offset towards higher concentrations (the sensor conductance increases with concentration). The strongest deviations from the calibration data were observed in the sensors operating on-demand. In this latter case, the shape of the pattern drastically changed after nine days of on-demand operation. Although the underlying mechanisms governing this drift are still not fully understood [58], it is thought that keeping the sensing surface hot promotes water desorption and cleans any existing organic deposits from the surface [59].

3.2. Prediction Error

PLS models were built separately for each sensor using the conductance patterns measured on Day 1. The optimum number of latent variables was from three to five, depending on the sensor. An example of the scoreplot of the orthogonalized model of one sensor from each operating mode is given in Figure 6a–c. In these plots, the x-axis represents the predictive component of the model and the y-axis captures the variance associated to the structured noise present in the sensor conductance patterns. Regardless of the operating mode, the five calibration concentrations were non-overlapping

in the score space, which indicates a good fit of the model to the data. One should note that all of the sensors were continuously powered before and during the calibration process, thus increasing their stability. The RMSEC (considering all sensors) was 0.35 $\pm$ 0.05 ppm.

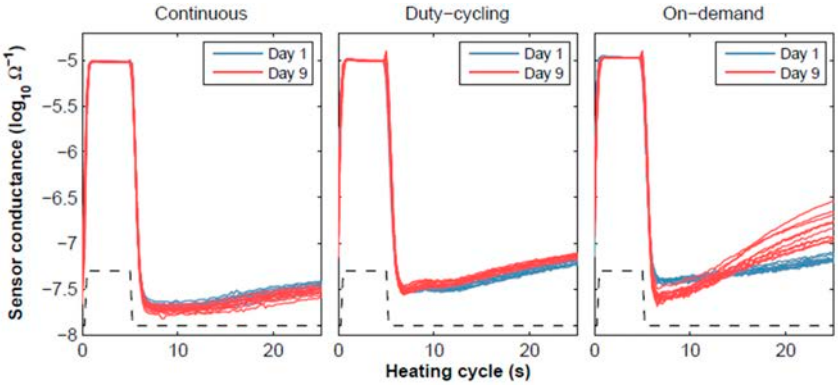


Figure 5. Drift in raw patterns for the three operating modes at 9 ppm of CO. The blue and red solid lines represent the sensor conductance measured during a full heating cycle taken on day 1 and day 9, respectively. The black dashed line represents the heater voltage (a.u.).

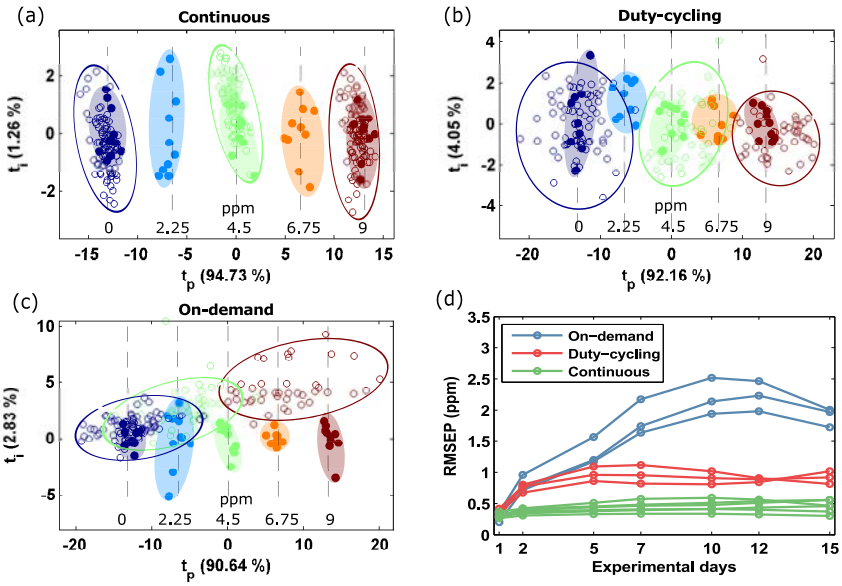


Figure 6. Score plot of the orthogonal Partial Least Squares (PLS) models obtained from one sensor of each group under continuous calibration. The title of the subplot indicates the target operating mode: (**a**) continuous, (**b**) duty-cycling, and (**c**) on-demand. The filled and open circles represent the calibration and validation observations, respectively. Hand drawn ellipses cluster observations of the same concentration level. The concentration (ppm) is indicated in the text at the bottom of the figure. (**d**) Temporal stability of each operating mode.

Immediately after calibration, the validation phase started, and the sensors were operated in the target operating mode. One would expect an increase in the prediction error introduced by the

discontinuous operation (duty-cycling and on-demand modes). The error would have two components: the bias and the variance. The scoreplot of the orthogonalized PLS model is an intuitive tool to visualize these two contributions to the error term. The scores of the validation samples were projected onto the latent space of the calibration samples (Figure 6a–c). Looking at the continuous sensors (Figure 6a), the validation scores remained centred around the calibration scores (i.e., small bias), but the dispersion mainly increased mainly in the vertical direction and, only slightly, in the horizontal axis. Because the vertical axis captures the structured noise which the multivariate model is able to reject, the prediction error increased slightly (Figure 6d) due to the horizontal component of the variability. When the sensors were operated in duty-cycling mode, there was still a small bias but the variability of the scores around the calibration data was shared among the *x*- and *y*-axis (Figure 6b). Because the *x*-axis captures the predictive direction of the model, the prediction error increases (Figure 6d). Thirdly, under on-demand operation the scores presented both high bias and high variance in the predictive direction (Figure 6c). This is the result of the drastic change in the shape of the sensor conductance patterns (Figure 5).

The prediction error as a function of the elapsed time since calibration is presented in Figure 6d. On the first validation day, all of the samples were used for model building. Since there were no validation samples, the first data point represents the fitting error (i.e., the RMSE computed on the fitting samples). Over the remaining days, the RMSEP was computed on the external validation samples. It is expected that a fitted model will perform worse on unseen data than on the fitting samples [60]. This produced an increase of the average prediction error from 0.35 ppm on Day 1 to 0.43 ppm on Day 2 (22%) even under continuous operation. This differential error of 0.06 ppm, known as the "shrinkage effect", cannot be attributed to the operating mode. For continuous operation, the average error subsequently remained stable at 0.45 ppm. The average RMSEP increase that could be attributed to the discontinuous operating modes (i.e., after discounting the shrinkage effect) was 75% (0.43 versus 0.75 ppm) for duty-cycling sensors and 97% (0.43 versus 0.85 ppm) for on-demand sensors. After the second day, the error of the duty-cycling sensors was largely stable at 0.9 ppm. Comparing this error to the error obtained under continuous powering (0.45 ppm), the trade-off between stability and power consumption can be quantified: reducing the power consumption by 90% increased the average RMSEP by only a factor of 2. On the other hand, the RMSEP of on-demand sensors kept growing for up to 9 days after calibration and then stabilized at 2.2 ppm. This represents a five-fold increase with respect to continuous power, which makes the calibration model unusable for accurate gas quantification.

3.3. Calibration in Discontinuous Mode

The calibration in continuous mode provided low RMSEC estimates, as the sensor response was stable throughout the calibration process. However, as we saw in the preceding section, the RMSEC was overoptimistic for the sensors that were discontinuously operated and, therefore, not representative of the detection capabilities of the system beyond the calibration data. This raises the question about the potential benefit of calibrating the sensors in the target operating mode instead of under continuous powering. To answer this question, we explored the performance when the discontinuous sensors were calibrated in the target operating mode.

Using the 30 samples (3 concentrations $\times$ 10 repetitions) corresponding to the second experimental day, in which the sensors were already operating in the target mode for more than 12 h, we built new PLS calibration models. The resulting RMSEC was 0.5 $\pm$ 0.15 ppm, which is 43% higher than in continuous calibration. The distribution of the validation samples in the scoreplot of the calibration samples is shown in Figure 7. Comparing these scoreplots to the scoreplots obtained in continuous calibration (Figure 5b–c), we can see that in the former case the bias is greatly reduced but the variability of the blanks has increased considerably. This means that when calibration and validation are performed both under discontinuous power, the resulting response patterns are not distorted, but their variability increases. The higher bias might possibly be explained by the lower number of

samples used in discontinuous calibration (30 rather than 50). With respect to continuous calibration, the average RMSEP increased 44% (0.9 to 1.3 ppm) in duty-cycling sensors and 18% (2.2 to 2.6 ppm) in on-demand sensors (Figure 8). The stabilization time was also higher in duty-cycling (2 versus 5 days) and on-demand sensors (9 versus 10 days). Therefore, we can conclude that discontinuous calibration did not improve the results obtained in continuous calibration. In addition, discontinuous calibration can be time-consuming for large duty-cycle periods. For example, assuming that 20 samples are used for calibration and the duty-cycle period is 10 min, the calibration process will take 3.3 h. In continuous mode, calibration is faster because samples can be measured consecutively.

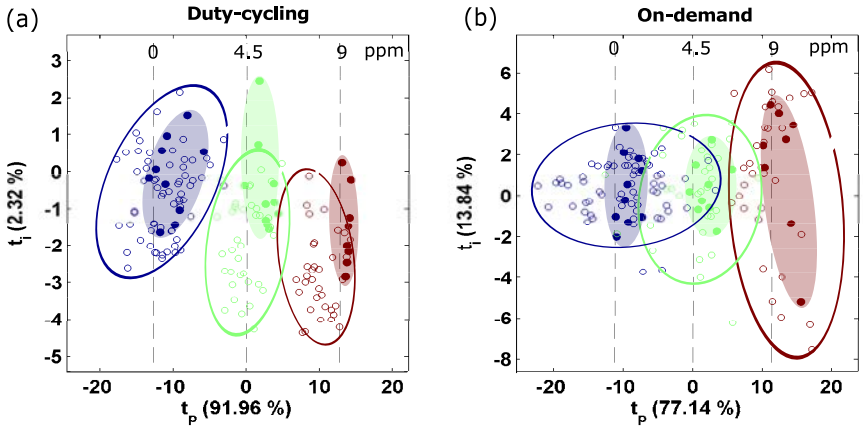


Figure 7. Discontinuous calibration. Score plot of the orthogonal PLS models obtained from one sensor of each group of discontinuously operated sensors. The title of the subplot indicates the target operating mode: (**a**) duty-cycling and (**b**) on-demand. The filled and open circles represent the calibration and validation observations, respectively. Hand drawn ellipses cluster observations of the same concentration level. The concentration (ppm) is indicated in the text at the top of the figure.

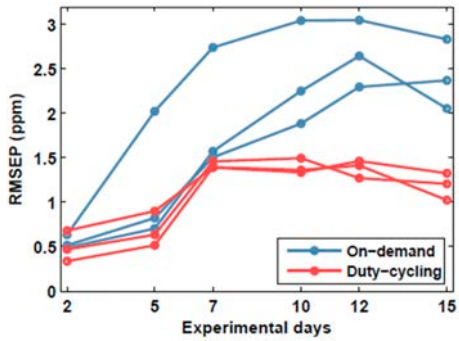


Figure 8. Temporal stability of duty-cycling and on-demand modes, after discontinuous calibration. RMSEP: root mean squared error in prediction.

4. Conclusions

MOX gas sensors are candidate devices for integration in mobile sensing devices, such as smartphones and wearables. Continuously powering a MOX sensor is not feasible in many battery-operated applications due to power consumption limitations or the intended intermittent operation of the device. This work compared two low-power operating modes (duty-cycling and on-demand) in the prediction of low concentrations of carbon monoxide under variations in ambient

humidity and temperature. The sensors were modulated in temperature to increase the selectivity to the target gas and reject chemical interferences. The sensors were calibrated in continuous operation and then benchmarked for two weeks against a group of continuously powered sensors. We found that on-demand operation produced a deformation of the sensor conductance patterns, which led to an increase in the prediction error by almost a factor of 5 as compared to continuous operation (2.2 versus 0.45 ppm). Applying a 10% duty-cycling operation in periods of 10 min reduced the prediction error to a factor of 2 (0.9 versus 0.45 ppm). This means that the proposed duty-cycling powering scheme saved up to 90% energy compared to the continuous operating mode, which could be advantageous for applications that do not require continuous and periodic measurements and that can tolerate slightly higher prediction errors. We found that continuous power during calibration produces better results than calibration under discontinuous operation.

In this work, we did not test other duty cycles or periods. We expect that increasing the duty cycle will further reduce the error at the cost of higher power consumption. Future work includes exploring what is the minimum duty cycle that would achieve errors comparable to continuously operated sensors.

Acknowledgments: This work was partially funded by the Spanish MINECO program under grants TEC2014-59229-R (SIGVOL), PCIN-2013-195 (SENSIBLE), and BES-2015-071698 (SEVERO-OCHOA). The Signal and Information Processing for Sensor Systems group is a consolidated Grup de Recerca de la Generalitat de Catalunya and has support from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2014-SGR-1445). This work has received support from the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya and the European Social Fund (ESF). Additional financial support has been provided by the Institut de Bioenginyeria de Catalunya (IBEC). IBEC is a member of the CERCA Programme/Generalitat de Catalunya.

Author Contributions: J.B. and S.M. conceived and designed the experiments; J.B. performed the experiments and analysed the data; J.B. wrote the paper; S.M. revised the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

In this appendix, we detail the computation of the power consumption in the Samsung Galaxy S7 smartphone, which hosts a 3000 mAh battery and operates at 3.7 V. The available energy (E) is therefore:

$$\mathrm{E} = \mathrm{I} \times \mathrm{V} = 3000\ \mathrm{mAh} \times 3.7\ \mathrm{V} = 11,100\ \mathrm{mWh} \quad \text{(A1)}$$

Assuming a battery life of 24 h (non-intensive use), the average power consumption (P) is given by Equation (A2):

$$\mathrm{P} = \mathrm{E} / \mathrm{T} = 11,000\ \mathrm{mWh} / 24\ \mathrm{h} = 462\ \mathrm{mW} \quad \text{(A2)}$$

Adding a TGS8100 sensor (which consumes 15 mW) to this smartphone, the total power consumption becomes 477 mW. Hence, the MOX sensor represents 3.14% of the total smartphone power. Recomputing the battery lifetime (T) using this power consumption yields:

$$\mathrm{T} = \mathrm{E} / \mathrm{P} = 11,000\ \mathrm{mWh} / 477\ \mathrm{mW} = 23.27\ \mathrm{h} \quad \text{(A3)}$$

The battery lifetime is reduced by 0.73 h = 45 min. Now, we consider the case in which the SGP sensor (Sensirion AG), which consumes 48 mW, is integrated into the smartphone. The power consumption of the full system (smartphone + sensor) will be (462 + 48) = 510 mW. Introducing P = 510 mW in Equation (A3) yields that the new battery lifetime would be 21.76 h. This represents a decrease in battery lifetime of 2.23 h.

References

1. Chansin, G.; Pugh, D. *Environmental Gas Sensors 2017–2027*; CISION: Cambridge, UK, 2017.
2. Pouter, J. *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies*; Pew Research Center: Washington, DC, USA, 2016.
3. Joshi, S.M. The sick building syndrome. *Indian J. Occup. Environ. Med.* **2008**, *12*, 61–64. [CrossRef] [PubMed]
4. Mendell, M.J. Indoor residential chemical emissions as risk factors for respiratory and allergic effects in children: A review. *Indoor Air* **2007**, *17*, 259–277. [CrossRef] [PubMed]
5. WHO. *WHO Guidelines for Indoor Air Quality: Selected Pollutants*; WHO: Geneva, Switzerland, 2010.
6. EPA. Criteria Air Pollutants. In *America's Children and the Environment*; EPA: Washington, DC, USA, 2015; pp. 1–22.
7. Kay, M.; Santos, J.; Takane, M. mHealth: New horizons for health through mobile technologies. *World Heal. Organ.* **2011**, *64*, 66–71.
8. Cavaliere, F.; Volpe, C.; Gargaruti, R.; Poscia, A.; Di Donato, M.; Grieco, G.; Moscato, U. Effects of acute hypoventilation and hyperventilation on exhaled carbon monoxide measurement in healthy volunteers. *BMC Pulm. Med.* **2009**, *9*, 51. [CrossRef] [PubMed]
9. Beg, M.F.S.; Alzoghaibi, M.A.; Abba, A.A.; Habib, S.S. Exhaled nitric oxide in stable chronic obstructive pulmonary disease. *Ann. Thorac. Med.* **2009**, *4*, 65. [PubMed]
10. Davies, S.; Spanel, P.; Smith, D. Quantitative analysis of ammonia on the breath of patients in end-stage renal failure. *Kidney Int.* **1997**, *52*, 223–228. [CrossRef] [PubMed]
11. Webb, A.; Al-Moundhri, M.; Souberbielle, B.E.; Norton, A.; Priest, K.; Ryan, C.; Mendes, R.; Ashley, S.; Smith, I.E.; O'Brien, M.E.R. A randomised study of SRL172 (Mycobacterium vaccae) in patients with advanced lung cancer treated with chemotherapy. *Br. J. Cancer* **1998**, *78*, 4.
12. Zhang, J.; Yao, X.; Yu, R.; Bai, J.; Sun, Y.; Huang, M.; Adcock, I.M.; Barnes, P.J. Exhaled carbon monoxide in asthmatics: a meta-analysis. *Respir. Res.* **2010**, *11*, 50. [CrossRef] [PubMed]
13. King, J.; Kupferthaler, A.; Frauscher, B.; Hackner, H.; Unterkofler, K.; Teschl, G.; Hinterhuber, H.; Amann, A.; Högl, B. Measurement of endogenous acetone and isoprene in exhaled breath during sleep. *Physiol. Meas.* **2012**, *33*, 413. [CrossRef] [PubMed]
14. Kundu, S.K.; Bruzek, J.A.; Nair, R.; Judilla, A.M. Breath acetone analyzer: diagnostic tool to monitor dietary fat loss. *Clin. Chem.* **1993**, *39*, 87–92. [PubMed]
15. Beard, E.; West, R. Pilot study of the use of personal carbon monoxide monitoring to achieve radical smoking reduction. *J. Smok. Cessat.* **2012**, *7*, 12–17. [CrossRef]
16. Delgado, M.K.; Huang, Y.; Wanner, K.; Goldberg, E.; Hemmons, J.; Spencer, E.; Wetherill, R. Test accuracy of smartphone-paired breathalysers: A validation study. *Inj. Prev.* **2017**, *23*, A15. [CrossRef]
17. Kim, Y.-S.; Hwang, I.-S.; Kim, S.-J.; Lee, C.-Y.; Lee, J.-H. CuO nanowire gas sensors for air quality control in automotive cabin. *Sens. Actuators B Chem.* **2008**, *135*, 298–303. [CrossRef]
18. Romain, A.C.; Nicolas, J. Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview. *Sens. Actuators B Chem.* **2010**, *146*, 502–506. [CrossRef]
19. Fine, G.F.; Cavanagh, L.M.; Afonja, A.; Binions, R. Metal oxide semi-conductor gas sensors in environmental monitoring. *Sensors* **2010**, *10*, 5469–5502. [CrossRef] [PubMed]
20. Borrego, C.; Costa, A.M.; Ginja, J.; Amorim, M.; Coutinho, M.; Karatzas, K.; Sioumis, T.; Katsifarakis, N.; Konstantinidis, K.; De Vito, S.; et al. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmos. Environ.* **2016**, *147*. [CrossRef]
21. Adam, G.; Lemaigre, S.; Goux, X.; Delfosse, P.; Romain, A.-C. Upscaling of an electronic nose for completely stirred tank reactor stability monitoring from pilot-scale to real-scale agricultural co-digestion biogas plant. *Bioresour. Technol.* **2015**, *178*, 285–296. [CrossRef] [PubMed]
22. Loutfi, A.; Coradeschi, S.; Mani, G.K.; Shankar, P.; Rayappan, J.B.B. Electronic noses for food quality: A review. *J. Food Eng.* **2015**, *144*, 103–111. [CrossRef]
23. Perera, A.; Pardo, A.; Barrettino, D.; Hierlermann, A.; Marco, S. Evaluation of fish spoilage by means of a single metal oxide sensor under temperature modulation. *Sens. Actuators B Chem.* **2010**, *146*, 477–482. [CrossRef]

24. Sekhar, P.K.; Brosha, E.L.; Mukundan, R.; Li, W.; Nelson, M.A.; Palanisamy, P.; Garzon, F.H. Application of commercial automotive sensor manufacturing methods for NOx/NH3 mixed potential sensors for on-board emissions control. *Sens. Actuators, B Chem.* **2010**, *144*, 112–119. [CrossRef]
25. Boon-Brett, L.; Bousek, J.; Moretto, P. Reliability of commercially available hydrogen sensors for detection of hydrogen at critical concentrations: Part II—Selected sensor test results. *Int. J. Hydrogen Energy* **2009**, *34*, 562–571. [CrossRef]
26. Hübert, T.; Boon-Brett, L.; Palmisano, V.; Bader, M.A. Developments in gas sensor technology for hydrogen safety. *Int. J. Hydrogen Energy* **2014**, *39*, 20474–20483. [CrossRef]
27. Kwor, R. *Carbon Monoxide Detectors, Carbon Monoxide Toxicity*; CRC: Boca Raton, FL, USA, 2000; pp. 61–82.
28. Tomchenko, A.A.; Harmer, G.P.; Marquis, B.T. Detection of chemical warfare agents using nanostructured metal oxide sensors. *Sens. Actuators B Chem.* **2005**, *108*, 41–55. [CrossRef]
29. Šundić, T.; Marco, S.; Perera, A.; Pardo, A.; Hahn, S.; Bârsan, N.; Weimar, U. Fuzzy inference system for sensor array calibration: prediction of CO and CH_4 levels in variable humidity conditions. *Chemom. Intell. Lab. Syst.* **2002**, *64*, 103–122. [CrossRef]
30. Kuske, M.; Padilla, M.; Romain, A.C.; Nicolas, J.; Rubio, R.; Marco, S. Detection of diverse mould species growing on building materials by gas sensor arrays and pattern recognition. *Sens. Actuators B Chem.* **2006**, *119*, 33–40. [CrossRef]
31. Righettoni, M.; Amann, A.; Pratsinis, S.E. Breath analysis by nanostructured metal oxides as chemo-resistive gas sensors. *Mater. Today* **2015**, *18*, 163–171. [CrossRef]
32. Wilson, A.D.; Baietto, M. Advances in electronic-nose technologies developed for biomedical applications. *Sensors* **2011**, *11*, 1105–1176. [CrossRef] [PubMed]
33. 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. [CrossRef]
34. Padilla, M.; Fonollosa, J.; Marco, S. *Improving the Robustness of Odour Sensing Systems by Multivariate Signal Processing*; IGI-Global: Hershey, PA, USA, 2013.
35. Ruhland, B.; Becker, T.; Müller, G. Gas-kinetic interactions of nitrous oxides with SnO2 surfaces. *Sens. Actuators B Chem.* **1998**, *50*, 85–94. [CrossRef]
36. Zampolli, S.; Elmi, I.; Cozzani, E.; Cardinali, G.C.; Scorzoni, A.; Cicioni, M.; Marco, S.; Palacio, F.; Gómez-Cama, J.M.; Sayhan, I.; et al. Ultra-low-power components for an RFID Tag with physical and chemical sensors. *Microsyst. Technol.* **2008**, *14*, 581–588. [CrossRef]
37. AMG AG. *AMS AS-MLV-P2 Air Quality Sensor*; AMS AG: Unterpremstätten, Austria, 2015.
38. Figaro Engineering. Inc. *Technical Information for Tgs2610*; Figaro Engineering. Inc.: Arlington Heights, IL, USA, 2012.
39. Clifford, P.K.; Tuma, D.T. Characteristics of semiconductor gas sensors II. Transient response to temperature change. *Sens. Actuators* **1982**, *3*, 255–281. [CrossRef]
40. Sayhan, I.; Helwig, A.; Becker, T.; Muller, G.; Elmi, I.; Zampolli, S.; Padilla, M.; Marco, S. Discontinuously Operated Metal Oxide Gas Sensors for Flexible Tag Microlab Applications. *IEEE Sens. J.* **2008**, *8*, 176–181. [CrossRef]
41. Wu, H.; Siegel, M. Odor-based incontinence sensor. In Proceedings of the 17th IEEE Instrumentation and Measurement Technology Conference, Baltimore, MD, USA, 1–4 May 2000; Volume 1, pp. 63–68.
42. Macias, M.M.; Orellana, C.J.G.; Velasco, H.M.G.; Manso, A.G.; Garzon, J.E.A.; Santamaria, H.S. Gas sensor measurements during the initial action period of duty-cycling for power saving. *Sens. Actuators B Chem.* **2017**, *239*, 1003–1009. [CrossRef]
43. Jelicic, V.; Oletic, D.; Sever, T.; Bilas, V. Evaluation of mox gas sensor transient response for low-power operation. In Proceedings of the 2015 IEEE Sensors Applications Symposium (SAS), Zadar, Croatia, 13–15 April 2015; pp. 1–5.
44. Rossi, M.; Brunelli, D. Ultra low power CH4 monitoring with wireless sensors. In Proceedings of the 2015 IEEE Sensors Applications Symposium (SAS), Baltimore, MD, USA, 3–6 November 2013; pp. 1–4.
45. Choi, S.; Kim, N.; Cha, H.; Ha, R. Micro sensor node for air pollutant monitoring: Hardware and software issues. *Sensors* **2009**, *9*, 7970–7987. [CrossRef] [PubMed]
46. Oletic, D.; Jelicic, V.; Antolovic, D.; Bilas, V. Energy-efficient atmospheric CO concentration sensing with on-demand operating MOX gas sensor. In Proceedings of the 2014 IEEE Sensors, Valencia, Spain, 2–5 November 2014; pp. 795–798.

47. Bicelli, S.; Depari, A.; Faglia, G.; Flammini, A.; Fort, A.; Mugnaini, M.; Ponzoni, A.; Vignoli, V.; Rocchi, S. Model and experimental characterization of the dynamic behavior of low-power carbon monoxide MOX sensors operated with pulsed temperature profiles. *IEEE Trans. Instrum. Meas.* **2009**, *58*, 1324–1332. [CrossRef]
48. Barsan, N.; Koziej, D.; Weimar, U. Metal oxide-based gas sensor research: How to? *Sens. Actuators B Chem.* **2007**, *121*, 18–35. [CrossRef]
49. Lee, A.P.; Reedy, B.J. Temperature modulation in semiconductor gas sensing. *Sens. Actuators B Chem.* **1999**, *60*, 35–42. [CrossRef]
50. Vergara, A.; Ramírez, J.L.; Llobet, E. Reducing power consumption via a discontinuous operation of temperature-modulated micro-hotplate gas sensors: Application to the logistics chain of fruit. *Sens. Actuators B Chem.* **2008**, *129*, 311–318. [CrossRef]
51. F.I.S. Inc. *Fis Gas Sensor SB-500-12*; F.I.S. Inc.: Jacksonville, FL, USA, 2017.
52. Conrad, T.; Hiry, P.; Schutze, A. PuMaH-a temperature control and resistance read-out system for microstructured gas sensors based on PWM signals. In Proceedings of the 2005 IEEE Sensors, Irvine, CA, USA, 31 October–3 November 2005.
53. Sensirion, A.G. Data Sheet SHT7x (SHT71, SHT75)-Humidity and Temperature Sensor IC. Available online: http//www.Sensirion.Com/en/products/humidity-Temperature/humidity-Sensor-sht75 (accessed on 12 November 2017).
54. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [CrossRef]
55. Wold, S.; Trygg, J.; Berglund, A.; Antti, H. Some recent developments in PLS modeling. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 131–150. [CrossRef]
56. Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16*, 119–128. [CrossRef]
57. Burgues, J.; Marco, S. Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors. *Anal. Chim. Acta.* **2018**, accepted.
58. Di Carlo, S.; Falasconi, M. Drift correction methods for gas chemical sensors in artificial olfaction systems: techniques and challenges. In *Advances in Chemical Sensors*; InTech: London, UK, 2012.
59. Korotcenkov, G.; Cho, B.K. Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey). *Sens. Actuators B Chem.* **2011**, *156*, 527–538. [CrossRef]
60. Everitt, B.; Skrondal, A. *The Cambridge Dictionary of Statistics*; Cambridge University Press: Cambridge, UK, 2002; Volume 106.

Chapter 5

Improving the Response Time of MOX Sensors

The low bandwidth of MOX sensors (<0.1 Hz) is a major hurdle for mobile robots tracking turbulent odor plumes, where detection of intermittent odor patches is key. Two common approaches in the MRO field to minimize the negative effects introduced by inherent low-pass filtering of MOX sensors is to drive the robot at low speeds or to use a stop-sense-go strategy, which increases the mission time considerably. In this chapter, we present a fast-response miniaturized e-nose composed of four naked MOX sensors and a digital band-pass filter that can boost the bandwidth of the system close to 1 Hz. The device was attached to a fast photo-ionization detector (330 Hz) to quantify the response time during exposure to turbulent gas plumes. The results indicate that the digital filter can improve the response time by at least a factor of 4, bringing new possibilities to mobile robot olfaction.

The content of the chapter is provided in Paper IV (shown below).

HIGH-BANDWIDTH E-NOSE FOR RAPID TRACKING OF TURBULENT PLUMES

Javier Burgués[1,2,*], *Luis F. Valdez*[3] *and Santiago Marco*[1,2]

[1] Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain
[2] Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028 Barcelona, Spain
[3] Bioelectronics Section, Department of Electrical Engineering, CINVESTAV, Mexico City 07360, Mexico
[*]Author to whom correspondence should be addressed.

ABSTRACT

The low bandwidth of metal oxide semiconductor (MOX) sensors (<0.1 Hz) is a major hurdle to gas source localization (GSL) in turbulent environments where detection of intermittent odor patches is key. We present a fast-response miniaturized electronic nose (Fast-eNose) composed of four naked MOX sensors and a digital band-pass filter that can boost the bandwidth of the system close to 1 Hz. The device was attached to a fast photo-ionization detector (330 Hz) to quantify the response time during exposure to turbulent gas plumes. The results indicate that the digital filter can improve the response time by at least a factor of 4, bringing new possibilities to mobile robot olfaction.

***Index Terms*—** MOX, signal processing, response time, CFD, gas plume, gas sensors

1. INTRODUCTION

A pre-requisite to successful olfactory search in turbulent environments is the detection of intermittent odor patches [1]. Insects such as the moth, which efficiently use odor plumes for mating and foraging can resolve ten pulses of pheromones in 1 s [1], which represents a bandwidth of at least 10 Hz. In contrast, most chemical sensors have sub-Hz bandwidths, which hinders the application of mobile robots to olfaction-related tasks [2]. Although fast-response photo-ionization detectors (PID) with a bandwidth of 330 Hz (miniPID 200B, Aurora Scientific, Canada) have been on the market for more than 15 years, they have been rarely used in mobile robots [3] probably due to the high cost, large volume, low specificity and insufficient sensitivity of this technology. Instead, low-cost MOX sensors are much more popular for this application due to their low cost, compact size, ease of use and high sensitivity to many gases. However, MOX sensors are slow (response time of 20-30 s) and not selective. While the lack of selectivity can be improved by combining multiple MOX sensors with different properties and pattern recognition algorithms [4], improving the response time is the main goal of this paper.

1.1. Related work

Previous research suggests that the response time of sensing systems based on MOX sensors can be improved by removing the sensor cap [1], by technology improvements [5], by novel system hardware designs [6,7] or by signal processing [8–10]. Removing the external housing of the sensor can increase the bandwidth up to 1 Hz [1], but the sensor might become less stable and more prone to damage. Vincent et al. [5] recently presented novel coatings of the sensing surface that can reduce the response time down to 10 s, which is insufficient to track chemical plumes in a wind tunnel (with concentration variations of 1 s). To reduce the effect of the sensor chamber on the system dynamics, Gonzalez-Jimenez et al. [6] designed a multi-chamber e-nose in which replicate MOX sensors stored in independent gas chambers are exposed to the gas in turns. The output of the system is the concatenation of the MOX responses during the rising phase, effectively ignoring the slow recovery phases. Batog et al. [7] demonstrated that the detection of a chemical trail on the ground can be improved by perforating the sensor housing and placing a small DC fan on top of it to create a recirculating airflow pattern. The addition of external elements, such as fans or electro-valves increase the cost, size, weight and power consumption.

Among the signal processing approaches, first order inverse dynamical models [9,10] or artificial neural networks [11] have been proposed to retrieve the fast excitation signal from the slow sensor response. Inverse modelling requires a complex setup to calibrate the model parameters (e.g. the time constants of an exponential) for each substance and sensor unit. This is normally done in controlled laboratory conditions where squared concentration pulses are delivered to the sensors. Then, it is unclear if the calibration model can work properly in a real scenario where turbulence chaotically delivers the chemicals to the sensors. In this sense, Schmuker et al. [12] recently obtained positive results in wind tunnel experiments using the band-pass filter proposed by Muezzinoglu et al. [10].

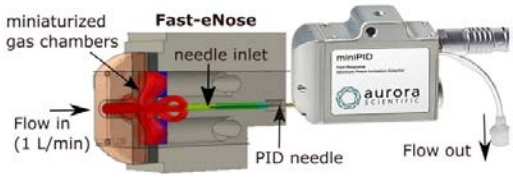


Figure 1. Schematic representation of the miniNose attached to the miniPID (not at scale).

1.2. Fast-response miniaturized e-nose (Fast-eNose)

We propose a fast-response miniaturized e-nose (Fast-eNose) specifically designed as an attachment to a miniPID (Fig. 1). The PID is used as a reference device to evaluate the dynamics of the MOX sensors and to draw the sample inside the e-nose (it has a pump that samples at 1 L/min).

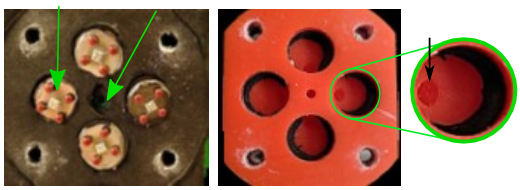


Figure 2. Photos of the (a) Naked MOX sensors; (b) Miniaturized sensor chambers.

2. MATERIALS AND METHODS

2.1. Fluidic design

The fluidic design is an assembly of 3D-printed pieces consisting of four small chambers and fluidic channels with a total combined volume of 1.3mL (Fig. 2). The fluidic channels split the inlet flow into four streams that fill the individual chambers in 75 ms (CFD simulation results), exit them from a different hole and are merged again into a single flow that enters the miniPID after 3 ms.

2.2. Electronics and signal processing

Four naked MOX sensors (i.e. without cap), two TGS 2602 and two TGS 2610 (Figaro Engineering, Japan) are placed inside the gas chambers (Fig. 2a). The sensors are powered at the nominal voltage (5 V) and their output is acquired at 1 kHz using a simple voltage divider ($R_L = 10\,k\Omega$) and a USB-6002 datalogger (National Instruments, USA). The PID was sampled synchronously with the MOX sensors. The MOX response is processed with a band-pass filter [10], computed as a moving average filter (window size of 20 ms) followed by a derivative, to further increase the bandwidth.

2.2. Experimental evaluation

To evaluate the response time of the fastNose, we generated turbulent plumes in an open environment using a pressurized air outlet (6.3 mm radius, 20 L/min) and a beaker (5 cm radius) filled with 200 mL of ethanol, which acts as a gas source (Fig. 3). The air outlet was placed at 150 cm to the sensing board and the beaker could be moved in the range 15-135 cm. The experimental protocol consists on placing the gas source at a random distance among the following ones [15, 45, 75, 105 and 135] cm and recording the signals of the MOX sensors and the miniPID for 5 minutes. Then, the similarity between the MOX sensor signals and the miniPID is computed using the cross-covariance, or cross-correlation of the mean centered signals

$$Cov_{xy}[m] = E\left[(\boldsymbol{x}[n] - \mu_{\boldsymbol{x}})(\boldsymbol{y}[n+m] - \mu_{\boldsymbol{y}})^T\right] \qquad \text{Eq. 1}$$

where $\mu_{\boldsymbol{x}}$ and $\mu_{\boldsymbol{y}}$ are the mean values of $\boldsymbol{x}$ and $\boldsymbol{y}$. The 99% confidence intervals are computed as 2.58 times the standard deviation of the covariance at negative lags, as these represent the unlikely situation of the MOX signal predicting future concentrations.

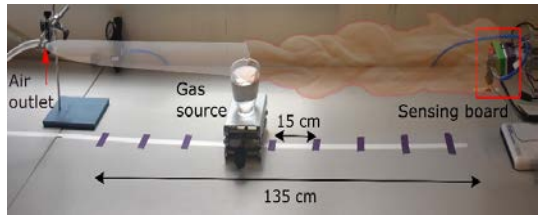


Figure 3. Generator of turbulent plumes.

3. RESULTS

The raw signals (Fig. 4) illustrate the long recovery time of the MOX sensor response and how fast concentration fluctuations are low-pass filtered by the sensor. The filtered signal is a much better reproduction of the ground truth (PID signal) at all tested distances. The cross-covariance (Fig. 5) allows an easy comparison between the dynamics of the MOX sensor and the miniPID. The MOX signal is a smoothed version of the PID signal, as indicated by significant correlation between the PID and lagged MOX responses. The slower the MOX response the longer is the correlation tail. Near the source, the correlation of the differentiated MOX signal disappears in about 2.3 s compared to 9.2 s for the raw signal. Since the correlation tail and the bandwidth are inversely proportional, the differentiator filter improves the bandwidth by a factor of 4. At longer distances to the source, the correlation falls very fast towards negative values but remains within the 99% confidence limits for the the hypothesis of null correlation. Taking the point where it reaches this band as the length of the correlation, we may claim that the digital differentiation improves the sensor bandwidth by a factor of 15.

4. DISCUSSION

MOX sensors can achieve an improved response time by combining specially designed fluidic connections and digital signal processing. The Fast-eNose can be used separately

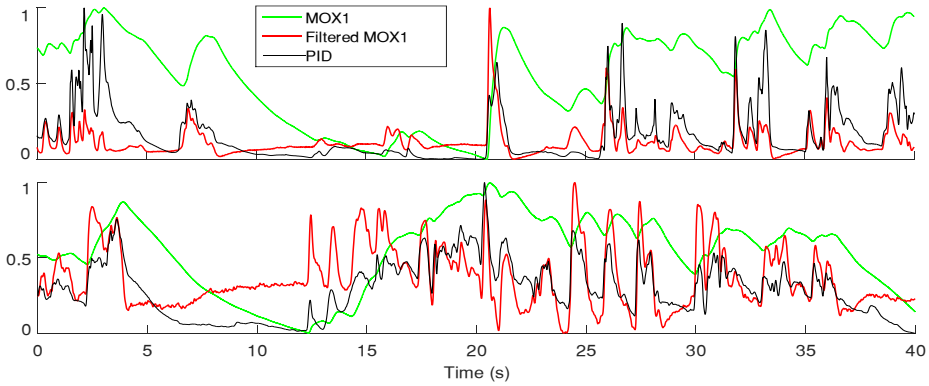


Figure 4. Raw and filtered signals of one MOX sensor (TGS 2610) at 15 cm (top) and 135 cm (bottom) to the source.

from the miniPID (using an alternative pump) or as a complement to it, extending the range of compounds detectable by the PID, as PIDs cannot detect molecules with high ionization potential, such as carbon monoxide (CO). GSL algorithms based on 'odor hits' (e.g. infotaxis [13]) can potentially increase their performance using the Fast-eNose.

An obvious follow-up research is to combine the output of the multiple MOX sensors to discriminate multiple gas sources, and testing the device in a longer distance range or on board of a mobile robot for GSL and mapping. Further work may also include the optimization of the digital differentiator to further reduce the time constant of the MOX sensor signals.

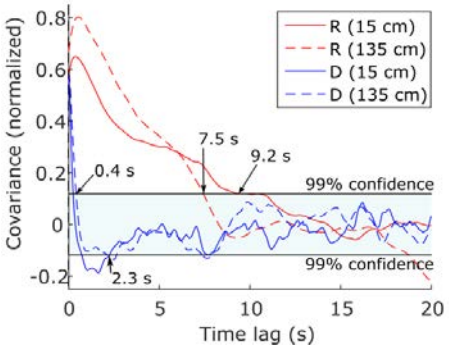


Figure 5. Cross-covariance between the miniPID and the MOX response (red) or the filtered response (blue) at 15 cm (solid line) and 135 cm (dashed line) to the source. The arrows point to the lag where the covariance becomes not statistically significant.

5. REFERENCES

[1] D. Martinez, L. Arhidi, et al., Using insect electroantennogram sensors on autonomous robots for olfactory searches, J. Vis. Exp. JoVE. (2014).

[2] V. Hernandez Bennetts, A.J. Lilienthal, et al., Mobile Robots for Localizing Gas Emission Sources on Landfill Sites: Is Bio-Inspiration the Way to Go?, 4 (2012) 20.

[3] J.K. Bailey, M.A. Willis, R.D. Quinn, A multi-sensory robot for testing biologically-inspired odor plume tracking strategies, in: Adv. Intell. Mechatronics. Proceedings, 2005 IEEE/ASME Int. Conf., 2005: pp. 1477–1481.

[4] J.W. Gardner, P.N. Bartlett, Performance definition and standardization of electronic noses, Sensors Actuators B Chem. 33 (1996) 60–67.

[5] T.A. Vincent, Y. Xing, et al., Investigation of the response of high-bandwidth MOX sensors to gas plumes for application on a mobile robot in hazardous environments, Sensors Actuators B Chem. 279 (2019) 351–360.

[6] J. Gonzalez-Jimenez, J.G. Monroy, et al., The Multi-Chamber Electronic Nose-An Improved Olfaction Sensor for Mobile Robotics, Sensors. 11 (2011) 6145–6464.

[7] P. Batog, A. Wołczowski, Odor markers detection system for mobile robot navigation, in: Procedia Eng., 2012: pp. 1442–1445.

[8] J. Fonollosa, A. Vergara, et al., Estimation of the Limit of Detection Using Information Theory Measures, Anal. Chim. Acta. 810 (2013) 2013–2015.

[9] J.G. Monroy, J. Gonźalez-Jimńez, J.L. Blanco, Overcoming the slow recovery of MOX gas sensors through a system modeling approach, Sensors (Switzerland). 12 (2012) 13664–13680.

[10] M.K. Muezzinoglu, A. Vergara, et al., Acceleration of chemo-sensory information processing using transient features, Sens Actuators, B Chem. 137 (2009) 507–512.

[11] J. Fonollosa, S. Sheik, et al., Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring, Sens Actuators B Chem. 215 (2015) 618–629.

[12] M. Schmuker, V. Bahr, et al., Exploiting plume structure to decode gas source distance using metal-oxide gas sensors, Sens Actuators, B Chem. 235 (2016) 636–646.

[13] M. Vergassola, et al., "Infotaxis" as a strategy for searching without gradients., Nature. 445 (2007) 406–409.

Chapter 6

Feature Extraction of MOX Sensor Signals for Gas Source Localization

Chemical sensor signals in response to a turbulent plume show high level of complexity and apparent randomness. There is no an accepted consensus on the best approach to extract valuable hidden information from these signals. In this chapter, we study which are the optimal features of a gas sensor signal for localizing a gas source in a turbulent environment. Specifically, we compare statistical descriptors directly extracted from the sensor response (mean, variance and maximum response) with transient features computed in the derivative of the response (the so-called "bouts").

A desirable behaviour of a signal feature for gas source localization is that its value is monotonically related to the gas source distance. In Paper V, we compare the different features based on the probability of monotonicity, defined as the number of experiments in which the the curve relating the feature value and the source distance is monotonic divided by the total number of experiments. The experimental setup includes a generator of turbulent gas plumes and a sensing board composed of three different MOX sensors. The plume is generated by passing a stream of synthetic air over a beaker filled with ethanol. The sensor signals are recorded for five minutes at four distances to the gas source (range 10-230 cm). We compute the signal features in a random chunk (of 30 s) of the measurement at each distance, and assess the monotonicity of the curve relating the feature value and the source distance. The process is repeated 100 times to obtain the probability of monotonicity. The main conclusion of the paper is that the maximum response is the most robust feature across the three studied sensors.

In Paper VI, we propose a method to extract transient signal features that are strongly correlated to the distance of the source under constant emission conditions. The paper follows-up on the research work “Exploiting plume structure to decode gas source distance using metal-oxide gas sensors” (Schmuker et al., Sens Actuator B-Chem, 2016). In that work, the authors propose a signal processing method to extract specific features from

the derivative of the response of MOX sensors, which are then used to predict the distance of a gas source. To compute the derivative of a noisy signal, they use an overly complicated digital filter with undesirable properties such as nonlinear phase or lack of causality, distorting the waveform of the signal and hindering real-time operation. The impact of key parameters of the algorithm, such as the length of the measurement window, or external factors, such as the wind speed, into the predictive performance was not studied.

In our paper, we improve Schmuker's algorithm by using a causal linear-phase finite impulse response (FIR) differentiator, optimizing all algorithm parameters in a multivariate way and validating the results against non-matching train and test wind speeds. This truly enables the use of bout-based features for gas source localization without using anemometry. We also study how the length of the measurement window affects the predictive performance. We demonstrate that certain combinations of parameters can reduce the prediction error to 8 cm (in a distance range of 1.45 m) improving the performance reported by Schmuker et al. in the same dataset by at least a factor of 2.5. An additional contribution is that we provide a MATLAB implementation of the improved algorithm and all of the analysis code used in the study.

Papers V and VI are shown below.

FEATURE EXTRACTION OF GAS SENSOR SIGNALS FOR GAS SOURCE LOCALIZATION

Javier Burgués[1,2,*] *and Santiago Marco*[1,2]

[1] Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain
[2] Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028 Barcelona, Spain
[*]Author to whom correspondence should be addressed.

ABSTRACT

This paper explores which signal features of a gas sensor are optimum for assessing the proximity to a gas source in an open environment. Specifically, we compare three statistical descriptors of the signal (mean, variance and maximum response) against the 'bout' frequency, a feature computed in the derivative of the response. The experimental setup includes a generator of turbulent plumes and a sensing board composed of three metal oxide (MOX) sensors of different types. The main conclusion is that the maximum response is the most robust feature across the three sensors. The 'bout' frequency can be very sensitive to an additional parameter (the noise threshold).

Index Terms— Gas source localization, gas sensors, signal processing, feature extraction, MOX, gas plume.

1. INTRODUCTION

MOX sensors are widely used in mobile robotic olfaction due to their low cost, small size, high sensitivity to many gases and ease of use. A key question is to determine which signal features are optimum for localizing a gas source in turbulent air flow conditions. Experimental results using MOX sensors in an unventilated indoor environment suggest that the magnitude of concentration fluctuations is a more reliable indicator of source proximity than the instantaneous or mean concentration [1-2]. However, the recovery time of MOX sensors (on the order of 10-30 s) is too slow to accurately measure the concentration fluctuations of a turbulent gas plume. To improve the response time, Schmuker et al. [3] proposed a signal processing method that extracts short time-scale features (the so-called 'bouts') from the derivative of the response. In wind tunnel experiments, the authors found that the frequency of 'bouts' with amplitude higher than a certain threshold linearly decreases with distance to the source, if the sensor is placed in the plume centerline. Based on these results, they claim that a robot could use the bout frequency to navigate along the plume centerline towards the point of release without anemometry. However, Vuka et al. [4] could not find a monotonic relationship between the bout frequency and the distance to the gas source in an open environment, probably because the noise threshold was not optimized for the sensor model they used.

The estimation of the noise level in the derivative to reduce false bout detections is a critical part of the bout algorithm that has not been studied in detail previously. Schmuker et al. proposed a threshold computed as the mean plus three standard deviations ($\mu+3\sigma$) of the amplitude of the bouts detected in the sensor baseline (i.e. in the absence of gas). However, they did not prove that such threshold is optimum for every sensor model.

In this work, we explore how different noise thresholds affect the performance of the bout frequency for localizing a gas source in an open environment, using three different MOX sensors. We propose modifications of the bout detection algorithm to detect bouts in real time. We finally compare the bout frequency with optimum threshold against the mean and variance of the response.

2. MATERIALS AND METHODS

2.1. Detection of 'bouts'

The bout detection algorithm [3] was modified by replacing the non-causal Gaussian filter with a first order infinite impulse response (IIR) filter, decoupling the derivative from the smoothing filter and smoothing the second derivative (Fig. 1). This enables real-time operation of the algorithm, reduces the number of parameters and slightly improves the segmentation of the bouts.

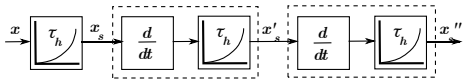


Figure 1. Bout detection algorithm.

The sensor response x is first smoothed using an IIR filter (with time constant τ). The 'bouts' are the rising edges of x_s', the smoothed derivative of x_s, which are delimited by two consecutive zero-crossings of the positive derivative of x_s', i.e. $x_s'' > 0$. The amplitude of a bout is x_s' at the end of the respective bout segment minus x_s' at the start of the same bout segment. Bouts with amplitude lower than a threshold (b_{thr}) are filtered out.

$$b_{\text{thr}} = \mu + 3\sigma \qquad \text{Eq. 1}$$

where μ and σ are the mean (first moment) and standard deviation (second moment), respectively, of the distribution of baseline bout amplitudes.

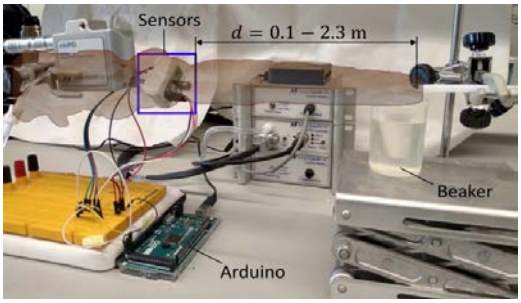


Figure 2. Generator of turbulent plumes and sensing platform.

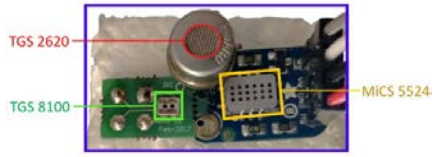


Figure 3. Sensing platform composed of three MOX sensors.

2.2. Experimental setup

The experimental setup comprises a sensing platform and a generator of turbulent plumes (Fig. 2). The sensing platform integrates 3 MOX sensors (Figaro TGS 8100, Figaro TGS 2620 and SGX MiCS 5524), which are fixed to a piece of expanded polystyrene (EPS) very close to each other (sensing area ≈1.9 cm^2) to ensure homogeneous exposure to the plume (Fig. 3). The sensors are powered by an Arduino Mega 2560 microcontroller and the responses are simultaneously acquired at 1 KHz by a USB-6002 datalogger (National Instruments) configured in differential mode. The plume generator, which consists of a stream of synthetic air (1 L/min) passing over a beaker filled with 200 mL of ethanol, was positioned at four distances (0.1-2.3 m) to the sensing unit and each measurement took 5 minutes.

3. RESULTS

The mean amplitude and variability of the recorded signals (Fig. 4) decrease with distance to the source. The MiCS 5524 sensor shows the highest sensitivity near the source but the lowest one far from the source. The mean response of the TGS 2620 at 2.3 m is clearly distinct from the baseline response but it is constant in the range 10-50 cm. The TGS 8100 achieves a tradeoff between both behaviours.

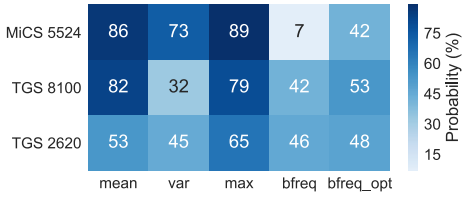
The amplitude threshold has a strong impact on the performance of the bout frequency (Fig. 5). If the threshold is too low, the relationships between bout frequency and distance is non-monotonic. If the threshold is too high, the sensitivity far from the source decreases to zero. The μ+3σ threshold might be too low for the MiCS 5524 sensor (non-monotonic behaviour), optimum for the TGS 8100 and too high for the TGS 2620 (low sensitivity). The optimum thresholds for the MiCS 5524 and TGS 2620 are one order of magnitude above and below μ+3σ, respectively (Table 1).

Table 1. Optimum bout parameters for each sensor.

	τ (ms)	b_{thr}[μ+3σ] (V/s)	b_{thr}[optimum] (V/s)
MiCS 5524	50	0.0021	0.075
TGS 8100	100	0.019	0.075
TGS 2620	200	0.0034	0.0003

To evaluate which features are more likely to produce monotonic behaviors, we split each 5-minute measurement into 10 chunks of 30 s (this is a reasonable measurement window for a mobile robot). Then, we compute the selected signal features (mean, variance, etc.) in random chunks taken from different distances and repeat this process 100 times to visualize all possible relationships between the selected feature and the distance to the source (Fig. 6). We can for example observe that the mean response is monotonic with high probability in the MiCS 5524 sensor but not in the TGS 2620. We repeated this process with the other features and summarized the results in Table 2. The mean and maximum response both achieve monotonicity with high probability (> 80%) in the MiCS 5524 and TGS 8100 but only with less than 65% in the TGS 2620. The bout frequency is monotonic in approximately 50% of the trials, but the μ+3σ threshold works particularly bad in the MiCS 5524 (7% probability), which can explain the findings of Vuka et al. Optimizing the threshold can improve the results but the results are still not comparable to the mean and maximum of the response.

Table 2. Probability of monotonic behaviour.



4. CONCLUSIONS

The maximum response is in average the most robust feature in the three studied sensors, followed by the mean response. The variance, which some authors consider a good indicator of source proximity, did not work well in our experiments. The μ+3σ threshold was too low in the MiCS 5524 sensor, leading to non-monotonic increasing behaviour. Replicate experiments in a larger distance range and with multiple units of each sensor shall be performed to confirm the results.

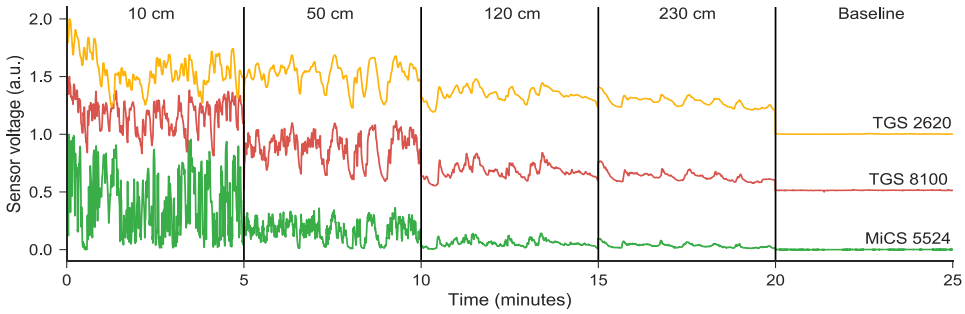


Figure 4. Sensor signals (normalized to unit range and shifted for better visualization) at multiple distances to the gas source.

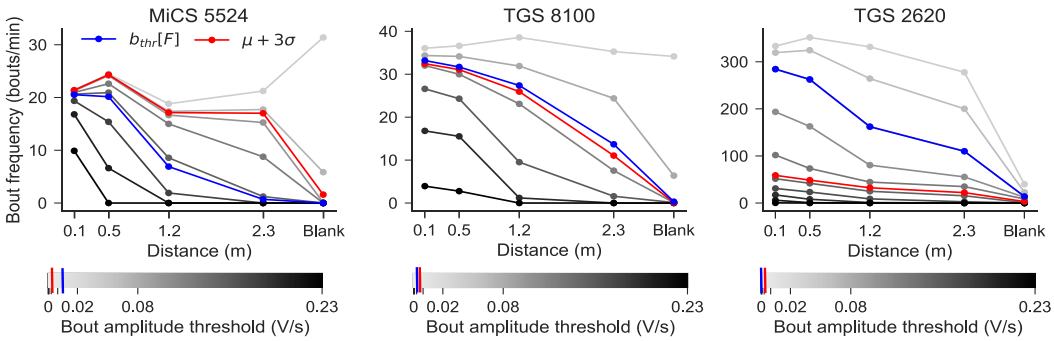


Figure 5. Bout frequency versus distance to the source for different amplitude thresholds (see color map). Each point is the average of ten windows of 30 s. The optimum and μ+3σ thresholds (Table 1) are highlighted in blue and red, respectively.

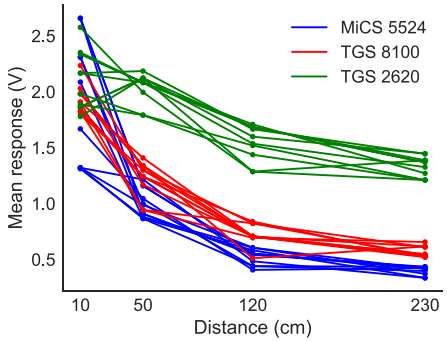


Figure 6. Mean response versus distance to the source in ten random trials.

5. REFERENCES

[1] Lilienthal, Achim J., et al. "Indicators of gas source proximity using metal oxide sensors in a turbulent environment." Proceedings of the IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics (Biorob), 2006, pp. 20-22.

[2] Ferri, Gabriele, et al. "SPIRAL: A novel biologically-inspired algorithm for gas/odor source localization in an indoor environment with no strong airflow." Robotics and Autonomous Systems, 57 (4), 2009, pp. 393-402.

[3] Schmuker, Michael, et al. "Exploiting plume structure to decode gas source distance using metal-oxide gas sensors." Sensors and Actuators B: Chemical, 235, 2016, pp. 636-646.

[4] Vuka, Mikel, et al. "Exploration and localization of a gas source with MOX gas sensors on a mobile robot—A Gaussian regression bout amplitude approach." ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), 2017, pp 1-3.

[5] Cabrita, Gonçalo, et al. "PlumeSim-player/stage plume simulator." Proceedings of the ICRA 2010 Workshop for Networked and Mobile Robot Olfaction in Natural, Dynamic Environments, Örebro, Sweden. 2010, pp 3-8.

Feature Extraction for Transient Chemical Sensor Signals in Response to Turbulent Plumes: Application to Chemical Source Distance Prediction

Javier Burgués [1, 2, *] and Santiago Marco [1,2]

[1] Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain
[2] Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028 Barcelona, Spain
* Correspondence: jburgues@ibecbarcelona.eu; Tel.: +34-934-029-070

Abstract: This paper describes the design of a linear phase low-pass differentiator filter with a finite impulse response (FIR) for extracting transient features of gas sensor signals (the so-called "bouts"). The detection of these bouts is relevant for estimating the distance of a gas source in a turbulent plume. Our current proposal addresses the shortcomings of previous 'bout' estimation methods, namely: (i) they were based in non-causal digital filters precluding real time operation, (ii) they used non-linear phase filters leading to waveform distortions and (iii) the smoothing action was achieved by two filters in cascade, precluding an easy tuning of filter performance. The presented method is based on a low-pass FIR differentiator, plus proper post-processing, allowing easy algorithmic implementation for real-time robotic exploration. Linear phase filters preserve signal waveform in the bandpass region for maximum reliability concerning both 'bout' detection and amplitude estimation. As a case study, we apply the proposed filter to predict the source distance from recordings obtained with metal oxide (MOX) gas sensors in a wind tunnel. We first perform a joint optimization of the cut-off frequency of the filter and the bout amplitude threshold, for different wind speeds, uncovering interesting relationships between these two parameters. We demonstrate that certain combinations of parameters can reduce the prediction error to 8 cm (in a distance range of 1.45 m) improving previously reported performances in the same dataset by a factor of 2.5. These results are benchmarked against traditional source distance estimators such as the mean, variance and maximum of the response. We also study how the length of the measurement window affects the performance of different signal features, and how to select the filter parameters to make the predictive models more robust to changes in wind speed. Finally, we provide a MATLAB implementation of the bout detection algorithm and all analysis code used in this study.

Keywords: Gas sensors; Differentiator; Low pass filter; Metal oxide semiconductor; MOX sensors; Signal processing; Feature extraction; Gas source localization; Robotics;

1. Introduction

While the spatial distribution of a time-average model of a chemical plume is often described by a Gaussian model [1], the dynamics of chemical concentration in the plume are complex. In the vicinity of the chemical source, the instantaneous concentration fluctuations depart strongly from Gaussianity and are characterized by strong bursts and intermittent periods of near-zero concentration. Yee et al. described that the concentration distribution of a turbulent plume fits well a Clipped-Gamma probability density function [2]. In the frequency domain, chemical plumes are wide-band random signals with substantial power spectral density (PSD) up to several KHz [3]. While it is accepted that insects are able to find hidden information from these complex signals for odor navigation purposes [4], the sub-Hz bandwidth of chemical sensor signals largely limits the efficacy of information retrieval.

One path to follow is to improve the dynamics of the sensor response by signal processing methods. While it is possible to model the direct dynamic behavior of chemical multisensory systems [5–7], for the purpose of real-time signal processing it is more convenient to use non-linear dynamic inverse filters [8,9]. In a very interesting approach, Di Lello et al. model the sensor dynamics as an Augmented Switching Linear Dynamical System [10] and use the Expectation Correction Inference

Method [11] to estimate the ground truth gas concentration. However, many of these attempts have been tested in laboratory conditions where the presence of a sensor chamber masks the underlying sensor dynamics. In few cases, this strategy has been used to compensate naked gas sensors for robotics applications [12]. However, currently there is a lack of consensus on the best procedure to extract information from slow-response chemical sensors for the task of chemical source localization.

Instead of pursuing the ambitious goal of improving the dynamic response of chemical sensors, a workaround is to extract dynamic features which are informative for the task to be completed. These dynamic features may appear related to the intermittent contact to gas/odor patches. For example, Schmuker et al. [13] recently proposed that short time-scale features (the so-called "bouts") extracted from the first derivative of the response of metal oxide (MOX) gas sensors can be exploited to predict the distance to a gas source in a turbulent environment. Before going into algorithmic details, at this point is sufficient to understand that the bouts are simply the rising edges of the "clean" derivative of the sensor response, and they appear when the sensor gets in contact with a chemical patch (Fig. 1). Estimating the distance of a gas source is important in applications such as environmental monitoring [14] or gas source localization by mobile robots [13,15,16]. In the latter case, a source distance estimate can be used by a mobile agent to navigate towards a gas source [13] or to declare that the source has been found [15,16]. The advantage of the bout approach is that the derivative is insensitive to changes in background concentration, providing a source distance estimate that is independent from slowly time-varying concentration fields.

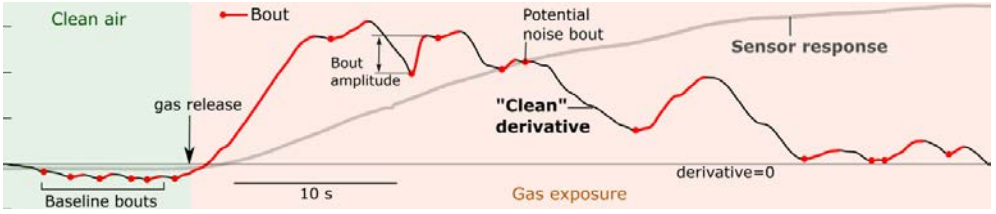


Figure 1. Schematic representation of the raw sensor response, its "clean" derivative and the detected "bouts" when the sensor is exposed to a chemical plume. The figure illustrates that the dynamics of the derivative are much faster than those of the raw response.

Differentiation is a risky signal processing operation that often leads to undesired amplification of high-frequency noise. To prevent signal to noise degradation, differentiation must be combined with low-pass filtering matched to the signal bandwidth. For that, Schmuker et al. employs a rather arbitrary cascading of low-pass digital filters with nonlinear phase, introducing unnecessary complexity into the algorithm and distorting the waveform of the output signal. Preserving the waveform is important because a thresholding decision must be made on the resulting waveform for bout detection. Therefore, preserving or recovering the "true" derivative of the sensor response is of utmost importance for the proper operation of the bout detector in terms of signal-detection theory. Bout detection errors will degrade the performance of posterior bout-based estimation algorithms.

In Schmuker's algorithm, the raw sensor response is first smoothed with a Gaussian filter, the resulting signal is differentiated, and the derivative is smoothed again using a first order infinite impulse response (IIR) filter [17]. Gaussian filters are ideal time response filters since they show the fastest step response under the condition of no overshooting [18]. They are mostly used in image processing, but seldomly in signal processing since they are non-causal and physically unrealizable. Beyond simple truncation, more sophisticated approaches to Gaussian filter approximation have been proposed for both IIR [19] and Finite Impulse Response (FIR) [20] filters. In any case, the intended advantages of the Gaussian filter are lost when in cascade with an IIR fitter. The disadvantage of IIR filters is the signal distortion due to nonlinear phase in the pass band. A further disadvantage of using two low-pass filters instead of one is that, for a desired filter response, two coupled parameters (the standard deviation σ of the Gaussian filter and the time constant τ of the IIR filter) with non-intuitive interpretation must be tuned.

Differentiation of noisy signals is a well-known signal processing problem [21]. Among other applications, differentiation is used in biomedical signal processing for the detection of QRS complexes in ECG signals [22]. The standard approach is to use a low pass differentiator (LPD), also known as linear-phase FIR differentiator, which can be designed from the selected specifications in the frequency domain using the Remez algorithm [23]. A linear-phase filter will delay all frequency components by the same amount, preserving the waveform of the input signal (assumed bandlimited and contained in the filter bandpass).

While computing the bouts, Schmuker et al. noticed that low-amplitude bouts produced by noise were as well detected in the baseline response of the sensor, i.e. in signals recorded in the absence of gas. In order to reduce false detections, Schmuker et al. filters out all bouts with amplitude below a threshold, defined as the mean plus three standard deviations of the amplitude of bouts detected in the sensor baseline. This statistical rule, known as the three-sigma rule, states that for a normal distribution 99.73% of the samples will fall within three standard deviations of the mean. Thus, it is empirically used by Schmuker et al. to treat 99.73% probability as near certainty that no bout in the sensor baseline will exceed the threshold. However, Burgués et al. [24] recently found empirically that the distribution of bout amplitudes in clean air does not follow a normal distribution. This was an expected result because the bout amplitude is bounded by zero (i.e. the bout amplitude is strictly positive). Two questions that immediately arise are: (1) is the three-sigma rule an optimal threshold? and (2) are these low-amplitude bouts a consequence of a non-optimal low-pass filtering? Instead of a heuristic selection of the bout amplitude threshold, we propose in this contribution a systematic optimization of this parameter aiming to extract the maximum information from the sensor signal.

In wind tunnel experiments, Schmuker et al. found that the number of above-threshold bouts detected by a MOX sensor placed in the centerline of a chemical plume monotonically increases with proximity to the source (distance range 25-145 cm). To quantify the source distance prediction error, the authors fitted a linear model relating the number of bouts detected in 3 minutes and the distance to the source and computed the average prediction error in cross-validation samples. Although the true relationship between bout count and source distance was non-linear, the linear model achieved an average error of 18 cm if the train and test wind speeds were the same. Based on these results, the authors claim that the number of bouts detected in a certain time window provides an accurate estimation of the distance to a gas source and propose a bout-based GSL strategy that does not require anemometry. However, they also found that the wind speed has a strong effect in the number of bouts detected by the sensor, which would notably increase the prediction error if the test and train wind speeds are different. To compensate that, they suggest incorporating wind speed measurements into the predictive models, which invalidates one of the main advantages of the algorithm (avoiding anemometry). In addition, the authors did not benchmark the results to traditional source distance estimators such as the mean [25], maximum [26] or variance [15,16,27] of the response, which may also perform well considering the long measurement windows.

In this paper, we propose a method to predict the distance to a gas source from the transient response of MOX sensor signals. As a case study, we use recordings from MOX sensors in a wind tunnel (same dataset used by Schmuker et al.). First, we optimize the algorithm parameters (bandpass frequency and the bout amplitude threshold) in a multivariate way, uncovering interesting relationships between these two parameters. We then find the optimum predictive model relating bout frequency and source distance and assess its performance in external validation samples, i.e. unseen during model fitting. The results are benchmarked against Schmuker's algorithm and traditional source distance estimators such as the mean, variance and maximum of the response. We demonstrate that certain combinations of parameters can largely improve the prediction performance reported by Schmuker et al. and make the predictive models more robust against changes in wind speed. We also study how the length of the measurement window affects the performance of different signal features. Finally, we provide a MATLAB implementation of the bout detection algorithm.

2. Signal Processing Methods

This section explains the exponentially weighted moving average (EWMA) filter, the ema_α operator and the bout computation algorithm, which is based on the ema_α operator.

2.1. Exponentially weighted moving average (EWMA)

The exponentially weighted moving average (EWMA) [28], also known as exponential moving average (EMA), is a first-order IIR filter commonly used to smooth temporal series. EWMA gives more importance to recent data by discounting older data in an exponential manner (Equation 1). At time $t = nT_s$, (being T_s the sampling interval) the smoothed value $y[n]$ is found by computing

$$y[n] = (1 - \alpha) \cdot y[n-1] + \alpha \cdot x[n] \tag{1}$$

where $x[n]$ is the observation at time $t = nT_S$, $y[n-1]$ is the previous output of the filter and the smoothing factor α ($0 < \alpha \leq 1$) controls the speed at which older responses are dampened. Values of α close to zero place most of the weight on past values of the signal, whereas α close to 1 give most importance to the current value, quickly forgetting old values. By expanding Equation 1, we can derive the output of the filter as a convolution of the input with the unit impulse response

$$y[n] = \sum_{k=0}^{n} h[k] \cdot x[n-k] = \alpha \sum_{k=0}^{n} (1-\alpha)^k \cdot x[n-k] \tag{2}$$

which illustrates the exponential behavior since the weights, $(1-\alpha)^k$ decrease geometrically. A meaningful way to specify α is by the half-life time (s), τ, of the exponential decay, which is the time at which the exponential weight $(1-\alpha)^k$ decays by one half

$$\alpha = 1 - \left(\frac{1}{2}\right)^{\frac{1}{\tau f_s}} \tag{3}$$

where f_s is the sampling frequency of x (Hz).

2.2. The ema_α operator

Muezzinoglu et al. [29] proposed a rough approximation for a linear inverse filter of the MOX sensor response to improve sensor dynamics, called the ema_α operator

$$y[n] = (1 - \alpha) \cdot y[n-1] + \alpha \cdot (x[n] - x[n-1]) \tag{4}$$

where the variables have the same meaning than in Equation 1. The ema_α operator is a connection in cascade of two linear time-invariant systems: first, taking the derivative of the input signal, i.e. $(x[n] - x[n-1])$; second, smoothing the derivative using an EWMA filter (Equation 1). The derivative of the response is the simplest inverse filter of a MOX sensor, which due to the slow chemical reactions happening on the sensor surface acts as an integrator of the input stimuli [13]. Smoothing is necessary because differentiation degrades the signal to noise ratio (SNR). As α approaches 1, the closer is the filtered signal to the derivative and the faster is the response. However, being close to the derivative means also higher noise. Thus, α governs a trade-off between response time and SNR. By embedding the computation of the derivative into the EWMA filter and naming Equation 4 the ema_α operator, Muezzinoglu et al. introduced a slight confusion in the terminology because Equation 1 is also known in the literature as the EMA filter.

2.3. Schmuker's (SMK) filter

To compute the smoothed derivative of a signal, Schmuker et al. [13] applies the ema_α operator to the smoothed sensor response x_s resulting from applying a Gaussian filter to the raw sensor response x. The procedure is illustrated in Figure 2. The choice of the Gaussian filter is rather arbitrary and introduces undesirable effects such as non-causality and introduces an extra parameter (the standard deviation σ) that needs to be optimized. Another problem of the Gaussian filter used by

Schmuker et al. is that different values of σ change the gain of the filter (see Fig. 11D in [13]) leading to potential misinterpretations of the effect of σ on the filtered signal. The description of Schmuker's filter in the original paper (i.e., Eqs. 1-4 in [13]) presents additional issues that are described in detail in the Appendix.

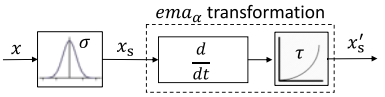


Figure 2. Flow diagram of the bout computation. The meaning of each symbol is given in the text.

2.4. Low pass differentiator (LPD) filter

As an alternative to Schmuker's filter, we propose a low pass differentiator (LPD) filter, also known as linear-phase FIR differentiator. An LPD filter differentiates the input signal at low frequencies and, at the same time, removes high frequency components from the signal. Operating with such a differentiator results in a constant non-zero group delay, which means that all frequency components are delayed by the same amount, preventing signal distortion. The design specifications of an LPD filter include the pass-band frequency (f_p), stop-band frequency (f_s), amplitude in pass-band (A_p), amplitude in stop-band (A_s) and the maximum allowable deviation or ripples between the frequency response and the desired amplitude of the output filter for each band. Given some filter specifications, the Parks-McClellan algorithm [23,30] provide filters with an optimal fit between the desired and actual frequency responses, and minimum order. For FIR differentiators, which have an amplitude characteristic proportional to frequency, the algorithm uses a special weighting technique so that the error at low frequencies is much smaller than at high frequencies.

2.5. Detection of 'bouts'

The bouts are the positive segments of the low-pass derivative, x'_s, and are found by looking for two consecutive zero-crossings of the positive derivative of this signal, i.e. $x''_s > 0$. The amplitude of a bout is defined as x'_s at the end of the respective bout segment minus x'_s at the start of the same bout segment. Bouts with very low amplitude, due to sensor noise, are usually interleaved with larger bouts produced by the plume. It is of practical interest to filter out bouts produced by noise, as they deteriorate the relationship between bout count and source distance. For that, Schmuker et al. use a fixed amplitude threshold

$$b_{thr} = \mu + 3\sigma \tag{5}$$

where μ and σ are the mean and standard deviation, respectively, of the distribution of amplitudes of the bouts detected in the sensor baseline (i.e. in clean air), since these are surely produced by noise if the underlying distribution follows a normal pattern.

3. Experimental

The goal of the experiments is to optimize the proposed algorithm and evaluate its performance for predicting the source distance in real data. For that, we use the wind tunnel dataset used by Schmuker et al. In this section, we provide an overview of the wind tunnel dataset and explain how we optimized the algorithm and analyzed the data.

3.1. Wind tunnel dataset

The wind tunnel dataset [31] contains recordings from nine gas sensor arrays inside a small wind tunnel where turbulent gas plumes of single compounds from ten possible chemicals (e.g., Acetaldehyde, Toluene, Carbon Monoxide) were created by injecting pressurized gas into one end of the tunnel and dragging it with an exhaust fan from the other end of the tunnel (Fig. 3). The gas source is a flexible nozzle that releases the selected gas at a constant flow rate of 320 sccm. The nozzle is connected to the outlet of three mass flow controllers (MFCs) that produce a certain concentration

from pressurized gas stored in cylinders. The nine sensing boards were always positioned along a line perpendicular to the wind direction, acquiring simultaneous measurements cross-wind. Depending on the experiment, the whole line of sensors was placed at a different distance to the gas source (range 0.25 - 1.45 m). In this way, the concentration within the wind tunnel was measured in 54 locations (9 x 6 grid). Each sensor array integrates eight MOX sensors (Several TGS 26XX models, Figaro Engineering Inc.) operated at the same, constant heater voltage (range 4.0-6.0 V). The sensor response was measured with a voltage divider (10 kΩ load resistor) and sampled with a 12-bit ADC. A total of 900 distinct experiments were performed by varying the distance to source (6 possible), gas (10 possible), sensor temperature (5 possible) and wind speed (3 possible). Each experiment was repeated 20 times. Different wind speeds (range 10-34 cm/s) were created by varying the rotational speed of the exhaust fan.

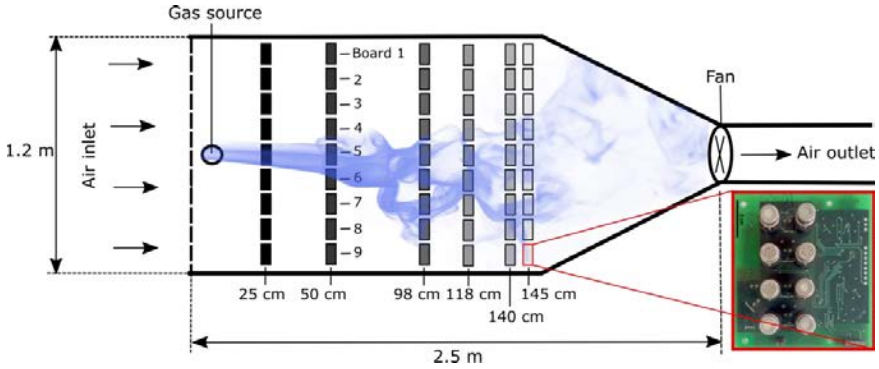


Figure 3. Schematic representation of the wind tunnel. The six measuring distances are coded in a gray scale and labelled P1 to P6. The inset on the bottom right corner shows a board with 8 MOX sensors (Several TGS models: 2600 (2x), 2602 (1x), 2610 (1x), 2611 (1x), 2612 (1x), 2620 (2x), Figaro Engineering Inc.). Adapted from [13].

In each experiment the following sequential procedure was performed: (i) Measure the baseline response of the sensors for 20 s in the absence of gas, (ii) Release the selected gas for 3 minutes, (iii) Circulate clean air for 1 minute to record the sensor recovery and (iv) Purge the wind tunnel by setting the fan at maximum speed (signals not recorded). Despite gas is released at $t = 20$ s, it does not arrive immediately to the sensors. The recorded signals show a transient behavior from $t = 20$ s to $t = [80, 110]$ s corresponding to the propagation and stabilization of the gas within the wind tunnel. The duration of the transient depends on the distance to the source (i.e., sensors closer to the source stabilize faster) and the wind speed (i.e., signals stabilize faster at high wind speeds). A common time frame where all signals are stable regardless of the distance and wind speed is from $t = 110$ s to $t = 200$ s. We denote this time frame (of length 90 s) as the "stable gas release" period.

3.2. Data analysis

We use the following data from the wind tunnel dataset: board #5 (located in the plume centerline), sensor #4 (TGS 2600), heater voltage of 6 V and Acetaldehyde gas. For this configuration, sensor recordings are available at 3 wind speeds, 6 distances to the source and 20 trials per distance/wind combination (360 experiments). The recordings captured at $d = 1.40$ m present some artifacts due to an imperfection of the wind tunnel (this was already pointed out by Vergara et al. and Schmuker et al.) so we excluded them from the analysis. Thus, the dataset that we use contains 300 experiments (3 wind speed x 5 distances x 20 trials). From each recording, we extract the region corresponding to the stable gas release period and compute four descriptors of the signals: the mean, variance, maximum response and bout frequency.

Regarding the computation of the bout frequency, we compare different filters and bout amplitude thresholds. Specifically, we use Schmuker's filter (Section 2.3) with default smoothing parameters (i.e., $\sigma = 0.3\,s$ and $\tau = 0.4\,s$) and 20 LPD filters (Section 2.4) with different pass-band

frequencies (f_{pass}) spanning the range [0.1, 2.0] Hz in steps of 0.1 Hz. We set $f_{\text{stop}} = f_{\text{pass}} + 0.1$ Hz, $A_{\text{pass}} = 1$, $A_{\text{stop}} = 0$ and a maximum deviation in both bands of -60 dB. For each of these filters, we extract the bouts according to the procedure described in Section 2.5 and, instead of applying the $\mu + 3\sigma$ threshold (Eq. 5), we compute the bout frequency using 500 thresholds (b_{thr}) spanning the range $[10^{-4}, 10^{1}]$ MS/s (S stands for Siemens) in logarithmic steps. The bout frequency (bouts/min) is the number of above-threshold bouts divided by the length (in minutes) of the measurement window. The bout frequency associated to each combination of f_{pass} and b_{thr} is considered a different signal feature. Thus, we have 10,500 bout-based features (500 thresholds x [20 LPD filters + 1 SMK filter]).

Using the bout-based features and the three statistical descriptors (mean, variance and maximum) we build a feature matrix $\mathbf{X}$ containing 300 rows (1 per experiment) and 10,503 columns (1 per feature). The matrix $\mathbf{X}$ is then split into 3 subsets $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, each one corresponding to a different wind speed. Each $\mathbf{X}_i$ is composed of 100 rows (5 distances x 20 trials) and 10,503 columns.

3.3. Fitting, selection and validation of predictive models

For each signal feature y (dependent variable) in $\mathbf{X}_i$, we want to find the model f that best fits the relationship $x = f(y)$, where x *is* the distance to the source (independent variable). Due to the limited number of levels (5) of the independent variable, we are forced to use simple empirical models with few degrees of freedom, such as linear, quadratic and cubic polynomials, one-term exponential and one-term power series. To find which of these models achieves the best prediction accuracy for each signal feature, we performed a combination of hold-out (external validation) and cross-validation (CV) (internal validation) procedure (Fig. 4).

The hold-out procedure is initially applied to split the dataset $\mathbf{X}_i$ into a 'train' and 'test' subsets using the first 14 trials (70% of the data) for training and the remaining 6 trials (30% of the data) for testing. The training set is used to find a suitable model for the data, and the test set is used to assess how well the selected model performs on unseen data. To find a suitable model for the data, we apply 5-fold CV in the 'train' subset which consists on breaking the data up into 5 partitions and then, 5 times in turn, using one partition for testing and the remaining ones for fitting each of the models. Then, we compute the RMSE (Eq. 6) of each model in each test partition and select the model that minimizes the average RMSE over the 5 folds. This model is refit using all training samples and validated against the hold-out test set by computing the RMSE in prediction (RMSEP).

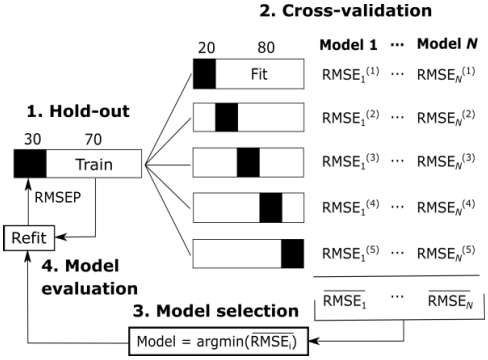


Figure 4. Model selection and validation by a combination of hold-out and 5-fold cross-validation (CV). The dataset is split into 'test' (30%) and 'train' (70%) partitions. The 'train' partition is used to evaluate the performance of different models by 5-fold CV. In each fold, different models are fit using 80% of the data and validated against the remaining 20% of the data. The model with minimum average RMSE across the 5 folds is refit using all training samples and validated against the hold-out subset.

To assess the performance of the predictive models under non-matching train and test wind speeds, the RMSEP is computed on hold-out partitions (i.e., last 6 trials) from datasets D_j with $i \neq j$.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (6)$$

where $\hat{y}_i$ is the distance (m) predicted by the model, y_i is the true distance (m) at which the measurements were taken and n is the number of validation samples.

3.4. Distance prediction using shorter measurement windows

We also studied the impact of using shorter measurement windows of 60, 30 and 10 s. Since the window is shorter than the length of the available signal (90 s) we need to select a chunk of the signal of the desired length. To minimize sampling bias, we randomly sample (with replacement) N segments of the desired length from the stable gas release period of the 300 available signals. We used values of N of 5, 10 and 20 for the window sizes of 60, 30 and 10 s, respectively. Then, for each wind speed, signal feature, window size and segment, we follow the procedure described in Section 3.3 to fit, select and validate a predictive model. The only difference is that now we have N RMSEP values associated to a given feature (one per segment), so we take the average RMSEP across all segments.

3.5. Spectral analysis

The power spectral density (S_{yy}) of the signal x will be approximated by the Periodogram [32].

3.6. Software and reproducibility

Data analysis was performed using MATLAB (version R2018b) and the following toolboxes: signal processing, and statistics and machine learning toolbox. Spectral analysis was performed using the function *fft*. LPD filters were designed using *firpmord* and *firpm* functions with '*differentiator*' argument. Schmuker filter was implemented in the time domain using the convolution of two impulse responses: a truncated Gaussian filter *G*, implemented using the function *gausswin(*N, α*)* with $N = 10\sigma_s$ and $\alpha = 5$, and a first order FIR differentiator *D* of the form [1, -1]. The output of this filter was the input of the EWMA filter *E*. Convolution was performed using function *conv*. The resulting filters were applied to new data using function *filter*. Curve fitting was performed using *polyfit* and evaluated on new data using *polyeval*.

All analysis code used in this study is freely available under an open source license at https://github.com/jburgues/DistPredMOX. The code allows the reader to reproduce the entire analysis, including recreation of most figures.

4. Results and Discussion

4.1. Raw signals

We analyzed the signals recorded by a MOX sensor placed in the plume centerline at several downwind distances from the gas source and three wind speeds (Fig. 5). Regardless of the wind speed and distance, the signals are characterized by a *plateau* of stable mean value (produced by the accumulation of gas within the wind tunnel) modulated by some fluctuations due to turbulence. The distance of the gas source and the wind speed have a strong influence on the mean intensity and the fluctuations of the signals: with increasing distance or wind speed, the intensity and variability of the signals decrease. At higher wind speeds, the chemical source experiences a larger dilution, resulting in lower average concentrations. At lower wind speeds, the mean, maximum and variance of the responses seem to be correlated with the distance to the gas source. As the wind speed increases, the differences between sampling locations diminish and it seems harder to predict the source distance from these features.

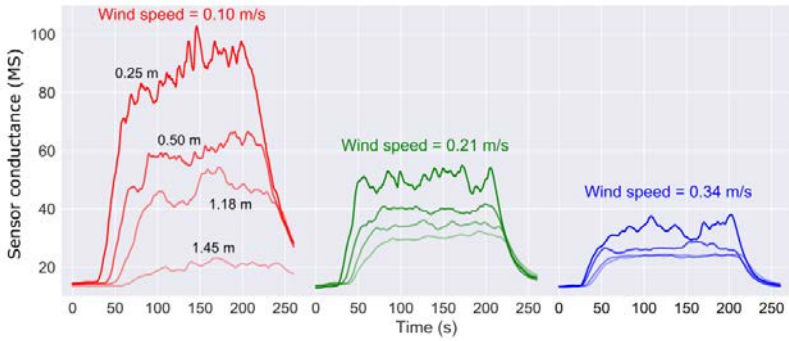


Figure 5. Responses of a TGS XX sensor (trial #1) at multiple downwind distances from the gas source and three wind speeds.

4.2. Distance prediction with mean, standard deviation and maximum of the response

Figure 6 shows the relationship between common statistical descriptors of the raw signal and the source distance at different wind speeds. The average response was best fit with logarithmic models at low wind speeds and with power models at high wind speed (Fig. 6a). The RMSEP degrades from 21 cm to 28 cm as wind speed increases, due to a decrease in sensitivity far from the source. The RMSEP is inflated by the abnormally low concentration measured at $d = 50$ cm, which does not agree with previously reported empirical models [33] describing the relationship between mean concentration and source distance, x, as a power law of the form $1/x$. A potential cause for the anomalous measurements at $d = 50$ cm is the high variance in wind direction measured by Vergara et al. at that sampling location (c.f. Fig. 3b in [31]), probably owing to the geometry of the wind tunnel.

The standard deviation of the response (Fig. 6b) is highly scattered in different trials, leading to a high RMSEP of around 26 cm regardless of the wind speed. The high scattering observed at the lowest wind speed may be a consequence of the low-energy turbulent mixing characteristic of environments with weak airflow [34]. The optimum model switches from linear to logarithmic to power with increasing wind speed.

Finally, the maximum of the response (Fig. 6c) has a similar behavior than the mean response but with lower RMSEP (20-25 cm), resulting the best estimator at all wind speeds.

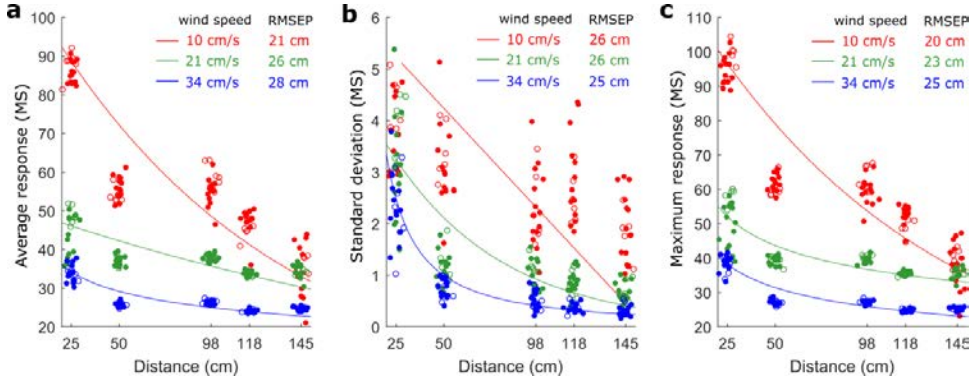


Figure 6. (**a**) mean, (**b**) variance and (**c**) maximum response vs. distance from source at different wind speeds in one of the folds of 5-fold cross-validation (CV). The fit and test samples for this fold are represented by solid circles and crosses, respectively. The solid lines represent the models with minimum $\overline{RMSE}$ over 5 folds.

4.3. Distance prediction with bouts

We now explore if the bout frequency can improve the previous results. The spectral analysis of the raw signals (Fig. 7a) indicate that the power content of the signal decays exponentially with

increasing frequency. The recordings at lower wind speeds or closer to the source contain more power in higher frequencies than the recordings at higher wind speed or longer distances. Thus, the range of frequencies in which the derivative can be applied with reasonable signal-to-noise ratio (SNR) depends on the wind speed and the distance to the source. From visual inspection, it seems that cut-off frequencies of 5 and 10 Hz could be appropriate for wind speeds of 34 and 10 cm/s, respectively. However, in practice these values are too high since computing the bouts requires taking the second derivative of the signals, which is extremely sensitive to noise. Thus, we limited the maximum pass-band frequency of our LPD filters to 2 Hz.

Figure 7b compares the frequency response of two Schmuker (SMK) filters and two low pass differentiator (LPD) filters with different specifications. SMK filters are characterized by a narrow pass-band region where the response of the filter approximates the ideal differentiator, and a wide transition band were neither the derivative is computed, nor the noise is fully rejected. Resultingly, the filtered signal will deviate from the ideal derivative in all frequency components corresponding to the transition band. Another disadvantage of SMK filter is that the design parameters σ and τ have no clear relationship to the standard filter design parameters such as the pass- and stop-band frequencies or the attenuation in the stop band. For example, the filter $\mathrm{SMK}_{0.3s,0.4s}$ (i.e., $\sigma = 0.3\ s$ and $\tau = 0.4\ s$) has pass- and stop-band frequencies of around 0.25 and 0.7 Hz, respectively, whereas in the filter $\mathrm{SMK}_{0.2s,0.2s}$ these frequencies are around 0.4 and 1.5 Hz. Considering that the power spectral density (PSD) of the recorded signals beyond 0.5 Hz is not negligible (especially at wind speed of 10 cm/s), the $\mathrm{SMK}_{0.3s,0.4s}$ filter is probably filtering out useful information contained in the signals.

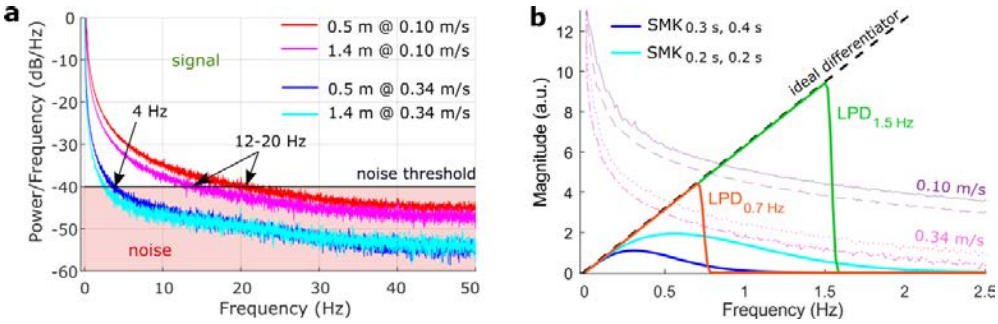


Figure 7. (**a**) Power spectral density (PSD) of the raw signals at different distances to the source and wind speeds. Each trace represents the average of 20 trials. The noise threshold is estimated from the PSD of the baseline responses; (**b**) Frequency response of two Schmuker filters ($\mathrm{SMK}_{\sigma,\tau}$) and two low pass differentiator filters (LPD_f) with different cut-off frequency. The signals of panel (**a**) are drawn in the background for visual reference. The ideal differentiator is depicted as a black dashed line.

In contrast, LPD filters maximize the region in which the derivative is computed and can provide narrow transition bands, resulting in signals with higher SNR. Comparing the $\mathrm{LPD}_{0.7\ \mathrm{Hz}}$ filter (i.e., $f_c = 0.7$ Hz) with the $\mathrm{SMK}_{0.3s,0.4s}$ filter, it can be seen that, with a similar stop-band frequency, the former one computes the derivative up to 0.7 Hz whereas the latter only up to 0.25 Hz. As an illustrative example, Figure 8 shows the output of both filters for a signal captured at $d = 0.98$ m and wind speed of 34 cm/s. Compared to the filtering produced by the SMK filter, the LPD produces a sharper signal with a similar level of noise and as twice more bouts (19 versus 38 bouts).

To evaluate if the higher number of bouts is advantageous for predicting the source distance, Figure 8 shows the regression between bout frequency and source distance for the two filters. When all detected bouts are used for regression (i.e., $b_{\mathrm{thr}} = 0$) the bout frequency decreases with proximity to the source, indicating that the noise (i.e. low-amplitude bouts) decreases with proximity to the source. In this case, both filters provide similar RMSEP values of around 20 cm, which is already 5 cm lower than the best error (25 cm) achieved for this wind speed with statistical descriptors of the signal. By increasing the bout amplitude threshold, b_{thr}, the slope of the regression changes from negative to positive, the scattering among trials decrease and the RMSEP decreases. For example,

using $b_{\mathrm{thr}} = \mu + 3\sigma = 0.06$ MS/s leads to a linear decay between bout frequency and source distance, yielding an RMSEP of 17 cm. Further increasing the threshold to 0.15 MS/s changes the functional relationship from linear to a power law and the RMSEP decreases to 14 cm. Since the threshold seems to have a clear impact into the prediction error, we will try to optimize it to achieve the minimum possible RMSEP.

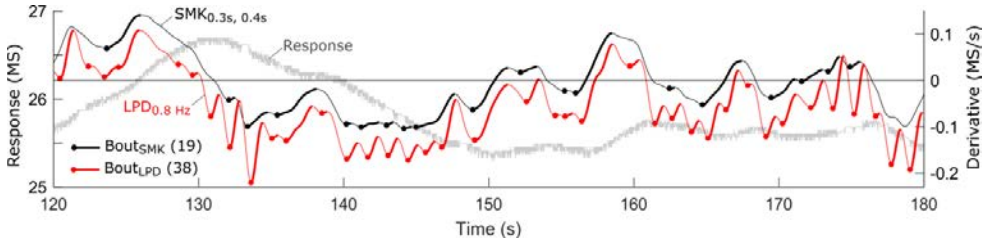


Figure 8. Filtered MOX-sensor signals at $d = 0.98$ m and wind speed of 34 cm/s using Schmuker (SMK) filter ($\sigma = 0.3\ s$ and $\tau = 0.4\ s$) and low pass differentiator (LPD) filter ($f_c = 0.8$ Hz). LPD signal has been shifted downwards by 0.01 MS/s for visualization purposes. "Bouts" in both signals are highlighted in black and red, respectively. The number in parenthesis indicates the number of detected bouts.

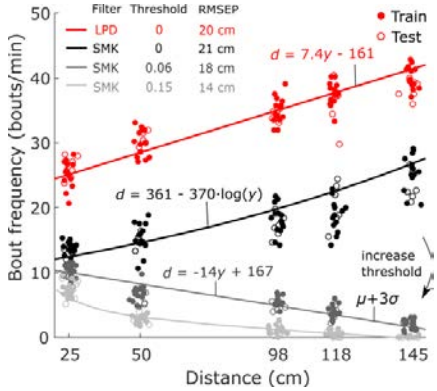


Figure 9. Bout frequency for 20 trials vs. distance from source using LPD filter ($f_c = 0.8$ Hz, $b_{thr} = 0$) and SMK filter ($\sigma = 0.3\ s$ and $\tau = 0.4\ s$) with different amplitude thresholds. The wind speed is 34 cm/s and window size is 90 s.

Figure 10 shows the RMSEP associated to different values of f_c and b_{thr}, for two different wind speeds. If b_{thr} is low or no threshold is used at all, f_c has a big impact on the RMSEP, varying in the range 16-40 cm. Filters with high f_c (i.e., 1.6-2.0 Hz) are optimal at low wind speed (Fig. 10a), whereas filters with f_c of 0.7-1.0 Hz are optimal at higher wind speeds (Fig. 10b), which matches our visual analysis of Figure 7b. By increasing b_{thr}, the RMSEP of certain filters can be reduced below 15 cm. For example, at low wind speed the threshold region 0.3-1.0 MS/s is optimum for all filters, and those with $f_c > 0.2$ Hz achieve RMSEP < 15 cm in this region. The absolute minimum RMSEP (7 cm) is obtained with $f_c = 1.9$ Hz and $b_{\mathrm{thr}} = 0.53$ MS/s. In contrast, the RMSEP of the SMK filter is always higher than 15 cm regardless of the value of b_{thr}. The $\mu + 3\sigma$ threshold (0.06 MS/s), located outside of the optimum threshold band, achieves an RMSEP of 23 cm.

As f_c increases (this is, the signal is less low-pass filtered) the RMSEP in the optimum region becomes more sensitive to b_{thr}. Looking at the reddish traces in Fig. 10a, we can observe that a small deviation from the optimum threshold produces a fast increase in RMSEP. The sensitivity is much lower at low f_c (bluish traces) where RMSEPs below 15 cm are obtained in nearly one decade around the optimum threshold value.

As wind speed increases, the optimum threshold band shifts towards smaller thresholds (Fig. 10b). At wind speed of 34 cm/s, the optimum threshold band 0.05-0.20 MS/s leads to RMSEP values below 15 cm for filters with $f_c < 0.7$ Hz. The minimum RMSEP (13 cm) is obtained with $f_c = 0.4$ Hz and $b_{thr} = 0.13$ MS/s. Despite the $\mu + 3\sigma$ threshold is located within the optimum threshold band, the SMK filter operated with this threshold provides a higher RMSEP of 18 cm. Low f_c values (bluish traces) which might appear as the worse choice if no threshold is used, provide the minimum RMSEP values if combined with a properly selected threshold. This illustrates that these two parameters are coupled and that simultaneous tuning of both of them is the means to achieve the best performance.

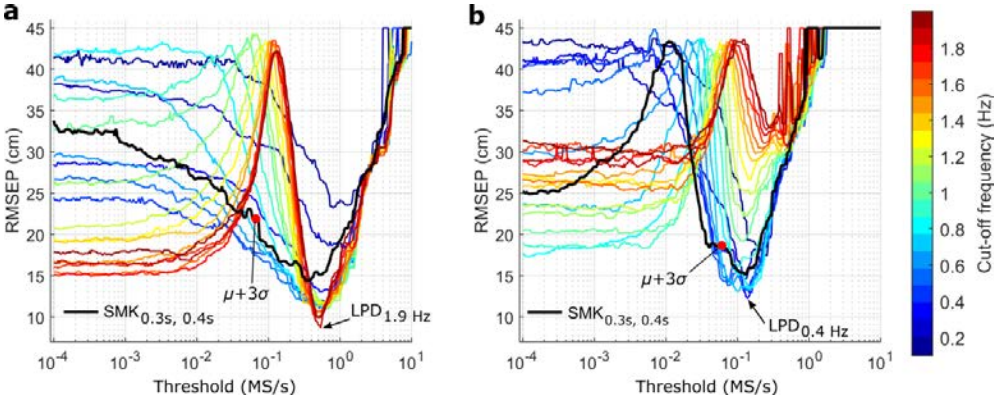


Figure 10. RMSEP of the bout frequency as a function of the bout amplitude threshold (log scale) and the cut-off frequency of the LPD filter, for two wind speeds: (**a**) 10 cm/s and (**b**) 34 cm/s. The RMSECV of Schmuker's filter (SMK_{σ_s,τ_h}) with default parameters is depicted as a black trace for visual reference and the $\mu + 3\sigma$ threshold is indicated by a red dot.

The observed shift in the optimum threshold band as the wind speed increases can be explained by looking at the histogram of bout amplitudes at different distances to the source (Fig. 11). At low wind speed, the overlapping between the baseline distribution (noise) and the gas distribution (signal) is low, so a high threshold can separate the noise from the signal. At higher wind speed, the histograms of bout amplitudes during gas exposure shift to the left while the baseline histogram remains fixed. Therefore, the threshold must "follow" the histogram shift or nearly no bouts will exceed the threshold far from the source. The optimum threshold lies as far to the right as possible in the baseline distribution while still capturing enough density of bouts at distant gas exposure.

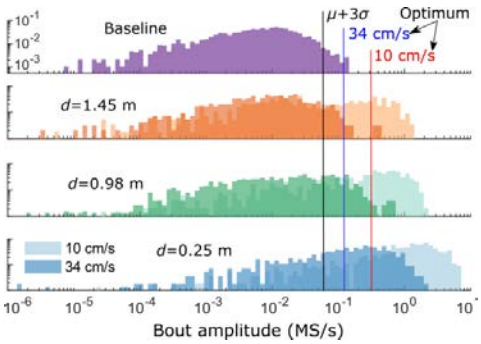


Figure 11. Histogram of bout amplitudes at various distances from the source and wind speeds. The width of the bins in the x-axis is distributed logarithmically. The y-axis represents the relative probability of each bin, computed as the number of elements in the bin divided by the number of samples. Both axes are displayed in logarithmic scale. The μ+3σ and optimum thresholds for each wind speed are depicted with vertical lines.

Since the baseline distribution does not resemble a normal pattern—instead it can be better approximated by a Beta function— the $\mu + 3\sigma$ threshold does not capture the expected 99.97% of the data, but a lower value (97.81%). In other words, approximately 2.2% (instead of 0.13%) of the bouts detected in clean air will be declared as "true bouts". This means that 2.2% of the detected bouts correspond to noise, increasing the scattering among different trials. This can be seen in Figure 8, by comparing the regression of the SMK filter with $b_{thr} = 0.06$ MS/s ($\mu + 3\sigma$) and $b_{thr} = 0.15$ MS/s (optimum for 34 cm/s). The higher threshold not only produces less scattering of the data but also produces a smoother increase in bout frequency as the distance to the source decreases, allowing a polynomial to accurately fit the data even at $d = 0.5$ m.

The optimum predictive model for each combination of f_c and b_{thr}, as determined in cross-validation, is shown in Figure 12. The optimum models can be roughly clustered in three main regions: power and logarithmic models in the upper left corner (high f_c and low b_{thr}), exponential models in the region $0.5\ \mathrm{MS/s} < b_{\mathrm{thr}} < 5\ \mathrm{MS/s}$, and linear models elsewhere. This explains why Schmuker et al. found that their filter, which has an approximate cut-off frequency of 0.6 Hz and a threshold of 0.06 MS/s, was best approximated by a linear model. However, if one considers the entire space of combinations, more complex models should be used for optimum performance. The best model for each f_c always lies in the leftmost edge of the exponential region, where quadratic models are chosen at low f_c and exponential models dominate at high f_c. A summary of the optimum filter parameters and predictive model for each wind speed is provided in Table 1.

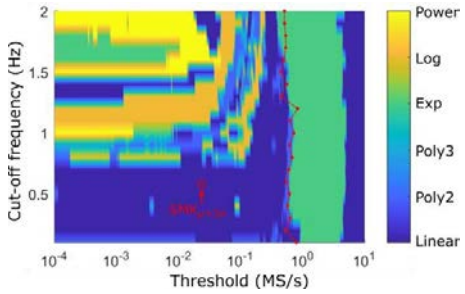


Figure 12. Optimum models, as found in cross-validation, as a function of the cut-off frequency (f_c) and the bout amplitude threshold (b_{thr}). The window size is 90 s and the wind speed 10 cm/s. The threshold that achieves the minimum RMSEP at each cut-off frequency is indicated by a red dot.

Wind (cm/s)	LPD [optimum]				SMK [optimum]			SMK [$\mu + 3\sigma$]		
	f_{cut} (Hz)	b_{thr} (MS/s)	Model	RMSEP (cm)	b_{thr} (MS/s)	Model	RMSEP (cm)	b_{thr} (MS/s)	Model	RMSEP (cm)
10	1.9	0.51	Exp	7	0.44	Linear	13	0.06	Log	22
21	1.0	0.21	Poly2	9	0.30	Exp	14	0.06	Log	22
34	0.7	0.11	Exp	10	0.15	Poly2	14	0.06	Linear	18

Table 1. Optimum filter parameters for different wind speeds. Wind speed in train is the same than in test.

4.4. Distance prediction using shorter measurement windows

Figure 13 shows the RMSEP curves of the $\mathrm{LPD}_{1.9\,\mathrm{Hz}}$ filter for different sizes of the measurement window (W). As it was expected, the RMSEP degrades with shorter measurement windows. Surprisingly, the degradation is consistent across all thresholds, resulting in an optimum threshold that is independent of W. Also, the RMSEP degrades non-linearly with W, since reducing W from 30 to 10 s (difference of 20 s) has a stronger impact into the RMSEP than reducing W from 90 to 30 s (difference of 60 s). Specifically, the relationship between the minimum RMSEP and W approximates a power law of the form $RMSEP = 80 \cdot W^{-0.26} - 17$. We observed similar trends at other cut-off frequencies and wind speeds, and for the SMK filter. This means that the optimum cut-off frequency is also independent of W. It mostly depends on the wind speed.

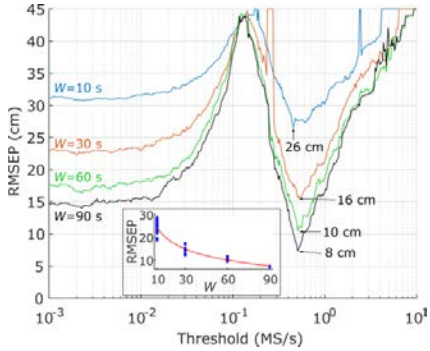


Figure 13. RMSEP of the bout frequency computed with the $\mathrm{LPD}_{1.9\,\mathrm{Hz}}$ filter, as a function of the bout amplitude threshold (log scale) and the window size (W). The inset shows a power series fit to the relationship between minimum RMSEP and W. The wind speed is 10 cm/s.

Figure 14 compares the minimum RMSEP of different signal features as a function of W. As can be seen, the mean and maximum response are nearly insensitive to W, probably due to the slow dynamics of the wind tunnel facility which produced a stable *plateau* of mean concentration (Fig. 5). The RMSEP of the standard deviation degrades at a linear rate of 1 cm for every 10 s of increase of W. Finally, the bout frequency is the most sensitive to W, exhibiting a power law dependency. Resultingly, at small window sizes the bout frequency provides worse performance than simpler statistical descriptors such as the mean or maximum response.

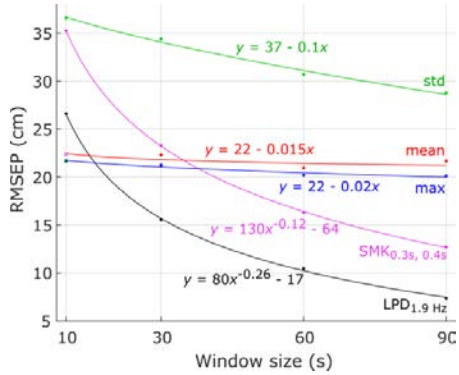


Figure 14. RMSEP of different signal features as a function of the window size. The wind speed is 10 cm/s.

4.5. Distance prediction under non-matching train and test wind speed

So far, we validated our models using training and test data captured at the same wind speed. If the train (w_{train}) and test wind speed (w_{test}) differ, the optimum predictive models when $w_{\mathrm{train}} = w_{\mathrm{test}}$ produce high RMSEP values when $w_{\mathrm{train}} \neq w_{\mathrm{test}}$. Figure 15 plots the RMSEP of the SMK filter as a function of the threshold and w_{test}, when $w_{\mathrm{train}} = 21$ cm/s. The optimum threshold for w_{train}, which produces an RMSEP of 15 cm when $w_{\mathrm{test}} = 0.21$ cm/s, leads to an average RMSEP across all possible wind speeds of 37 cm. Similarly, the $\mu + 3\sigma$ threshold produces errors as high as 62 cm if $w_{\mathrm{test}} = 34$ cm/s. There is an optimum threshold region (0.019 – 0.033 MS/s) in which the prediction error at different wind speeds is balanced, providing average RMSEPs of around 30 cm. In the LPD filter, the same cut-off frequency (0.6 Hz) is optimum for all w_{train}, and the optimum threshold region ranges from 0.035 to 0.050 MS/s. Table 2 summarizes the optimum filter values for each filter and w_{train}.

Wind (cm/s)	LPD [optimum]				SMK [optimum]			SMK [3sigma]		
	f_{cut} (Hz)	b_{thr} (MS/s)	Model	RMSEP (cm)	b_{thr} (MS/s)	Model	RMSEP (cm)	b_{thr} (MS/s)	Model	RMSEP (cm)
10	0.6	0.035	Linear	31	0.019	Log	33	0.06	Log	56
21	0.6	0.050	Linear	29	0.033	Linear	28	0.06	Exp	40
34	0.6	0.041	Linear	30	0.021	Linear	30	0.06	Linear	36

Table 2. Optimum filter parameters when the train and test wind speed are different. The train wind speed is indicated in the first column. The other columns are optimum values considering the average of all possible test wind speeds. The window size is 90 s.

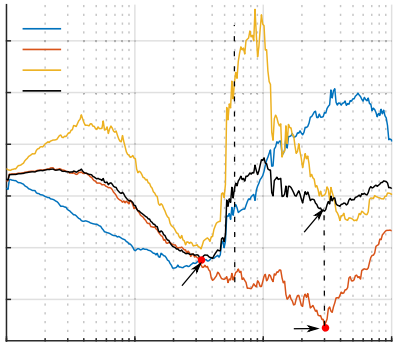


Figure 15. Sensitivity of the bout frequency to the test wind speed. The figure shows the RMSEP of the bout frequency versus threshold, using window size of 90s, training wind speed of 10 cm/s and different test wind speeds. LPD 2.0 filter (black) and SMK filter (red).

5. Conclusions

We have experimentally demonstrated that, in the wind tunnel facility and assuming the same constant emission rate during training and test, the proposed algorithm can exploit transient features of gas sensor signals to predict the gas source distance with high accuracy. This algorithm significantly reduces the prediction error, as compared to previously reported algorithms, over a wide range of amplitude thresholds. The main finding of the study is that the cut-off frequency (f_c) of the low-pass differentiator filter and the bout amplitude threshold (b_{thr}) are coupled parameters, which means that both must be tuned simultaneously to find the global optimum in terms of prediction error. The optimum values of f_c and b_{thr} strongly depend on the wind speed, with low f_c and b_{thr} being suitable for scenarios with high wind speed and vice-versa. Intuitively, this says that the cut-off frequency of the filter should match the bandwidth of the signal, which in this dataset is lower at high wind speeds. If the cut-off frequency is reduced, the derivative will be noisier and the detected bouts will be smaller in amplitude. This is probably the reason why both parameters are coupled and why smaller amplitude thresholds improve performance in this case.

The optimization methodology presented in this paper is intuitive and should be applicable to a wide range of different scenarios, including those that do not satisfy the assumptions of constant emission rate or non-matching train and test emission rates. For example, we used the methodology to find optimum values in the challenging scenario in which the train and test wind speed do not match. In the likely case that the test wind conditions are unknown, our results suggest that lower prediction errors will be obtained if f_c and b_{thr} are small and the models are trained at medium wind speed. We also optimized the algorithm for different sizes of the measuring window, and we found that the optimum values of both parameters do not change substantially but the minimum possible error increases as a power law of the window size. These results encourage us to think that the same

procedure can be applied, for example, to optimize the algorithm for the case of non-constant emission rate. Expressing the low-pass differentiator filter with a single parameter (f_c) was a key step for reducing the complexity of the optimization process and aiding in the visualization of the results.

A second question addressed by this work is whether the optimized bout frequency outperforms other estimators of source distance such as the mean, maximum or variance of the sensor response. We found that if the measurement window is larger than 30 s, the bout frequency always provides lower prediction errors. However, the mean and maximum of the signal are more accurate if the measurement window is as short as 10 s. This raises the question as to what extent the bout frequency provides any advantage over these simpler estimators in applications such as gas source localization by mobile robots, where rapid response time is critical. The variance of the sensor response—which some authors consider a reliable estimator of source proximity—did not work well in this dataset. The reason for the good performance of the bout frequency (assuming a long enough measurement window) is the consistency of its value among different trials, probably owing to the insensitivity of the derivative to changes in background concentration.

Author Contributions: conceptualization, J.B.; methodology, J.B.; software, J.B.; validation, J.B. and S.M.; formal analysis, J.B.; investigation, J.B.; resources, S.M.; data curation, J.B.; writing—original draft preparation, J.B.; writing—review and editing, J.B. and S.M.; visualization, J.B.; supervision, S.M.

Funding: This research was funded by Spanish MINECO, grant numbers BES-2015-071698 (Severo-Ochoa) and TEC2014-59229-R (SIGVOL). CERCA Programme / Generalitat de Catalunya

Acknowledgments: We would like to acknowledge, the Departament d''Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2017 SGR 1721); the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya; and the European Social Fund (ESF). Additional financial support has been provided by the Institut de Bioenginyeria de Catalunya (IBEC). IBEC is a member of the CERCA Programme/Generalitat de Catalunya.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

The description of Schmuker's filter in the original paper (i.e., Eqs. 1-4 in [13]) contain two typos that hinder the implementation of the algorithm by other authors. First, the smoothed response (x_s) is differentiated before applying the ema_α transformation (Eq. 2 in [13]) (this step has been omitted in the diagram shown in Fig. 2), which is unnecessary because the ema_α transformation already differentiates the input signal. This extra derivative was probably included by error, due to the ambiguous terminology used by Muezzinoglu et al. [29] to define the ema_α operator. The source code published by Schmuker et al. is however not affected by this error because they apply the ema_α operator by calling the *Python* function *ewma*, which provides the functionality of the EWMA filter only (i.e. it does not differentiate the input signal). Similarly, the definition of α as a function of τ_{half} in Schmuker's paper (Eq. 4 in [13]) contains a typo (compare it to the correct formulation, i.e. Eq. 3 of the current paper). Despite the source code published by the authors is not affected by this error (because such definition is implemented by the *Python* function *pandas.ewma)* this typo has difficulted replication of the algorithm by other authors [35].

References

1. Beychok, M.R. *Fundamentals of stack gas dispersion*; M.R. Beychok, 1994; ISBN 0964458802.
2. Yee, E.; Wang, B.-C.; Lien, F.-S. Probabilistic model for concentration fluctuations in compact-source plumes in an urban environment. *Boundary-layer Meteorol.* **2009**, *130*, 169–208.
3. Jones, A.R.; Thomson, D.J. Simulation of Time Series of Concentration Fluctuations in Atmospheric Dispersion Using a Correlation-distortion Technique. *Boundary-Layer Meteorol.* **2006**, *118*, 25–54.

4. Jacob, V.; Monsempès, C.; Rospars, J.-P.; Masson, J.-B.; Lucas, P. Olfactory coding in the turbulent realm. *PLOS Comput. Biol.* **2017**, *13*, e1005870.

5. Davide, F.A.M.; Natale, C. Di; D'Amico, A.; Hierlemann, A.; Mitrovics, J.; Schweizer, M.; Weimar, U.; Göpel, W.; Marco, S.; Pardo, A. Dynamic calibration of QMB polymer-coated sensors by Wiener kernel estimation. *Sensors Actuators B. Chem.* **1995**, *27*, 275–285.

6. Marco, S.; Pardo, A.; Davide, F.A.M.; Di Natale, C.; D'Amico, A.; Hierlemann, A.; Mitrovics, J.; Schweizer, M.; Weimar, U.; Göpel, W. Different strategies for the identification of gas sensing systems. *Sensors Actuators, B Chem.* **1996**, *34*, 213–223.

7. Yamanaka, T.; Ishida, H.; Nakamoto, T.; Moriizumi, T. Analysis of gas sensor transient response by visualizing instantaneous gas concentration using smoke. *Sensors Actuators, A Phys.* **1998**, *69*, 77–81.

8. Pardo, A.; Marco, S.; Samitier, J. Nonlinear inverse dynamic models of gas sensing systems based on chemical sensor arrays for quantitative measurements. *Ieee Trans. Instrum. Meas.* **1998**, *47*, 644–651.

9. Fonollosa, J.; Sheik, S.; Huerta, R.; Marco, S. Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring. *Sensors Actuators B Chem.* **2015**, *215*, 618–629.

10. Di Lello, E.; Trincavelli, M.; Bruyninckx, H.; De Laet, T. Augmented switching linear dynamical system model for gas concentration estimation with MOX sensors in an open sampling system. *Sensors (Basel).* **2014**, *14*, 12533–12559.

11. Barber, D.; Ch, D.B. *Expectation Correction for Smoothed Inference in Switching Linear Dynamical Systems 1. Switching Linear Dynamical System*; 2006; Vol. 7;.

12. Monroy, J.G.; Gonźalez-Jimenez, J.; Blanco, J.L. Overcoming the slow recovery of MOX gas sensors through a system modeling approach. *Sensors (Switzerland)* **2012**, *12*, 13664–13680.

13. Schmuker, M.; Bahr, V.; Huerta, R. Exploiting plume structure to decode gas source distance using metal-oxide gas sensors. *Sensors Actuators, B Chem.* **2016**, *235*, 636–646.

14. Liu, Y.; Yu, Q.; Huang, Z.; Ma, W.; Zhang, Y. Identifying Key Potential Source Areas for Ambient Methyl Mercaptan Pollution Based on Long-Term Environmental Monitoring Data in an Industrial Park. *Atmosphere (Basel).* **2018**, *9*, 501.

15. Lilienthal, A.J.; Duckett, T.; Ishida, H.; Werner, F. Indicators of gas source proximity using metal oxide sensors in a turbulent environment. In Proceedings of the Proceedings of the First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006, BioRob 2006; 2006; Vol. 2006, pp. 733–738.

16. Lilienthal, A.; Ulmer, H. Learning to detect proximity to a gas source with a mobile robot. *Intell. Robot. …* **2004**, 1444–1449.

17. Parks, T.W.; Burrus, C.S. *Digital Filter Design. Topics in Digital Signal Processing*; John Wiley & Sons, New York, 1987;

18. Blinchikoff, H.; Zverev, A. *Filtering in the Time and Frequency Domains*; Scitech Publishing Inc.:

Raleigh, 2001; ISBN 1884932177.

19. Young, I.T.; Van Vliet, L.J. Recursive implementation of the Gaussian filter. *Signal Processing* **1995**, *44*, 139–151.

20. Rau, R.; McClellan, J.H. Efficient approximation of Gaussian filters. *IEEE Trans. Signal Process.* **1997**, *45*, 468–471.

21. Dutta Roy, S.C.; Kumar, B. 6 Digital differentiators. *Handb. Stat.* **1993**, *10*, 159–205.

22. Kazanavicius, E.; Gircys, R.; Vrubliauskas, A.; Lugin, S. Mathematical methods for determining the foot point of the arterial pulse wave and evaluation of proposed methods. *Inf. Technol. Control* **2005**, *34*.

23. McClellan, J.; Parks, T.W.; Rabiner, L. A computer program for designing optimum FIR linear phase digital filters. *IEEE Trans. Audio Electroacoust.* **1973**, *21*, 506–526.

24. Burgués, J.; Hernández, V.; Lilienthal, A.J.; Marco, S. Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping. *Sensors* **2019**, *19*, 478.

25. Lilienthal, A.; Duckett, T. Building gas concentration gridmaps with a mobile robot. *Rob. Auton. Syst.* **2004**, *48*, 3–16.

26. Purnamadjaja, A.H.; Russell, R.A. Pheromone communication in a robot swarm: Necrophoric bee behaviour and its replication. *Robotica* **2005**, *23*, 731–742.

27. Lilienthal, A.J.; Reggente, M.; Trinca, M.; Blanco, J.L.; Gonzalez, J. A statistical approach to gas distribution modelling with mobile robots - The Kernel DM+V algorithm. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), St. Louis, MO, USA, 10–15 October; 2009; pp. 570–576.

28. Roberts, S.W. Control Chart Tests Based on Geometric Moving Averages. *Technometrics* **1959**, *1*, 239–250.

29. Muezzinoglu, M.K.; Vergara, A.; Huerta, R.; Rulkov, N.; Rabinovich, M.I.; Selverston, A.; Abarbanel, H.D.I. Acceleration of chemo-sensory information processing using transient features. *Sensors Actuators, B Chem.* **2009**, *137*, 507–512.

30. Parks, T.; McClellan, J. Chebyshev approximation for nonrecursive digital filters with linear phase. *IEEE Trans. Circuit Theory* **1972**, *19*, 189–194.

31. Vergara, A.; Fonollosa, J.; Mahiques, J.; Trincavelli, M.; Rulkov, N.; Huerta, R. On the performance of gas sensor arrays in open sampling systems using Inhibitory Support Vector Machines. *Sensors Actuators B Chem.* **2013**, *185*, 462–477.

32. Bartlett, M.S. Periodogram Analysis and Continuous Spectra. *Biometrika* **1950**, *37*, 1.

33. Webster, D.R.; Rahman, S.; Dasi, L.P. Laser-induced fluorescence measurements of a turbulent plume. *J. Eng. Mech.* **2003**, *129*, 1130–1137.

34. Ferri, G.; Caselli, E.; Mattoli, V.; Mondini, A.; Mazzolai, B.; Dario, P. SPIRAL: A novel biologically-inspired algorithm for gas/odor source localization in an indoor environment with no strong airflow. *Rob. Auton. Syst.* **2009**, *57*, 393–402.

35. Vuka, M.; Schaffernicht, E.; Schmuker, M.; Bennetts, V.H.; Amigoni, F.; Lilienthal, A.J. Exploration and localization of a gas source with MOX gas sensors on a mobile robot-A Gaussian regression bout amplitude approach. *ISOEN 2017 - ISOCS/IEEE Int. Symp. Olfaction Electron. Nose, Proc.* **2017**, 3–5.

Chapter 7

Gas Distribution Mapping and Gas Source Localization using MOX sensors

In this chapter, we address gas distribution mapping (GDM) and gas source localization (GSL) tasks in indoor environments using MOX sensors.

In Paper VII, we use a 3D grid of MOX sensors to visualize the temporal evolution of the gas distribution produced by an evaporating gas source placed at different locations in an office room (36 m^2) including variations in height, release rate and air flow profiles. We also compared different features of the MOX sensor signals (mean, variance and bout frequency) for predicting the true source location (TSL) in each experiment, considering different lengths of the measurement window. The analysis of the data revealed strongly time-varying and counter-intuitive gas distribution patterns that disprove some assumptions commonly held in the MRO field. For example, gas dispersion in the absence of strong airflow was dominated by convective or recirculating air currents from the HVAC unit, as demonstrated by the progressive gas accumulation towards the ceiling observed in most experiments. Resultingly, ground-level measurements were not useful for pinpointing the TSL and measurements at the height of the gas source were required. The variance and the bout frequency yielded in average lower localization errors than the mean (0.5 m and 1.1 m lower, respectively). The bout frequency performed better than the variance in experiments with strong air flow conditions (localization errors 0.9 m lower) but was much more sensitive to the window size.

In Paper VIII, a nano-drone equipped with a MOX sensor is used for GSL and GDM in a large robotics laboratory (160 m^2) containing obstacles such as tables, bookshelves and partitions. The gas source was placed in challenging positions for the drone, for example hidden in the ceiling of the room or inside a power outlet box. The drone navigates the environment following a predefined 3D sweeping path, which takes less than 3 minutes in total. The high complexity of the proposed scenario (size, obstacles and source location), the 3D navigation strategy and the reduced mission time are key differences with respect to

previous works using nano-drones [72,75], where the experimental scenario was as simple as a chemical hood [75] or the drone only took measurements at ground level [72]. In contrast to previously reported gas distribution mapping experiments in which long measurement times are needed at predefined [28] or adaptively chosen sampling locations [165], we demonstrate that a rough approximation of a gas distribution map can be quickly obtained with concentration measurements acquired in motion. The "bout" approach yielded on average a higher localization accuracy (1.38 m error) than using the instantaneous gas sensor response (2.05 m error).

In both papers, we dedicated a substantial portion of the effort to the processing of the MOX sensor signals, including calibration, linearization, LOD estimation and "bout" detection. Regarding calibration, we specifically focused on matching the calibration conditions (temperature, humidity and concentration range) to those of the test scenario to ensure the validity of the calibration models and the derived figures of merit (e.g. the LOD). For example, we assessed the sensor noise—critical for LOD estimation—through preliminary blank measurements acquired by the drone in the target scenario, instead of relying on the variability of the blanks in the calibration testbench. We also optimized the calibration concentration range and applied a log-log transformation to fit a linear model to the calibration data, thus enabling the computation of the LOD using the standard equations. Regarding the detection of "bouts", we simplified the original algorithm [121], adapted it for real-time operation and optimized the noise threshold. We found that the 3-sigma threshold recommended by Schmuker et al. [121] was too low in most scenarios, yielding an average localization error (4.23 m) three times higher than the one obtained when the threshold was optimized for each scenario (1.38 m).

A detailed description of both works is provided in Papers VII and VIII (shown below).

Gas distribution mapping and source localization using a 3D grid of metal oxide semiconductor sensors

Javier Burgués [1, 2, *], Victor Hernández [3], Achim J. Lilienthal [3] and Santiago Marco [1,2]

[1] Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain
[2] Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028 Barcelona, Spain
[3] AASS Mobile Robot Olfaction Lab, Örebro University, Örebro, Sweden
* Correspondence: jburgues@ibecbarcelona.eu; Tel.: +34-934-029-070

Abstract: The difficulty to obtain ground truth (i.e. empirical evidence) about how a gas disperses in an environment is one of the major hurdles in the field of mobile robotic olfaction (MRO), impairing our ability to develop efficient gas source localization strategies and to validate gas distribution maps produced by autonomous mobile robots. Previous ground truth measurements of gas dispersion have been mostly based on expensive tracer optical methods or 2D chemical sensor grids deployed only at ground level. With the ever-increasing trend towards gas-sensitive aerial robots, 3D measurements become necessary to characterize the environment prior to the development of source localization and mapping algorithms for these platforms. This paper presents ten different experiments performed with a 3D grid of 27 metal oxide semiconductor (MOX) sensors to visualize the temporal evolution of 3D gas distribution produced by an evaporating ethanol source placed at different locations in an office room, including variations in height, release rate and air flow profiles. We also studied which features of the MOX sensor signals are optimal for predicting the source location in each case, considering different lengths of the measurement window. We found strongly time-varying and counter-intuitive gas distribution patterns that disprove some assumptions commonly held in the MRO field. In most cases, measurements at the height of the gas source are key to accurately pinpoint its location. We make the dataset and the code publicly available to enable the community to develop, validate, and compare new approaches related to gas sensing in complex environments.

Keywords: Mobile robots; Artificial olfaction; Gas sensors; Signal Processing; Sensor networks; Gas dispersion; Gas source localization; Gas distribution mapping; Metal oxide semiconductor sensors

1. Introduction

The difficulty to obtain ground truth (i.e. empirical evidence) about how a gas disperses in an environment is one of the major hurdles in the field of mobile robotic olfaction (MRO), impairing our ability to develop efficient gas source localization (GSL) strategies and to validate gas distribution maps produced by autonomous mobile robots. For example, terrestrial robots equipped with in-situ gas sensors have been extensively studied for GSL under the assumption that gases heavier than air, such as ethanol, will flow along ground level [1–3]. However, simply using the density of the pure substance as the only parameter governing gas dispersion is dangerous, as it is known that atmospheric conditions (outdoors) or thermal sources or recirculating air currents from forced ventilation systems (indoors) may have a stronger impact into gas dispersion than the density of the pure gas [4]. Similarly, researchers conducting GSL and gas distribution mapping (GDM) experiments usually wait a small amount of time (few minutes) after the gas is released to let the gas spread and stabilize in the environment. If the robot starts exploring the environment when the gas distribution is not stable, the experimental results can depend on the time frame in which the experiments were performed [5]. For example, Purnamadjaja and Russell [6] waited 3 minutes before mapping the gas distribution produced by an evaporating alcohol mixture in a small unventilated

indoor arena (9 m^2). However, Lilienthal et al. [7] found that the gas distribution produced by a similar gas source in a similar environment became stable in approximately 30 minutes.

Any physical process is better understood if the changes of the variables involved in such process can be observed by visual inspection. The most straightforward method to make the gas flow pattern visible is to add small tracer particles (e.g., dust, smoke, dye) to the carrier flow and video-record the resulting clouds or plumes of particles [15]. Quantification of individual particles can be achieved by particle image velocimetry (PIV) [16], particle tracking velocimetry (PTV) [17] or laser-induced fluorescence (LIF) [12,18,19]. Tracer methods can provide a 3D representation of gas dispersion, but the required flow seeding methods and optical instruments for quantification are expensive and require a complex setup process that hinders its practical application [24]. Further, tracer methods assume that the tracer particles will have a similar motion to the target chemical when it is suspended on the fluid, which is not necessarily true.

Two alternative experimental methods are available for visualizing the gas flow in the experimental scenario without adding tracers: optical remote sensing (ORS) or chemical sensor grids (CSGs). ORS coupled with computed tomography (CT) is an old technique to visualize 2D gas distributions over room-sized areas [8–11]. ORS instruments such as the open-path Fourier transform infrared spectrometer (OP-FTIR) [12] or the tunable diode laser absorption spectroscoper (TDLAS) [13] are first used to accurately measure the total number of gas molecules in multiple beams of radiation. The path-integrated measurements are then converted by means of CT into a representation of the concentration distribution in the sampling plane [14]. Apart from the high cost of ORS instruments, mirrors and retroreflectors should be carefully placed over the test site to avoid unexpected optical effects that would otherwise lead to measurement errors [8].

An inexpensive, alternative solution to ORS is to use point chemical sensors, e.g. metal oxide semiconductor (MOX) sensors, to perform spatially distributed gas concentration measurements [15–21]. Indeed, measurements obtained with chemical sensors are often used to validate gas concentration maps obtained by ORS [11]. The reduced area covered by each measurement (≈0.5 cm^2) and the time-varying nature of gas dispersion in most natural environments requires deploying a grid format of sensors (i.e., a CSG). To build a gas distribution map using a CSG, the environment is discretized into cells and a chemical sensor is placed in the center of each cell. The sensor responses are recorded for a certain amount of time (typically 2-8 minutes) and a map of the gas distribution is created by plotting the average response [7,22–24] or the maximum response [6,17,25] in each cell. The resulting maps can be smoothed by bicubic [24], linear [26], kriging [11] or Gaussian [7] interpolation. Only in scenarios where the gas distribution is expected to be time-invariant, such as wind tunnels with a continuous gas source [17,24], sequential measurements of a single sensor (e.g., using a robot) can replace a CSG. The advantage of CSGs is that measurements are performed simultaneously in space and time, providing a visualization of the temporal evolution of the gas distribution. With an increasing area, establishing a dense CSG would involve an arbitrarily high number of fixed gas sensors, which poses problems such as cost and lack of flexibility. Because of this, CSGs have been mostly restricted to 1D [20,21] and 2D [15–19] measurements in small arenas.

With the ever-increasing trend towards gas-sensitive aerial robots [27–29], 3D measurements seem necessary to characterize the environment prior to the development of source localization and mapping algorithms for these platforms. However, the three-dimensionality of the environment has been only considered by few works [30–32], and mostly using sequential measurements with a robot. For example, Russell et al. [30] mapped the 3D gas distribution produced by an ethanol gas source in a ventilated indoor arena using a mobile robot with a MOX sensor mounted on an extendable vane. Although the gas source was placed on the floor and ethanol is heavier than air, most gas concentration was found near the ceiling. This supports the idea that dispersion was dominated by convective air currents instead of by the density of the pure substance. Similarly, Reggente et al. [31] used a terrestrial robot equipped with three gas sensors mounted at different heights (0.2, 0.4 and 0.6 m) to acquire 3D measurements in an indoor room. They proposed a multivariate Gaussian function to interpolate between measurements and produce smooth 3D maps of the mean and variance of the concentration. The only work using a 3D CSG was presented by De Vito et al. [32], where 4 wireless

MOX sensors where placed in different locations of a small glass box (0.36 m^3 volume). In this case, the CSG was just a proof of concept of their proprietary wireless technology.

Once a gas distribution map has been obtained, a relevant question is whether the map can be used to pinpoint the gas source location. Lilienthal et al. [7] observed that the concentration maximum estimate (CME), this is the cell with maximum value in the mean concentration map, is a good estimate of the source location when the gas distribution has a circular shape with a peak at the source location. If the gas distribution resembles instead an elongated plume, fitting a gas dispersion model with gas concentration measurements [33] or using the divergence operator [34] can improve the localization accuracy. Later, it was found in 2D [35] and 3D [31] experiments that the maximum cell in a variance map, i.e. the variance maximum estimate (VME) often provides a more accurate indication of the source location than the CME because the variance is less sensitive to saturation. The good localization performance of the CME and VME has been confirmed by Di Rocco et al. [36] and Hernandez et al. [37].

A relatively unexplored characteristic of turbulent odor plumes for source localization using CSGs is the odor intermittency [38]. The odor intermittency, or frequency of odor patches, is a key directional cue for animal chemo-orientation, as demonstrated by several authors [26,38–40]. For example, Haverkamp et al. [26] mapped the odor intermittency inside a wind tunnel using a fast photo-ionization detector (PID) to analyze the odor tracking behaviour of the hawkmoth *Manduca sexta*. Their results highlight the importance of the odor encounter rate for M. sexta to navigate towards a flower in the absence of informative visual cues. Using a CSG to monitor the temporal evolution of the odor intermittency in environments with time-varying airflow could provide valuable information for developing bioinspired plume tracking algorithms for mobile robots that need to work in these conditions. However, this is a priori not feasible due to the slow response time of the low-cost chemical sensors typically used to make up CSGs. A possible solution to overcome this issue is to use the signal processing method recently proposed by Schmuker et al. [41] to extract specific features from the derivative of a MOX sensor (the so-called 'bouts') which are produced by individual plume filaments. A 'bout' is declared when the derivative of the sensor response continuously rises at a fast rate, and the frequency of these 'bouts' is correlated to the distance to a gas source (at least in 2D wind tunnel experiments) [41]. At this time it is unclear if the bout frequency is also correlated to the distance to a gas source in a 3D environment with time-varying airflow.

This paper, which is an extended version of our previous conference paper [42], presents a 3D grid of MOX sensors capable of reconstructing a 3D chemical image of the sensed environment. In contrast to the work by Vito et al. [32], our network is denser (27 nodes instead of 4), the environment is more realistic (an office room instead of a gas chamber) and the set of experiments is richer. The first contribution of this paper is the dataset itself, that we make publicly available. Experimental datasets are fundamental to validate GSL and GDM algorithms, and collecting them is an expensive task. For this reason, the number of available datasets is quite limited and only available in simple 2D environments such as wind tunnels [17]. The second and main contribution is the analysis of indoor gas dispersion patterns (3D) resulting from ten experiments in which an evaporating gas source released ethanol for 90 minutes, including variations in height, release rate and air flow profiles. The third contribution is to study which type of map (mean, variance or bout frequency) is better for estimating the source location in each experiment. In this context, we propose a novel estimator, the bout frequency maximum estimate (BME) which is defined as the cell with maximum value in a bout frequency map, and compare it against the CME and VME.

2. Materials and Methods

2.1. Test environment and MOX sensor grid

A 30 m^2 office room was used as the test environment (Figure 1). The volume of the room (78 m^3) was divided into a 3 × 3 × 3 grid (27 cells) of approximately 2.9 m^3/cell and a MOX sensor was placed in the center of each cell. The 27 MOX sensors (Models: TGS 2600, 2602, 2610, 2611 and 2620, Figaro Engineering Inc., Osaka, Japan) were selected based on availability in the lab and were calibrated to

compensate differences in sensitivity and baseline (see Supplementary Material). Each sensor was mounted on a conditioning board that integrates a voltage divider (load resistor = 68 kΩ) to read out the sensor output (see inset on Fig. 1). The 27 sensors were divided into four groups of 9, 6, 6 and 6 sensors, respectively, where each group of sensors was connected to a power source and a custom processing node that acquired the analog output signals. Each processing node integrates an Arduino Mega microcontroller with a WiFi shield (Arduino AG, Turin, Italy) and a temperature/humidity sensor (DHT22, Adafruit Industries, New York, NY, USA) with an accuracy of 0.2 °C and 2% r.h (see inset on Fig. 1). The sensor signals together with temperature and humidity measurements in different locations of the room were sent to a central computer every 0.5 s via WiFi. A 2D ultrasonic anemometer with 0.01 m/s resolution and 2% accuracy (Model: WindSonic, Gill Instruments Ltd., Hampshire, UK) was used to measure the wind direction and speed in specific locations of the room.

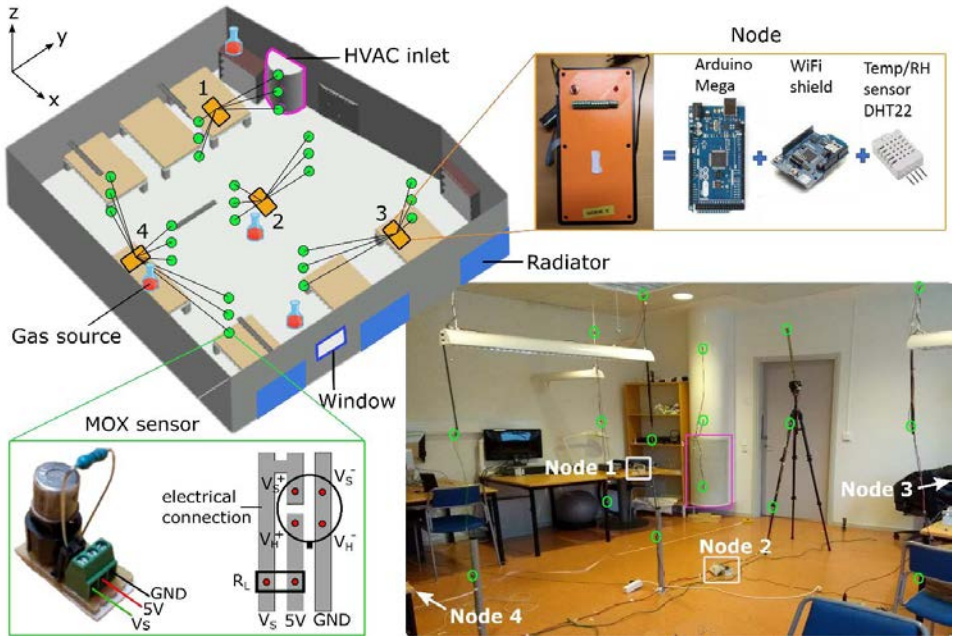


Figure 1. CAD drawing and photo of the test room, highlighting the MOX sensors (green circles), processing nodes (orange rectangles), HVAC inlet (pink rectangle), window and radiators.

2.2. *MOX measuring circuit*

MOX sensors are based on the changes in sensor resistance (R_S) produced by gas exposure. To measure R_S, a voltage (V_C) of 5 V is applied across the sensor electrodes and the load resistor (R_L) connected in series. The voltage V_S (V) measured across the load resistor can be related to R_s $(k\Omega)$ by applying the voltage divider equation

$$R_S = R_L \cdot \frac{V_C - V_S}{V_S} \quad (1)$$

2.2. *Gas source and experiments*

Ten independent experiments were performed over the course of three weeks (Fall 2017) to study the dispersion patterns of a single gas source placed at several locations of the room, including variations in height (h), release rate and air flow profiles (Table 1). The gas source was either a large glass beaker (10 cm diameter, 700 mL capacity) filled with 350 mL of ethanol or a small beaker (3 cm diameter, 100 mL capacity) filled with 80 mL of ethanol. Varying the cross-section of the container was used to increase the strength of the source, as suggested by Lilienthal et al. [43]. The gas source was

normally placed on top of a table ($h = 0.9$ m) near the south wall of the room (Fig. 2a). In Experiment 3, it was placed on the floor near a radiator to test the effect of heat in the gas release (Fig. 2b). In Experiment 4, it was placed on top of a bookshelf ($h = 2.0$ m) and the window was open (Fig. 2c). A fish tank pump was used in Experiment 5 to further increase the release rate. In Experiment 7, the beaker was placed on top of a blackboard ($h = 2.2$ m) and a DC fan placed behind it was used to disperse the chemicals. In Experiment 10, the source was placed on a tripod ($h = 1.75$ m) in the middle of the room.

Table 1. List of experiments.

	Gas source			Airflow	
Experiment	Location	(x, y, z)	Beaker size	Speed	Source
1	Table	(2.7, 0.5, 0.9)	Small	Weak	HVAC
2	Table	(2.7, 0.5, 0.9)	Large	Weak	HVAC
3	Floor	(5.9, 1.6, 0.0)	Large	Weak	HVAC
4	Elevated	(0.1, 4.5, 2.0)	Large	Weak	Window
5	Floor	(2.7, 2.7, 0.0)	Large*	Weak	HVAC
6	Table	(2.7, 0.5, 0.9)	Small	Strong	DC Fan
7	Elevated	(2.7, 0.1, 2.2)	Small	Strong	DC Fan
8	Table	(2.7, 0.5, 0.9)	Small	Strong	Tower fan (I)
9	Table	(2.7, 0.5, 0.9)	Large	Strong	Tower fan (I)
10	Elevated	(2.7, 3.1, 1.7)	Large	Strong	Tower fan (III)

* Bubbler was used to increase the release rate

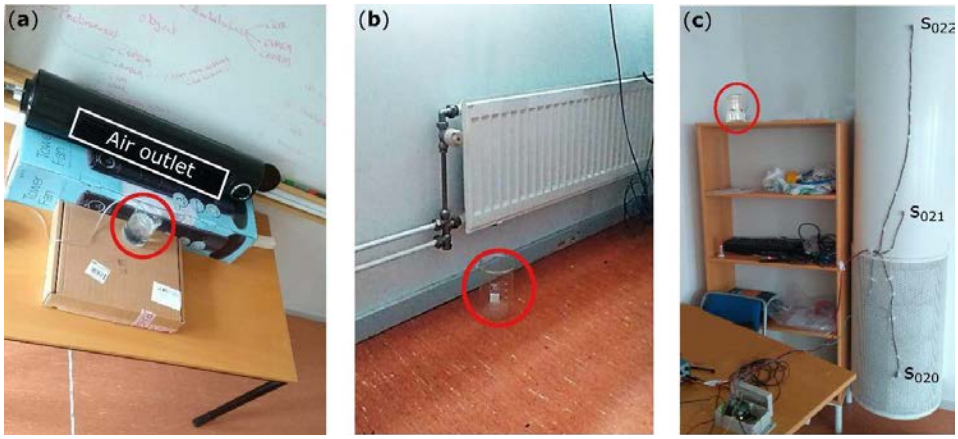


Figure 2. Gas source placement on (**a**) Top of table (Exp. 1); (**b**) Near radiator (Exp. 3); (**c**) Top of bookshelf (Exp. 4).

The first five experiments were performed without strong artificial airflow in the room, since these are the conditions found in most indoor environments (especially if windows and doors are closed). Strong airflow was introduced in the last five experiments by using two types of fan. A 12V DC fan placed behind the source was used in Experiments 6 and 7, whereas a 45W tower fan (Model: ND-668R-42, SEEWILLOPTOELECTRONICS TECHNOLOGY CO., Guangdong, China) with three selectable speeds was used in Experiments 8-10. The tower fan was placed in horizontal position tilted slightly (approx. 10º) towards the ceiling (Fig. 2a). The air flow in both cases was measured using a 2D anemometer placed in the middle of the room (h = 1.6 m). The effect of certain experimental variables can be analyzed by comparing pairs of experiments in which only one variable changes. For example, the effect of the beaker size on gas dispersion under weak airflow can be assessed by comparing

Experiments 1 and 2, and under strong airflow by comparing Experiments 8 and 9. Similarly, the impact of the fan type can be analyzed by comparing Experiments 6 and 8.

An experiment is divided into three phases: baseline recording (10′), gas release (90′) and room cleaning (60′). During the first 10 minutes of the experiment, the baseline of the sensors is recorded in the absence of gas. The gas release starts by pouring ethanol into the beaker and it is maintained for 90 minutes. After that, the cleaning process consists on taking the beaker out of the room, opening the window and door and turning on the fan at maximum power.

2.3. Sensor calibration

Each gas sensor was calibrated individually in a custom testbench (Fig. 3) in which dynamic gas mixtures of ethanol and synthetic air were generated and delivered to a small gas chamber (48 cm^3 internal volume) containing the sensor under test. The gas generation system combines a permeation oven (Model: OHG-4, Owlstone Inc, Connecticut, USA) and a custom evaporation system to produce concentrations below and above 25 ppm, respectively. The following concentrations were prepared using the two methods described above: 0, 1, 2.5, 4, 6, 10, 12, 15, 20, 25, 50, 75, 100, 200, 333, 500, 750 and 1000 ppm. The samples were humidified to 30 ± 1 % r.h. using the bubbler method [21] and the temperature inside the chamber was kept constant at 23 ± 1.5 °C (these are typical values in the test room). Humidity, temperature and pressure inside the gas chamber were monitored using an environmental sensor (Model: BME680, Bosch Sensortec GmbH, Reutlingen, Germany). The concentration was measured using a PID. The sensor response V_s (V) was recorded at 10 Hz by a digital datalogger (Model: USB-6002, National Instruments, Austin, TX, USA). The calibration process starts by flushing the gas chamber with clean air at 1000 ml/min for 5 minutes. Then, randomized gas mixtures are introduced into the chamber for 2 minutes to let the sensor response reach the steady-state value. The last 5 s of the measurement are averaged to produce the steady-state value V_s^* (V), which is then converted to resistance by applying Equation 1. The sensor resistance is divided by the resistance in clean air R_0 $(k\Omega^{-1})$ to correct potential baseline changes. The calibration curves of the 27 MOX sensors were obtained by spline interpolation of the normalized resistance values in the steady-state.

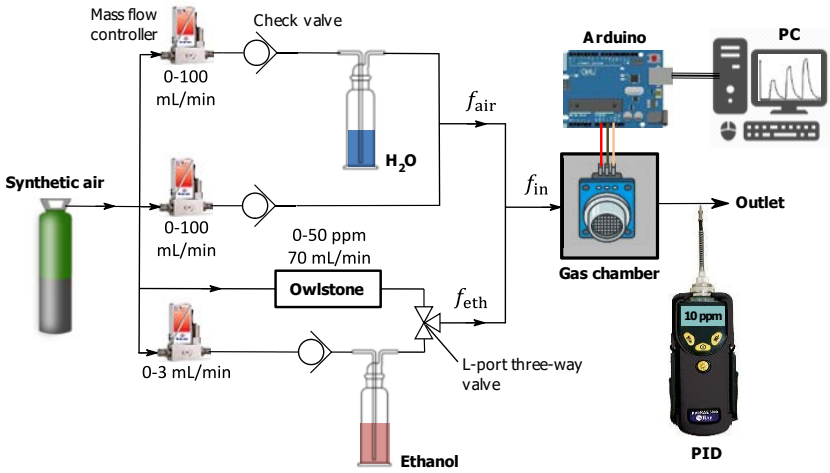


Figure 3. Schematic diagram of the test bench used for sensor calibration.

2.4. Gas distribution maps

The calibrated sensor signals, $s_{x,y,z}(t)$, are processed to provide spatial representations (i.e. a map) of the average concentration, variance of the concentration or bout frequency in the room during a given time interval $t \in \left[T - \frac{W}{2}, T + \frac{W}{2}\right]$, where T is the center of the interval and W the window size. Figure 4 shows a diagram of how to compute a mean concentration map. The mean concentration $\mu_{x,y,z}$ (ppm)

at every sensor location (x, y, z) is computed during the given interval and the resulting values are grouped by height (z) to produce three 2D maps, M_z (3 x 3). The three maps are independently interpolated using a triangulation-based cubic method (MATLAB function *griddata*) and a grid resolution of 1 cm. The maps are then rotated in 3D and vertically stacked. Interpolation occurs only within the bounding box of the sensors, i.e. measurements are not extrapolated, so the maps show an empty area between the sensors and the walls of the room. In our experiments, we used values of $T = \{15, 30, 60, 90\}$ minutes and $W = 5$ minutes. The impact of the window size on the gas distribution maps is studied in Section 3.4, by reducing the measurement window to 30 seconds.

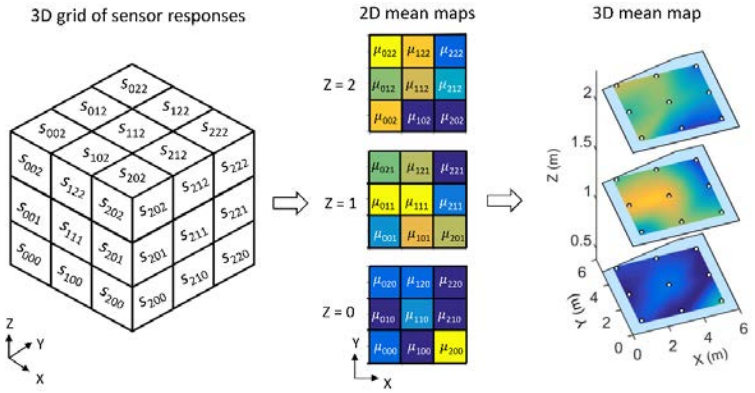


Figure 4. Computation of a mean concentration map (3D). The calibrated sensor responses, $s_{x,y,z}(t)$, are averaged in the time interval $t \in \left[t - \frac{T}{2}, t + \frac{T}{2}\right]$ to produce $\mu_{x,y,z}$, where $x, y, z \in \{0,1,2\}$ are the coordinates of a sensor in the grid. The mean values are grouped by height to produce three 2D maps, which are independently interpolated and vertically stacked.

The variance and bout frequency maps are computed in a similar way, replacing $\mu_{x,y,z}$ by the variance of the response, $\sigma^2_{x,y,z}$, or the bout frequency, $B_{x,y,z}$. To compute $B_{x,y,z}$, the bouts of each sensor signal $s_{x,y,z}(t)$ are extracted using Schmuker's algorithm [41], using as smoothing parameters $\sigma_{\text{smooth}} = 2.5$ s and $\tau_{\text{half}} = 2.5$ s. Regarding the bout amplitude threshold (b_{thr}) necessary to filter out low-amplitude bouts, we used a constant threshold for all sensors instead of applying the 3-sigma criterion proposed by Schmuker et al. The constant threshold was optimized in each experiment to minimize the average localization error (Section 2.5) across different timestamps. Maps obtained by estimation of b_{thr} using the 3-sigma criterion are shown in Section 3.5.

2.5. Source localization accuracy

The gas source localization error (m) is quantified by the Euclidean distance between the cell corresponding to the maximum of the map (i.e., CME, VME or BME) and the true source location (TSL). As a result of the interpolation method that we used, the maximum of a gas distribution map will be always located in one of the sampling points, which imposes a lower bound in the localization error.

3. Results

3.1. Temperature, humidity and airflow in the test room

The environmental parameters in the test room were continuously monitored to ensure that the MOX sensors were operated within the range of variation covered in calibration, and therefore, the calibration models were valid. Stable temperature and humidity values around 21-24 ºC and 25-35 % r.h., respectively, were measured in different locations of the room (except when the window was open). The side of the room near Node 4 was 1 ºC warmer and 2 % r.h. less humid than the rest of the room. When the window was open (Experiment 4), temperature decreased by 2-3 ºC (experiments were

carried out in Fall), humidity increased by 4 % r.h and wind speeds up to 0.2 m/s were measured near the window inside the room. The DC fan generated a wide air beam (approx. 90º), with speeds between 0.05 and 0.2 m/s measured at the height of the fan and 3 meters in front of it. The tower fan at the lowest speed produced a highly directional air flow (approx. 40º) with speed between 0.6 and 1.1 m/s. The HVAC released a small airflow with speed below 1 cm/s.

3.2. Gas distribution patterns under weak airflow

3.2.1. Gas source on table and different beaker size (Experiments 1 and 2)

The gas source was placed on top of a table ($h = 0.9$ m) in Experiments 1 and 2, and the size of the beaker was the only variable that changed between both experiments. In both experiments, the mean gas distribution maps show a gas cloud that initially develops near the source and progressively occupies the middle-upper heights of the room (Figure 5). After approximately 1 hour of release, the gas cloud disperses towards the opposite side of the room and the gas distribution becomes relatively stable. The average concentration inside the cloud is only 1 ppm higher than the surrounding background concentration, which reaches maximum average values of 1.5 ppm after 90 min of release. Accumulation near the ceiling is more evident in Experiment 2, probably due to the larger beaker size. The floor was the area of the room with the lowest concentration, which was usually below 1 ppm. The CME coincided with the TSL frequently during the first hour of release in Experiment 1 and only intermittently during the first 40 minutes of release in Experiment 2. In both cases, the performance of the CME degraded towards the end of the experiments. The average localization error considering the whole release period is 2.0 m in Experiment 1 and 3.9 m in Experiment 2 (the minimum possible error is 0.8 m in both cases).

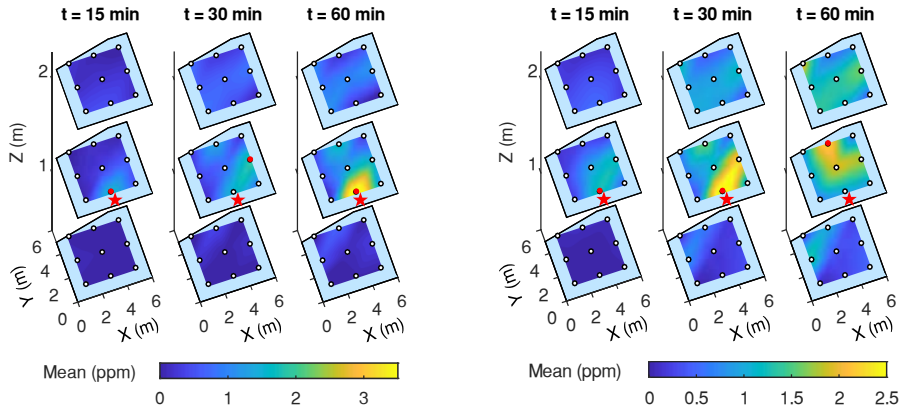


Figure 5. Mean concentration maps at several timestamps in (**a**) Experiment 1 (small beaker) and (**b**) Experiment 2 (large beaker). The gas source is denoted by a red star, the CME by a red dot and the MOX sensors by white dots.

The raw signals in both experiments show high intermittency of the concentration in two specific locations: in front of the source (s_{101}) and near one of the walls of the room (s_{211}). In Experiment 2 (Fig. 6), concentration peaks of 5-25 ppm occurred frequently near the source during the first 40 minutes of release whereas in the second half of the experiment the concentration dropped below 1 ppm for extended periods of time (e.g. $t \in [50, 70]$ minutes). Near s_{211}, concentration peaks are less intense (2-10 ppm) but more consistent along the whole experiment. The remaining sensors show a steady increase in concentration after 10-15 minutes of release, never exceeding concentration values of 2 ppm. A slightly different trend is observed in Experiment 1 (small beaker) because the high concentration fluctuations near the source last for 60 minutes instead of 40, and the concentration near s_{211} fluctuates less intensely.

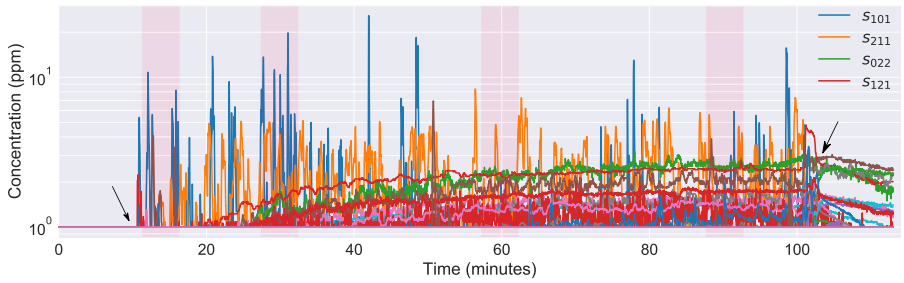


Figure 6. Instantaneous concentration (all sensor responses stacked together) in Experiment 2.

The strong fluctuations of the instantaneous concentration near the source allow a better estimation of the source location through maps representing the concentration variance (Figure 7a) or the bout frequency (Figure 7b). Whereas a gradient in the variance maps is observed at the height of the source and close to it, the bout frequency maps show a larger gradient also visible above the source, extending the detection area from a plane to a volume around the source. In none of the maps the source could be detected from ground-level measurements. The VME and BME in Experiment 1 coincide with the TSL along the entire experiment, yielding the minimum possible localization error (0.8 m) at all timestamps (Fig. 7c). In Experiment 2, the VME coincides with the TSL mostly before $t = 50$ min (when the variability near s_{101} is high). Afterwards, it shifts towards s_{211} and the localization error increases (Fig. 7d). The average error considering the whole experiment is 2.1 m. The bout frequency is less sensitive to the high variability of s_{211} and the BME is located either in front of the source (s_{101}) or on top of it (s_{102}), leading to lower localization errors (average of 1.3 m).

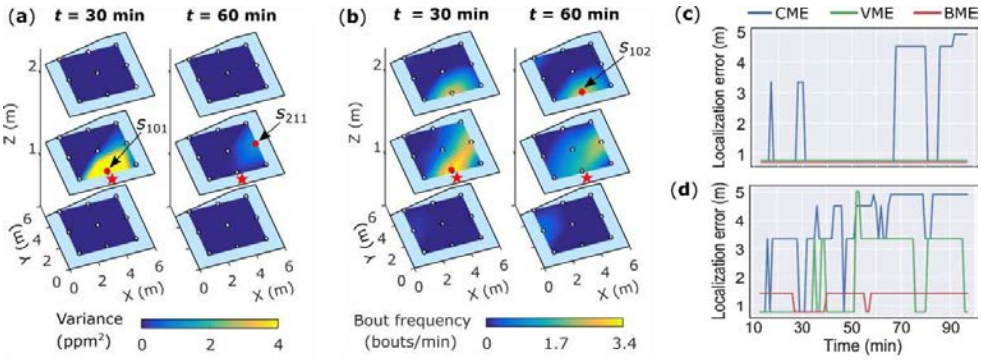


Figure 7. (**a**) Variance and (**b**) Bout frequency maps ($b_{thr} = 0.01$ ppm/s) at $t = [30, 60]$ min in Experiment 2. The gas source is denoted by a red star, the VME and BME by a red dot and the MOX sensors by white dots; (**c-d**) Localization error in (**c**) Experiment 1 and (**d**) Experiment 2. The minimum possible error is 0.8 m.

3.2.3. Gas source near a heat source (Experiment 3)

The gas source was placed on the floor next to a radiator to determine if gas dispersion was affected by convective air currents and how challenging would be to localize a gas source in these conditions. The mean concentration maps (Fig. 8a) reveal that gas progressively accumulates in the upper-middle heights of the room, hindering the localization of the gas source through the CME at all studied timestamps (mean localization error = 4.6 m). The variance (Fig. 8b) and bout frequency maps (Fig. 8c) yielded only slightly better localization errors (average of 4.0 m).

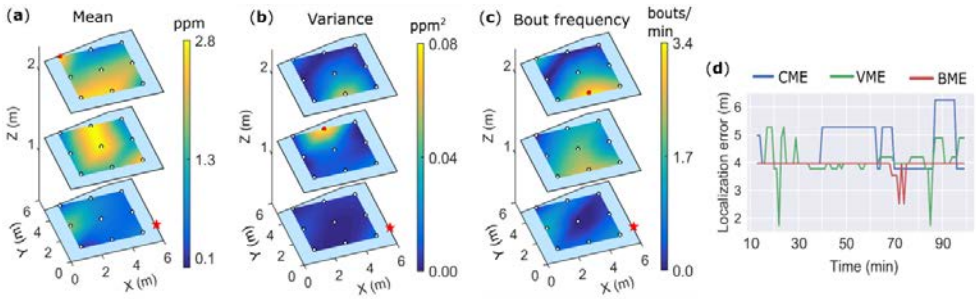


Figure 8. Experiment 3. (**a**) Mean, (**b**) Variance and (**c**) Bout frequency maps (b_{thr} = 0.005 ppm/s) at t = 90 min. The gas source is denoted by a red star, the CME by a red dot and the MOX sensors by white dots; (**d**) Localization error (the minimum possible error is 0.9 m).

3.2.4. Gas source with high release rate (Experiment 4)

The gas source was placed on the floor in the middle of the arena and a fish tank pump was used to increase the release rate. In less than 5 minutes, instantaneous concentrations one order of magnitude higher than in previous experiments were measured in the room (Fig. 9). The mean concentration map at t = 60 min indicates that gas concentration on the floor (≈150 ppm) is three times higher than near the ceiling (≈50 ppm), probably because the molecular weight of the released gas dominated dispersion over convective air currents. The sensor that registers the highest mean concentration and variance is not the closest one to the source (s_{110}) but the one located downwind of the source (s_{210}). Although no artificial airflow was induced in this experiment, the source was located on the path of the air flow released by the HVAC inlet, which dispersed the gas towards the east side of the room. Also, the closest sensor to the gas source (s_{110}) was held on a vertical support that could partially block the amount of gas reaching the sensor surface. This can explain why the CME, VME and BME were not found in the center of the room (i.e. near the source), and consequently the high localization errors (2.3 m).

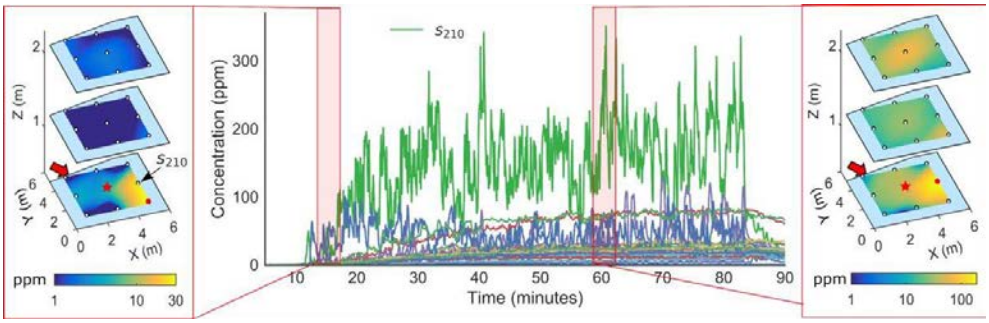


Figure 9. Experiment 4. Instantaneous concentration and mean concentration maps at t = 15 min and t = 60 min, in log scale. In the maps, the gas source is denoted by a red star, the airflow direction by a red arrow, the CME by a red dot and the MOX sensors by white dots.

3.3. Gas distribution patterns under strong unidirectional airflow

3.3.6. Source on table and different types of fan and beaker (Experiments 6, 8 and 9)

Placing a fan behind the source produces gas dispersion patterns radically different than the ones observed previously. For instance, in Experiment 6 (where a DC fan was used) the sensor signals located within the main airflow path (e.g., s_{111} and s_{121}) are characterized by a plateau of stable mean concentration with strong fluctuations due to turbulence (Fig. 10). Sensors outside the main airflow path (e.g. s_{112}) show instead a steady increase in concentration that may take up to 40 minutes to reach a steady state. The mean concentration maps show a plume that develops in the

positive y-direction at the height of the source since the beginning of the release. The plume origin seems to be in the middle of the room instead of near the source, probably because the strong airflow cooled down the sensing layer of the closest sensor to the fan (s_{101}). Remember that a cold MOX sensor is a worse electrical conductor than a hot sensor, i.e. it has higher electrical resistance. Since resistance decreases with concentration in n-type MOX sensors, an increase in resistance translates into a lower measured concentration.

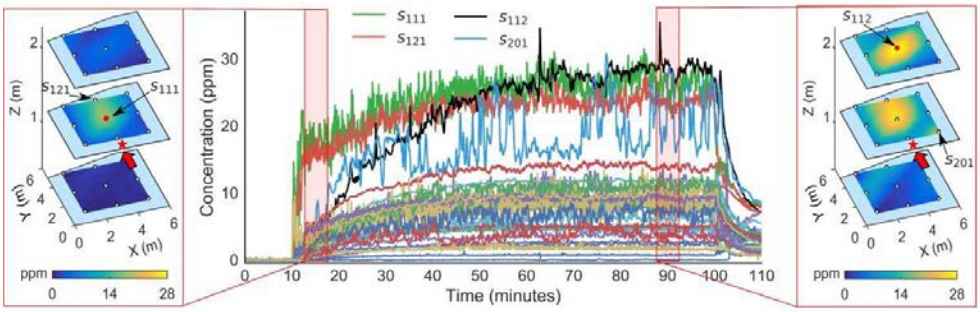


Figure 10. Experiment 6. Instantaneous concentration and mean concentration maps at $t = 15$ min (left inset) and $t = 90$ min (right inset). In the maps, the gas source is denoted by a red star, the airflow direction by a red arrow, the CME by a red dot and the MOX sensors by white dots.

The plume discontinuity near the source together with the progressive accumulation of gas near the ceiling hinders the localization of the source using the CME (mean error = 2.65 m). The variance maps (Fig. 11a) yield a slightly better localization performance (mean error = 2.36 m) but cannot accurately pinpoint the TSL because sensors outside of the plume (e.g. s_{201}) registered higher variance than the sensor in front of the source (s_{101}). For example, at $t = 90$ min the concentration variance of s_{201} (19.2 ppm²) is one order of magnitude higher than the variance of s_{101} (1.7 ppm²). The high variance of s_{201} at this timestamp may be caused by a lasting contact with a gas patch whereas s_{101} presents multiple short fluctuations (i.e. bouts) produced by the plume (Fig. 12). This explains why the bout frequency maps (Fig. 11b) can better reconstruct the gas plume and further reduce the localization error (mean error = 1.63 m).

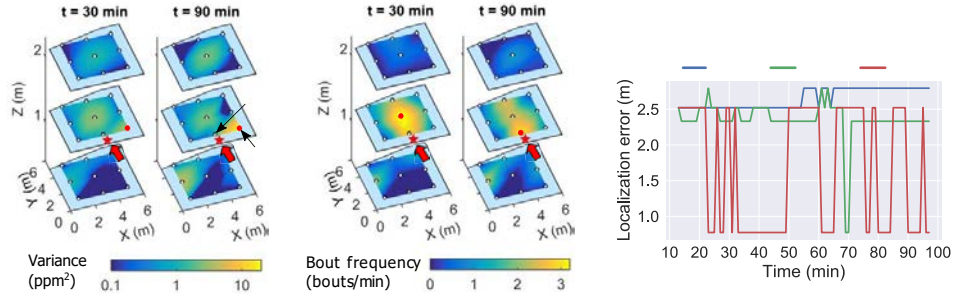


Figure 11. Experiment 6. (**a**) Variance and (**b**) Bout frequency maps ($b_{thr} = 0.045$ ppm/s) at $t = 30$ min and $t = 90$ min. The gas source is denoted by a red star, the airflow direction by a red arrow, the BME and VME by a red dot and the MOX sensors by white dots. (**c**) Localization error (minimum possible error is 0.8 m).

Similar qualitative results were obtained for different types of fan and beaker. For example, in Experiment 8 (tower fan) the mean concentration maps have similar shape than those of Experiment 6, but the maximum mean concentration after 90 minutes of release (11 ppm) is three times lower than in Experiment 6 (28 ppm) despite the tower fan is more powerful than the DC fan. Since the DC fan is more compact than the tower fan, it was placed closer to the headspace of the beaker and this may increase the evaporation rate of the substance. On the other hand, using the large beaker in combination with the tower fan (Experiment 9) yielded the highest average concentrations (78 ppm)

among all experiments with artificial airflow. This indicates that the size of the beaker and the type of fan can strongly affect the evaporation rate of the substance. Regarding GSL, the CME, VME and BME produced errors of 2.5-3.3, 2.6-2.9 and 1.7-2.3 m, respectively, in these experiments.

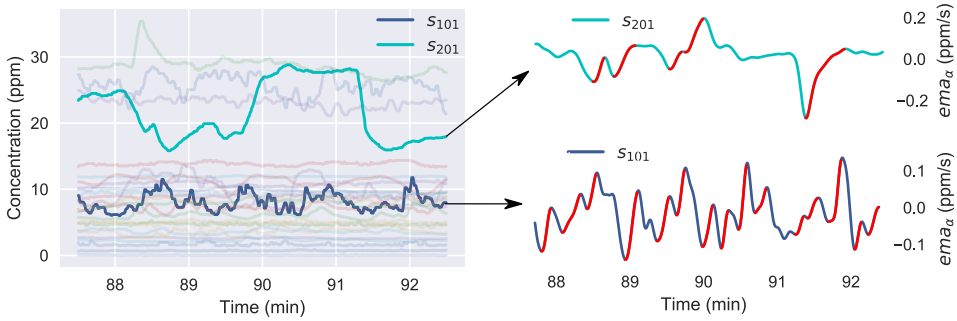


Figure 12. (**a**) Instantaneous concentration of sensors s_{201} and s_{101} between $t = 87.5$ and 92.5 min in Experiment 6. The variance of each signal is indicated in the plot; (**b-c**) ema_α filtered signals of sensors s_{201} and s_{101}, with bouts above a threshold of 0.045 ppm/s marked in red. The number of detected bouts is indicated.

3.3.7. Elevated source and DC Fan (Experiment 7)

In this experiment, the gas source was placed at a height of 2.2 m in front of a DC fan that generated unidirectional air flow. At the beginning of the release, a buoyant gas cloud develops at the height of the source and, as the experiment progresses, the gas cloud progressively bends downwards (Fig. 13a). This is probably because the fan is not powerful enough to maintain a steady height airflow. After 45 min of release, most gas is accumulated 3-5 m downwind of the source and 1 m below it. Consequently, the CME accurately estimates the TSL only immediately after the release and the error increases afterwards (mean error = 4.8 m). The variance maps (Fig. 13b) show the opposite trend, with the VME located far away from the source at the beginning of the experiment and getting closer to the source afterwards (mean error = 3.4 m). The bout frequency maps (Fig. 13c) present two local maxima in most timestamps (one of them matching the TSL), thus leading to the lowest localization errors (mean error = 1.3 m).

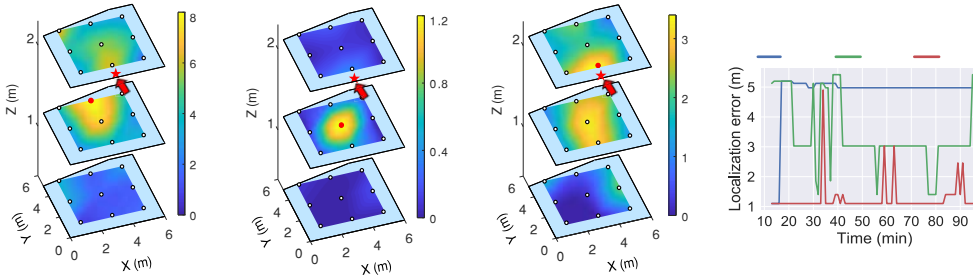


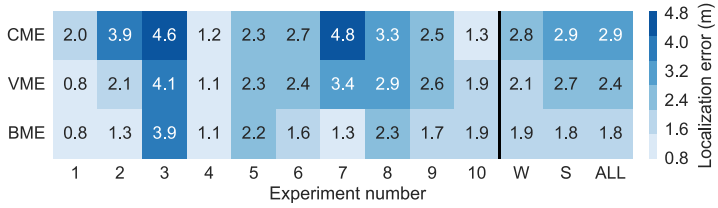
Figure 13. Experiment 7. (**a**) Mean, (**b**) Variance and (**c**) Bout frequency maps ($b_{thr} = 0.01$ ppm/s) at $t = 60$ min. The gas source is denoted by a red star, the airflow direction by a red arrow, the CME, BME and VME by a red dot and the MOX sensors by white dots. (**d**) Localization error (minimum possible error is 1.1 m).

3.3.9. Overall localization performance

The average localization error in each experiment is summarized in Table 2, together with the average values corresponding to weak (W) and strong (S) airflow conditions. It should be recalled that the minimum possible localization error ranges between 0.8 and 1.1 m. The worst localization performance considering the three estimators occurs in Experiment 3, because the source was placed next to a heat source which created buoyant dispersion patterns. The best performance is obtained in

Experiment 4, when the source was placed on one corner of the room and the wind blew against that corner. In weak airflow conditions, the VME and BME perform equally well (1.9-2.1 m error) and both better than the CME (2.8 m error). Under strong airflow, the BME is clearly the best estimator (1.8 m error), followed by the VME (2.7 m error) and the CME (2.9 m error). Considering all scenarios, the BME achieves the lowest localization error (1.8 m) and the CME the largest one (2.9 m).

Table 2. Average localization error (m) in each experiment (columns 1-10), in weak airflow experiments (W), strong airflow experiments (S) and all experiments (ALL).



3.4. *Effect of the window size on the gas distribution maps*

In this section, we explore the impact of reducing the window size (ΔT) from 5 min to 30 s on the mean, variance and bout frequency maps. For that, we plot the temporal evolution of the residual localization error of each estimator in different experiments (Fig. 14), computed as the difference between the localization error using $\Delta T = 5$ min and $\Delta T = 30$ s. The obtained curves indicate that the impact of ΔT in the maps strongly depends on the experiment, and changes over time. For example, reducing the window size negatively affects the performance of the CME during the first hour of Experiment 1 (weak airflow) but can improve the performance afterwards (Fig. 14a). In the same experiment, the VME and BME always degrade at small ΔT. A completely different trend is observed in Experiment 5 (Fig. 14b). Neither one of the three estimators is sensitive to ΔT during the first 30 minutes of release. Afterwards, the VME and BME improve their performance by reducing the window size. In Experiment 7 (strong airflow), the performance of the BME strongly degrades for small ΔT while the variance slightly improves (Fig. 14c).

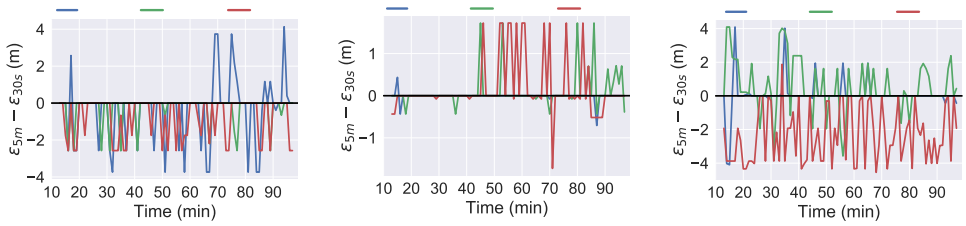


Figure 14. Impact of the window size (ΔT) on the source localization error (ε). The three panels show the temporal evolution of the residual localization error (m), computed as the difference between the localization error using $\Delta T = 5$ min and $\Delta T = 30$ s, in (**a**) Experiment 1, (**b**) Experiment 5 and (**c**) Experiment 7. Positive values of the residual error indicate improvement in localization error when $\Delta T = 30$ s and vice-versa.

The reason why the BME strongly depends on ΔT is illustrated using the bout frequency maps and signals of Experiment 1 at $T = 30$ min (Fig. 15). At this timestamp, the BME coincides with the TSL if $\Delta T = 5$ min but not if $\Delta T = 30$ s (Fig. 15a). In the first case, 16 bouts are detected by the closest sensor to the source (s_{101}) whereas at most 11 bouts are detected by any other sensor (e.g., s_{102}). However, if the window size is too short, the same number of bouts (1) is detected by s_{101} and s_{102}, leading to binary bout frequency maps with multiple local maxima. The reason why s_{101} only detects a single bout when $\Delta T = 30$ s is that the timeframe $T = [29.5, 30.5]$ min (see black rectangle in Fig. 15b) coincides with a period of inactivity near the source after a series of concentration peaks. Resultingly, only a single long bout corresponding to the sensor recovery is detected by s_{101}. This indicates that a window size of 30 s is probably too short for bout detection in weak airflow conditions because 'void' periods longer than 30 s may occur frequently.

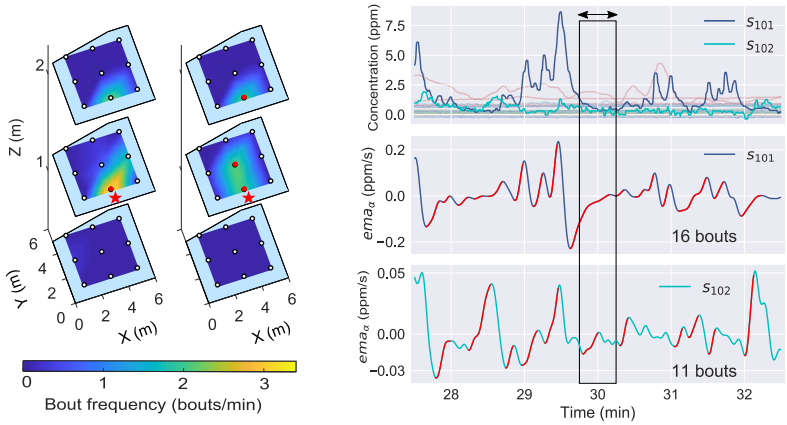


Figure 15. (**a**) Effect of the measurement window (ΔT) on the bout frequency maps at time $T = 30$ min in Experiment 1. The gas source is denoted by a red star, the BME by a red dot and the MOX sensors by white dots; (**b**) (from top to bottom) Instantaneous concentration (highlighting the response of sensors s_{101} and s_{102} over the rest), ema_α filtered signals of sensors s_{101} and s_{102}, with bouts above threshold of 0.01 ppm/s marked in red. The black rectangle highlights the region corresponding to a window size of 30 s.

3.5. Effect of the bout amplitude threshold

The bout frequency maps presented so far were computed using a constant b_{thr} instead of applying the 3σ criteria proposed by Schmuker et al. The former approach requires a priori knowledge of the TSL (for optimizing the threshold) whereas the latter method only requires sensor exposure to clean air during several minutes, which is a more feasible requirement. However, maps obtained with the 3σ approach might not clearly indicate the TSL whereas maps obtained with the optimized constant b_{thr} continuously present its maximum value near the TSL (Fig. 16b).

The reason why the 3σ threshold may not work properly can be explained by looking at the raw signals corresponding to the period $T = [27.5, 32.5]$ min in Experiment 4 (Fig. 16). While the closest sensor to the source (s_{022}) clearly shows the highest mean value and variability among all sensors (Fig. 16a), the highest bout frequency is obtained by a sensor (s_{121}) located further away from the source (Fig. 16b). In the same 5-minute period, 8 bouts are detected by s_{022} (Fig. 16d) whereas 12 bouts are detected by s_{121} (Fig. 16e). The 3σ threshold neglects the fact that the bout amplitude, which is in average 2.6 times larger in s_{022} (0.13 ppm/s) than in s_{121} (0.05 ppm/s) also provides valuable information regarding the TSL. Since the 3σ threshold is estimated in the absence of gas, it cannot predict the bout amplitude during gas exposure. For example, thresholds between 0.001 and 0.04 were obtained using the 3σ criteria, whereas a constant threshold of 0.05 ppm/s was the optimum value to minimize the localization error.

4. Discussion

Monitoring the 3D gas distribution produced by an evaporating chemical substance heavier than air has revealed that assumptions commonly held in the MRO field, such as that ethanol will accumulate on the floor or that the lack of induced airflow produces a stable gas distribution [7], do not necessarily hold in every scenario. The slow evaporation rate of ethanol at room temperature led to mean concentrations of only few ppm in most locations of the room after 90 minutes of release. At such low concentrations, the composition of the ambient air did not change enough to affect the buoyancy of the gas and dispersion was fully dominated by convective or recirculating air currents. This is demonstrated by the progressive gas accumulation towards the ceiling observed in most experiments, especially when the gas source was placed next to a heat source. Resultingly, ground-level gas distributions were not useful for pinpointing the TSL in most experiments.

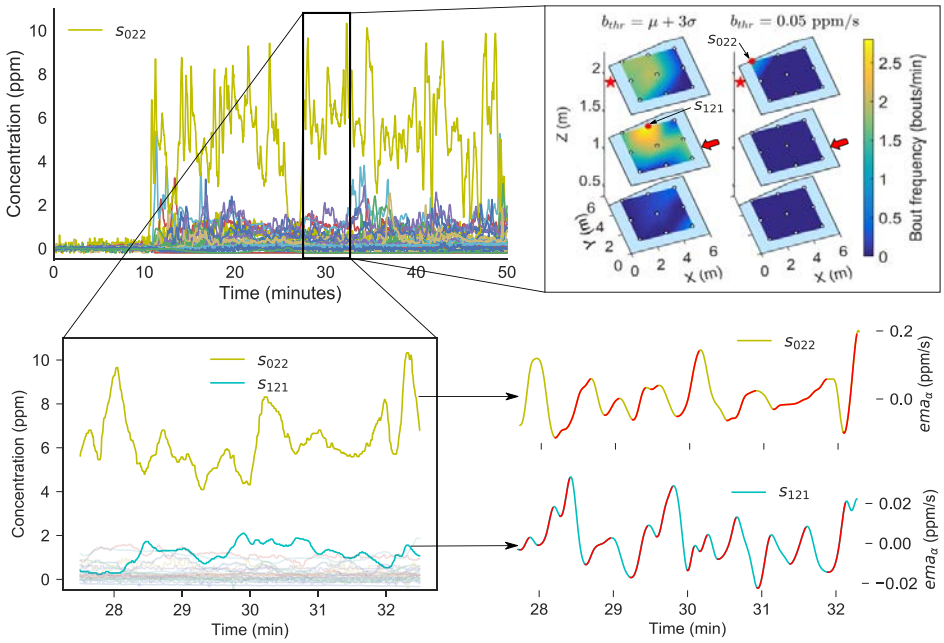


Figure 16. Effect of the bout amplitude threshold (b_{thr}) on the bout frequency maps of Experiment 4. (**a**) Instantaneous concentration of all sensors in the first 50 minutes of experiment; (**b**) Bout frequency map computed at $T = 30$ min using $b_{\text{thr}} = \mu + 3\sigma$ (left) or $b_{\text{thr}} = 0.05$ ppm/s (right). The gas source is denoted by a red star, the airflow direction by a red arrow, the BME by a red dot and the MOX sensors by white dots; (**c**) Zoom-in of the signals between $T = 27.5$ and 32.5 min, highlighting the response of sensors s_{121} and s_{022}; (**d-e**) ema_α filtered signals of sensors s_{022} and s_{121}, with bouts larger than $\mu + 3\sigma$ (ppm/s) marked in red.

The high variability observed near the source in the absence of induced airflow seems to contradict the findings reported by Wandel et al. [43], in which the gas distribution captured by a terrestrial robot in an unventilated indoor environment was stable over several hours and across several trials. Similarly, Lilienthal et al [7] observed stable gas distributions in 10 to 25 minutes. This discrepancy might be due to the different characteristics of the gas source and the environment (e.g. natural air currents) or, more likely, due to the inherent difficulty in assessing stability of a gas distribution with a mobile robot. For example, Lilienthal et al. [7] considered the mapped distribution to be stable when the grid cell with maximum value became spatially invariant, which does not necessary imply that the gas distribution is time-invariant.

These observations suggest that an evaporating chemical source may not be the ideal method to simulate a gas leak in weak airflow environments, since it poses unnecessary challenges to mobile robots such as low concentrations similar to the LOD of MOX sensors [44,45] even near the source, buoyant dispersion patterns and instable gas distributions. Increasing the size of the beaker was not an effective method to increase the release rate, as similar gas distributions were observed regardless of the beaker size. Instead, using a bubbler led to dense gas dispersion patterns and rapid stabilization of the gas distribution, which are more favorable conditions for terrestrial robots. Placing a fan behind the gas source led to stable gas distributions with a clearly visible plume. The air flow produced by the fan had nonetheless a negative effect on the observed gas distribution maps when the fan was placed close to any of the sensors in the grid, reducing the concentration measured by such sensor and "disconnecting" the plume from the source location. Actively controlling the heater temperature is the most obvious solution for this problem, which would also benefit mobile robots equipped with MOX sensors (as they may experience a similar problem).

Regarding GSL, the VME and BME yield in average lower localization errors than the mean (0.5 m and 1.1 m lower, respectively). The BME notably better than the VME in experiments with strong air flow conditions (localization errors 0.9 m lower) and enables the detection of the source from above. The VME requires measurements taken at the height of the source but is less sensitive to the window size than the BME, which requires long measurement windows to produce reliable results. Applying a constant threshold for computing the bout frequency worked better than applying the 3-sigma rule, which suggests that the bout amplitude encodes important information for GSL. However, how to optimize the threshold without knowing the TSL remains an open question to be explored in future works. One idea could be to use a multi-threshold approach, i.e., building bout frequency maps with different thresholds and searching for consistent source predictions over multiple threshold values. It is worth noticing that further analysis can be done on the generated dataset, for example to optimize the bout parameters (i.e., measurement window and amplitude threshold) or to test other estimators of source proximity, such as the mean bout amplitude or the maximum concentration. To achieve this and to enable other further follow-up research we made the dataset and the code used through this work available under the GNU General Public License v3.0.

The CSG used in this work can only provide a rough estimation of the gas concentration distribution due to the low spatial resolution, limitations of MOX technology and technical considerations related to the installation of the sensors. Despite the gas distribution maps are interpolated to smooth the representation, interpolation cannot predict the instantaneous concentration outside of the measurement points due to turbulence. Even at the sensor nodes, the low response time of MOX sensors produce considerable integration of successive measurements. Thus, the interpolated gas distribution maps are only valid when averaging over some time interval. The response time of MOX sensors can be improved by simply removing the cap [46], however directly exposing the sensing surface to the ambient atmosphere may introduce additional problems such as higher sensitivity to air flow or lower stability of the response. Another alternative is to use the smoothed derivative of the sensor response as an estimate of the instantaneous concentration [47], but this requires tuning additional parameters such as the smoothing factor.

The orientation of the sensors with respect to the main air flow and the type of support used to hold them at different heights might also affect the results, because MOX sensors are sensitive to air flow and the holder might block part of the gas reaching the sensor surface. A virtual 3D sensor made of multiple sensors arranged at different orientations and some kind of sensor fusion algorithm [48] could be a solution to this problem. The grid resolution establishes a lower bound in the localization error, determined by the distance between the source and the closest sensor to it. The error can obviously be reduced by increasing the network density, which requires considerable deployment effort and poses problems related to the heat emitted by many sensor heaters (which would affect gas dispersion). Possible options to reduce the power dissipation include low power operation of the sensors or using MEMS sensors with power consumption of few mW [49]. The deployment time can be speed up by using wireless modules [50] with factory-calibrated sensors.

Supplementary Materials: The dataset and code used in this article are available online at https://github.com/jburgues/Orebro3DSEN, under the GNU General Public License v3.0.

Author Contributions: conceptualization, J.B. and V.H.; methodology, J.B. and V.H.; software, J.B. and V.H.; validation, V.H., A.L. and S.M.; formal analysis, J.B.; investigation, J.B.; resources, A.L. and S.M.; data curation, J.B.; writing—original draft preparation, J.B.; writing—review and editing, all authors; visualization, J.B.; supervision, V.H., A.L. and S.M.

Funding: This work was partially funded by the Spanish MINECO program, under grants BES-2015-071698 (SEVERO-OCHOA) and TEC2014-59229-R (SIGVOL), and supported within H2020-ICT by the European Commission under grant agreement number 645101 (SmokeBot). The Institute of Bioengineering of Catalonia (IBEC) is a member of the CERCA Programme/Generalitat de Catalunya.

Acknowledgments: CERCA Programme / Generalitat de Catalunya. The Signal and Information Processing for Sensor Systems group is a consolidated Grup de Recerca de la Generalitat de Catalunya and has support from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2014-SGR-1445).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kowadlo, G.; Russell, R. A. Robot odor localization: A taxonomy and survey. *Int. J. Rob. Res.* **2008**, *27*, 869–894, doi:10.1177/0278364908095118.
2. Lilienthal, A. J.; Loutfi, A.; Duckett, T. Airborne chemical sensing with mobile robots. *Sensors* **2006**, *6*, 1616–1678, doi:10.3390/s6111616.
3. Ishida, H.; Wada, Y.; Matsukura, H. Chemical sensing in robotic applications: A review. *IEEE Sens. J.* 2012, *12*, 3163–3173.
4. Drager - Risk Management Program *Gas Dispersion*; 2013;
5. Lilienthal, A.; Ulmer, H. Learning to detect proximity to a gas source with a mobile robot. *Intell. Robot. …* **2004**, 1444–1449, doi:10.1109/IROS.2004.1389599.
6. Purnamadjaja, A. H.; Russell, R. A. Pheromone communication in a robot swarm: Necrophoric bee behaviour and its replication. *Robotica* **2005**, *23*, 731–742, doi:10.1017/S0263574704001225.
7. Lilienthal, A.; Duckett, T. Building gas concentration gridmaps with a mobile robot. *Rob. Auton. Syst.* **2004**, *48*, 3–16, doi:10.1016/j.robot.2004.05.002.
8. Drescher, A. C.; Park, D. Y.; Yost, M. G.; Gadgil, A. J.; Levine, S. P.; Nazaroff, W. W. Stationary and time-dependent indoor tracer-gas concentration profiles measured by OP-FTIR remote sensing and SBFM-computed tomography. *Atmos. Environ.* **1997**, *31*, 727–740, doi:10.1016/S1352-2310(96)00221-X.
9. Fischer, M. L.; Price, P. N.; Thatcher, T. L.; Schwalbe, C. A.; Craig, M. J.; Wood, E. E.; Sextro, R. G.; Gadgil, A. J. Rapid measurements and mapping of tracer gas concentrations in a large indoor space. *Atmos. Environ.* **2001**, *35*, 2837–2844, doi:10.1016/S1352-2310(01)00081-4.
10. Todd, L. A.; Ramanathan, M.; Mottus, K.; Katz, R.; Dodson, A.; Mihlan, G. Measuring chemical emissions using open-path Fourier transform infrared (OP-FTIR) spectroscopy and computer-assisted tomography. *Atmos. Environ.* **2001**, *35*, 1937–1947.
11. Park, D. Y.; Fessier, J. A.; Yost, M. G.; Levine, S. P. Tomographic reconstruction of tracer gas concentration profiles in a room with the use of a single OP-FTIR and two iterative algorithms: ART and PWLS. *J. Air Waste Manage. Assoc.* **2000**, *50*, 357–370.
12. Simonds, M.; Xiao, H.; Levine, S. P. Optical remote sensing for air pollutants-review. *Am. Ind. Hyg. Assoc. J.* **1994**, *55*, 953–965.
13. Lackner, M. Tunable diode laser absorption spectroscopy (TDLAS) in the process industries - A review. *Rev. Chem. Eng.* 2007, *23*, 65–147.
14. Brooks, R. A.; Di Chiro, G. Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging. *Phys. Med. Biol.* **1976**, *21*, 689.
15. Murai, A.; Yoshimoto, K.; Takemura, R.; Matsukura, H.; Ishida, H. Robotic gas source localization assisted by active airflow generation. In *2015 IEEE SENSORS - Proceedings*; 2015.
16. Zakaria, S. M. M. S.; Visvanathan, R.; Kamarudin, K.; Yeon, A. S. A.; Shakaff, A. Y. M.; Zakaria, A.; Kamarudin, L. M. Development of a scalable testbed for mobile olfaction verification. *Sensors (Switzerland)* **2015**, *15*, 30894–30912, doi:10.3390/s151229834.
17. Vergara, A.; Fonollosa, J.; Mahiques, J.; Trincavelli, M.; Rulkov, N.; Huerta, R. On the performance of gas sensor arrays in open sampling systems using Inhibitory Support Vector Machines. *Sensors Actuators B Chem.* **2013**, *185*, 462–477, doi:10.1016/j.snb.2013.05.027.
18. Yamakawa, S.; Yamaguchi, A. Visualization of Methanol Gas Concentration Distribution Using a Fiber-Optic Sensor Array with Dye Coating. *Trans. Soc. Instrum. Control Eng.* **1995**, *31*, 1273–1278.

19. Hiranaka, Y.; Yamasaki, H. Locating gas sources using a gas sensor array. *Anal. Sci.* **1991**, *7*, 1565–1568.
20. Ferri, G.; Caselli, E.; Mattoli, V.; Mondini, A.; Mazzolai, B.; Dario, P. SPIRAL: A novel biologically-inspired algorithm for gas/odor source localization in an indoor environment with no strong airflow. *Rob. Auton. Syst.* **2009**, *57*, 393–402, doi:10.1016/j.robot.2008.07.004.
21. Shepherd, R.; Beirne, S.; Lau, K. T.; Corcoran, B.; Diamond, D. Monitoring chemical plumes in an environmental sensing chamber with a wireless chemical sensor network. *Sensors Actuators B Chem.* **2007**, *121*, 142–149.
22. Ishida, H.; Kagawa, Y.; Nakamoto, T.; Moriizumi, T. Odor-source Localization In Clean Room By Autonomous Mobile Sensing System. *Proc. Int. Solid-State Sensors Actuators Conf. - TRANSDUCERS '95* **1995**, *1*, 115–121, doi:10.1109/SENSOR.1995.717349.
23. Ishida, H.; Nakamoto, T.; Moriizumi, T. Remote sensing of gas/odor source location and concentration distribution using mobile system. *Sensors Actuators, B Chem.* **1998**, doi:10.1016/S0925-4005(98)00036-7.
24. Pyk, P.; Bermúdez i Badia, S.; Bernardet, U.; Knüsel, P.; Carlsson, M.; Gu, J.; Chanie, E.; Hansson, B. S.; Pearce, T. C.; Verschure, P. F. M. J. An artificial moth: Chemical source localization using a robot based neuronal model of moth optomotor anemotactic search. *Auton. Robots* 2006, *20*, 197–213.
25. Purnamadjaja, A. H.; Russell, R. A. Congregation Behaviour in a Robot Swarm Using Pheromone Communication. *Proceeding Aust. Conf. Robot. Autom.* **2005**, 1–7.
26. Haverkamp, A.; Bing, J.; Badeke, E.; Hansson, B. S.; Knaden, M. Innate olfactory preferences for flowers matching proboscis length ensure optimal energy gain in a hawkmoth. *Nat. Commun.* **2016**, *7*, doi:10.1038/ncomms11644.
27. Burgués, J.; Hernández, V.; Lilienthal, A. J.; Marco, S. Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping. *Sensors* **2019**, *19*, 478.
28. Rossi, M.; Brunelli, D. Gas Sensing on Unmanned Vehicles: Challenges and Opportunities. In *2017 New Generation of CAS (NGCAS), Genova, Italy, 6–9 September*; 2017; pp. 117–120.
29. Neumann, P. P.; Asadi, S.; Lilienthal, A. J.; Bartholmai, M.; Schiller, J. H. Autonomous gas-sensitive microdrone: Wind vector estimation and gas distribution mapping. *IEEE Robot. Autom. Mag.* **2012**, *19*, 50–61, doi:10.1109/MRA.2012.2184671.
30. Russell, R. A. Tracking chemical plumes in 3-dimensions. *2006 IEEE Int. Conf. Robot. Biomimetics, ROBIO 2006* **2006**, 31–36, doi:10.1109/ROBIO.2006.340274.
31. Reggente, M.; Lilienthal, A. J. The 3D-kernel DM+V/W algorithm: Using wind information in three dimensional gas distribution modelling with a mobile robot. *Proc. IEEE Sensors* **2010**, 999–1004, doi:10.1109/ICSENS.2010.5690924.
32. De Vito, S.; Fattoruso, G.; Liguoro, R.; Oliviero, A.; Massera, E.; Sansone, C.; Casola, V.; Di Francia, G. Cooperative 3D Air Quality assessment with wireless chemical sensing networks. *Procedia Eng.* **2011**, *25*, 84–87, doi:10.1016/j.proeng.2011.12.021.
33. Lilienthal, A.; Streichert, F.; Zell, A. Model-based shape analysis of gas concentration gridmaps for improved gas source localisation. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*; 2005; pp. 3564–3569.
34. Cabrita, G.; Marques, L. Divergence-based odor source declaration. *2013 9th Asian Control Conf. ASCC 2013* **2013**, doi:10.1109/ASCC.2013.6606390.
35. Lilienthal, A. J.; Reggente, M.; Trinca, M.; Blanco, J. L.; Gonzalez, J. A statistical approach to gas distribution modelling with mobile robots - The Kernel DM+V algorithm. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), St. Louis, MO, USA, 10–15 October*; 2009; pp. 570–576.

36. Di Rocco, M.; Reggente, M.; Saffiotti, A. Gas source localization in indoor environments using multiple inexpensive robots and stigmergy. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*; 2011; pp. 5007–5014.
37. Hernandez Bennetts, V.; Lilienthal, A. J.; Neumann, P. P.; Trincavelli, M. Mobile Robots for Localizing Gas Emission Sources on Landfill Sites: Is Bio-Inspiration the Way to Go? **2012**, *4*, 20, doi:10.3389/fneng.2011.00020.
38. Murlis, J. Odor Plumes And How Insects Use Them. *Annu. Rev. Entomol.* **1992**, *37*, 505–532, doi:10.1146/annurev.ento.37.1.505.
39. Mafra-Neto, A.; Cardé, R. T. Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. *Nature* **1994**, *369*, 142.
40. Weissburg, M. J.; Zimmer-Faust, R. K. Odor plumes and how blue crabs use them in finding prey. *J. Exp. Biol.* **1994**, *197*, 349–375.
41. Schmuker, M.; Bahr, V.; Huerta, R. Exploiting plume structure to decode gas source distance using metal-oxide gas sensors. *Sensors Actuators, B Chem.* **2016**, *235*, 636–646, doi:10.1016/j.snb.2016.05.098.
42. Burgués, J.; Hernandez, V.; Lilienthal, A. J.; Marco, S. 3D Gas Distribution with and without Artificial Airflow: An Experimental Study with a Grid of Metal Oxide Semiconductor Gas Sensors. In *Multidisciplinary Digital Publishing Institute Proceedings*; 2018; Vol. 2, p. 911.
43. Lilienthal, A.; Zell, A.; Wandel, M.; Weimar, U. Sensing odour sources in indoor environments without a constant airflow by a mobile robot. In *2001 IEEE International Conference on Robotics and Automation (ICRA), Seoul, Korea, 21–26 May*; 2001; Vol. 4, pp. 1–6.
44. Burgués, J.; Jimenez-Soto, J. M.; Marco, S. Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models. *Anal. Chim. Acta* **2018**, *1013*, 13–25, doi:10.1016/j.aca.2018.01.062.
45. Burgués, J.; Marco, S. Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors. *Anal. Chim. Acta* **2018**, *1019*, 49–64, doi:10.1016/j.aca.2018.03.005.
46. Martinez, D.; Arhidi, L.; Demondion, E.; Masson, J.-B.; Lucas, P. Using insect electroantennogram sensors on autonomous robots for olfactory searches. *J. Vis. Exp. JoVE* **2014**.
47. Muezzinoglu, M. K.; Vergara, A.; Huerta, R.; Rulkov, N.; Rabinovich, M. I.; Selverston, A.; Abarbanel, H. D. I. Acceleration of chemo-sensory information processing using transient features. *Sensors Actuators, B Chem.* **2009**, *137*, 507–512, doi:10.1016/j.snb.2008.10.065.
48. Klein, L. A.; Klein, L. A. *Sensor and data fusion concepts and applications*; Spie optical engineering press Bellingham (WA), 1993; Vol. 2;.
49. Burgués, J.; Marco, S. Low power operation of temperature-modulated metal oxide semiconductor gas sensors. *Sensors (Switzerland)* **2018**, *18*, doi:10.3390/s18020339.
50. De Vito, S.; Fattoruso, G.; Liguoro, R.; Oliviero, A.; Massera, E.; Sansone, C.; Casola, V.; Di Francia, G. Cooperative 3D air quality assessment with wireless chemical sensing networks. *Procedia Eng.* **2011**, *25*, 84–87.

sensors

MDPI

Article

Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping

Javier Burgués [1,2,*], Victor Hernández [3], Achim J. Lilienthal [3] and Santiago Marco [1,2]

1 Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain; smarco@ibecbarcelona.eu
2 Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Marti i Franqués 1, 08028 Barcelona, Spain
3 AASS Mobile Robot Olfaction Lab, Örebro University, SE 70182 Örebro, Sweden; victor.hernandez@oru.se (V.H.); achim@lilienthals.de (A.J.L.)
* Correspondence: jburgues@ibecbarcelona.eu; Tel.: +34-934-029-070

Received: 11 December 2018; Accepted: 22 January 2019; Published: 24 January 2019

Abstract: This paper describes the development and validation of the currently smallest aerial platform with olfaction capabilities. The developed Smelling Nano Aerial Vehicle (SNAV) is based on a lightweight commercial nano-quadcopter (27 g) equipped with a custom gas sensing board that can host up to two in situ metal oxide semiconductor (MOX) gas sensors. Due to its small form-factor, the SNAV is not a hazard for humans, enabling its use in public areas or inside buildings. It can autonomously carry out gas sensing missions of hazardous environments inaccessible to terrestrial robots and bigger drones, for example searching for victims and hazardous gas leaks inside pockets that form within the wreckage of collapsed buildings in the aftermath of an earthquake or explosion. The first contribution of this work is assessing the impact of the nano-propellers on the MOX sensor signals at different distances to a gas source. A second contribution is adapting the 'bout' detection algorithm, proposed by Schmuker et al. (2016) to extract specific features from the derivative of the MOX sensor response, for real-time operation. The third and main contribution is the experimental validation of the SNAV for gas source localization (GSL) and mapping in a large indoor environment (160 m^2) with a gas source placed in challenging positions for the drone, for example hidden in the ceiling of the room or inside a power outlet box. Two GSL strategies are compared, one based on the instantaneous gas sensor response and the other one based on the bout frequency. From the measurements collected (in motion) along a predefined sweeping path we built (in less than 3 min) a 3D map of the gas distribution and identified the most likely source location. Using the bout frequency yielded on average a higher localization accuracy than using the instantaneous gas sensor response (1.38 m versus 2.05 m error), however accurate tuning of an additional parameter (the noise threshold) is required in the former case. The main conclusion of this paper is that a nano-drone has the potential to perform gas sensing tasks in complex environments.

Keywords: robotics; signal processing; electronics; gas source localization; gas distribution mapping; gas sensors; drone; UAV; MOX sensor; quadcopter

1. Introduction

Thanks to recent advances in micro-technology, manufacturers of unmanned aerial vehicles (UAVs), or drones, have been able to develop miniaturized flying platforms; with insect-sized aircrafts expected in the future [1] (Figure 1). A micro-UAV (MAV or μUAV) has a length between 15 cm and 100 cm and a weight between 50 g and 2 kg [2]. A nano air vehicle (NAV) or nano-drone is extremely small, with a wingspan lower than 15 cm, and weighs less than 50 g [2]. If compared to piloted aircrafts

or larger UAVs, MAVs and NAVs can fly at low altitudes (i.e., below 150–200 m) over small geographic or site-specific areas on a real-time basis at affordable operational costs [3]. MAVs equipped with gas detection systems and/or sampling bags have been already used in the fields of environmental monitoring [4–9], volcanic gas sampling [10–14], localization of fugitive emissions [15,16], early fire detection [17,18], precision agriculture [19–21], landfill monitoring [22–24], disaster response [25] and mine blasting [26], among others [27,28].

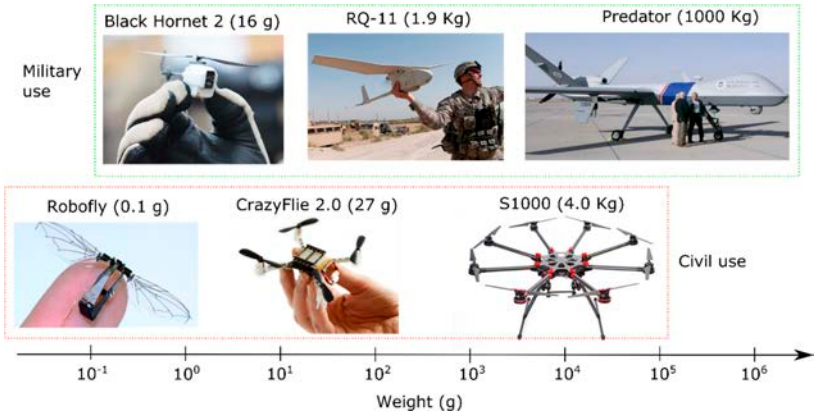


Figure 1. Overview of the UAV landscape, from insect-sized drones to military aircrafts, classified according to the approximate weight and size. The graphic shows the large range of UAV sizes, which spans seven orders of magnitude.

The tiny form-factor and maneuverability of NAVs allow sensing of hazardous environments inaccessible to terrestrial robots and bigger drones, can fly over areas being unobserved and are not a hazard for humans, enabling their use in public areas or inside buildings. Providing a NAV with olfaction capabilities is now possible due to miniaturization and low-cost fabrication of gas sensors. An odor-sensitive nano-drone can be used in a myriad of applications that range from environmental monitoring to search and rescue, leak detection, chemical, biological, radiological and nuclear (CBRN) defense, explosive finding, among others. For example, in the aftermath of an earthquake or explosion it is important to search for victims and hazardous gas leaks inside pockets that form within the wreckage of collapsed buildings. A nano-drone could navigate such scenarios much faster than a terrestrial robot, passing through confined spaces that preclude human entry, evading obstacles or large gaps and sampling the space in three dimensions (3D).

1.1. Related Work on Gas-Sensitive Nanodrones

Two experimental works [29,30] already explored the viability of nano-drones for gas sensing tasks. Rossi et al. [29] performed preliminary indoor experiments using a CrazyFlie 2.0 nano-drone equipped with a metal-oxide semiconductor (MOX) gas sensor. The authors found that the air drawn around the airframe strongly affected the sensor response, resulting in useless signals. They evaluated several mechanical solutions to keep the sensor out of the region of airflow created by the nano-drone, but the drone became uncontrollable because of inertia problems. The adopted solution was to operate the drone in the so-called "butterfly" mode, in which a human pilot lands the drone in the proximity of the source and halts the motors to take a measurement. In this way, the sensor signals are not affected by the rotors but, at the same time, the 3D sensing capabilities of the drone are not used, and the approach might not scale well to large scenarios.

Fahad et al. [30] equipped the same nano-drone with a chemically sensitive field effect transistor (CS-FET) sensor for hydrogen (H_2) detection. The test environment was a chemical hood in which H_2 was released from the upper part of the hood and the drone ascended from the bottom of the hood to

the area near the source (h = 60 cm), aided by high tension strings. The sensor response increased as the drone approached the source, reaching its maximum value after hovering (i.e., levitating) near the source for 40 s. The drone was merely used as a proof-of-concept demonstration of the proprietary gas sensor developed by the authors. The above works suggest that a nano-drone might be used for gas source localization (GSL), however the experimental scenarios were extremely simple.

1.2. Experimental Evaluation of Gas-Sensitive Nanodrones

The main problems for performing large-scale experiments in complex environments using nano-drones are related to the limited on-board resources and difficulty to control the platform due to inertia and stability issues. Taking as an example the CrazyFlie 2.0 (CF2) quadcopter, the tiny 240 mAh battery delivers power for up to 7 min of flight and 15 grams of payload, which means that only lightweight and power-efficient sensors can be mounted on board. Self-localization and obstacle avoidance—required for autonomous navigation—are hard to accomplish because laser scanners, for example, are too heavy for the nano-drone payload. Autonomous hovering of a CF2 equipped only with a tiny RGB camera has been achieved in indoor experiments [31], although the camera consumes all available payload and reduces the flight time to 3.5 min. The Global Positioning System (GPS) can be used for localization outdoors where, however, nano-drones can often not be controlled stably due to their low inertia and strong wind. Navigation in indoor areas can be achieved through motion capture systems (MOCAPs) [32] or radio frequency (RF)-based systems [33]. MOCAPs offer high accuracy (1 mm error) but are expensive, typically only cover small volumes and require line-of-sight (LOS) between the cameras and the drone. RF-based systems are cheaper, have a larger coverage area, do not necessarily require LOS but are less accurate (10 cm error). In many realistic scenarios, deploying an external localization system might not be possible (e.g., in a disaster situation) and the drone would have to navigate autonomously or remotely controlled by an operator.

Gas sensing tasks are also subject to additional constraints, as they must be executed in the short time limited by the battery capacity and relying exclusively on one or two chemical sensors. It should be noted that most research on GSL is based on terrestrial robots, which can be running for hours, perform long measurements of 1–2 min at each sampling location and possibly use selective sensors (e.g., TDLAS, OGI cameras, e-noses) and anemometers. Large drones can fly for 20–25 min and be equipped with the same technology as terrestrial robots. Nano-drones are therefore subject to unprecedented constraints because a stop-sense-go strategy would only allow for at most 10 measurements (of 30 s each), the limited number of sensors hinder rejecting chemical interferences and the absence of anemometry prevents assessing the wind direction, which is a key parameter for GSL. During its operation, the drone can also not fly too fast or the relevant structures of the chemical plume may become blurred due to the slow response time of the sensors [34,35]. Lilienthal et al. [35] pointed out that the gas distribution mapped by a terrestrial robot may be slightly shifted as compared to the real distribution, due to the memory effect of MOX sensors.

1.3. Gas Source Localization

Gas source localization (GSL) is a key task for gas-sensitive robots that consists in identifying the point of release of a hazardous gas. GSL strategies can be divided into three groups: reactive plume tracking (bioinspired), plume modelling and gas distribution mapping (GDM) strategies [36] (Figure 2). Bioinspired algorithms attempt to track the gas plume along its entire length, mimicking the excellent odor plume tracking capabilities of some flying insects. At this time it is unclear whether bioinspired reactive behaviours have better performance than other approaches based on statistical inference from cumulative readings [37–39]. According to [11], the bioinspired reactive behaviors that researchers have implemented on mobile robots are modelled too simple to cope with complex environments, gas sensors used are too slow to resolve plume features in a milliseconds scale and mobile robots are not agile enough for performing insect-like reactive movements. Besides of that, bioinspired algorithms often require wind measurements (anemometry) and real-time obstacle detection, which hinders their use in nano-drones.

Plume modelling algorithms [40–43] assume a mathematical model for the plume, such as Gaussian shaped plumes [44] or filament/particle based models [45–47], and use local measurements of concentration and wind to fit the model and estimate the source location, which is usually a parameter of the model. The practical applicability of plume modelling methods is limited because they tend to make overly simplifying assumptions (e.g., that the wind field is stable, spatially uniform and measurable), often require a-priori information such as the source release rate in Gaussian models [40], or are sensitive to meta-parameters such as the odor detection threshold in filament-based models [41] or the probability of particle encounter as a function of distance to the source in particle models [42]. The Gaussian model also assumes that the exploration area does not contain obstacles or walls which could otherwise distort the plume. Further, long time-averaging might be required to observe a Gaussian plume [48,49] or to estimate the particle density in a certain region, which slows down plume modelling approaches for GSL.

GDM approaches use sensor measurements to first build a map of the gas distribution in the environment, which is then used to estimate the source location. Maps reflecting the instantaneous concentration [50], the mean concentration [35,38], the variance of the concentration [38,51] or the number of odor hits (which are over-threshold segments in the sensor response) [52] have been successfully used for GSL. To build a gas distribution map, the path of the robot should roughly cover the entire search area, typically moving along a predefined trajectory [35,38,50], although adaptative approaches have been proposed [53]. When the map is based on statistical properties of the gas distribution (e.g., mean or variance), long measurements (typically 30 s and more) are often carried out at each sampling location. This stop-and-sense strategy is not suitable for UAVs, particularly for nano-drones, since any hovering stop quickly drains the battery [15]. GDM algorithms are less efficient than bioinspired and plume modelling algorithms (in terms of distance travelled by the robot) but do not rely on unrealistic assumptions or require wind information or plume parameters. They are compatible with slower robots and sensors and the resulting gas distribution map can be used for other purposes beyond GSL.

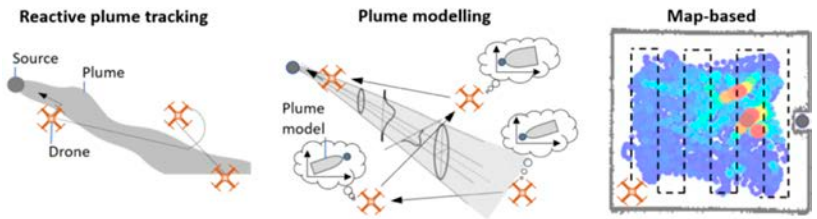


Figure 2. Gas source localization strategies. (**left**) Reactive plume tracking; (**center**) Plume modelling; (**right**) Map-based.

MOX sensors are probably the most suitable gas sensing technology for nano-drones due to the reduced size (few cm^2), low power consumption (some models (few mW)) [54] and simplicity of the conditioning electronics. Ishida et al. [55] observed that the maximum response of a MOX sensor often corresponds to the approximate location of a gas source if the sensor readings are acquired in motion. Such a correlation was never observed if the concentration measurements were collected with a stop-sense-go strategy. This behaviour, which was previously reported by Atema et al. [56] and confirmed by Farah and Duckett [57], was attributed to the long recovery time of MOX sensors. If a MOX sensor is exposed to two consecutive gas patches, the response to the second stimulus will occur when the sensor has not yet recovered from the first exposure. The overall response to the second patch will be higher than if the sensor had been already fully recovered from the first patch. Since the local density of gas patches tends to be higher near the source, the encounter rate with gas patches is higher if the robot is moving. Thus, it is plausible that Rossi et al. [36] and Luo et al. [50] obtained a good estimate of source location in outdoor experiments using the instantaneous response of a MOX sensor mounted on a micro drone (≈800 g). In the latter case, the authors were able to build a 3D gas

distribution map of a relatively large outdoor environment (10×16 m^2) and localize the source in less than 10 min. For that, they used a predefined flight path consisting of two 2D rectangular sweepings at different heights (0.3 m and 1 m), without stopping at predefined locations for measuring. The cell with maximum value of the gas distribution map coincided approximately with the true source location.

The results of Luo et al. [50] also show that the density of odor hits in a map is correlated to the source location. An odor hit is typically declared when the instantaneous concentration exceeds a certain threshold. Although the term 'odor hit' is widely used in the literature to indicate contact between the gas and the sensor, strictly speaking it would be more precise to refer to these events as 'gas hits' or 'plume hits' since most gases being released in reported experiments are odorless for humans. Nonetheless, Thomas Lochmatter [11] argues that the definition of odor as a gas that humans can smell can be extended to robots. In this sense, terms such as "odor" can also be used instead of "gas". Odor hits are supposedly caused by contact with individual patches of the plume and there are indications that insects use similar features to orient rapidly in turbulent plumes [58,59]. Well-known GSL algorithms such as Pang and Farrell's method [41] or Infotaxis [42] model the plume as a sequence of chemical filaments/particles and use odor hits to localize the source. Detecting odor hits with a MOX sensor is a challenging task due to the long recovery time (in the order of 10–30 s) compared to the temporal resolution of the chemical stimuli (in the order of ms). Several research groups attempted to improve the response time of MOX sensors by novel hardware designs [60,61] or by signal processing, using inverse dynamical models [34,62,63], artificial neural networks [64] or extracting specific features [65]. Hardware methods are not appropriate for nano-drones, as they usually increase the size, weight and power consumption of the system. Signal processing methods are more suitable. Schmuker et al. [65] proposed a method to extract short time-scale features (called 'bouts') from the derivative of the MOX sensor response that could be caused by contact with individual filaments of the plume. In wind tunnel experiments, the authors found that the frequency of these 'bouts' (as detected with MOX sensors) is strongly correlated to the distance of a gas source: the higher the bout frequency, the closer the sensor to the gas source. The proposed algorithm uses a threshold to filter out low-amplitude bouts—produced by sensor noise—that would otherwise lead to meaningless correlations. It was also found that the variance of the bout frequency (measured across multiple trials) indicates whether the detector is in the plume centerline (low variance) or slightly lateral from it (high variance). These features suggest a plume-tracking GSL strategy in which the robot first tries to locate the plume centerline by monitoring the variance of the bout frequency and then approaches the source by moving in the direction of increasing bout frequency. The advantages of this method are that anemometry is not required and, since bouts are detected in the derivative of the signal, the algorithm is not very sensitive to changes in the background concentration or to differences between individual gas sensors. The sensitivity of the algorithm to the threshold used to discard noise-induced bouts has not been studied yet, but it might have a large impact on the results. Since this method has not been experimentally validated beyond a wind tunnel, it remains to be shown if meaningful gradients of both bout frequency and its variance can be obtained in real scenarios.

1.4. Proposed Smelling Nano Aerial Vehicle (SNAV)

In this work, we propose a Crazyflie 2.0 nano-drone equipped with a MOX sensor for gas source localization in large indoor environments. We calibrate the sensor to compensate the non-linear response, obtain measurements in concentration units and to estimate the limit of detection (LOD). We assess the impact of the propellers on the MOX sensor signals at different distances of a chemical source. We then compare two GSL strategies, one based on the instantaneous response and the other one based on the bout frequency in two experiments where the source is placed challenging positions for the drone. We show that proper selection of the bout amplitude threshold is critical for good localization performance. We also demonstrate that a 3D gas distribution map of an environment of 160 m^2 can be built in less than 3 min using the proposed platform and the source can be accurately localized from the map.

2. Materials and Methods

2.1. Nano-Drone and Gas Sensors

Among all commercial nano-drones, we selected the CF2 (Bitcraze AB, Malmö, Sweden) due to its low cost, reduced dimensions (10 × 10 cm) and open hardware/software architecture. Weighing only 27 g, it has a maximum recommended payload of 15 g and is capable of up to 7 min of continuous flight. The main microcontroller (µC) is an ARM 32-bit STM32F405 Cortex-M4, which runs an open source real-time operating system (FreeRTOS). The CF2 communicates with a ground station (PC with USB radio antenna) over the 2.4 GHz ISM radio band in up to 1 km range line-of-sight (LOS). The CF2 periodically sends over this link internal variables of the system (sensor measurements, position, rotor speed, battery level, user-defined variables, etc.) and receives commands from the base station such as waypoints or parameter updates. An expansion port—accessible from top and bottom of the drone—provides the user with access to certain µC pins (4 × GPIO, 3 × PWM), power lines (GND, VCC 3.0 V, VBAT 3.7 V) and communication buses (I2C, SPI, 2 × UART). New hardware can be added, which can be driven with 3.0 V logic and is easily integrated into the platform by soldering it into an extension board called a deck, which can be connected to the expansion port. An installed deck is automatically detected and initialized by the CF2 at startup, without having to modify the stock firmware. Only the deck driver needs to be programmed, a separate piece of code that specifies the functionality of the deck in FreeRTOS language.

Among all commercial nano-drones, we selected the CF2 (Bitcraze AB, Malmö, Sweden) due to its low cost, reduced dimensions (10 × 10 cm) and open hardware/software architecture. Weighing only 27 g, it has a maximum recommended payload of 15 g and is capable of up to 7 minutes of continuous flight. The main microcontroller (µC) is an ARM 32-bit STM32F405 Cortex-M4, which runs an open source real-time operating system (FreeRTOS). The CF2 communicates with a ground station (PC with USB radio antenna) over the 2.4 GHz ISM radio band in up to 1 km range line-of-sight (LOS). The CF2 periodically sends over this link internal variables of the system (sensor measurements, position, rotor speed, battery level, user-defined variables, etc.) and receives commands from the base station such as waypoints or parameter updates. An expansion port—accessible from top and bottom of the drone—provides the user with access to certain µC pins (4 × GPIO, 3 × PWM), power lines (GND, VCC 3.0 V, VBAT 3.7 V) and communication buses (I2C, SPI, 2 × UART). New hardware can be added, which can be driven with 3.0 V logic and is easily integrated into the platform by soldering it into an extension board called a deck, which can be connected to the expansion port. An installed deck is automatically detected and initialized by the CF2 at startup, without having to modify the stock firmware. Only the deck driver needs to be programmed, a separate piece of code that specifies the functionality of the deck in FreeRTOS language.

A custom deck (i.e., a printed circuit board), named the MOX deck, was developed to interface two MOX gas sensors to the CF2 (Figure 3). The deck contains two sockets for 4-pin Taguchi-type (TGS) gas sensors, a temperature/humidity sensor (SHT25, Sensirion AG, Stäfa, Switzerland), a dual-channel digital potentiometer (AD5242BRUZ1M, Analog Devices, Norwood, MA, USA) and two MOSFET p-type transistors (NX2301P, NEXPERIA, Nimega, NL). We selected the TGS 8100 sensor (Figaro Engineering Inc., Osaka, Japan) due to its compatibility with 3.0 V logic, power consumption of only 15 mW (the lowest in the market as of June 2016) and miniaturized form factor (MEMS). Since the sensor heater uses 1.8V, we included two transistors (one per sensor) to reduce the applied power by means of pulse width modulation (PWM). The MOX read-out circuit (Figure 4) is a voltage divider connected to the µC's analog-to-digital converter (ADC). The voltage divider is powered at 3.0 V and the load resistor (R_L) can be set dynamically by the potentiometer (from 60 Ω to 1 MΩ in steps of 3.9 kΩ). In the current work, we only used one of the two sensors and R_L was fixed to 70 kΩ. This load resistor value is selected to operate the voltage divider near its mid-range, where the sensitivity is maximum (according to the expected concentrations for the gas source described in Section 2.2.2).

A custom deck (i.e., a printed circuit board), named the MOX deck, was developed to interface two MOX gas sensors to the CF2 (Figure 3). The deck contains two sockets for 4-pin Taguchi-type (TGS) gas sensors, a temperature/humidity sensor (SHT25, Sensirion AG, Stäfa, Switzerland), a dual-channel digital potentiometer (AD5242BRUZ1M, Analog Devices, Norwood, MA, USA) and two MOSFET p-type transistors (NX2301P, NEXPERIA, Nimega, NL). We selected the TGS 8100 sensor (Figaro Engineering Inc., Osaka, Japan) due to its compatibility with 3.0 V logic, power consumption of only 15 mW (the lowest in the market as of June 2016) and miniaturized form factor (MEMS). Since the sensor heater uses 1.8V, we included two transistors (one per sensor) to reduce the applied power by means of pulse width modulation (PWM). The MOX read-out circuit (Figure 4) is a voltage divider connected to the µC's analog-to-digital converter (ADC). The voltage divider is powered at 3.0 V and the load resistor (R_L) can be set dynamically by the potentiometer (from 60 Ω to 1 MΩ in steps of 3.9 kΩ). In the current work, we only used one of the two sensors and R_L was fixed to 70 kΩ. This load resistor value is selected to operate the voltage divider near its mid-range, where the sensitivity is maximum (according to the expected concentrations for the gas source described in Section 2.2.2).

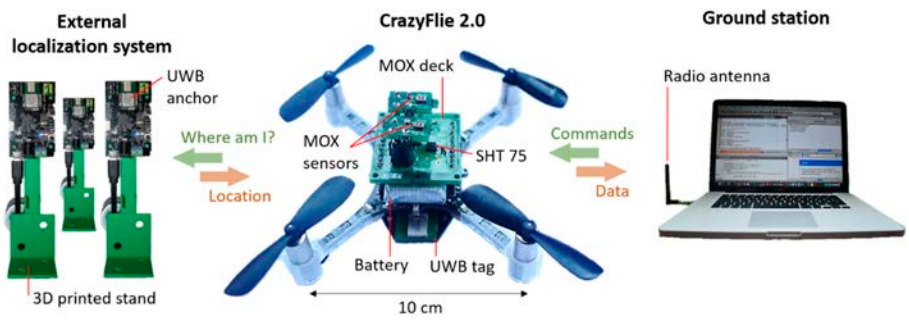


Figure 3. The CrazyFlie 2.0 equipped with the MOX deck and the UWB tag (**center**) gets 3D position from an external localization system composed of six ultra-wideband anchors (**left**). The location and sensor data are communicated to the ground station (**right**) over the 2.4 GHz ISM band.

The initialization task of the MOX deck driver configures the PWM, initializes the SHT25 sensor, sets the wiper position of both channels of the potentiometer and adds the MOX readout registers to the list of variables that are continuously logged and transmitted to the base station. The goal of PWM

the list of variables that are continuously logged and transmitted to the base station. The goal of PWM is to convert from 3.0 V to 1.8 V required by the MOX heater. For that, the PWM frequency is set to 8.4 KHz and the duty cycle (DC) to 36% according to

$$DC = \frac{P_{avg}}{P_{peak}} = \frac{V_{avg}^2/R}{V_{peak}^2/R} = \frac{1.8^2}{3.0^2} = 36\%, \quad (1)$$

where P_{avg} is the average power delivered to the sensor, P_{peak} is the peak power of the PWM signal and $P = V^2/R$ (Joule's first law combined with Ohm's law) is used to convert from power to voltage. The duty cycle is the fraction of time that the transistor delivers power to the sensor.

The main task of the deck driver reads the MOX sensor output voltage and the temperature/humidity values from the SHT25 and sends them to the ground station at 10 Hz.

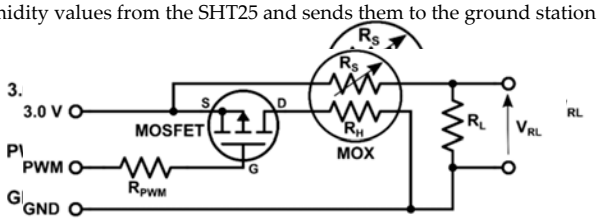


Figure 4. Schematic of the conditioning electronic circuit for each MOX sensor in the MOX deck, using PWM for powering and a voltage divider for read-out.

2.2. Experimental Arena, Gas Source and External Localization System

All experiments were performed in a large robotics laboratory (160 m^2 × 2.7 m height) at Örebro University (Sweden). The laboratory is divided into three connected areas (R1–R3) of 132 m^2 and a contiguous room (R4) of 28 m^2 (Figure 5). The ventilation system of the laboratory was not modified for the experiments and all windows and doors were kept closed.

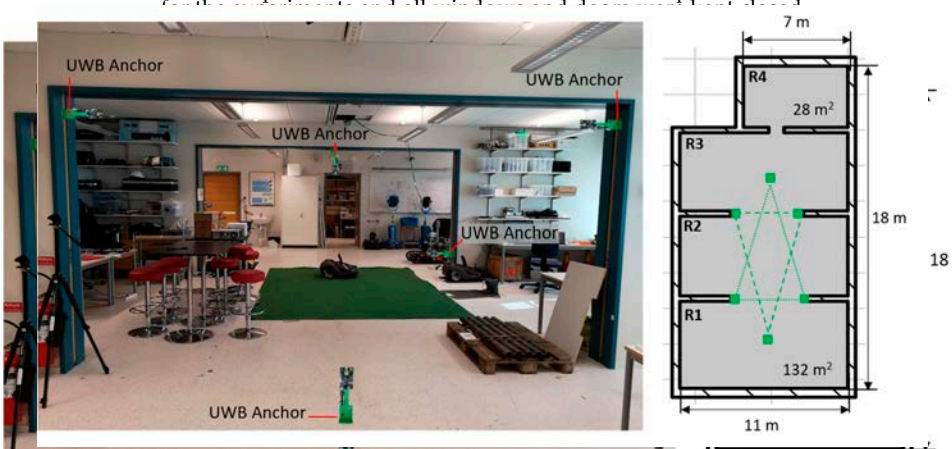


Figure 5. Experimental arena. (**left**) Frontal picture; (**right**) Schematic top view. The green squares indicate the position of the UWB anchors, which are positioned along two inverted triangles (green lines).

To obtain the 3D position of the drone, we used an external localization system (Loco positioning system, Bitcraze AB) [66] based on ultra-wide band (UWB) radio transmitters. The system is composed of six anchors that are positioned in known locations of the room and one tag that is fixed to the drone. The anchors were placed in the central area of the laboratory, shaped in two inverted triangles (below and above the flight area), as recommended by the manufacturer. The tag on the

continuously sends short high frequency radio messages to the anchors and estimates its relative position to them based on the timestamps of transmitted and received messages. The accuracy in the estimated position is approximately 10 cm if the tag is within the space delimited by the anchors and there is line of sight (LOS) between the anchors and the tag [67]. In the absence of these conditions, the system may still work but with degraded accuracy.

continuously sends short high frequency radio messages to the anchors and estimates its relative position to them based on the timestamps of transmitted and received messages. The accuracy in the estimated position is approximately 10 cm if the tag is within the space delimited by the anchors and there is line of sight (LOS) between the anchors and the tag [67]. In the absence of these conditions, the system may still work but with degraded accuracy.

A gas leak was emulated by placing a small beaker filled with 200 mL of ethanol 96% (Sigma–Aldrich, Germany) in different locations of the arena (Figure 6). Ethanol was used because it is non-toxic and easily detectable by MOX sensors. Three experiments were carried out to check the viability of the proposed system for GSL in complex environments. In the first experiment, the gas source was placed on top of a table (height = 1 m) in the small room (R4). In the second experiment, the source was placed inside the suspended ceiling (height = 2.7 m) near the entrance to the lab (R1). Since the piping system of the lab runs through the suspended ceiling, the gas source could represent a leak in one of the pipes. In these two experiments, a 12 V DC fan (Model: AD0612HB-A70GL; ADDA Corp., Taiwan) placed behind the beaker facilitated the dispersion of the chemicals in the environment, creating a plume. In the third experiment, the source was placed inside a power outlet box (height = 0.9 m) and a fish tank bubbler was used to increase the evaporation rate. The source in this location could simulate the early stages of an electrical fire (most of them are caused by faulty electrical outlets) where volatile organic compounds (VOCs) are released into the environment. The three experiments started five minutes after setting up the source and turning on the DC fan or the bubbler.

A gas leak was emulated by placing a small beaker filled with 200 mL of ethanol 96% (Sigma–Aldrich, Germany) in different locations of the arena (Figure 6). Ethanol was used because it is non-toxic and easily detectable by MOX sensors. Three experiments were carried out to check the viability of the proposed system for GSL in complex environments. In the first experiment, the gas source was placed on top of a table (height = 1 m) in the small room (R4). In the second experiment, the source was placed inside the suspended ceiling (height = 2.7 m) near the entrance to the lab (R1). Since the piping system of the lab runs through the suspended ceiling, the gas source could represent a leak in one of the pipes. In these two experiments, a 12 V DC fan (Model: AD0612HB-A70GL; ADDA Corp., Taiwan) placed behind the beaker facilitated the dispersion of the chemicals in the environment, creating a plume. In the third experiment, the source was placed inside a power outlet box (height = 0.9 m) and a fish tank bubbler was used to increase the evaporation rate. The source in this location could simulate the early stages of an electrical fire (most of them are caused by faulty electrical outlets) where volatile organic compounds (VOCs) are released into the environment. The three experiments started five minutes after setting up the source and turning on the DC fan or the bubbler.

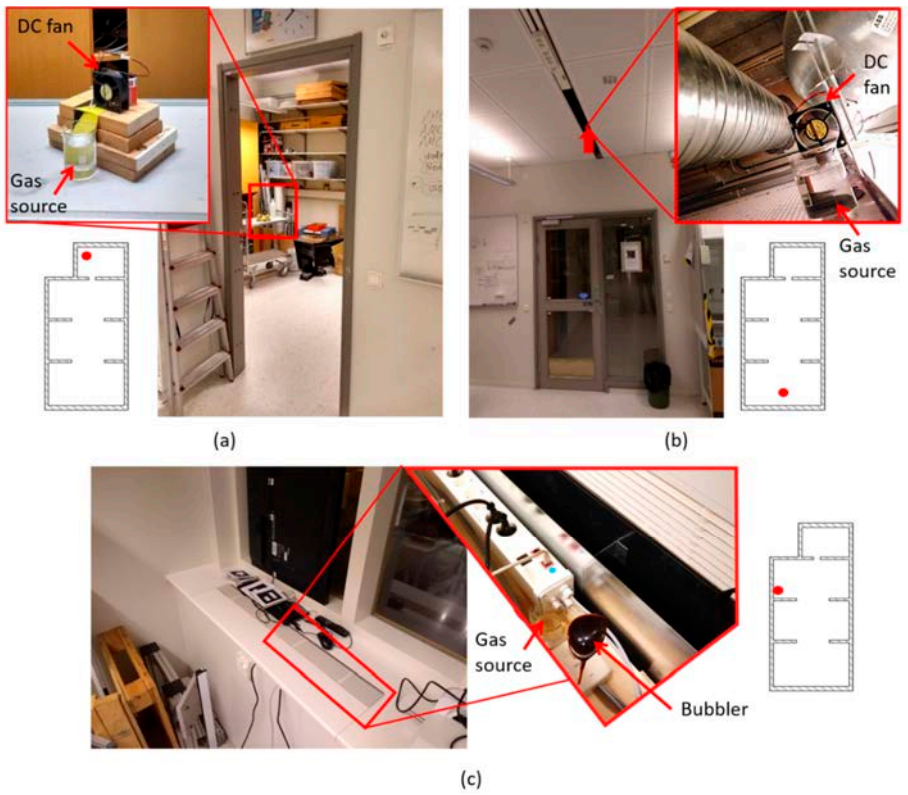


Figure 6. Gas source location in the three experiments. (**a**) Experiment 1: inside small room; (**b**) Experiment 2: hidden in suspended ceiling; (**c**) Experiment 3: hidden in a power outlet box.

Figure 6. Gas source location in the three experiments. (**a**) Experiment 1: inside small room; (**b**) Experiment 2: hidden in suspended ceiling; (**c**) Experiment 3: hidden in a power outlet box.

2.3. Gas Sensor Calibration and Limit of Detection (LOD) Estimation

2.3. Gas Sensor Calibration and Limit of Detection (LOD) Estimation

The MOX sensor was calibrated under laboratory conditions, to compensate the non-linear response and obtain measurements in concentration units. Ethanol concentrations up to 50 ppm were generated using the permeation method [68], humidified to 30 % r.h. and delivered in random order at 70 mL/min to the chamber containing the sensor under test. The uncertainty of the concentration reaching the gas chamber was determined by propagation of the main sources of error, namely the permeation rate, the oven temperature and the MFCs. Uncertainty values ranged from 50 ppb (at 1 ppm of ethanol) to 1 ppm (at 50 ppm of ethanol), which represents a relative uncertainty of 2–5%. The concentration range and humidity level for calibration samples were selected based on previous experience by the authors performing similar experiments in the same test environment.

To estimate the LOD (ppm), we used the simplified LOD formula [69]:

$$\mathrm{LOD} = \frac{3.3 \times s_0}{\hat{A}}, \tag{2}$$

where s_0 is the estimated standard deviation of blank measurements (assuming homoscedasticity and normality) and $\hat{A}$ is the estimated slope of the calibration graph (assuming linearity). To account for the expected variability between the calibration setup and the test environment [70], we estimated s_0 from a preliminary exploration of the target scenario in the absence of the gas. The LOD was used during the gas source localization experiments to remove false alarms, by setting to zero any measured concentration below the LOD.

2.4. Detection of 'Bouts'

To compute the 'bouts' from the MOX response, we adapted the signal processing pipeline proposed by Schmuker et al. [65]. The goal of this algorithm is to extract the rising edges of the smoothed derivative of the sensor response, which are called 'bouts'. Schmuker's algorithm is based on a non-causal Gaussian smoothing filter that prevents real-time bout detection, includes an unnecessary derivative, embeds the computation of the derivative within the smoothing filter (which may lead to potential implementation errors) and does not smooth the second derivative of the signal where the 'bouts' are segmented. The derivative that converts s to x in [65] is unnecessary because the ema_α transformation already differentiates the input signal. The source code published by Schmuker et al. is not affected by this error because the transformation is implemented by calling the Python function *pandas.ewma()*, which provides the functionality of an EWMA filter (i.e. it does not perform the derivative). We replaced the Gaussian smoothing filter with a causal (realizable) exponentially weighted moving average (EWMA) filter, removed the unnecessary derivative, decoupled the derivative from the EWMA filter and smoothed the second derivative. We chose the EWMA filter because it is causal, easy to implement and is the same filter used by Schmuker et al. to smooth the derivative in their bout computation algorithm. The proposed bout computation pipeline is represented in Figure 7.

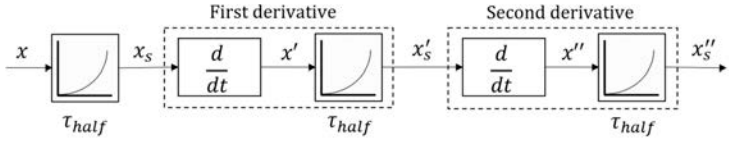


Figure 7. Flowdiagram of the improved bout computation. The meaning of each symbol is given in the text.

The sensor response x is first smoothed using a EWMA low pass filter to remove high-frequency noise. At time t, the smoothed value $x_s(t)$ is found by computing

$$x_s(t) = (1-\alpha) \cdot x_s(t-1) + \alpha \cdot x(t), \tag{3}$$

where $x(t)$ is the observation at time t, $x_s(t-1)$ is the previous output of the filter and the smoothing factor α $(0 < \alpha \leq 1)$ controls the speed at which older responses are dampened. The

where $x(t)$ is the observation at time t, $x_s(t-1)$ is the previous output of the filter and the smoothing factor α ($0 < \alpha \leq 1$) controls the speed at which older responses are dampened. The smoothing factor α in the EWMA filter is equivalent to the cut-off frequency of a low-pass filter. For example, Pashami et al. [71] found out that $\alpha = 0.9$ represents a cut-off frequency of 0.44 Hz. It is convenient to express the smoothing factor as a function of the half-life time τ_{half} (s)

$$\alpha = 1 - \exp\left(\frac{\log 0.5}{\tau_{half} \cdot f_s}\right), \tag{4}$$

where f_s is the sampling frequency (Hz) of x. The smoothed response is differentiated and smoothed again to increase the signal-to-noise ratio (SNR), producing x'_s. The 'bouts' are the rising edges of x'_s, which are delimited by two consecutive zero-crossings in the positive derivative of x'_s, i.e., $x''_s > 0$. The amplitude of a bout is defined as x'_s at the end of the respective bout segment minus x'_s at the start of the same bout segment. To remove low-amplitude bouts produced by noise, Schmuker et al. propose to estimate the noise threshold (b_{thr}) using the 3-sigma rule

$$b_{thr} = \mu + 3\sigma, \tag{5}$$

where μ and σ are the estimated mean and standard deviation, respectively, of the distribution of amplitudes of bouts detected in the sensor baseline (i.e., in the absence of gas). Bouts with amplitude lower than b_{thr} are filtered out. The algorithm that we propose reduces the number of parameters from three to two (τ_{half} and b_{thr}), which we estimate after a preliminary exploration of the target scenario in the absence of gas (i.e., using the signals corresponding to blank measurements).

2.5. Effect of Rotors on MOX Sensor Signals

Previous work using the CF2 indicate that turbulence generated by the propellers may severely affect the MOX sensor signals [29]. To evaluate this effect in our platform, we performed a set of measurements near a gas source under two conditions: rotors switched on or rotors switched off. The drone was placed on a height-adjustable stand—designed to minimize its interference with the rotors' airflow—that could be moved around the source (Figure 8a). Using a stand is necessary to perform measurements when the rotors are switched off, and we also used it to perform measurements with the rotors switched on. The gas source was an open ethanol bottle (Figure 8b) and measurements were performed above the source (vertical distance between 25 and 65 cm) and in front of it (at 50 cm). At each location, the sensor response was recorded for 25 min (using an external battery), first with the rotors switched off and, after cleaning the room, measurements were repeated with the four rotors spinning at 10,000 rpm (this is typical for hovering).

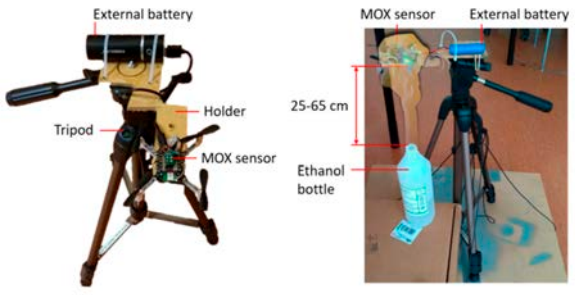


Figure 8. Setup for assessing the effect of the rotors on the MOX sensor signals. (**a**) Top view of the stand used to hold the drone at different heights while minimizing interference with the rotors air flow; (**b**) Photo of an experiment with the drone placed 25 cm above an ethanol bottle (gas source), overlaid with an illustration of a gas cloud.

2.6. Gas Source Localization Strategies

Two gas source localization strategies, one based on the instantaneous response and the other one based on odor hits are evaluated using the nano-drone. In both cases, the drone was sent to fly along a predefined sweeping path consisting of two 2D rectangular sweepings at different heights (0.9 m and 1.8 m), collecting measurements in motion (Figure 9). These two heights divide the vertical space of the lab in three parts of equal size. Flying first at a lower altitude minimizes the impact of the propellers' downwash in the gas distribution. For safety reasons, the trajectory is designed to ensure enough clearance around obstacles and walls, and people working inside the laboratory were told to remain in their seats during the experiments. The ground station communicates the flight path to the drone as a sequence of (x, y, z) waypoints, with a target flight speed of 1.0 m/s.

As the drone navigates the environment, it reports the instantaneous concentration and its location to the ground station. At the end of the exploration, the ground station uses all the received information to compute a 3D map of the instantaneous response (first strategy) and the bout frequency (second strategy). The location of the gas source is estimated in both cases as the cell of the map with maximum value. We will also discuss the viability of both methods for real-time plume tracking, assuming that the drone would follow the gradient of instantaneous concentration or the gradient of bout frequency (computed using a sliding window of 5 s).

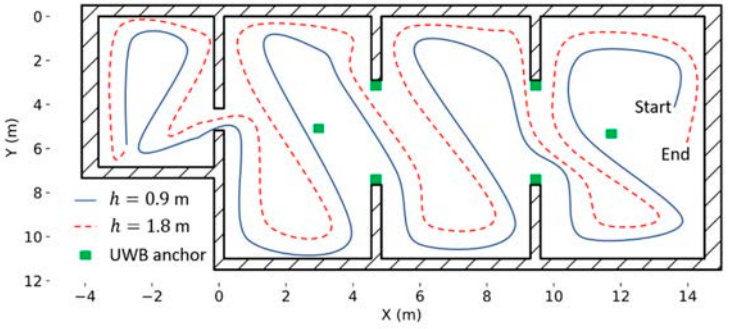


Figure 9. Predefined navigation strategy based on zig-zag sweeping at two heights (0.9 and 1.8 m). The green squares indicate the location of the UWB anchors.

3. Results

3.1. Calibration, LOD and Optimum Parameters for Bout Detection

A preliminary exploration of the test environment in the absence of gas was performed to estimate the sensor noise and the optimum bout parameters (Figure 10A). The raw response was smoothed using an EWMA filter with a smoothing factor $\tau_{half} = 0.25$ s. The noise can be approximated by a Gaussian distribution with mean value of 2.047 $M\Omega^{-1}$ and standard deviation of 0.013 $M\Omega^{-1}$ (Figure 10B). The observed variability was used in combination with the calibration line (Figure 11A) to estimate the LOD (Equation (2)), assuming homoscedasticity. The calibration line behaved linearly in the range 1–50 ppm after applying the logarithm to both concentration and response. The bout amplitude threshold (b_{thr}) was estimated as 0.04 ppm/s by applying Equation (5) to the amplitude of the bouts detected in the calibrated blank signals (Figure 11B).

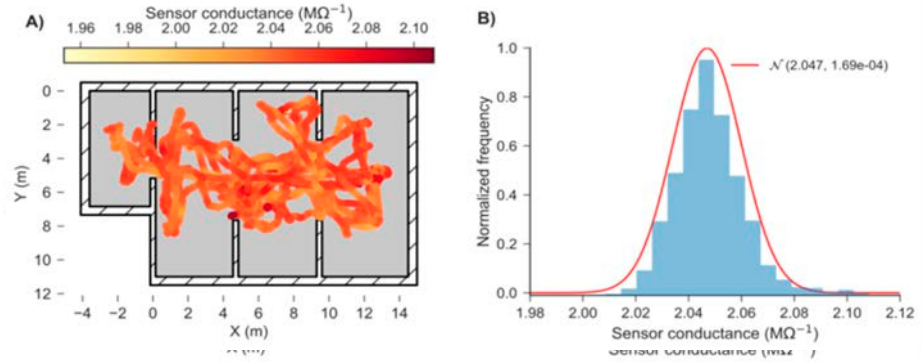


Figure 10. (**A**) 2D map of MOX sensor response during 15 minutes of random exploration of the target area without gas; (**B**) Histogram of blank readings, with a Gaussian curve $N(\mu, \sigma^2)$ superimposed.

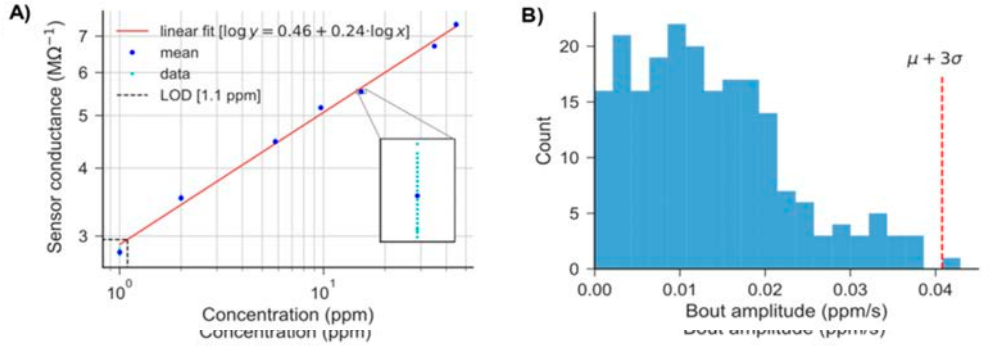


Figure 11. (**A**) Calibration line in the range 1–50 ppm (log-log plot), with blank variability superimposed at each concentration level (see inset). The LOD is estimated using Equation (2); (**B**) Histogram of amplitudes of bouts detected in the calibrated blank signals. b_{thr} (Equation (5)) is indicated by a red dashed vertical line.

3.2. Effect of Propulsion on MOX Signals

When the propellers were switched off (Figure 12A), the concentration fluctuations due to the gas evaporating from the ethanol bottle were clearly reflected on the on-board sensor signals at 25 cm above the source (green trace) and, to less extent, at 50 cm in front of the source and (blue trace). At 65 cm above the source (yellow trace), the sensor does not seem to detect the source except after opening the bottle ($t = 2$ min). When the same measurements were repeated with the propellers switched on (Figure 12B), the fluctuations of the signals at 25 cm above the source became less intense but more frequent. This can be better observed in Figure 13, where the bouts detected with the propellers switched on are as twice as frequent than when the propellers are switched off (Table 1). This effect is even more noticeable at 65 cm above the source because the bout frequency increases from 0.48 to 7.74 bouts/min by switching on the propellers, improving the detection of the source. The propellers produced nonetheless a negative effect in the signals acquired in front of the source, reducing the bout frequency by a factor of two.

When the propellers were switched off (Figure 12A), the concentration fluctuations due to the gas evaporating from the ethanol bottle were clearly reflected on the on-board sensor signals at 25 cm above the source (green trace) and, to less extent, at 50 cm in front of the source and (blue trace). At 65 cm above the source (yellow trace), the sensor does not seem to detect the source except after opening the bottle ($t = 2$ min). When the same measurements were repeated with the propellers switched on (Figure 12B), the fluctuations of the signals at 25 cm above the source became less intense but more frequent. This can be better observed in Figure 13, where the bouts detected with the propellers switched on are as twice as frequent than when the propellers are switched off (Table 1). This effect is even more noticeable at 65 cm above the source because the bout frequency increases from 0.48 to 7.74 bouts/min by switching on the propellers, improving the detection of the source. The propellers produced nonetheless a negative effect in the signals acquired in front of the source, reducing the bout frequency by a factor of two.

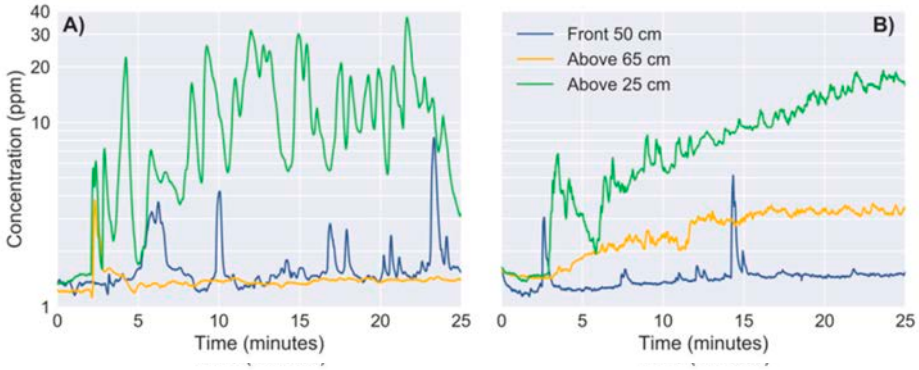


Figure 12. Sensor signals (log scale) near an evaporating source. (**A**) Propellers switched off; (**B**) Propellers switched on. The ethanol bottle is opened at $t = 2$ min.

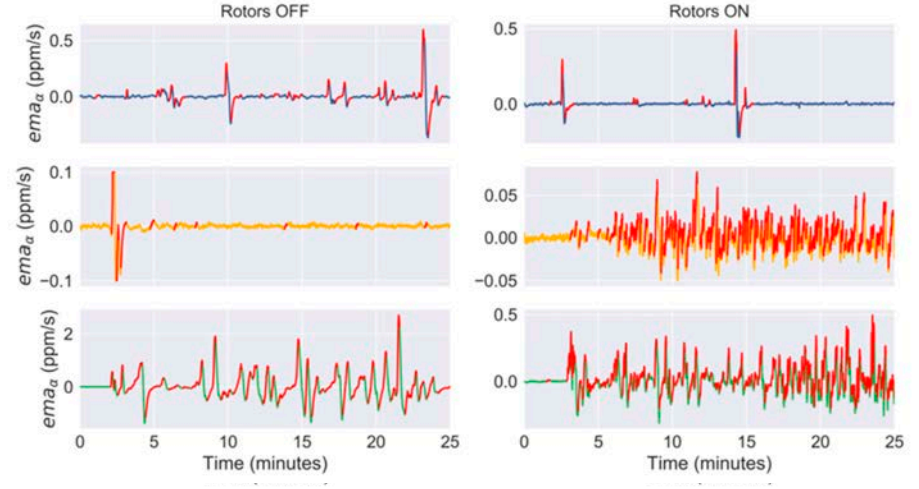


Figure 13. Smoothed derivative (i.e., x' in Figure 7) of the sensor signals at 50 cm in front of the source (blue line), 65 cm above the source (yellow line) and 25 cm above the source (green line). Bouts with amplitude higher than $\mu + 3\sigma$ are highlighted in red. In the left column, the propellers are switched off whereas in the right column they are switched on. The ethanol bottle is opened at $t = 2$ min.

Table 1. Characterization of MOX signals at different distances of the source under two conditions: propellers switched on or off.

Distance	Propellers	Mean (ppm)	Variance (ppm^2)	Bout Frequency (Bouts/min)	Bout Amplitude (ppm/s)
Above 25 cm	OFF	10.05	60.46	2.52	0.89
	ON	9.22	29.97	7.69	0.084
Above 65 cm	OFF	1.39	0.053	1.48	0.027
	ON	2.67	0.53	4.74	0.015
Front 50 cm	OFF	1.68	0.59	1.13	0.10
	ON	1.45	0.12	0.47	0.101

The steady increase of the average concentration and higher bout frequency observed when the drone was sampling above the source with the rotors switched on can be explained based on how the propellers interact with the gas surrounding the drone. The propellers generate a downwash (i.e., a downward airflow) that acts as an opposing force to the gas moving upwards by convection,

breaking the laminar flow into a turbulent gas cloud that spreads around the drone's frame (Figure 14A), probably increasing the background concentration. When the gas is spread around the drone, the propellers drag the gas patches towards the MOX sensor (Figure 14B) and this probably increases the bout frequency. Therefore, from this study we can conclude that the propellers severely change the gas distribution near the drone, which in turn affects the MOX signals. However, it may still be possible to extract relevant features for gas source localization (e.g., the bout frequency).

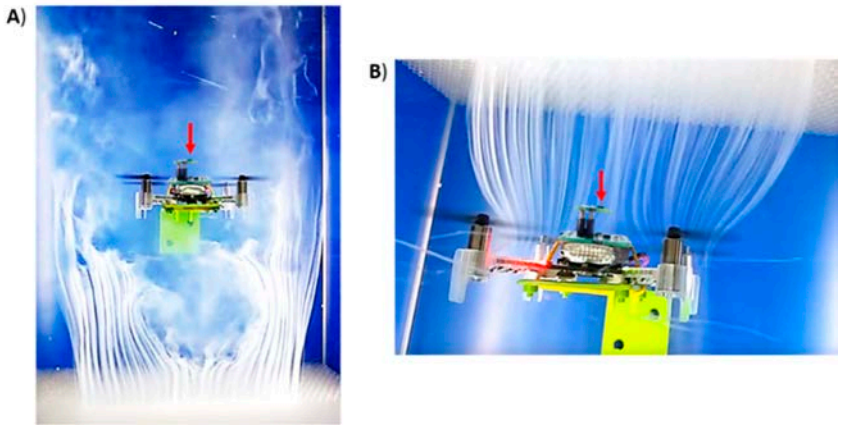


Figure 14. Aerodynamics of Crazyflie 2.0 when the four rotors are spinning, visualized using a Dekabreeze wind tunnel (Courtesy of Bitcraze AB). The drone is fixed to one of the walls of the tunnel using a 3D printed stand and dry ice fog is emitted from (**A**) below the drone or (**B**) above the drone. It shows the downwash of the propellers and how part of the fog reaches the MOX gas sensor (red arrow). The MOX deck has been overlaid to the original images for visual clarity.

3.3. Experiment 1: Localization of a Source 17 m Away from the Starting Point

In the first experiment, the drone took off near the entrance of the lab (R1), 17 meters downwind of a gas source located in the other end of the laboratory (R4). Flying along a predefined trajectory that took 2.75 min, the drone acquired measurements to build a map of the instantaneous response (Figure 15A). From this map it is evident that the gas source must be in R4, because the maximum concentration (35 ppm) was found there while concentrations below 5 ppm were measured in the rest of the lab. The gas plume can be outlined from the obtained bout map, which shows the highest bout density (25 bouts/min) also in R4. Most bouts were detected during the first part of the exploration (Figure 15C), when the drone was flying at the same height of the source (Figure 15B). Nonetheless, multiple bouts were also detected flying above the source in R4 ($t \approx 110$–120 s). The cells corresponding to the maximum concentration and maximum bout frequency were found at 0.94 and 1.16 m of the true source location, respectively.

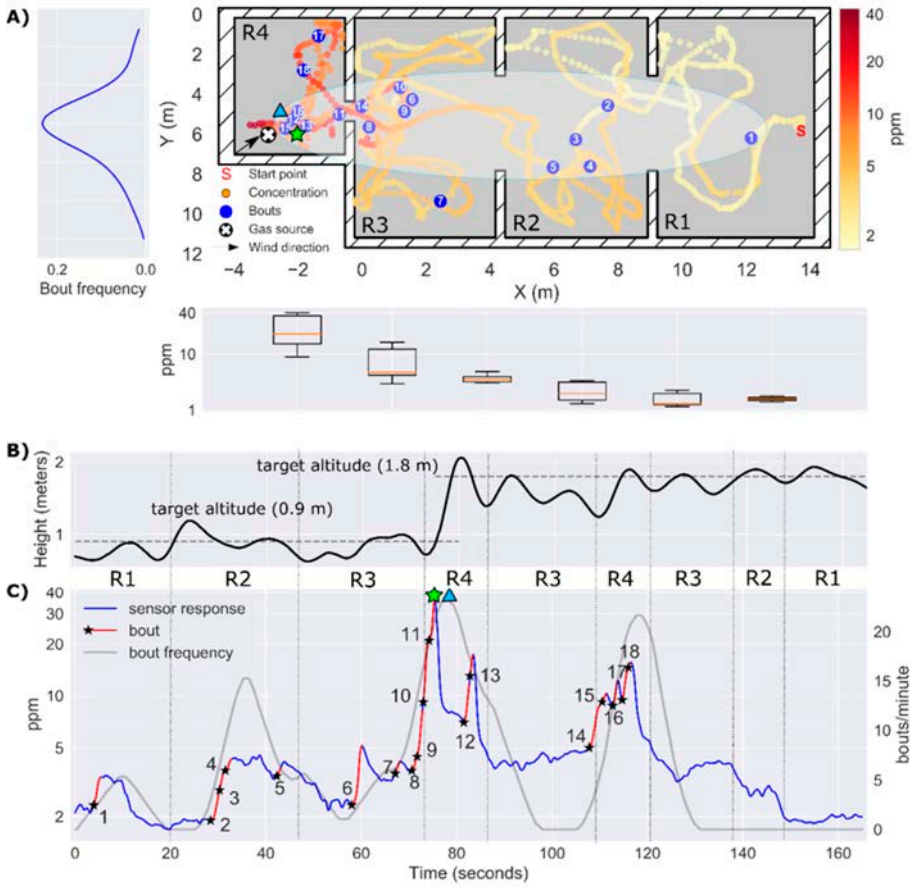


Figure 15. Results of Experiment 1. (**A**) 2D map of the instantaneous concentration (ppm) in log scale, with bouts represented by blue circles (b_{thr} = 0.52 ppm/s). A hand-drawn ellipse outlines the approximate plume shape based on the location of bouts. The average bout frequency along the y-axis is shown in the leftmost panel. The box plot below the map represents the instantaneous concentration along the x-axis; (**B**) Trajectory of the drone along the z-axis. (**C**) Temporal evolution of the instantaneous concentration (ppm) on a log scale, with detected bouts highlighted in red (the black star indicates the start of a bout). The identifiers R1–R4 between panels (**B**) and (**C**) indicate the area of the map in which the drone is flying at each moment. The maximum instantaneous concentration and the maximum bout frequency are indicated by a green star and a blue triangle, respectively.

For real-time plume tracking, following the gradient of instantaneous concentration would not easily lead the drone towards the source, since the concentrations measured in R1 are only slightly higher than the LOD of the sensor (1.1 ppm). Although it is true that the instantaneous concentration increases when the drone crosses the plume (e.g., bouts #1, #2 and #6), it is not clear when the drone exits the plume due to the increase in background concentration and slow sensor recovery. In this sense, the bout frequency behaves like a more sensitive version of the instant response, exhibiting abrupt changes when the drone enters/exits the plume. This can be seen at t = 85 s, when the drone exits R4 and enters R3. The bout frequency sharply drops to zero, but the instantaneous concentration remains flat at 5 ppm. Another example is found by comparing both features when the drone crosses the door that connects R3 to R4 flying at h = 0.9 m (t = 75 s) or at h = 1.8 m (t = 110 s). In the first case, the instantaneous concentration jumps from 4 to 35 ppm whereas in the latter case the step is only from 5 to 10 ppm. On the other hand, the bout frequency shows a similar increase (0 to 23–25

5 to 10 ppm. On the other hand, the bout frequency shows a similar increase (0 to 23–25 bouts/min) in both cases. This could mean that the bout frequency, which simply counts events regardless of their amplitude is more robust for detecting the source when its height is unknown.

To obtain the bout map shown in Figure 15A, we had to increase the bout threshold (b_{thr}) from 0.04 ppm/s (computed using Equation (5)) to 0.52 ppm/s (determined from visual inspection). The effect of the bout amplitude threshold on the bout map is show in Figure 16. If the threshold is set too low (blue circles), slight variations of the background concentration can trigger bout detections that are not necessarily produced by the gas plume, resulting in a sparse distribution of bouts that hinders the localization of the source (localization error of 4.32 m). This variability in the background concentration may be caused by the downwash of the drone, which breaks the plume and disperses fragments of the plume in the room. On the other hand, if the threshold is set too high (green circles) it is virtually guaranteed that every detected bout is produced by the plume. However, in this case, bouts appear only near the source, which reduces the distance at which the source can be detected. Only when the threshold is suitably estimated (Figure 15A), the bout map resembles an elongated plume that can be used to move towards the source from a long distance. As of now, we are not aware of a systematic approach to properly set this threshold without prior information about the source intensity. One idea could be to use a multi-threshold approach, i.e., building bout maps with different thresholds and searching for consistent source predictions over multiple threshold values.

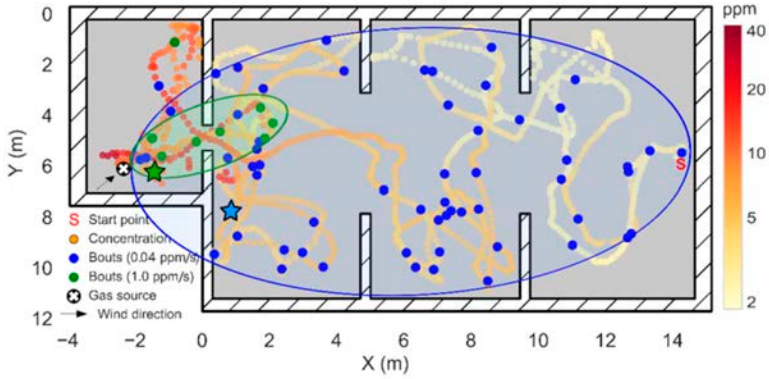


Figure 16. Effect of the bout amplitude threshold in the results of Experiment 1. The blue circles represent bouts with amplitude higher than 0.04 ppm/s ($\mu + 3\sigma$ threshold), and the green circles represent bouts with amplitude higher than 0.10 ppm/s. In each case, a hand-drawn ellipse outlines the approximate plume shape based on the location of the bouts. The green and blue stars indicate the source location estimate in each case, according to the maximum bout frequency.

3.4. Experiment 2: Localization of a Source Hidden in the Suspended Ceiling ($h = 2.7$ m)

In this experiment, the source was located just above the starting point of the exploration, hidden in the suspended ceiling (Figure 17). The resulting maximum concentration in the test room was measured when the drone flew at $h = 1.8$ m, highlighting the importance of sampling in 3D for localization and mapping of elevated gas sources. However, since the source is presumably not directly exposed to the environment, concentrations below 3 ppm were found in most locations of the room, which complicates the GSL task. The $\mu + 3\sigma$ threshold yielded better results than in the previous experiment, outlining a feasible plume shape according to the true gas source location and predominant wind direction (Figure 18A). The instantaneous response map and the bouts map suggest that the gas source is in the division between R1 and R2, which represents a localization error of 4.0 and 3.31 m, respectively. The localization error can be reduced to 2.22 m by increasing the bout threshold from 0.04 ppm/s to 0.18 ppm/s (Figure 19). In this case, the lower inter-bout interval between bouts #3 and #4 (as compared to bouts #1 and #2) leads to a higher bout frequency near the source.

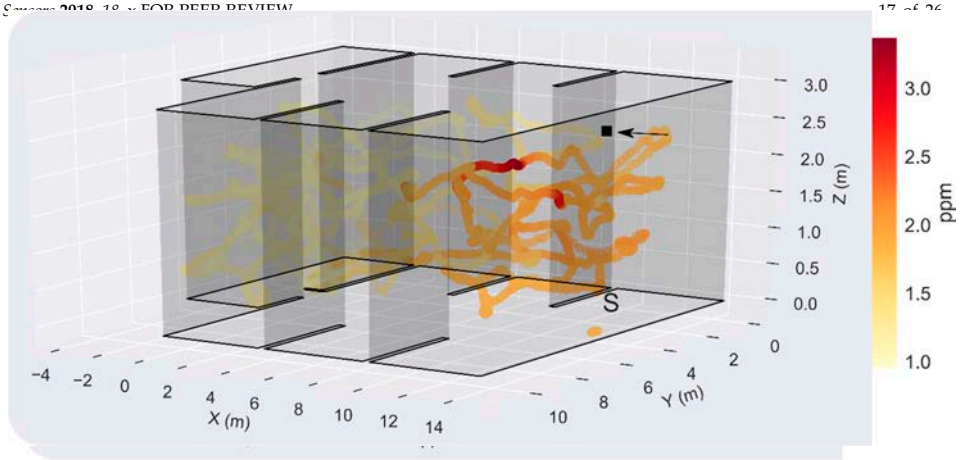


Figure 17. 3D map of the instantaneous concentration (ppm) in Experiment 2. The black square indicates the gas source location $(x,y,z) = (14.0, 5.2, 2.7)$ m, the black arrow the wind direction (positive x-axis) and the letter 'S' the starting point of the drone $(x,y,z) = (13.5, 5.2, 0.0)$ m.

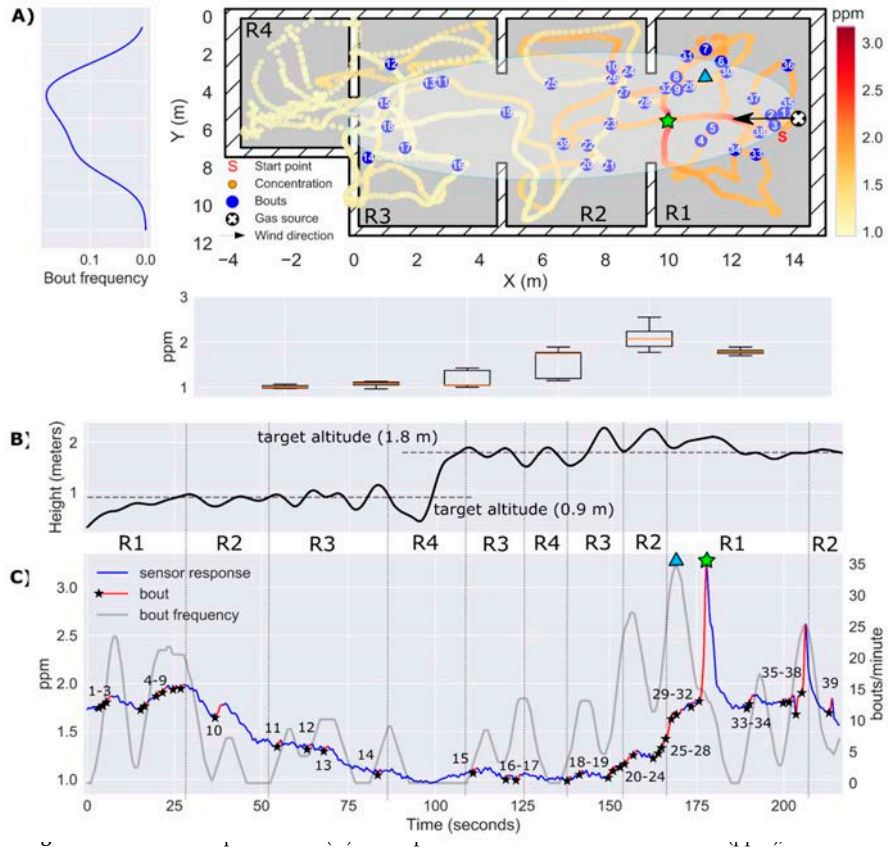


Figure 18. Results of Experiment 2: (**A**) 2D map of the instantaneous concentration (ppm), with odor hits represented by blue circles (b_{thr} = 0.04 ppm/s). A hand-drawn ellipse outlines the approximate plume shape based on the location of bouts. The average bout frequency along the x-axis is shown in the panel on the left. The box plots below the map represents the instantaneous concentration along the x-axis; (**B**) Drone trajectory in the z-axis. (**C**) Temporal evolution of the instantaneous

concentration (ppm), with detected bouts highlighted in red (the black star indicates the start of the bout). The bout frequency (gray line) is computed using a sliding window of 5 s. The identifiers R1–R4 between panels (**B**) and (**C**) indicate the area of the map in which the drone is flying at each moment. The maximum instantaneous concentration and the maximum bout frequency are indicated by a green star and a blue triangle, respectively.

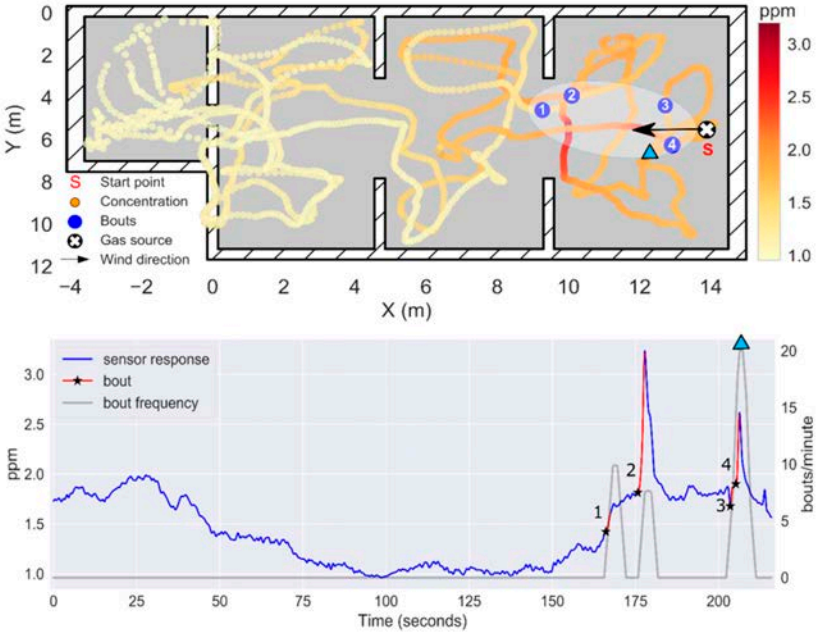


Figure 19. Results of Experiment 2 when b_{thr} is increased to 0.18 ppm/s. (**top**) 2D map of the instantaneous concentration (ppm), with odor hits represented by blue circles. A hand-drawn ellipse outlines the approximate plume shape based on the location of bouts. (**bottom**) Temporal evolution of the instantaneous concentration (ppm), with detected bouts highlighted in red (the black star indicates the start of the bout). The bout frequency (gray line) is computed using a sliding window of 5 s. The maximum bout frequency is indicated by a blue triangle.

3.5. Experiment 3: Localization of a Source Hidden Inside a Power Outlet Box (h = 0.9 m)

In this experiment, the gas source was placed inside a power outlet box and a bubbler was used to increase the release rate (no artificial airflow). The gas distribution map (Figure 20A) shows a local accumulation of gas in R3 near the source (max concentration = 8 ppm), whereas in the rest of the map the gas concentration is below 1 ppm. The bout map does not resemble the elongated plume observed in previous experiments probably due to the absence of induced airflow. During the first part of the exploration at $h = 0.9$ m, the measured concentration was constantly below the LOD and no bouts were detected (Figure 20C), which represents a clear challenge for a robot attempting to reactively track the plume. Only when the drone approached the source ($d < 3$ m) the first bouts were detected (bouts #1–3). The remaining bouts and the maximum concentration were found during the second pass at $h = 1.8$ m, which may indicate a buoyant gas dispersion due to convective air currents produced by heat inside the power outlet box. The location of the source in the instantaneous concentration map and using the bout frequency were very accurate in the x-y plane (error of 0.7 m), but the bout frequency yielded a lower localization error (0.77 m) than the instantaneous response (1.22 m) because the estimated height of the source in the first case (1.2 m) was closer to the true value (0.9 m) than in the second case (1.7 m).

In this experiment, the gas source was placed inside a power outlet box and a bubbler was used to increase the release rate (no artificial airflow). The gas distribution map (Figure 20A) shows a local accumulation of gas in R3 near the source (max concentration = 8 ppm), whereas in the rest of the map the gas concentration is below 1 ppm. The bout map does not resemble the elongated plume observed in previous experiments probably due to the absence of induced airflow. During the first part of the exploration at $h = 0.9$ m, the measured concentration was constantly below the LOD and no bouts were detected (Figure 20C), which represents a clear challenge for a robot attempting to reactively track the plume. Only when the drone approached the source ($d < 3$ m) the first bouts were detected (bouts #1–3). The remaining bouts and the maximum concentration were found during the second pass at $h = 1.8$ m, which may indicate a buoyant gas dispersion due to convective air currents produced by heat inside the power outlet box. The location of the source in the instantaneous concentration map and using the bout frequency were very accurate in the x-y plane (error of 0.7 m), but the bout frequency yielded a lower localization error (0.77 m) than the instantaneous response (1.22 m) because the estimated height of the source in the first case (1.2 m) was closer to the true value (0.9 m) than in the second case (1.7 m).

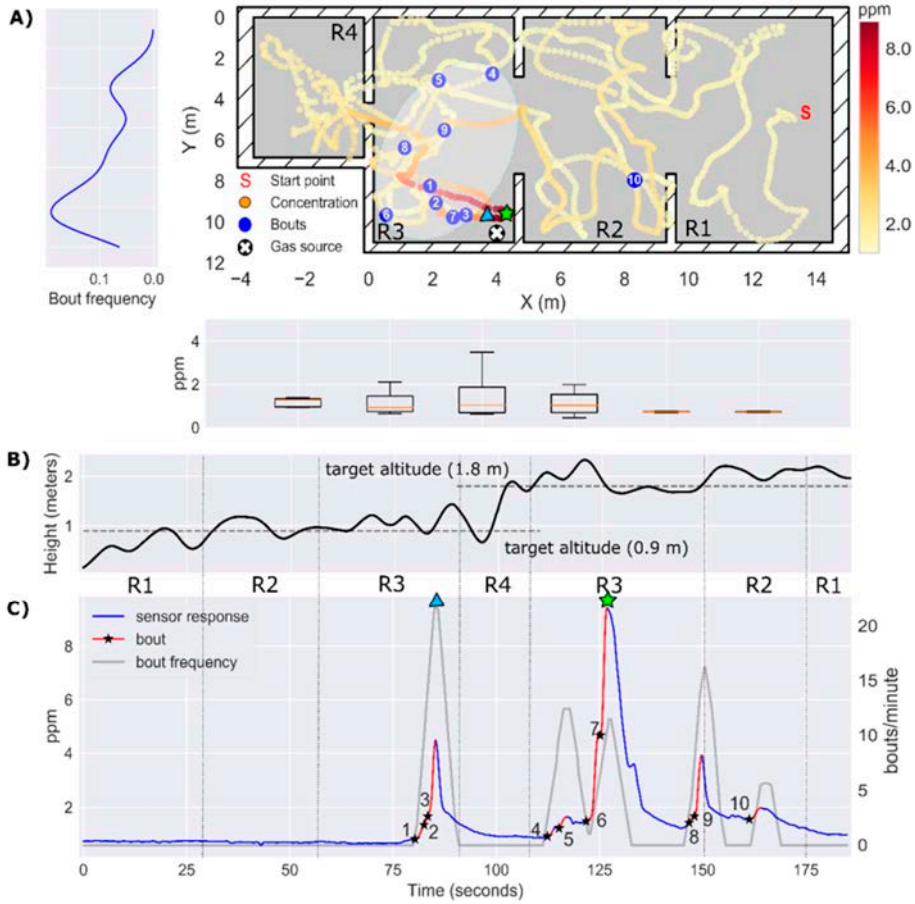


Figure 20. Results of Experiment 3. (**A**) 2D map of the instantaneous concentration (ppm), with bouts represented by blue circles (b_{thr} = 0.20 ppm/s). A hand-drawn ellipse outlines the approximate plume shape based on the location of bouts. The average bout frequency along the y-axis is shown in the leftmost panel. The box plot below the map represents the instantaneous concentration along the x-axis; (**B**) Drone trajectory in the z-axis. (**C**) Temporal evolution of the instantaneous concentration (ppm), with detected bouts highlighted in red (the black star indicates the start of the bout). The bout frequency (gray line) is computed using a sliding window of 5 s. The identifiers R1–R4 between panels (**B**) and (**C**) indicate the area of the map in which the drone is flying at each moment. The maximum of the instantaneous concentration and the bout frequency are indicated by a green star and a blue triangle, respectively.

3.6. Overall Localization Results

A summary of the localization results is given in Table 2. The bout frequency with optimum noise threshold achieved comparable results to the instantaneous concentration, except in Experiment 2 where the former achieved better performance (2.2 m versus 4.0 m error). The bout frequency ($\mu + 3\sigma$ threshold) produced relatively high localization errors.

Table 2. Gas source localization error (m) in the three experiments, using the instantaneous concentration, the bout frequency with $\mu + 3\sigma$ threshold or the bout frequency with optimum threshold.

Experiment	Instantaneous	Bout Frequency	Bout Frequency

Table 2. Gas source localization error (m) in the three experiments, using the instantaneous concentration, the bout frequency with $\mu + 3\sigma$ threshold or the bout frequency with optimum threshold.

Experiment	Instantaneous Concentration	Bout Frequency (μ+3σ)	Bout Frequency (Optimum Threshold)
1	0.94	4.32	1.16
2	4.0	3.31	2.22
3	1.22	5.07	0.77
Mean	2.05	4.23	1.38

4. Discussion

Our results suggest that a gas-sensitive NAV can be used for gas source localization and mapping in large indoor environments. The two GSL strategies compared in this paper require either building a map of the gas distribution or a map of the bouts. In contrast to previous works in which long measurement times are needed at predefined or adaptively chosen sampling locations, we have demonstrated that a rough approximation of such maps can be obtained in very short time with concentration measurements acquired in motion. Both maps seem coherent with respect to the true source location and wind direction, and not only enable the detection of the source with relatively small localization errors but also provide a rich visual interpretation of the gas distribution, especially if the bouts are overlaid on the gas distribution map. We adapted the bout detection algorithm for real-time operation and optimized the noise threshold, discussing how it affects the results in each experiment. The instantaneous gas distribution provides a more intuitive representation of the gas source location whereas the bout map can outline the gas plume without requiring anemometry.

The experiments presented in this paper also demonstrate that the air flow generated by the propellers greatly affects the gas sensor signals, but key information for GSL can still be extracted by signal processing methods. This appears to contradict the results reported by Rossi et al. [29] who claimed that the signals of a MOX sensor on board of a Crazyflie 2.0 are useless if the propellers are switched on. It is difficult to reason about the causes of this discrepancy because Rossi et al. do not show nor describe the experimental arena in which their experiments were performed, including the type of gas source and its location relative to the drone position. If they attempted to measure the gas coming out from an evaporating chemical source placed right below the drone, we have seen in this work that the downwash of the drone prevents most of the gas reaching the sensor surface and this produces raw signals with much less fluctuations than they would have if the same measurements were carried out with the propellers turned off (see green trace in Figure 12). We have also seen that such characteristic fluctuations can be still extracted from the smoothed derivative of the response (Figure 13). Another possible reason for this difference may be that different MOX sensors were used (MiCS 5525 for CO detection in [29] and Figaro TGS 8100 in this work) or they were operated differently. While we operate the sensor in continuous mode, Rossi et al. apply duty-cycling to the sensor heater to reduce the power consumption, which is known to degrade the signal quality [54]. Other reasons may be the signal processing carried out, the test gas concentrations or the range of variation of environmental conditions. If the gas concentrations are close to the LOD of the sensor, any turbulence created by the propellers or slight changes in ambient temperature or humidity may hinder distinguishing the response to the gas from the noise [72].

We dedicated a substantial portion of the effort to calibrate the MOX sensor in similar conditions to the test scenario and to find a suitable linear calibration model that enables the computation of the LOD using the standard equations. The log-log transformation linearized the MOX response in the range 0–50 ppm and the LOD was estimated as 1.1 ppm. This is something not usually done in the field of mobile robot olfaction since calibration of gas sensors is time-consuming and requires expensive dedicated equipment. The sensor signals are often uncalibrated [73–76] or only scaled to unit range [77–80], which produces maps that represent the sensor response instead of the absolute concentration. Response maps are not easily interpretable due to the non-linear sensor response

and are sensor-dependent; i.e., two response maps obtained with different gas sensors may show notable differences even if the underlying gas distribution is similar. Scaling is a linear operation that cannot remove non-linearities of the original signals. Some authors assume a linear response within a certain concentration range [75,76]; however, for most gas sensing technologies (e.g., MOX, gasFETs or thermoelectric sensors) this is only true within a narrow range of concentrations (e.g., 0–10 ppm) often exceeded by the typical concentrations obtained in real experiments.

Most GSL algorithms require a detection threshold that is used to differentiate when a given sensor reading is produced by the gas plume or by background noise. Algorithms based on 'odor hits' require such a threshold to binarize the sensor response or to filter out low-amplitude bouts produced by noise. The rapid decay in chemical concentration with increasing distance from the source combined with incorrectly estimated detection thresholds may lead to a reduced area where the plume can effectively be detected (if the threshold is too high) or a high number of false alarms (if the threshold is too low). This has not been reported in the literature as an issue in GSL experiments because the concentration was often unrealistically high and the arena too small, i.e., the robot was very close to the source during the whole experiment. In this work, we challenged the drone by placing the source up to 17 m from the starting point or hiding it in the suspended ceiling or inside a power outlet box, leading to concentrations near the LOD of the sensor. It seems that for proper estimation of the noise threshold it is critical to have some prior information regarding the source intensity. A threshold computed exclusively from measurements in the absence of gas (e.g., $\mu + 3\sigma$ threshold) may only work when the concentrations in the room are small. If the background concentration in the room increases, for example due to gas dispersed by the propellers of the drone, the threshold must be increased to reduce false bout detections. This suggests a multi-threshold approach, i.e., building bout maps with different thresholds and searching for consistent source predictions over multiple threshold values.

An obvious future direction for nano-drones is to achieve autonomous navigation, i.e., without resorting to external localization systems. This requires three-dimensional obstacle detection and self-localization, which is not an easy task whatsoever because there are many degrees of freedom and the low payload does often not allow for range sensors. In this line, the Fast Lightweight Autonomy (FLA) program (https://www.darpa.mil/program/fast-lightweight-autonomy) from the Defense Advanced Research Projects Agency (DARPA) is exploring novel perception methods and algorithms for small autonomous UAVs to fly at speeds up to 20 m/s in cluttered environments with no remote pilot, no communication links, no GPS guidance and no pre-programmed map of the area. In phase 2 of the program (June 2018), successful tests were performed with drones weighing around 2 kg equipped with a single camera, flying through a narrow window into a building, searching rooms, creating a 3-D map of the interior and exiting the building through an open doorway. In this paper we have demonstrated that a gas sensor could complement the camera in this application, adding a gas distribution layer to the acquired map of the environment which could be eventually analyzed by an operator to identify potential leak sources or victims.

Before reaching this goal, the SNAV shall be tested in more complex scenarios, for example in environments with chemical interferences or multiple gas sources. In this context, it may be required to use multiple MOX sensors, integrate air flow information or implement more advanced localization/mapping algorithms, for example bioinspired reactive plume tracking or BASED maps of bouts [81,82] (i.e., using bouts as events or as weighted events).

Author Contributions: Conceptualization, J.B. and S.M.; methodology, J.B. and V.H.; software, J.B.; validation, V.H., A.L. and S.M.; formal analysis, J.B.; investigation, J.B.; resources, A.L. and S.M.; data curation, J.B.; writing—original draft preparation, J.B.; writing—review and editing, all authors; visualization, J.B.; supervision, A.L. and S.M.

Funding: This research was funded by Spanish MINECO, grant numbers BES-2015-071698 (Severo-Ochoa) and TEC2014-59229-R (SIGVOL).

Acknowledgments: CERCA Programme/Generalitat de Catalunya. The Signal and Information Processing for Sensor Systems group is a consolidated Grup de Recerca de la Generalitat de Catalunya and has support from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2014-SGR-1445).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ma, K.Y.; Chirarattananon, P.; Fuller, S.B.; Wood, R.J. Controlled flight of a biologically inspired, insect-scale robot. *Science* **2013**, *340*, 603–607. [CrossRef]
2. Hassanalian, M.; Abdelkefi, A. Classifications, applications, and design challenges of drones: A review. *Prog. Aerosp. Sci.* **2017**, *91*, 99–131. [CrossRef]
3. Everaerts, J. The use of unmanned aerial vehicles (uavs) for remote sensing and mapping. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 1187–1192.
4. Khan, A.; Schaefer, D.; Tao, L.; Miller, D.J.; Sun, K.; Zondlo, M.A.; Harrison, W.A.; Roscoe, B.; Lary, D.J. Low power greenhouse gas sensors for unmanned aerial vehicles. *Remote Sens.* **2012**, *4*, 1355–1368. [CrossRef]
5. Berman, E.S.F.; Fladeland, M.; Liem, J.; Kolyer, R.; Gupta, M. Greenhouse gas analyzer for measurements of carbon dioxide, methane, and water vapor aboard an unmanned aerial vehicle. *Sens. Actuators B Chem.* **2012**, *169*, 128–135. [CrossRef]
6. Carrozzo, M.; De Vito, S.; Esposito, E.; Salvato, M.; Formisano, F.; Massera, E.; Di Francia, G.; Veneri, P.D.; Iadaresta, M.; Mennella, A. UAV Intelligent Chemical Multisensor Payload for Networked and Impromptu Gas Monitoring Tasks. In Proceedings of the 2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace), Rome, Italy, 20–22 June 2018; pp. 112–116.
7. Chang, C.C.Y.; Wang, J.L.; Chang, C.C.Y.; Liang, M.C.; Lin, M.R. Development of a multicopter-carried whole air sampling apparatus and its applications in environmental studies. *Chemosphere* **2016**, *144*, 484–492. [CrossRef]
8. Xie, T.; Liu, R.; Hai, R.T.; Hu, Q.H.; Lu, Q. UAV platform based atmospheric environmental emergency monitoring system design. *J. Appl. Sci.* **2013**, *13*, 1289–1296. [CrossRef]
9. Rossi, M.; Brunelli, D. Autonomous Gas Detection and Mapping With Unmanned Aerial Vehicles. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 765–775. [CrossRef]
10. McGonigle, A.J.S.; Aiuppa, A.; Giudice, G.; Tamburello, G.; Hodson, A.J.; Gurrieri, S. Unmanned aerial vehicle measurements of volcanic carbon dioxide fluxes. *Geophys. Res. Lett.* **2008**, *35*. [CrossRef]
11. Shinohara, H. Composition of volcanic gases emitted during repeating Vulcanian eruption stage of Shinmoedake, Kirishima volcano, Japan. *Earth Planets Space* **2013**, *65*, 667–675. [CrossRef]
12. Rüdiger, J.; Tirpitz, J.L.; de Moor, J.M.; Bobrowski, N.; Gutmann, A.; Liuzzo, M.; Ibarra, M.; Hoffmann, T. Implementation of electrochemical, optical and denuder-based sensors and sampling techniques on UAV for volcanic gas measurements: Examples from Masaya, Turrialba and Stromboli volcanoes. *Atmos. Meas. Tech.* **2018**, *11*, 2441–2457. [CrossRef]
13. Mori, T.; Hashimoto, T.; Terada, A.; Yoshimoto, M.; Kazahaya, R.; Shinohara, H.; Tanaka, R. Volcanic plume measurements using a UAV for the 2014 Mt. Ontake eruption the Phreatic Eruption of Mt. Ontake Volcano in 2014 5. Volcanology. *Earth Planets Space* **2016**, *68*. [CrossRef]
14. Astuti, G.; Giudice, G.; Longo, D.; Melita, C.D.; Muscato, G.; Orlando, A. An overview of the "volcan project": An UAS for exploration of volcanic environments. *J. Intell. Robot. Syst. Theory Appl.* **2009**. [CrossRef]
15. Neumann, P.P.; Kohlhoff, H.; Hüllmann, D.; Lilienthal, A.J.; Kluge, M. Bringing Mobile Robot Olfaction to the next dimension—UAV-based remote sensing of gas clouds and source localization. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 3910–3916.
16. Golston, L.M.; Aubut, N.F.; Frish, M.B.; Yang, S.; Talbot, R.W.; Gretencord, C.; McSpiritt, J.; Zondlo, M.A. Natural gas fugitive leak detection using an unmanned aerial vehicle: Localization and quantification of emission rate. *Atmosphere* **2018**, *9*, 333. [CrossRef]
17. Krüll, W.; Tobera, R.; Willms, I.; Essen, H.; Von Wahl, N. Early forest fire detection and verification using optical smoke, gas and microwave sensors. *Procedia Eng.* **2012**, *45*, 584–594. [CrossRef]

18. Merino, L.; Caballero, F.; Martínez-de Dios, J.R.; Ferruz, J.; Ollero, A. A cooperative perception system for multiple UAVs: Application to automatic detection of forest fires. *J. F. Robot.* **2006**, *23*, 165–184. [CrossRef]
19. Pfeifer, J.; Khanna, R.; Constantin, D.; Popovic, M.; Galceran, E.; Walter, A.; Siegwart, R.; Liebisch, F. Towards automatic UAV data interpretation. In Proceedings of the International Conference of Agricultural Engineering 2016, At Aahus, Denmark, 26–29 June 2016.
20. Roldán, J.J.; Joossen, G.; Sanz, D.; del Cerro, J.; Barrientos, A. Mini-UAV based sensory system for measuring environmental variables in greenhouses. *Sensors* **2015**, *15*, 3334–3350. [CrossRef]
21. Pobkrut, T.; Eamsa-Ard, T.; Kerdcharoen, T. Sensor drone for aerial odor mapping for agriculture and security services. In Proceedings of the 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 28 June–1 July 2016. [CrossRef]
22. Lega, M.; Napoli, R.M.A. A new approach to solid waste landfills aerial monitoring. *WIT Trans. Ecol. Environ.* **2008**, *109*, 193–199. [CrossRef]
23. Allen, G.; Hollingsworth, P.; Kabbabe, K.; Pitt, J.R.; Mead, M.I.; Illingworth, S.; Roberts, G.; Bourn, M.; Shallcross, D.E.; Percival, C.J. The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots. *Waste Manag.* **2018**. [CrossRef]
24. Emran, B.J.; Tannant, D.D.; Najjaran, H. Low-altitude aerial methane concentration mapping. *Remote Sens.* **2017**, *9*, 823. [CrossRef]
25. Daniel, K.; Dusza, B.; Lewandowski, A.; Wietfeld, C. Airshield: A system-of-systems muav remote sensing architecture for disaster response. In Proceedings of the 2009 3rd Annual IEEE Systems Conference, Vancouver, BC, Canada, 23–26 March 2009; pp. 196–200. [CrossRef]
26. Alvarado, M.; Gonzalez, F.; Fletcher, A.; Doshi, A. Towards the development of a low cost airborne sensing system to monitor dust particles after blasting at open-pit mine sites. *Sensors* **2015**, *15*, 19667–19687. [CrossRef] [PubMed]
27. Villa, T.; Gonzalez, F.; Miljievic, B.; Ristovski, Z.; Morawska, L. An Overview of Small Unmanned Aerial Vehicles for Air Quality Measurements: Present Applications and Future Prospectives. *Sensors* **2016**, *16*, 1072. [CrossRef] [PubMed]
28. Pajares, G. Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs). *Photogramm. Eng. Remote Sens.* **2015**, *81*, 281–330. [CrossRef]
29. Rossi, M.; Brunelli, D. Gas Sensing on Unmanned Vehicles: Challenges and Opportunities. In Proceedings of the 2017 New Generation of CAS (NGCAS), Genova, Italy, 6–9 September 2017; pp. 117–120.
30. Fahad, H.M.; Shiraki, H.; Amani, M.; Zhang, C.; Hebbar, V.S.; Gao, W.; Ota, H.; Hettick, M.; Kiriya, D.; Chen, Y.; et al. Room temperature multiplexed gas sensing using chemical-sensitive 3. 5-nm-thin silicon transistors. *Sci. Adv.* **2017**, *3*, 1–9. [CrossRef] [PubMed]
31. Dunkley, O.; Engel, J.; Sturm, J.; Cremers, D. Visual-Inertial Navigation for a Camera-Equipped 25 g Nano-Quadrotor. In Proceedings of the IROS Aerial Open Source Robotics Workshop, Chicago, IL, USA, 14–18 September 2014; pp. 4–5.
32. Preiss, J.A.; Honig, W.; Sukhatme, G.S.; Ayanian, N. Crazyswarm: A large nano-quadcopter swarm. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 3299–3304.
33. Farid, Z.; Nordin, R.; Ismail, M. Recent advances in wireless indoor localization techniques and system. *J. Comput. Networks Commun.* **2013**, *2013*, 185138. [CrossRef]
34. Monroy, J.G.; González-Jiménez, J.; Blanco, J.L. Overcoming the slow recovery of MOX gas sensors through a system modeling approach. *Sensors* **2012**, *12*, 13664–13680. [CrossRef]
35. Lilienthal, A.; Duckett, T. Building gas concentration gridmaps with a mobile robot. *Rob. Auton. Syst.* **2004**, *48*, 3–16. [CrossRef]
36. Kowadlo, G.; Russell, R.A. Robot odor localization: A taxonomy and survey. *Int. J. Rob. Res.* **2008**, *27*, 869–894. [CrossRef]
37. Lochmatter, T. *Bio-Inspired and Probabilistic Algorithms for Distributed Odor Source Localization using Mobile Robots*; École polytechnique fédérale de Lausanne (EPFL): Lausanne, Switzerland, 2010; Volume 4628.
38. Hernandez Bennetts, V.; Lilienthal, A.J.; Neumann, P.P.; Trincavelli, M. Mobile Robots for Localizing Gas Emission Sources on Landfill Sites: Is Bio-Inspiration the Way to Go? *Front. Neuroeng.* **2012**, *4*. [CrossRef]

39. Lilienthal, A.; Duckett, T. Experimental analysis of gas-sensitive Braitenberg vehicles. *Adv. Robot.* **2004**, *18*, 817–834. [CrossRef]
40. Ishida, H.; Nakamoto, T.; Moriizumi, T. Remote sensing of gas/odor source location and concentration distribution using mobile system. *Sens. Actuators B Chem.* **1998**, *49*, 52–57. [CrossRef]
41. Pang, S.; Farrell, J.A. Chemical plume source localization. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2006**, *36*, 1068–1080. [CrossRef]
42. Vergassola, M.; Villermaux, E.; Shraiman, B.I. "Infotaxis" as a strategy for searching without gradients. *Nature* **2007**, *445*, 406–409. [CrossRef] [PubMed]
43. Pomareda, V.; Magrans, R.; Jiménez-Soto, J.M.; Martínez, D.; Tresánchez, M.; Burgués, J.; Palacín, J.; Marco, S. Chemical source localization fusing concentration information in the presence of chemical background noise. *Sensors* **2017**, *17*, 904. [CrossRef] [PubMed]
44. Turner, D.B. *Workbook of Atmospheric Dispersion Estimates: An Introduction to Dispersion Modeling*; CRC Press: Boca Raton, FL, USA, 1994.
45. Farrell, J.A.; Murlis, J.; Long, X.; Li, W.; Cardé, R.T. Filament-based atmospheric dispersion model to achieve short time-scale structure of odor plumes. *Environ. Fluid Mech.* **2002**, *2*, 143–169. [CrossRef]
46. Thomson, D.J. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.* **1987**, *180*, 529–556. [CrossRef]
47. Falkovich, G.; Gawędzki, K.; Vergassola, M. Particles and fields in fluid turbulence. *Rev. Mod. Phys.* **2001**, *73*, 913–975. [CrossRef]
48. Sutton, O.G. The problem of diffusion in the lower atmosphere. *Q. J. R. Meteorol. Soc.* **1947**, *73*, 257–281. [CrossRef]
49. Bakkum, E.A.; Duijm, N.J. *Vapour Cloud Dispersion*; CPR E: London, UK, 1997; Volume 14.
50. Luo, B.; Meng, Q.H.; Wang, J.Y.; Sun, B.; Wang, Y. Three-dimensional gas distribution mapping with a micro-drone. In Proceedings of the 2015 34th Chinese Control Conference (CCC), Hangzhou, China, 28–30 July 2015; pp. 6011–6015. [CrossRef]
51. Lilienthal, A.J.; Reggente, M.; Trinca, M.; Blanco, J.L.; Gonzalez, J. A statistical approach to gas distribution modelling with mobile robots—The Kernel DM+V algorithm. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), St. Louis, MO, USA, 10–15 October 2009; pp. 570–576.
52. Hayes, A.T.; Martinoli, A.; Goodman, R.M. Distributed odor source localization. *IEEE Sens. J.* **2002**, *2*, 260–271. [CrossRef]
53. Neumann, P.P.; Asadi, S.; Lilienthal, A.J.; Bartholmai, M.; Schiller, J.H. Autonomous gas-sensitive microdrone: Wind vector estimation and gas distribution mapping. *IEEE Robot. Autom. Mag.* **2012**, *19*, 50–61. [CrossRef]
54. Burgués, J.; Marco, S. Low power operation of temperature-modulated metal oxide semiconductor gas sensors. *Sensors* **2018**, *18*, 339. [CrossRef] [PubMed]
55. Lilienthal, A.; Zell, A.; Wandel, M.; Weimar, U. Sensing odour sources in indoor environments without a constant airflow by a mobile robot. In Proceedings of the 2001 IEEE International Conference on Robotics and Automation (ICRA), Seoul, Korea, 21–26 May 2001; Volume 4, pp. 1–6.
56. Atema, J. Chemical signals in the marine environment: Dispersal, detection, and temporal signal analysis. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 62–66. [CrossRef] [PubMed]
57. Farah, A.; Duckett, T. Reactive Localisation of an Odour Source by a learning Mobile Robot. In Proceedings of the Second Swedish Workshop on Autonomous Robotics, Stockholm, Sweden, 11–12 October 2002; pp. 29–38.
58. Weissburg, M.J.; Dusenbery, D.B.; Ishida, H.; Janata, J.; Keller, T.; Roberts, P.J.W.; Webster, D.R. A multidisciplinary study of spatial and temporal scales containing information in turbulent chemical plume tracking. *Environ. Fluid Mech.* **2002**, *2*, 65–94. [CrossRef]
59. Webster, D.R.; Weissburg, M.J. Chemosensory guidance cues in a turbulent chemical odor plume. *Limnol. Oceanogr.* **2001**, *46*, 1034–1047. [CrossRef]
60. Gonzalez-Jimenez, J.; Monroy, J.G.; Blanco, J.L. The multi-chamber electronic nose-an improved olfaction sensor for mobile robotics. *Sensors* **2011**, *11*, 6145–6164. [CrossRef] [PubMed]
61. Batog, P.; Wołczowski, A. Odor markers detection system for mobile robot navigation. *Procedia Eng.* **2012**, *47*, 1442–1445. [CrossRef]

62. Marco, S.; Pardo, A.; Davide, F.A.M.; Di Natale, C.; D'Amico, A.; Hierlemann, A.; Mitrovics, J.; Schweizer, M.; Weimar, U.; Göpel, W. Different strategies for the identification of gas sensing systems. *Sens. Actuators B Chem.* **1996**, *34*, 213–223. [CrossRef]
63. Pardo, A.; Marco, S.; Samitier, J.; Davide, F.A.M.; Di Natale, C.; D'Amico, A. Dynamic measurements with chemical sensor arrays based on inverse modelling. In Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Brussels, Belgium, 4–6 June 1996; Volume 1.
64. Fonollosa, J.; Sheik, S.; Huerta, R.; Marco, S. Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring. *Sens. Actuators B Chem.* **2015**, *215*, 618–629.
65. Schmuker, M.; Bahr, V.; Huerta, R. Exploiting plume structure to decode gas source distance using metal-oxide gas sensors. *Sens. Actuators B Chem.* **2016**, *235*, 636–646. [CrossRef]
66. Bitcraze, A.B. Getting Started with the Loco Positioning System. Available online: https://www.bitcraze.io/getting-started-with-the-loco-positioning-system/ (accessed on 7 July 2018).
67. *DecaWave DWM1000 Datasheet*; DecaWave: Dublin, Ireland, 2016.
68. Nelson, G. *Gas Mixtures: Preparation and Control*; CRC Press: Boca Raton, FL, USA, 1992.
69. Currie, L.A. Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995). *Pure Appl. Chem.* **1995**, *67*, 1699–1723. [CrossRef]
70. Burgués, J.; Jimenez-Soto, J.M.; Marco, S. Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models. *Anal. Chim. Acta* **2018**, *1013*, 13–25. [CrossRef] [PubMed]
71. Pashami, S.; Lilienthal, A.J.; Trincavelli, M. Detecting changes of a distant gas source with an array of MOX gas sensors. *Sensors* **2012**, *12*, 16404–16419. [CrossRef] [PubMed]
72. Burgués, J.; Marco, S. Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors. *Anal. Chim. Acta* **2018**, *1019*, 49–64. [CrossRef] [PubMed]
73. Shakaff, A.; Yeon, A.; Kamarudin, K.; Visvanathan, R. Gas Source Localization via Behaviour Based Mobile Robot and Weighted Arithmetic Mean Gas Source Localization via Behaviour Based Mobile Robot and Weighted Arithmetic Mean. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *318*, 012049. [CrossRef]
74. Li, J.G.; Sun, B.; Zeng, F.L.; Liu, J.; Yang, J.; Yang, L. Experimental study on multiple odor sources mapping by a mobile robot in time-varying airflow environment. In Proceedings of the Chinese Control Conference (CCC), Chengdu, China, 27–29 July 2016; pp. 6032–6037.
75. Kowadlo, G.; Russell, R.A. Using naïve physics for odor localization in a cluttered indoor environment. *Auton. Robots* **2006**, *20*, 215–230. [CrossRef]
76. Lilienthal, A.J.; Loutfi, A.; Duckett, T. Airborne chemical sensing with mobile robots. *Sensors* **2006**, *6*, 1616–1678. [CrossRef]
77. Reggente, M.; Lilienthal, A.J. Three-dimensional statistical gas distribution mapping in an uncontrolled indoor environment. *AIP Conf. Proc.* **2009**, *1137*, 109–112.
78. Pashami, S.; Lilienthal, A.J.; Schaffernicht, E.; Trincavelli, M. TREFEX: Trend estimation and change detection in the response of MOX gas sensors. *Sensors* **2013**, *13*, 7323–7344. [CrossRef]
79. Loutfi, A.; Coradeschi, S.; Lilienthal, A.J.; Gonzalez, J. Gas distribution mapping of multiple odour sources using a mobile robot. *Robotica* **2009**, *27*, 311–319. [CrossRef]
80. Lilienthal, A.; Reimann, D.; Zell, A. Gas Source Tracing with a Mobile Robot Using an Adapted Moth Strategy. *Auton. Mob. Syst.* **2003**, 150–160. [CrossRef]
81. Lilienthal, A.; Trincavelli, M.; Schaffernicht, E. It's always smelly around here! Modeling the Spatial Distribution of Gas Detection Events with BASED Grid Maps. In Proceedings of the 15th International Symposium on Olfaction and Electronic Nose (ISOEN 2013), Daegu, Korea, 2–5 July 2013; Volume 27.
82. Schaffernicht, E.; Trincavelli, M.; Lilienthal, A.J. Bayesian Spatial Event Distribution Grid Maps for Modeling the Spatial Distribution of Gas Detection Events. *Sens. Lett.* **2014**, *12*, 1142–1146. [CrossRef]

Chapter 8

Conclusions

The main conclusion of this thesis is that a nano-drone equipped with MOX sensors can be used for gas source localization and mapping in large indoor environments. Other conclusions that we reached during the development and test of the platform are detailed below:

- A rough approximation of the gas distribution can be obtained in very short time with concentration measurements acquired in motion.
- The bout frequency yields higher localization accuracy than the instantaneous concentration (average error of 1.55 m and 2.45 m, respectively).
- The log-log transformation linearizes the response of the TGS 8100 sensor in the range 0-50 ppm, enabling the computation of the LOD with the standard formulas.
- Regarding the bout amplitude threshold,
  - If it is too low, the high number of false alarms hinders source localization.
  - If it is too high, the plume can only be detected near the source.
  - When correctly estimated, a narrow plume can be detected at 15 m distance from the source.
- Turbulence generated by the propellers affects the gas sensor signals, but key features for gas source localization can still be extracted by signal processing methods.
- To localize an elevated source is necessary to sample the space in 3D.

The analysis of gas distribution patterns using a MOX sensor grid (Chapter 7) reveled that:

- Convective or recirculating air currents may have a stronger effect on the gas distribution than the density of the released substance.
- A gas heavier than air (e.g., ethanol) does not necessarily accumulate on the floor.
- The gas distribution might not be stable in the absence of strong air flow.

- Strong unidirectional airflow produces a stable gas distribution.
- The density of the gas is a relevant factor for gas dispersion only at high release rates.
- The installation of the sensors (e.g. orientation and type of support) with respect to the main airflow stream might affect the results.

Regarding gas source localization using a grid of MOX sensors (Chapter 7),

- The bout frequency and the variance yield lower localization errors than the mean (average difference of 0.9 and 0.5 m, respectively).
- The bout frequency outperforms the variance under strong air flow (average difference of 1.0 m).
- Measuring at the height of the source is critical for successful gas source localization, especially when using the variance or mean concentration.
- The bout frequency allows detecting the source from above.
- The performance of the bout frequency strongly depends on the noise threshold and the size of the measurement window.
- The variance is less sensitive to the window size than the bout frequency.

Regarding the use of the bout frequency for predicting the distance to a gas source (Chapter 6),

- The cut-off frequency of the low-pass differentiator filter and the bout amplitude threshold are coupled parameters, which means that both must be tuned simultaneously to find the global optimum in terms of prediction error.
- The optimum parameter values strongly depend on the wind speed, with low values being suitable for scenarios with high wind speed and vice-versa.
- The bout frequency with optimized parameters outperforms the mean, variance and maximum response when the measuring window is large enough.
- The three-sigma threshold is not an optimum threshold, probably because the distribution of amplitudes of baseline bouts does not follow a Gaussian distribution.
- The optimum bout amplitude threshold is specific to the sensor model.

The response time of MOX sensors can be improved by a factor of 10 by miniaturization of the gas chamber, removing the sensor cap and filtering the response with a low pass differentiator filter (Chapter 5).

Regarding low power operation of temperature-modulated MOX sensors (Chapter 4),

- 10% duty-cycling operation of 10-min periods reduces the prediction error 2.5 times with respect to on-demand operation (2.2 versus 0.9 ppm) and saves up to 90% energy as compared to the continuous operating mode.

- On-demand operation increases the prediction error by almost a factor of 5 as compared to continuous operation (2.2 versus 0.45 ppm).
- Applying continuous power during calibration yields lower prediction errors (15-30 %) than calibration in the target powering mode.

A series of conclusions were reached regarding the use of temperature-modulated MOX sensors for quantification of low concentrations of CO under variable humidity conditions (Chapter 3):

- The optimum measurement point ($t = 6 - 7$ s) differs from manufacturer recommendations ($t = 25$ s).
- The optimum measurement point reduces the LOD from 4.0 to 2.3 ppm and increases the reliability of the estimate from 80% to 95%.
- The LOD can be further reduced to 0.73 ppm by using multivariate calibration (PLS).
- PLS models are simpler than MLPCR models, for the same level of prediction error.
- The underlying assumptions of the LOD formula can be satisfied by reducing the calibration range, applying logarithmic transformation to the signal and/or the concentration and optimizing the measurement point.

An obvious future direction for gas source localization using nano-drones is to achieve autonomous navigation. This requires three-dimensional obstacle detection and self-localization, which is not an easy task whatsoever because there are many degrees of freedom and the low payload does often not allow for range sensors. The lack of selectivity and high limit of detection of current gas sensors, as compared to biological chemoreceptors, are also major limitations for gas source localization in realistic scenarios characterized by large exploration areas and chemical interferences. In this context, it may be worth it to investigate the use of multiple MOX sensors (to improve selectivity), use a pump to "sniff" the air surrounding the drone (improve limit of detection) or implement more advanced localization/mapping algorithms.

Chapter 9

Resumen en castellano

Esta tesis representa un paso adelante en el uso de sensores semiconductores de óxido de estaño (MOX) en diversos entornos, desde medidas estáticas en una cámara de gas hasta medidas dinámicas con un robot en entornos abiertos. Debido a su baja selectividad y estabilidad, los sensores MOX se utilizan en productos comerciales principalmente como detectores de altas concentraciones de gases tóxicos, por ejemplo, en alarmas de monóxido de carbono (CO) o en detectores de límite inferior de explosividad (LEL). Sin embargo, muchas aplicaciones interesantes de estos sensores requieren cuantificación de bajas concentraciones; desde robots para localización de fuentes químicas hasta dispositivos biomédicos para detección de biomarcadores en muestras de aliento. En estas aplicaciones, el limite de detección (LOD) es una figura de mérito fundamental que mide la capacidad del sensor para medir bajas concentraciones del gas objetivo en presencia de ruido químico/eléctrico. El principal problema es que los sensores MOX presentan características incompatibles con las fórmulas para el cálculo del LOD definidas por la International Union of Pure and Applied Chemistry (IUPAC), las cuales asumen que la respuesta es univariante, lineal y que los errores residuales están distribuidos de manera Gaussiana y homocedástica (varianza constante).

El primer objetivo de esta tesis es desarrollar métodos para superar los retos asociados al cálculo del LOD en sensores MOX, particularmente en sensores modulados en temperatura. La modulación en temperatura es una manera comúnmente aceptada de incrementar la selectividad de un sensor, mediante la aplicación de una onda de potencia periódica en el calentador del sensor. De hecho, es el modo de operación recomendado por los fabricantes de sensores MOX para la detección de CO. Pese a que la respuesta de un sensor modulado en temperatura es multivariante, por simplicidad para el usuario los fabricantes recomiendan medir únicamente en un punto del ciclo de temperatura. El primer método que proponemos (Paper I) consiste por tanto en optimizar el punto de medida de manera que el error de predicción sea lo más bajo posible y se cumplan las hipótesis subyacentes a la fórmula del LOD. Para esto último, definimos una batería de tests estadísticos que comprueban que un determinado punto de medida cumple las suposiciones de linealidad, normalidad y homocedasticidad. Para ilustrar la metodología, utilizamos siete unidades del sensor FIS SB-500-12 en un escenario de detección de CO con variaciones controladas de

humedad (interferencia química), representativo de varias aplicaciones reales. En este sensor, el modo de operación recomendado combina un perfil de temperatura cuadrado (alta temperatura durante 5 s seguido de baja temperatura durante 20 s) y medida univariante de la respuesta al final del ciclo ($t = 25$ s). En ese punto, la sensibilidad al CO es máxima en un rango alto de concentraciones (p.ej. 20-300 ppm) pero no está claro que sea el punto óptimo en el rango de bajas concentraciones (p.ej. 0-10 ppm) relevante para el cálculo del LOD. De hecho, nuestros experimentos confirman que los errores de predicción en dicho punto son mayores que en otros puntos del ciclo y las suposiciones de la IUPAC sólo se cumplen con una probabilidad del 80%. El punto óptimo, que se encuentra situado al principio del escalón de baja temperatura (t = 7 s), disminuye el LOD desde 4.0 ppm (punto de medida recomendado) hasta 2.3 ppm y aumenta la probabilidad de cumplir las hipótesis subyacentes (y por tanto la fiabilidad de la estimación) al 95%. Para cumplir la hipótesis de linealidad tuvimos que transformar logarítmicamente la señal. Los modelos de calibración en el punto óptimo (y por tanto el LOD) son estables al menos durante dos semanas (incremento de error de predicción < 90 ppb), si se corrige la línea de base de manera periódica. La manera más simple de realizar dicha corrección requiere exponer los sensores a aire limpio, lo cual podría no ser trivial en ciertas aplicaciones. En cualquier caso, este estudio demuestra que sensores MOX linealizados, previa optimización del punto de trabajo y con ajustes periódicos de la línea de base, pueden ser usados en aplicaciones que requieran un umbral de detección del orden de 3 ppm.

Utilizar toda la información contenida en el patrón de respuesta multivariante, en lugar de un único punto de medida, parece una estrategia inmediata para reducir el error de predicción y, consecuentemente, mejorar el LOD de sensores modulados en temperatura. Debido a que el patrón de respuesta presenta ruido heterocedástico y correlacionado, no está claro cual es el mejor modelo de calibración multivariante para estos datos. Por un lado, los métodos de máxima verosimilitud (p.ej. MLPCR), son óptimos desde un punto de vista estadístico porque incluyen información explícita sobre la incertidumbre de la medida durante el proceso de calibración. Sin embargo, no hay referencias disponibles en la literatura sobre la idoneidad de MLPCR para la calibración de sensores MOX modulados en temperatura. Por otro lado, PLS es el modelo de calibración de facto para este tipo de datos, debido a su inherente robustez ante la multicolinearidad y rechazo de sustancias interferentes. Nuestros experimentos (Paper II) demuestran que PLS y MLPCR son capaces de alcanzar niveles similares de error (0.4 ppm), pero los modelos PLS son más simples. Este comportamiento—que fue originalmente reportado en una comparativa entre PCR y PLS y después confirmado para MLPCR—está relacionado con el hecho de que las variables latentes de PLS consideran la correlación con el vector de etiquetas (y), proporcionando espacios predictivos más parsimoniosos que PCR, cuyas componentes principales están únicamente basadas en la matriz de predictores (X). PLS filtra de manera efectiva el ruido en X sin conocimiento explícito de la incertidumbre en las medidas, mediante la reducción del peso asociado a áreas del patrón con alto error de medida (estimado en réplicas) si dichas áreas no están correlacionadas con el vector de etiquetas. Esto sucede, por ejemplo, en las inmediaciones de las transiciones de temperatura (p.ej. $t = 5$ s) debido al mayor error de medida. Por otro lado, las variables con mayor poder predictivo están situadas al principio del ciclo de baja temperatura (i.e. $t = 6-7$ s), lo que coincide con el punto óptimo de medida encontrado en el estudio univariante.

La estimación de figuras de mérito básicas como la sensibilidad, selectividad o LOD en calibración multivariante es actualmente un problema sin resolver. En el caso del LOD, el mayor problema son las discrepancias fundamentales entre la aproximación analítica, basada en propagación de errores, y la aproximación empírica, basada en predicciones sobre muestras reales. En el primer caso, las incertidumbres instrumentales y de las concentraciones de calibración se propagan a través del modelo para estimar la varianza de la concentración predicha, dando lugar a un rango de LODs que hace explícito que la composición de la muestra (debido a diferentes niveles de interferentes) afecta al LOD. En el segundo caso, las predicciones sobre muestras de calibración o (idealmente) validación se tratan como una variable subrogada sobre la que se aplican las fórmulas del LOD univariantes, dando lugar a un único LOD que podría considerarse como un promedio representativo de las muestras utilizadas. Puesto que el método analítico requiere conocimiento a priori de los errores instrumentales (incluyendo todas las fuentes de error) y esta información no está habitualmente disponible, los métodos empíricos son los más usados en la práctica. Pese a que el método empírico basado en la señal analítica neta (NAS) como variable subrogada está aceptado por la IUPAC, el LOD obtenido varía en función del método que se utilice para calcular el NAS. Para solucionar este problema, hemos desarrollado un método empírico (Paper II) para estimar el LOD de manera unívoca en modelos PLS, basado en la aplicación de las fórmulas univariantes a los scores de la primera variable latente del modelo ortogonalizado (sólo existe una manera de calcular dichos scores). La ortogonalización del modelo condensa en la primera componente toda la varianza del bloque X relacionada exclusivamente con el vector de etiquetas y, no sólo asegurando que el LOD se calcule teniendo en cuenta todo el poder predictivo del modelo sino facilitando también la visualización de la estructura del modelo en un único gráfico 2D, independientemente del numero de componentes del modelo PLS original. Usando este método, obtuvimos un LOD medio de 0.77 ppm (considerando los siete sensores) que aproximadamente representa una mejora de un factor 5 respecto al modo de operación recomendado por el fabricante (4.0 ppm) y un factor 3 respecto al modelo univariante optimizado (2.3 ppm). Este valor está contenido dentro del rango de LODs obtenidos mediante el método analítico cuando la complejidad del modelo es adecuada, llegando por tanto a un acuerdo entre ambos métodos. La estabilidad de los modelos de calibración PLS, calculada en muestras de validación externa durante dos semanas, es similar a la obtenida en los modelos univariantes optimizados. Estos resultados sugieren que sensores MOX modulados en temperatura y calibrados con modelos PLS tienen potencial para ser utilizados en aplicaciones que requieran umbrales de detección de 1 ppm.

Tradicionalmente, los sensores MOX se han operado mediante alimentación ininterrumpida (incluso aunque no estén midiendo) para aumentar la estabilidad de la respuesta. Es conocido que, tras encender el sensor, la respuesta entra en un régimen transitorio que puede durar desde pocos segundos a varias horas en función del tiempo que el sensor haya estado apagado. Mantener un sensor continuamente alimentado no siempre es posible, sobre todo en instrumentos portables (p.ej. robots detectores de olor), debido a limitaciones en el consumo de energía o por la operación intermitente del dispositivo. Por otro lado, operar el sensor bajo demanda (o sea, encenderlo únicamente para hacer medidas puntuales) es muy eficiente a nivel energético pero una medida estable sólo es posible pasado un cierto tiempo. Para alcanzar un compromiso entre consumo y estabilidad, proponemos

un modo de operación de bajo consumo (Paper III) que consiste en encender el sensor durante 1 minuto cada 10 minutos (ciclo de trabajo del 10%) cuando no se está midiendo para evitar largos periodos de desconexión y, al mismo tiempo, ahorrar un 90% de energía respecto al modo continuo. Tras comparar el rendimiento de los tres modos de alimentación (continuo, ciclo 10% y bajo demanda) en predicciones de CO con condiciones de humedad variable durante dos semanas, observamos que el modo propuesto reduce el error de predicción, en media, 2.5 veces con respecto a la operación bajo demanda (0.9 versus 2.2 ppm) y duplica el error del modo continuo (0.9 versus 0.45 ppm). También observamos que, independientemente del modo de operación, resulta beneficioso calibrar el sensor en modo continuo ya que los errores de predicción son un 15-30% más pequeños que si se calibra el sensor en un modo de bajo consumo.

El tiempo de respuesta de los sensores MOX (del orden de varios segundos) es demasiado lento para extraer las características de alta frecuencia de una pluma turbulenta, que facilitarían la localización de fuentes químicas en entornos reales. De entre todas las opciones disponibles para mejorar el tiempo de respuesta, en esta tesis se explora en detalle una técnica de procesado de señal (Schmuker et al., 2016) que aprovecha la rápida dinámica de la derivada de la respuesta. Experimentos realizados en un túnel de viento demuestran que el número de flancos de subida de la derivada filtrada, denominados "bouts", disminuye de manera lineal a medida que el sensor se aleja del origen de la pluma. Esto permitiría estimar la distancia a una fuente química más rápidamente que utilizando la concentración media, lo cual puede requerir ventanas de medida de hasta 10 minutos para obtener una estimación fiable. La ventaja de esta técnica es que no requiere hardware adicional—algo deseable en plataformas robóticas con poca carga útil—ni complicados procesos de calibración en comparación con otras técnicas de procesado de señal como los modelos dinámicos inversos. Utilizando un detector de fotoionización (PID) con alto ancho de banda (330 Hz) hemos correlacionado, por primera vez, los "bouts" de un sensor MOX con las fluctuaciones de concentración de una pluma turbulenta. Durante estos experimentos, observamos que bouts no producidos por la pluma pueden aparecer durante la recuperación del sensor o simplemente debido a ruido. Eliminar dichos bouts "falsos", caracterizados por baja amplitud y larga duración, es crítico para obtener una relación monótonamente creciente entre el número de bouts y la proximidad a la fuente. En este sentido, el umbral basado en la regla de las tres desviaciones estándar (i.e. 3-sigma), propuesto por Schmuker et al., podría no ser óptimo ya que (entre otras cosas) está basado en una suposición Gaussiana que hemos comprobado no se cumple en la práctica.

Nuestros experimentos revelan que el umbral óptimo es específico al modelo de sensor, el estimador de distancia a la fuente utilizado (p.ej. frecuencia de bouts, amplitud media de bouts, ...) y la intensidad de la pluma. La amplitud media de bouts es un indicador robusto en los tres sensores MOX probados, ya que crece monótonamente con proximidad a la fuente, independientemente del umbral escogido. Sin embargo, es muy poco sensible lejos de la fuente a menos que el umbral sea muy alto (al menos un orden de magnitud superior al umbral 3-sigma), lo cual requiere largas ventanas de medida que permitan capturar suficientes bouts como para obtener una estimación fiable de la amplitud media. Por otro lado, la frecuencia de bouts es monótonamente creciente y sensible lejos de la fuente, sin necesidad de utilizar umbrales tan altos (excepto en el sensor MiCS 5524). En este sensor, el umbral tiene que estar un orden de magnitud por encima de 3-sigma para

producir el comportamiento deseado, lo cual podría explicar el nulo poder predictivo de la frecuencia de bouts reportado por Vuka et al. en experimentos realizados con sensores de la familia MiCS, operados (probablemente) con el umbral 3-sigma. En los sensores de la familia TGS (Figaro Engineering Inc), el umbral 3-sigma y valores más bajos sí dan buenos resultados. Por ejemplo, el sensor TGS 2620 se vuelve mucho más sensible en el rango de distancias probado (0-2.3 m) al reducir el umbral un orden de magnitud por debajo de 3-sigma. Un problema que se abre tras esta investigación es cómo optimizar de manera sistemática dicho umbral sin tener que exponer los sensores a la pluma, lo cual podría no ser posible en ciertas aplicaciones.

La conclusión del estudio anterior es que la frecuencia de bouts, tras aplicar el umbral óptimo, es un mejor estimador de proximidad a la fuente que la amplitud media de bouts. La siguiente pregunta es si este estimador también es superior a estimadores de proximidad conocidos como la media o la varianza de la concentración. Nuestros resultados sugieren que la frecuencia de bouts es superior a la varianza cuando el sensor está relativamente lejos de la fuente (d>1.2 m), mientras que la varianza es más sensible a distancias cortas (d<1.2 m). La concentración media, que algunos autores consideran un estimador no fiable de proximidad a la fuente, es el mejor estimador (ligeramente más sensible que la frecuencia de bouts) en el sensor TGS 2620 pero no en los otros. Esto puede ser debido al comportamiento conocido del encapsulado del sensor (de mucho mayor tamaño en el TGS 2620 que en los otros dos sensores tipo MEMS), que filtra paso bajo las fluctuaciones de la pluma, dando lugar quizás a una mejor estimación de la concentración media que con encapsulados más pequeños. Una ventaja teórica de utilizar la concentración media o la varianza es que no se requiere ningún umbral para su cómputo. Sin embargo, un umbral sí es necesario en la práctica ya que seguir un gradiente de cualquiera de estas variables para moverse en dirección a la fuente puede dar lugar a muchas falsas alarmas debido al ruido, especialmente si no hay ninguna fuente de gas en el área de exploración.

Las conclusiones alcanzadas respecto al umbral de ruido y al rendimiento de diferentes estimadores de proximidad están basadas en estudios unidimensionales (1D) en los que el sensor está situado en el eje central de la pluma. Sin embargo, en una aplicación real el robot puede estar en cualquier posición respecto a la pluma (p.ej. fuera de la pluma) y, en el caso de robots aéreos, también por encima o por debajo de la pluma. Por consiguiente, para el desarrollo de algoritmos de localización de fuentes que funcionen en tres dimensiones (3D) es necesario explorar el comportamiento de estos estimadores en estudios 3D. Para ello, desplegamos una red tridimensional de sensores MOX en una sala de pruebas (6 x 5 m2) donde una fuga de gas es emulada mediante un recipiente con etanol y un ventilador. Los 10 experimentos realizados, de 1.5 horas de duración cada uno, incluyen variaciones de la posición y la altura de la fuente, la velocidad de evaporación y la velocidad del viento.

Tras monitorizar la evolución temporal de la distribución de gas en dichos experimentos, descubrimos que ciertas suposiciones comúnmente aceptadas en el campo de olfacción con robots móviles (MRO) podrían no ser ciertas en todos los escenarios. Por ejemplo, muchos autores asumen que gases más densos que el aire (p.ej. etanol) se acumulan en el suelo, justificando así el uso de robots terrestres. Sin embargo, en nuestros experimentos esto sólo sucedió cuando incrementamos la velocidad de evaporación (por medio de un burbujeador), llegando a concentraciones medias de 100-150 ppm en la sala. Sin el burbujeador, la baja volatilidad del etanol a temperatura ambiente da lugar a concentraciones

medias de pocos ppm, demasiado bajas como para alterar la flotabilidad del aire de la sala. La acumulación progresiva de gas en el techo de la habitación que observamos en varios experimentos sin burbujeador sugiere que la dispersión podría estar dominada por las corrientes de convección o recirculación generadas por el sistema de climatización. Otra suposición de ciertos investigadores es que la ausencia de un flujo dominante de aire produce una distribución de gas estable, lo cual no se cumple en nuestros experimentos (al menos, en las inmediaciones de la fuente de gas).

Respecto a los estimadores de proximidad de fuente, parece fundamental tomar medidas a la altura de la fuente para poder localizarla. Este requisito no es tan estricto cuando se utiliza la frecuencia de bouts, ya que si hay un flujo de aire dominante es posible detectar la fuente a una altura ligeramente superior a la de la fuente. Usando este estimador, es importante seleccionar un tamaño de ventana lo suficientemente grande y optimizar el umbral de ruido. Tal y como se observó en los estudios univariantes, para optimizar dicho umbral es necesario exponer los sensores a la pluma o conocer la intensidad de la fuente. La frecuencia de bouts y la varianza tienen un rendimiento similar en ausencia de viento, mientras que el primero es superior en condiciones de mucho viento (diferencia en error de localización de 1.0 m). Ambos son más precisos que la concentración media (errores de localización 0.5-0.9 m menores), independientemente del nivel de viento. Futuras líneas de investigación en este campo podrían ir encaminadas a estudiar el rendimiento de diferentes estimadores de proximidad en presencia de interferencias químicas, por ejemplo, en un escenario con varias fuentes de gas.

Cuando los bouts son utilizados por un robot para seguir una pluma de gas hacia la fuente, éstos tienen que ser detectados en tiempo real para responder inmediatamente a ellos. Una limitación práctica del algoritmo de detección de bouts es que implementa de filtros de suavizado que utilizan datos futuros, es decir, no son realizables. Para solucionar este problema, proponemos una serie de modificaciones para simplificar el algoritmo y adaptarlo para operación en tiempo real. La idea es sustituir los filtros no-causales por filtros causales (realizables) que utilizan el mismo factor de suavizado en las múltiples aplicaciones del filtro, reduciendo así el número de parámetros requerido por el algoritmo. El resultado es que la detección de bouts se puede conseguir con retrasos de sólo 25-65 ms. El algoritmo fue testeado en un nano-drone equipado con un sensor MOX, encargado de localizar una fuente química en una sala de grandes dimensiones (160 m^2). A pesar de que las turbulencias generadas por los rotores afectan significativamente la señal del sensor, pudimos obtener un mapa aproximado de la distribución del gas en la sala y de los bouts en menos de 3 minutos, tomando medidas en movimiento. El uso de un nano-drone para este propósito y el hecho de tomar medidas en movimiento diferencia nuestro trabajo de trabajos previos, donde lo robots típicamente se detienen para tomar medidas, lo cual aumenta la duración de la misión considerablemente. La construcción del mapa de concentración o de bouts en entornos de mayores dimensiones podría acelerarse utilizando un grupo de robots que exploren colaborativamente diferentes partes del escenario.

References

[1] Faruk Dincer, Mustafa Odabasi, and Aysen Muezzinoglu. Chemical characterization of odorous gases at a landfill site by gas chromatography-mass spectrometry. *Journal of Chromatography A*, 1122(1-2):222–229, 2006. (Cited on page 1.)

[2] John E. Amoore and Earl Hautala. Odor as an aid to chemical safety: Odor thresholds compared with threshold limit values and volatilities for 214 industrial chemicals in air and water dilution. *Journal of Applied Toxicology*, 3(6):272–290, 1983. (Cited on page 1.)

[3] Oliver W Johnson and Philip S Williams. Combustible gas analysis apparatus, 1935. (Cited on page 2.)

[4] Victor Manuel Hernandez Bennetts, Achim J. Lilienthal, Ali Abdul Khaliq, Victor Pomareda Sese, and Marco Trincavelli. Towards real-world gas distribution mapping and leak localization using a mobile robot with 3d and remote gas sensing capabilities. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 2335–2340, 2013. (Cited on pages 2 and 31.)

[5] Maximilian Lackner. Tunable diode laser absorption spectroscopy (TDLAS) in the process industries - A review, 2007. (Cited on pages 2, 36, and 58.)

[6] Mickey B. Frish, Richard T. Wainner, Joy Stafford-Evans, Byron D. Green, Mark G. Allen, S. Chancey, J. Rutherford, G. Midgley, and P. Wehnert. Standoff Sensing of Natural Gas Leaks: Evolution of the Remote Methane Leak Detector (RMLD). In *Photonic Applications Systems Technologies Conference*, page JThF3, 2005. (Cited on pages 2 and 37.)

[7] M.H.E. Larcombe and J.R. Halsall. Robotics in nuclear engineering, 1984. (Cited on page 2.)

[8] M W Siegel. Olfaction metal oxide semiconductor gas sensors and neural networks. In *Traditional and non-traditional robotic sensors*, pages 143–157. Springer, 1990. (Cited on page 2.)

[9] Vincenzo Genovese, Paolo Dario, R Magni, and L Odetti. Self organizing behavior and swarm intelligence in a pack of mobile miniature robots in search of pollutants.

In *Intelligent Robots and Systems, 1992., Proceedings of the 1992 lEEE/RSJ International Conference on*, volume 3, pages 1575–1582. IEEE, 1992. (Cited on pages 2 and 4.)

[10] P Mächler. Detection technologies for anti-personnel mines, 1995. (Cited on page 2.)

[11] James P. Trevelyan, Sung-Chul Kang, and William R. Hamel. Robotics in Hazardous Applications. In *Springer Handbook of Robotics*, pages 1101–1126. Springer, 2008. (Cited on page 2.)

[12] Charles W. Gardner, Rachel Wentworth, Patrick J. Treado, Parag Batavia, and Gary Gilbert. Remote chemical biological and explosive agent detection using a robot-based Raman detector. *Proceedings of SPIE*, 6962:69620T–69620T–10, 2008. (Cited on page 2.)

[13] Curtis M Humphrey and Julie A Adams. Robotic Tasks for CBRNE Incident Response. *Advanced Robotics*, 2009. (Cited on page 2.)

[14] Achim J. Lilienthal, Amy Loutfi, and Tom Duckett. Airborne chemical sensing with mobile robots. *Sensors*, 6(11):1616–1678, 2006. (Cited on pages 3, 9, 20, 22, 40, 57, and 62.)

[15] Andrew Russell, David Thiel, and Alan Mackay-Sim. Sensing odour trails for mobile robot navigation. In *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*, pages 2672–2677. IEEE, 1994. (Cited on page 3.)

[16] Reimundo Deveza, David Thiel, Andrew Russell, and Alan Mackay-Sim. Odor sensing for robot guidance. *The International Journal of Robotics Research*, 13(3):232–239, 1994. (Cited on page 3.)

[17] R.a. Russell, D. Thiel, R. Deveza, and A. Mackay-Sim. A robotic system to locate hazardous chemical leaks. *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, 1:556–561, 1995. (Cited on pages 3, 15, 16, and 18.)

[18] Achim J. Lilienthal, Matteo Reggente, Marco Trinca, Jose Luis Blanco, and Javier Gonzalez. A statistical approach to gas distribution modelling with mobile robots - The Kernel DM+V algorithm. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), St. Louis, MO, USA, 10–15 October*, pages 570–576, 2009. (Cited on pages 3 and 30.)

[19] Samuel Soldan, Gero Bonow, and Andreas Kroll. RoboGas Inspector - A mobile robotic system for remote leak sensing and localization in large industrial environments: Overview and first results. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 1(PART 1):33–38, 2012. (Cited on pages 3, 4, and 31.)

[20] R. Rozas, J. Morales, and D. Vage. Artificial smell detection for robotic navigation, 1991. (Cited on pages 3, 14, 15, and 44.)

[21] H Ishida, T Nakamoto, and T Moriizumi. Fundamental study of mobile system for smelling-object localization using plural gas sensors. *Proceedings of the Society for Instrument and Control Engineers, Kanazawa, Japan*, pages 767–768, 1993. (Cited on pages 3, 4, 14, and 15.)

[22] H. Ishida, K. Suetsugu, T. Nakamoto, and T. Moriizumi. Study of autonomous mobile sensing system for localization of odor source using gas sensors and anemometric sensors. *Sensors and Actuators: A. Physical*, 45(2):153–157, 1994. (Cited on pages 3, 17, and 25.)

[23] H. Ishida, Y. Kagawa, T. Nakamoto, and T. Moriizumi. Odor-source Localization In Clean Room By Autonomous Mobile Sensing System. *Proceedings of the International Solid-State Sensors and Actuators Conference - TRANSDUCERS '95*, 1:115–121, 1995. (Cited on pages 3, 15, 17, 18, and 59.)

[24] T. L. Thatcher, D. J. Wilson, E. E. Wood, M. J. Craig, and R. G. Sextro. Pollutant dispersion in a large indoor space: Part 1 - Scaled experiments using a water-filled model with occupants and furniture. *Indoor Air*, 14(4):258–271, 2004. (Cited on pages 4, 10, and 58.)

[25] Hiroshi Ishida, Takamichi Nakamoto, Toyosaka Moriizumi, Timo Kikas, and Jiri Janata. Plume-tracking robots: a new application of chemical sensors. *Biol. Bull.*, 200(2):222–226, 2001. (Cited on pages 4 and 58.)

[26] Kristine A. Justus, John Murlis, Chris Jones, and Ring T. Cardé. Measurement of odor-plume structure in a wind tunnel using a photoionization detector and a tracer gas. *Environmental Fluid Mechanics*, 2(1-2):115–142, 2002. (Cited on pages 4 and 59.)

[27] Gideon Kowadlo and R. Andrew Russell. Robot odor localization: A taxonomy and survey. *International Journal of Robotics Research*, 27(8):869–894, 2008. (Cited on pages 4, 13, and 15.)

[28] Victor Hernandez Bennetts, Achim J. Lilienthal, Patrick P. Neumann, and Marco Trincavelli. Mobile Robots for Localizing Gas Emission Sources on Landfill Sites: Is Bio-Inspiration the Way to Go? *Frontiers in Neuroengineering*, 4, 2012. (Cited on pages 4, 5, 11, 30, 31, 37, 61, 62, 63, and 162.)

[29] Thomas Lochmatter. *Bio-Inspired and Probabilistic Algorithms for Distributed Odor Source Localization using Mobile Robots*. PhD thesis, École polytechnique fédérale de Lausanne (EPFL): Lausanne, Switzerland, 2010. (Cited on pages 4, 18, 30, 44, and 61.)

[30] Achim Lilienthal and Tom Duckett. Experimental analysis of gas-sensitive Braitenberg vehicles. *Advanced Robotics*, 18(8):817–834, 2004. (Cited on pages 4 and 28.)

[31] Victor Hernandez Bennetts, Achim J Lilienthal, Ali Abdul Khaliq, V\ictor Pomareda Sesé, and Marco Trincavelli. Gasbot: A mobile robotic platform for methane leak detection and emission monitoring. In *Workshop on Robotics for Environmental Monitoring at the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012. (Cited on pages 4 and 37.)

[32] Achim Lilienthal, Marco Trincavelli, and Erik Schaffernicht. It's always smelly around here! Modeling the Spatial Distribution of Gas Detection Events with BASED Grid Maps. In *15th International Symposium on Olfaction and Electronic Nose (ISOEN 2013), Daegu, Korea, 2-5 July*, volume 27, 2013. (Cited on page 4.)

[33] Matteo Reggente, Alessio Mondini, Gabriele Ferri, Barbara Mazzolai, Alessandro Manzi, Matteo Gabelletti, Paolo Dario, and Achim J. Lilienthal. The DustBot system: Using mobile robots to monitor pollution in pedestrian area. *Chemical Engineering Transactions*, 23:273–278, 2010. (Cited on page 5.)

[34] Hiroshi Ishida, Yuta Wada, and Haruka Matsukura. Chemical sensing in robotic applications: A review, 2012. (Cited on pages 5 and 22.)

[35] Matthew Dunbabin and Lino Marques. Robots for environmental monitoring: Significant advancements and applications. *IEEE Robotics and Automation Magazine*, 19(1):24–39, 2012. (Cited on page 5.)

[36] R. Andrew Russell. Robotic location of underground chemical sources. *Robotica*, 22(1):109–115, 2004. (Cited on pages 5 and 40.)

[37] R. Andrew Russell. A ground-penetrating robot for underground chemical source location. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, pages 1879–1884, 2005. (Cited on pages 5 and 16.)

[38] R. Andrew Russell. CRABOT: A biomimetic burrowing robot designed for underground chemical source location. *Advanced Robotics*, 25(1-2):119–134, 2011. (Cited on page 5.)

[39] R. Camilli and H. F. Hemond. NEREUS/Kemonaut, a mobile autonomous underwater mass spectrometer. *TrAC - Trends in Analytical Chemistry*, 2004. (Cited on page 5.)

[40] J.a. Farrell, Shuo Pang Shuo Pang, and Wei Li Wei Li. Chemical plume tracing via an autonomous underwater vehicle. *IEEE Journal of Oceanic Engineering*, 30(2):428–442, 2005. (Cited on pages 5, 18, and 45.)

[41] T R Consi, F Grasso, D Mountain, and J Atema. Explorations of turbulent odor plumes with an autonomous underwater robot. *The Biological Bulletin*, 189(2):231–232, 1995. (Cited on pages 5 and 15.)

[42] Mohammad A. Khan, Mark A. Zondlo, and David J. Lary. Open-Path Greenhouse Gas Sensor for UAV applications. *Conference on Lasers and Electro-Optics 2012*, 1:JTh1L.6, 2012. (Cited on page 5.)

[43] Elena S F Berman, Matthew Fladeland, Jimmy Liem, Richard Kolyer, and Manish Gupta. Greenhouse gas analyzer for measurements of carbon dioxide, methane, and water vapor aboard an unmanned aerial vehicle. *Sensors and Actuators, B: Chemical*, 169:128–135, 2012. (Cited on page 5.)

[44] T. Watai, T. Machida, N. Ishizaki, G. Inoue, T. Watai, T. Machida, N. Ishizaki, and G. Inoue. A Lightweight Observation System for Atmospheric Carbon Dioxide Concentration Using a Small Unmanned Aerial Vehicle. *Journal of Atmospheric and Oceanic Technology*, 23(5):700–710, 2006. (Cited on pages 5 and 33.)

[45] T. Xie, Rui Liu, R. T. Hai, Q. H. Hu, and Q. Lu. UAV platform based atmospheric environmental emergency monitoring system design. *Journal of Applied Sciences*, 13:1289–1296, 2013. (Cited on page 5.)

[46] C. E. Corrigan, G. C. Roberts, M. V. Ramana, D. Kim, and V. Ramanathan. Capturing vertical profiles of aerosols and black carbon over the Indian Ocean using autonomous unmanned aerial vehicles. *Atmos. Chem. Phys.*, 8(3):737–747, 2008. (Cited on page 5.)

[47] M Rossi and D Brunelli. Autonomous Gas Detection and Mapping With Unmanned Aerial Vehicles. *IEEE Transactions on Instrumentation and Measurement*, 65(4):765–775, 2016. (Cited on pages 5 and 30.)

[48] Chih Chung Yuan Chang, Jia Lin Wang, Chih Chung Yuan Chang, Mao Chang Liang, and Ming Ren Lin. Development of a multicopter-carried whole air sampling apparatus and its applications in environmental studies. *Chemosphere*, 144:484–492, 2016. (Cited on page 5.)

[49] M Carrozzo, S De Vito, E Esposito, M Salvato, F Formisano, E Massera, G Di Francia, P Delli Veneri, M Iadaresta, and A Mennella. UAV Intelligent Chemical Multisensor Payload for Networked and Impromptu Gas Monitoring Tasks. In *2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace), Rome, Italy, 20-22 June*, pages 112–116. IEEE, 2018. (Cited on pages 5 and 30.)

[50] A. J S McGonigle, A. Aiuppa, G. Giudice, G. Tamburello, A. J. Hodson, and S. Gurrieri. Unmanned aerial vehicle measurements of volcanic carbon dioxide fluxes. *Geophysical Research Letters*, 35(6), 2008. (Cited on page 5.)

[51] Patrick P. Neumann. *Gas source localization and gas distribution mapping with a micro-drone*. PhD thesis, University of Berlin, 2013. (Cited on page 5.)

[52] H. Shinohara. Composition of volcanic gases emitted during repeating Vulcanian eruption stage of Shinmoedake, Kirishima volcano, Japan. *Earth, Planets and Space*, 65(6):667–675, 2013. (Cited on page 5.)

[53] Toshiya Mori, Takeshi Hashimoto, Akihiko Terada, Mitsuhiro Yoshimoto, Ryunosuke Kazahaya, Hiroshi Shinohara, and Ryo Tanaka. Volcanic plume measurements using a UAV for the 2014 Mt. Ontake eruption the Phreatic Eruption of Mt. Ontake Volcano in 2014 5. Volcanology. *Earth, Planets and Space*, 68(1), 2016. (Cited on page 5.)

[54] Julian Rüdiger, Jan Lukas Tirpitz, J. Maarten De Moor, Nicole Bobrowski, Alexandra Gutmann, Marco Liuzzo, Martha Ibarra, and Thorsten Hoffmann. Implementation of electrochemical, optical and denuder-based sensors and sampling techniques on UAV for volcanic gas measurements: Examples from Masaya, Turrialba and Stromboli volcanoes. *Atmospheric Measurement Techniques*, 11(4):2441–2457, 2018. (Cited on page 5.)

[55] G. Astuti, G. Giudice, D. Longo, C. D. Melita, G. Muscato, and A. Orlando. An overview of the "volcan project": An UAS for exploration of volcanic environments. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 2009. (Cited on page 5.)

[56] Patrick P Neumann, Harald Kohlhoff, Dino Hüllmann, Achim J Lilienthal, and Martin Kluge. Bringing Mobile Robot Olfaction to the next dimension—UAV-based remote sensing of gas clouds and source localization. In *2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, Singapore, 29 May-3 June*, pages 3910–3916. IEEE, 2017. (Cited on pages 5, 32, and 37.)

[57] Levi M. Golston, Nicholas F. Aubut, Michael B. Frish, Shuting Yang, Robert W. Talbot, Christopher Gretencord, James McSpiritt, and Mark A. Zondlo. Natural gas fugitive leak detection using an unmanned aerial vehicle: Localization and quantification of emission rate. *Atmosphere*, 9:333, 2018. (Cited on page 5.)

[58] Wolfgang Krüll, Robert Tobera, Ingolf Willms, Helmut Essen, and Nora Von Wahl. Early forest fire detection and verification using optical smoke, gas and microwave sensors. *Procedia Engineering*, 45(December 2012):584–594, 2012. (Cited on page 5.)

[59] Luis Merino, Fernando Caballero, J. R. Martínez-de Dios, Joaquín Ferruz, and Aníbal Ollero. A cooperative perception system for multiple UAVs: Application to automatic detection of forest fires. *Journal of Field Robotics*, 23:165–184, 2006. (Cited on page 5.)

[60] Johannes Pfeifer, Raghav Khanna, C Dragos, Marija Popovic, Enric Galceran, Norbert Kirchgessner, Achim Walter, Roland Siegwart, and Frank Liebisch. Towards automatic UAV data interpretation for precision farming. In *CIGR-AgEng conference. Aarhus, Denmark*, 2016. (Cited on page 5.)

[61] Theerapat Pobkrut, Tanthip Eamsa-Ard, and Teerakiat Kerdcharoen. Sensor drone for aerial odor mapping for agriculture and security services. *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2016, Chiang Mai, Thailand, 28 June–1 July*, pages 1–5, 2016. (Cited on page 5.)

[62] Juan Jesús Roldán, Guillaume Joossen, David Sanz, Jaime del Cerro, and Antonio Barrientos. Mini-UAV based sensory system for measuring environmental variables in greenhouses. *Sensors (Switzerland)*, 15(2):3334–3350, 2015. (Cited on page 5.)

[63] M. Lega and R. M.A. Napoli. A new approach to solid waste landfills aerial monitoring. *WIT Transactions on Ecology and the Environment*, 109:193–199, 2008. (Cited on page 5.)

[64] Grant Allen, Peter Hollingsworth, Khristopher Kabbabe, Joseph R. Pitt, Mohammed I. Mead, Samuel Illingworth, Gareth Roberts, Mark Bourn, Dudley E. Shallcross, and Carl J. Percival. The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots, 2018. (Cited on page 5.)

[65] Bara J. Emran, Dwayne D. Tannant, and Homayoun Najjaran. Low-altitude aerial methane concentration mapping. *Remote Sensing*, 9:823, 2017. (Cited on page 5.)

[66] Kai Daniel, Bjoern Dusza, Andreas Lewandowski, and Christian Wietfeld. Airshield: A system-of-systems muav remote sensing architecture for disaster response. *2009 3rd Annual IEEE Systems Conference, Vancouver, BC, Canada, 23–26 March*, 15:196–200, 2009. (Cited on page 5.)

[67] Miguel Alvarado, Felipe Gonzalez, Andrew Fletcher, and Ashray Doshi. Towards the development of a low cost airborne sensing system to monitor dust particles after blasting at open-pit mine sites. *Sensors (Switzerland)*, 15:19667–19687, 2015. (Cited on page 5.)

[68] Tommaso Villa, Felipe Gonzalez, Branka Miljievic, Zoran Ristovski, and Lidia Morawska. An Overview of Small Unmanned Aerial Vehicles for Air Quality Measurements: Present Applications and Future Prospectives. *Sensors*, 16(7):1072, 2016. (Cited on page 5.)

[69] Gonzalo Pajares. Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs). *Photogrammetric Engineering & Remote Sensing*, 81(4):281–330, 2015. (Cited on page 5.)

[70] Patrick P. Neumann, Victor Hernandez Bennetts, Achim J. Lilienthal, Matthias Bartholmai, and Jochen H. Schiller. Gas source localization with a micro-drone using bio-inspired and particle filter-based algorithms. *Advanced Robotics*, 27(9):725–738, jun 2013. (Cited on pages 6, 16, 17, and 18.)

[71] Bing Luo, Qing Hao Meng, Jia Ying Wang, and Ming Zeng. A Flying Odor Compass to Autonomously Locate the Gas Source, 2017. (Cited on pages 6 and 24.)

[72] Maurizio Rossi and Davide Brunelli. Gas Sensing on Unmanned Vehicles: Challenges and Opportunities. In *2017 New Generation of CAS (NGCAS), Genova, Italy, 6–9 September*, number Dii, pages 117–120, 2017. (Cited on pages 6, 8, and 162.)

[73] M. Hassanalian and A. Abdelkefi. Classifications, applications, and design challenges of drones: A review. *Progress in Aerospace Sciences*, 91:99–131, 2017. (Cited on page 7.)

[74] Javier Burgués, Victor Hernández, Achim J Lilienthal, and Santiago Marco. Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping. *Sensors*, 19(3):478, 2019. (Cited on page 8.)

[75] Hossain Mohammad Fahad, Hiroshi Shiraki, Matin Amani, Chuchu Zhang, Vivek Srinivas Hebbar, Wei Gao, Hiroki Ota, Mark Hettick, Daisuke Kiriya, Yu-ze Chen, Yu-lun Chueh, and Ali Javey. Room temperature multiplexed gas sensing using chemical-sensitive 3 . 5-nm-thin silicon transistors. *Science advances*, 3(March):1–9, 2017. (Cited on pages 8 and 162.)

[76] Oliver Dunkley, Jakob Engel, Jürgen Sturm, and Daniel Cremers. Visual-Inertial Navigation for a Camera-Equipped 25 g Nano-Quadrotor. In *IROS Aerial Open Source Robotics Workshop, Chicago, IL, USA, 14–18 September*, pages 4–5, 2014. (Cited on page 9.)

[77] James A. Preiss, Wolfgang Honig, Gaurav S. Sukhatme, and Nora Ayanian. Crazyswarm: A large nano-quadcopter swarm. In *2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, Singapore, 29 May–3 June*, pages 3299–3304, 2017. (Cited on page 9.)

[78] Zahid Farid, Rosdiadee Nordin, and Mahamod Ismail. Recent advances in wireless indoor localization techniques and system, 2013. (Cited on page 9.)

[79] Victor L Streeter, E Benjamin Wylie, and Keith W Bedford. Fluid mechanics. WCB, 1998. (Cited on page 10.)

[80] E. U. Finlayson, A. J. Gadgil, T. L. Thatcher, and R. G. Sextro. Pollutant dispersion in a large indoor space. Part 2: Computational fluid dynamics predictions and comparison with a scale model experiment for isothermal flow. *Indoor Air*, 14(4):272–283, 2004. (Cited on page 10.)

[81] A. Lilienthal, A. Zell, M. Wandel, and U. Weimar. Sensing odour sources in indoor environments without a constant airflow by a mobile robot. In *2001 IEEE International Conference on Robotics and Automation (ICRA), Seoul, Korea, 21–26 May*, volume 4, pages 1–6, 2001. (Cited on pages 10 and 29.)

[82] Achim Lilienthal and Tom Duckett. Building gas concentration gridmaps with a mobile robot. *Robotics and Autonomous Systems*, 48(1):3–16, 2004. (Cited on pages 10 and 30.)

[83] Firooz Rasouli and Ted A. Williams. Application of dispersion modeling to indoor gas release scenarios. *Journal of the Air and Waste Management Association*, 45(3):191–195, 1995. (Cited on page 10.)

[84] D Bruce Turner. *Workbook of atmospheric dispersion estimates: an introduction to dispersion modeling*. CRC press: Boca Raton, FL, USA, 1994. (Cited on pages 11 and 20.)

[85] E A Bakkum and N J Duijm. Vapour cloud dispersion. *Yellow Book, CPR E*, 14:3, 1997. (Cited on pages 11 and 15.)

[86] O. G. Sutton. The problem of diffusion in the lower atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 73(317-318):257–281, 1947. (Cited on page 11.)

[87] D. R. Webster and M. J. Weissburg. Chemosensory guidance cues in a turbulent chemical odor plume. *Limnology and Oceanography*, 46(5):1034–1047, 2001. (Cited on pages 11 and 15.)

[88] John P Crimaldi, Megan B Wiley, and Jeffrey R Koseff. The relationship between mean and instantaneous structure in turbulent passive scalar plumes. *Journal of Turbulence*, 3:N14, 2002. (Cited on page 11.)

[89] J O Hinze. Turbulence, (1975). *New York*, 1959. (Cited on pages 11 and 20.)

[90] Jay A. Farrell, John Murlis, Xuezhu Long, Wei Li, and Ring T. Cardé. Filament-based atmospheric dispersion model to achieve short time-scale structure of odor plumes. *Environmental Fluid Mechanics*, 2(1-2):143–169, 2002. (Cited on pages 11, 12, and 20.)

[91] J. D. Wilson and B. L. Sawford. Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. *Boundary-Layer Meteorology*, 78(1-2):191–210, 1996. (Cited on page 11.)

[92] Joel H Ferziger and Milovan Peric. *Computational methods for fluid dynamics*. Springer Science & Business Media, 2012. (Cited on page 12.)

[93] Peter Constantin and Ciprian Foias. *Navier-stokes equations*. University of Chicago Press, 1988. (Cited on page 12.)

[94] A Stamou and I Katsiris. Verification of a CFD model for indoor airflow and heat transfer. *Building and Environment*, 41(9):1171–1181, 2006. (Cited on page 13.)

[95] David C Wilcox and Others. *Turbulence modeling for CFD*, volume 2. DCW industries La Canada, CA, 1998. (Cited on page 13.)

[96] S M Tauseef, D Rashtchian, and S A Abbasi. CFD-based simulation of dense gas dispersion in presence of obstacles. *Journal of Loss Prevention in the Process Industries*, 24(4):371–376, 2011. (Cited on page 13.)

[97] Marco Pontiggia, Marco Derudi, M Alba, Marco Scaioni, and Renato Rota. Hazardous gas releases in urban areas: assessment of consequences through CFD modelling. *Journal of Hazardous Materials*, 176(1-3):589–596, 2010. (Cited on page 13.)

[98] Gonçalo Cabrita, Pedro Sousa, and Lino Marques. Player / Stage Simulation of Olfactory Experiments. In *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings*, 2010. (Cited on pages 13, 53, and 60.)

[99] Javier Monroy, Victor Hernandez-Bennetts, Han Fan, Achim Lilienthal, and Javier Gonzalez-Jimenez. GADEN: A 3D gas dispersion simulator for mobile robot olfaction in realistic environments. *Sensors (Switzerland)*, 17(7):1–16, 2017. (Cited on page 13.)

[100] Kai Song, Qi Liu, and Qi Wang. Olfaction and hearing based mobile robot navigation for odor/sound source search. *Sensors*, 11(2):2129–2154, 2011. (Cited on pages 14, 40, and 59.)

[101] Owen Holland and Chris Melhuish. Some adaptive movements of animats with single symmetrical sensors. *From Animals to Animats*, 4(6):55–64, 1996. (Cited on page 14.)

[102] Benjamin Russell and Herschel Rabitz. Common foundations of optimal control across the sciences: Evidence of a free lunch. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2017. (Cited on page 14.)

[103] R Andrew Russell. Chemical source location and the robomole project. In *Proceedings Australian Conference on Robotics and Automation*. Citeseer, 2003. (Cited on page 14.)

[104] W F J M Engelhard. Methods for the Calculation of Physical Effects. *Yellow Book*, pages 6.1–6.132, 1997. (Cited on page 15.)

[105] Giulio Sandini, Giovanni Lucarini, and Marco Varoli. Gradient driven self-organizing systems. In *IROS*, volume 93, pages 24–26. Citeseer, 1993. (Cited on page 15.)

[106] Agenor Mafra-Neto and Ring T Cardé. Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. *Nature*, 369(6476):142, 1994. (Cited on page 15.)

[107] M. J. Weissburg, D. B. Dusenbery, H. Ishida, J. Janata, T. Keller, P. J.W. Roberts, and D. R. Webster. A multidisciplinary study of spatial and temporal scales containing information in turbulent chemical plume tracking. *Environmental Fluid Mechanics*, 2(1-2):65–94, 2002. (Cited on page 15.)

[108] Mark A. Willis and Edmund A. Arbas. Odor-modulated upwind flight of the sphinx moth, Manduca sexta L. *Journal of Comparative Physiology A*, 169(4):427–440, 1991. (Cited on page 15.)

[109] Edwin A. Cowen and Keith B. Ward. Chemical plume tracing. *Environmental Fluid Mechanics*, 2(1-2):1–7, 2002. (Cited on page 15.)

[110] A.J. Rutkowski, S. Edwards, M.A. Willis, R.D. Quinn, and G.C. Causey. A robotic platform for testing moth-inspired plume tracking strategies. *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, pages 3319–3324 Vol.4, 2004. (Cited on pages 16 and 18.)

[111] H. Ishida, K. Hayashi, M. Takakusaki, T. Nakamoto, T. Moriizumi, and R. Kanzaki. Odour-source localization system mimicking behaviour of silkworm moth, 1995. (Cited on pages 15, 22, and 27.)

[112] Pawel Pyk, Sergi Bermúdez i Badia, Ulysses Bernardet, Philipp Knüsel, Mikael Carlsson, Jing Gu, Eric Chanie, Bill S. Hansson, Tim C. Pearce, and Paul F.M.J. Verschure. An artificial moth: Chemical source localization using a robot based neuronal model of moth optomotor anemotactic search, 2006. (Cited on pages 15, 18, 19, 58, and 59.)

[113] Adam T. Hayes, Alcherio Martinoli, and Rodney M. Goodman. Distributed odor source localization. *IEEE Sensors Journal*, 2(3):260–271, 2002. (Cited on pages 15, 16, 17, 30, 43, and 45.)

[114] Ra Russell. Using volatile chemicals to help locate targets in complex environments. *Proceedings of the . . .* , 2000. (Cited on pages 15 and 44.)

[115] R. Andrew Russell, Alireza Bab-Hadiashar, Rod L. Shepherd, and Gordon G. Wallace. A comparison of reactive robot chemotaxis algorithms. *Robotics and Autonomous Systems*, 45(2):83–97, 2003. (Cited on pages 15 and 16.)

[116] Matthew E. Staymates, William A. MacCrehan, Jessica L. Staymates, Roderick R. Kunz, Thomas Mendum, Ta Hsuan Ong, Geoffrey Geurtsen, Greg J. Gillen, and Brent A. Craven. Biomimetic Sniffing Improves the Detection Performance of a 3D Printed Nose of a Dog and a Commercial Trace Vapor Detector. *Scientific Reports*, 6, 2016. (Cited on page 16.)

[117] Hiroshi Ishida, Takashi Ushiku, Shigeki Toyama, Haruki Taniguchi, and Toyosaka Moriizumi. Mobile robot path planning using vision and olfaction to search for a gas source. In *Proceedings of IEEE Sensors*, 2005. (Cited on pages 16 and 17.)

[118] Ji-Gong Li, Qing-Hao Meng, Yang Wang, and Ming Zeng. Odor source localization using a mobile robot in outdoor airflow environments with a particle filter algorithm. *Autonomous Robots*, 30(3):281–292, 2011. (Cited on pages 16, 18, and 40.)

[119] A. H. Purnamadjaja and R Andrew Russell. Congregation Behaviour in a Robot Swarm Using Pheromone Communication. *Proceeding of the Australian Conference on Robotics and Automation*, pages 1–7, 2005. (Cited on pages 16, 57, and 59.)

[120] Sepideh Pashami, Achim J. Lilienthal, and Marco Trincavelli. Detecting changes of a distant gas source with an array of MOX gas sensors. *Sensors (Switzerland)*, 12(12):16404–16419, 2012. (Cited on page 16.)

[121] Michael Schmuker, Viktor Bahr, and Ramón Huerta. Exploiting plume structure to decode gas source distance using metal-oxide gas sensors. *Sensors and Actuators, B: Chemical*, 235:636–646, 2016. (Cited on pages 16, 22, 44, 53, 56, and 162.)

[122] Gabriele Ferri, Emanuele Caselli, Virgilio Mattoli, Alessio Mondini, Barbara Mazzolai, and Paolo Dario. SPIRAL: A novel biologically-inspired algorithm for gas/odor source localization in an indoor environment with no strong airflow. *Robotics and Autonomous Systems*, 57(4):393–402, 2009. (Cited on pages 17, 18, 28, 29, 40, 41, 45, 57, and 60.)

[123] Nicole Voges, Antoine Chaffiol, Philippe Lucas, and Dominique Martinez. Reactive Searching and Infotaxis in Odor Source Localization. *PLoS Computational Biology*, 10(10), 2014. (Cited on pages 17 and 22.)

[124] Thomas Lochmatter and Alcherio Martinoli. Tracking Odor Plumes in a Laminar Wind Field with Bio-inspired Algorithms. In *Springer Tracts in Advanced Robotics*, volume 54, pages 473–482, 2009. (Cited on pages 17 and 18.)

[125] J. H. Belanger and E. A. Arbas. Behavioral strategies underlying pheromone-modulated flight in moths: Lessons from simulation studies. *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*, 183(3):345–360, 1998. (Cited on page 18.)

[126] Lino Marques, Urbano Nunes, Anıbal T. de Almeida, Kai Song, Qi Liu, and Qi Wang. Olfaction-based mobile robot navigation. *Thin Solid Films*, 418(1):51–58, 2002. (Cited on pages 18, 20, 21, 35, and 44.)

[127] David J. Harvey, Tien Fu Lu, and Michael A. Keller. Comparing insect-inspired chemical plume tracking algorithms using a mobile robot. *IEEE Transactions on Robotics*, 24(2):307–317, 2008. (Cited on page 18.)

[128] Hiroshi Ishida, Hidenao Tanaka, Haruki Taniguchi, and Toyosaka Moriizumi. Mobile robot navigation using vision and olfaction to search for a gas/odor source. *Autonomous Robots*, 20(3):231–238, 2006. (Cited on page 18.)

[129] Luís Osório, Gonçalo Cabrita, and Lino Marques. Mobile Robot Odor Plume Tracking Using Three Dimensional Information. *European Conference on Mobile Robots*, pages 1–6, 2011. (Cited on pages 18, 22, 23, 24, 40, 44, and 57.)

[130] Achim J. Lilienthal, Tom Duckett, Hiroshi Ishida, and Felix Werner. Indicators of gas source proximity using metal oxide sensors in a turbulent environment. In *Proceedings of the First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006, BioRob 2006*, volume 2006, pages 733–738, 2006. (Cited on page 18.)

[131] Qing-hao Meng Jun-cai Li, Fei Li Fei Li, Ming Zeng Ming Zeng, Q. H Meng, and Jun-Cai Li Jun-Cai Li. Mobile Robots Odor Localization with an Improved Ant Colony Algorithm. In *2006 IEEE International Conference on Robotics and Biomimetics*, number 1, pages 959–964, 2006. (Cited on page 18.)

[132] Goncalo Cabrita and Lino Marques. Divergence-based odor source declaration. *2013 9th Asian Control Conference, ASCC 2013*, 2013. (Cited on page 18.)

[133] Yu Teng Wei, Qing Hao Meng, Ya Qi Jing, Ying Jie Liu, and Ming Zeng. A Portable Odor-Tracing Instrument. *IEEE Transactions on Instrumentation and Measurement*, 2016. (Cited on pages 18, 19, and 24.)

[134] D. J. Thomson. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180(2):529–556, 1987. (Cited on pages 20 and 21.)

[135] G. Falkovich, K. Gawędzki, and M. Vergassola. Particles and fields in fluid turbulence. *Reviews of Modern Physics*, 73(4):913–975, 2001. (Cited on pages 20 and 21.)

[136] Daniel Grünbaum and Mark A. Willis. Spatial memory-based behaviors for locating sources of odor plumes. *Movement Ecology*, 3(1), 2015. (Cited on page 20.)

[137] Hiroshi Ishida, Takamichi Nakamoto, and Toyosaka Moriizumi. Remote sensing of gas/odor source location and concentration distribution using mobile system. *Sensors and Actuators, B: Chemical*, 49:52–57, 1998. (Cited on pages 20 and 59.)

[138] Takashi Ushiku, Nozomi Satoh, Hiroshi Ishida, and Shigeki Toyama. Estimation of gas-source location using gas sensors and ultrasonic anemometer. In *Proceedings of IEEE Sensors*, pages 420–423, 2006. (Cited on pages 20 and 21.)

[139] Yuichiro Fukazawa and Hiroshi Ishida. Estimating gas-source location in outdoor environment using mobile robot equipped with gas sensors and anemometer. *Proceedings of IEEE Sensors*, pages 1721–1724, 2009. (Cited on pages 20 and 21.)

[140] Jorge Sánchez-Sosa, Juan Castillo-Mixcóatl, Georgina Beltrán-Pérez, and Severino Muñoz-Aguirre. An Application of the Gaussian Plume Model to Localization of an Indoor Gas Source with a Mobile Robot. *Sensors*, 18(12):4375, 2018. (Cited on pages 20, 21, and 40.)

[141] Jay A. Farrell, Shuo Pang, and Wei Li. Plume Mapping via Hidden Markov Methods. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 33(6):850–863, 2003. (Cited on pages 21 and 44.)

[142] Shuo Pang and Jay A. Farrell. Chemical plume source localization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 36(5):1068–1080, 2006. (Cited on pages 21, 43, and 44.)

[143] Víctor Pomareda, Rudys Magrans, J.M. Juan M. Jiménez-Soto, Dani Martínez, Marcel Tresánchez, Javier Burgués, Jordi Palacín, and Santiago Marco. Chemical source localization fusing concentration information in the presence of chemical background noise. *Sensors (Switzerland)*, 17(4), 2017. (Cited on pages 21, 34, and 44.)

[144] Massimo Vergassola, Emmanuel Villermaux, and Boris I Shraiman. 'Infotaxis' as a strategy for searching without gradients. *Nature*, 445(7126):406–409, 2007. (Cited on pages 21, 43, and 45.)

[145] Mikel Vuka, Erik Schaffernicht, Michael Schmuker, Victor Hernandez Bennetts, Francesco Amigoni, and Achim J. Lilienthal. Exploration and localization of a gas source with MOX gas sensors on a mobile robot-A Gaussian regression bout amplitude approach. *ISOEN 2017 - ISOCS/IEEE International Symposium on Olfaction and Electronic Nose, Proceedings*, pages 3–5, 2017. (Cited on pages 22 and 44.)

[146] Hiroshi Ishida, Akito Kobayashi, Takamichi Nakamoto, and Toyosaka Moriizumi. Three-dimensional odor compass. *IEEE Transactions on Robotics and Automation*, 15(2):251–257, 1999. (Cited on pages 22, 25, and 40.)

[147] H. Ishida, K. Yoshikawa, and T. Moriizumi. Three-dimensional gas-plume tracking using gas sensors and ultrasonic anemometer. *Proceedings of IEEE Sensors, 2004.*, pages 1175–1178, 2004. (Cited on pages 22, 23, and 25.)

[148] R. Andrew Russell. Tracking chemical plumes in 3-dimensions. *2006 IEEE International Conference on Robotics and Biomimetics, ROBIO 2006*, pages 31–36, 2006. (Cited on pages 22, 23, 25, 26, 44, 59, 60, and 62.)

[149] Matteo Reggente and Achim J. Lilienthal. The 3D-kernel DM+V/W algorithm: Using wind information in three dimensional gas distribution modelling with a mobile robot. *Proceedings of IEEE Sensors*, pages 999–1004, 2010. (Cited on pages 22 and 60.)

[150] Hiroshi Ishida. Blimp robot for three-dimensional gas distribution mapping in indoor environment. In *AIP Conference Proceedings*, volume 1137, pages 61–64, 2009. (Cited on pages 23 and 24.)

[151] Jorge M Soares, Ali Marjovi, Jonathan Giezendanner, Anil Kodiyan, A Pedro Aguiar, M Pascoal, and Alcherio Martinoli. Towards 3-D Distributed Odor Source Localization : An Extended Graph-Based Formation Control Algorithm for Plume Tracking. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016. (Cited on page 23.)

[152] Kok Seng Eu and Kian Meng Yap. Chemical plume tracing: A three-dimensional technique for quadrotors by considering the altitude control of the robot in the casting stage. *International Journal of Advanced Robotic Systems*, 15(1), 2018. (Cited on pages 25 and 26.)

[153] Adam J. Rutkowski, Roger D. Quinn, and Mark A. Willis. A sensor fusion approach to odor source localization inspired by the pheromone tracking behavior of moths. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 4873–4878, 2007. (Cited on pages 26 and 27.)

[154] Achim Lilienthal, Denis Reimann, and Andreas Zell. Gas Source Tracing with a Mobile Robot Using an Adapted Moth Strategy. *Autonome Mobile Systeme*, pages 150–160, 2003. (Cited on pages 27, 28, 40, and 44.)

[155] J. Atema. Chemical signals in the marine environment: dispersal, detection, and temporal signal analysis. *Proceedings of the National Academy of Sciences*, 92(1):62–66, 1995. (Cited on page 29.)

[156] A. Farah and T Duckett. Reactive Localisation of an Odour Source by a learning Mobile Robot. *Second Swedish Workshop on Autonomous Robotics, Stockholm, Sweden, 11–12 October*, pages 29–38, 2002. (Cited on page 29.)

[157] Michael R Wandel, Achim Lilienthal, Tom Duckett, Udo Weimar, and Andreas Zell. Gas Distribution in Unventilated Indoor Environments Inspected by a Mobile Robot. *Proceedings of the IEEE International Conference on Advanced Robotics (ICAR 2003)*, pages 507–512, 2003. (Cited on page 30.)

[158] Anies Hannawati Purnamadjaja and R. Andrew Russell. Pheromone communication in a robot swarm: Necrophoric bee behaviour and its replication. *Robotica*, 23(6):731–742, 2005. (Cited on pages 30, 40, and 59.)

[159] Amy Loutfi, Silvia Coradeschi, Achim J. Lilienthal, and Javier Gonzalez. Gas distribution mapping of multiple odour sources using a mobile robot. *Robotica*, 27(2):311–319, 2009. (Cited on pages 30, 35, and 40.)

[160] Matteo Reggente and Achim J. Lilienthal. Using local wind information for gas distribution mapping in outdoor environments with a mobile robot. In *Proceedings of IEEE Sensors*, pages 1715–1720, 2009. (Cited on page 30.)

[161] Yuta Wada, Marco Trincavelli, Yuichiro Fukazawa, and Hiroshi Ishida. Collecting a Database for Studying Gas Distribution Mapping and Gas Source Localization with Mobile Robots. *International Conference on Advanced Mechatronics*, pages 183—-188, 2010. (Cited on pages 30 and 34.)

[162] Ali Marjovi, João Gonçalo Nunes, Lino Marques, and Aníbal De Almeida. Multi-robot exploration and fire searching. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, 2009. (Cited on page 30.)

[163] Mirbek Turduev, Gonçalo Cabrita, Murat Kirtay, Veysel Gazi, and Lino Marques. Experimental studies on chemical concentration map building by a multi-robot system using bio-inspired algorithms. *Autonomous Agents and Multi-Agent Systems*, 28(1):72–100, 2014. (Cited on pages 30 and 40.)

[164] Ali Marjovi and Lino Marques. Multi-robot odor distribution mapping in realistic time-variant conditions. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 3720–3727, 2014. (Cited on page 30.)

[165] Patrick P. Neumann, Sahar Asadi, Achim J. Lilienthal, Matthias Bartholmai, and Jochen H. Schiller. Autonomous gas-sensitive microdrone: Wind vector estimation and gas distribution mapping. *IEEE Robotics and Automation Magazine*, 19(1):50–61, 2012. (Cited on pages 30 and 162.)

[166] Bing Luo, Qing Hao Meng, Jia Ying Wang, Biao Sun, and Ying Wang. Three-dimensional gas distribution mapping with a micro-drone. *Chinese Control Conference (CCC), Hangzhou, China, 28–30 July*, 2015-Septe:6011–6015, 2015. (Cited on pages 30, 31, and 40.)

[167] Hiroshi Ishida, Takamichi Nakamoto, and Toyosaka Moriizumi. Remote Sensing and Localization of Gas/Odor Source and Distribution Using Mobile Sensing System. *International Conference on Solid State Sensors and Actuators*, 49(1):4–7, 1997. (Cited on pages 31 and 40.)

[168] Dominique Martinez and Laurent Perrinet. Cooperation between vision and olfaction in a koala robot. *Report on the 2002 Workshop on Neuromorphic Engineering*, pages 51–53, 2002. (Cited on page 31.)

[169] A. Loutfi, S. Coradeschi, L. Karlsson, and M. Broxvall. Putting olfaction into action: using an electronic nose on a multi-sensing mobile robot. *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, 1:337–342, 2004. (Cited on page 31.)

[170] Gideon Kowadlo, David Rawlinson, R. Andy Russell, and Ray Jarvis. Bi-modal search using complementary sensing (olfaction/vision) for odour source localisation. In *IEEE International Conference on Robotics and Automation (ICRA), Orlando, FL, USA, 15–19 May*, volume 2006, pages 2041–2046, 2006. (Cited on pages 31 and 40.)

[171] Marco Trincavelli, Matteo Reggente, Silvia Coradeschi, Hiroshi Ishida, Amy Loutfi, and Achim J Lilienthal. Towards environmental monitoring with mobile robots. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 2210–2215, 2008. (Cited on page 31.)

[172] Ping Jiang, Qing Hao Meng, and Ming Zeng. Mobile robot gas source localization via top-down visual attention mechanism and shape analysis. In *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, pages 1818–1823, 2010. (Cited on page 31.)

[173] Victor Hernandez Bennetts, Erik Schaffernicht, Todor Stoyanov, Achim J. Lilienthal, and Marco Trincavelli. Robot assisted gas tomography - Localizing methane leaks in outdoor environments. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 6362–6367, 2014. (Cited on page 31.)

[174] Werner Baetz, Andreas Kroll, and Gero Bonow. Mobile robots with active IR-optical sensing for remote gas detection and source localization. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 2773–2778, 2009. (Cited on page 31.)

[175] G. Bonow and A. Kroll. Gas leak localization in industrial environments using a TDLAS-based remote gas sensor and autonomous mobile robot with the Tri-Max method. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 987–992, 2013. (Cited on page 31.)

[176] H Nanto and J R Stetter. Introduction to Chemosensors. In *Handbook of Machine Olfaction: Electronic Nose Technology*, chapter 4, pages 79–104. John Wiley & Sons, 2004. (Cited on page 32.)

[177] Javier G. Monroy, Jose Luis Blanco, and Javier Gonzalez-Jimenez. Time-variant gas distribution mapping with obstacle information. *Autonomous Robots*, 40(1):1–16, 2016. (Cited on page 34.)

[178] Julian W Gardner and Philip N Bartlett. Performance definition and standardization of electronic noses. *Sensors and Actuators B: Chemical*, 33(1-3):60–67, 1996. (Cited on page 34.)

[179] J. W. Gardner and P. N. Bartlett. Electronic noses. Principles and applications. *Measurement Science and Technology*, 11(7):1087, 1999. (Cited on page 34.)

[180] E. L. Hines, Pascal Boilot, Julian W. Gardner, and Mario A. Gongora. Pattern analysis for electronic noses. In *Handbook of Machine Olfaction*, pages 133–160. John Wiley & Sons, 2003. (Cited on page 34.)

[181] M. Castro, B. Kumar, J. F. Feller, Z. Haddi, A. Amari, and B. Bouchikhi. Novel e-nose for the discrimination of volatile organic biomarkers with an array of carbon nanotubes (CNT) conductive polymer nanocomposites (CPC) sensors. *Sensors and Actuators, B: Chemical*, 159(1):213–219, 2011. (Cited on page 35.)

[182] Marco Trincavelli, Silvia Coradeschi, and Amy Loutfi. Online classification of gases for environmental exploration. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pages 3311–3316, 2009. (Cited on page 35.)

[183] Marco Trincavelli and Amy Loutfi. Feature selection for gas identification with a mobile robot. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 2852–2857, 2010. (Cited on page 35.)

[184] Victor Bennetts, Erik Schaffernicht, Victor Pomareda, Achim Lilienthal, Santiago Marco, and Marco Trincavelli. Combining Non Selective Gas Sensors on a Mobile Robot for Identification and Mapping of Multiple Chemical Compounds. *Sensors*, 14(9):17331–17352, 2014. (Cited on pages 35 and 57.)

[185] Victor Hernandez Bennetts, Erik Schaffernicht, Victor Pomareda Sesé, Achim J. Lilienthal, and Marco Trincavelli. A novel approach for gas discrimination in natural environments with open sampling systems. *Proceedings of IEEE Sensors*, 2014-Decem(December):2046–2049, 2014. (Cited on page 35.)

[186] Ji Gong Li, Biao Sun, Fan Lin Zeng, Jia Liu, Jing Yang, and Li Yang. Experimental study on multiple odor sources mapping by a mobile robot in time-varying airflow environment. In *Chinese Control Conference (CCC), Chengdu, China, 27–29 July*, volume 2016-Augus, pages 6032–6037, 2016. (Cited on pages 35 and 40.)

[187] Javier G. Monroy and Javier Gonzalez-Jimenez. Gas classification in motion: An experimental analysis. *Sensors and Actuators, B: Chemical*, 240:1205–1215, 2017. (Cited on page 35.)

[188] Han Fan, Victor Hernandez Bennetts, Erik Schaffernicht, and Achim J. Lilienthal. A cluster analysis approach based on exploiting density peaks for gas discrimination with electronic noses in open environments. *Sensors and Actuators, B: Chemical*, 2018. (Cited on page 35.)

[189] Jane Hodgkinson and Ralph P. Tatam. Optical gas sensing: A review, 2013. (Cited on page 35.)

[190] FLIR. Gas detection. Technical report, FLIR Systems, 2009. (Cited on page 36.)

[191] Arvind P. Ravikumar, Jingfan Wang, and Adam R. Brandt. Are Optical Gas Imaging Technologies Effective for Methane Leak Detection? *Environmental Science and Technology*, 51(1):718–724, 2017. (Cited on page 36.)

[192] D. F. Swinehart. The Beer-Lambert Law. *Journal of Chemical Education*, 39(7):333, 1962. (Cited on page 37.)

[193] Marco Trincavelli, Victor Hernandez Bennetts, and Achim J. Lilienthal. A least squares approach for learning gas distribution maps from a set of integral gas concentration measurements obtained with a TDLAS sensor. *Proceedings of IEEE Sensors*, 2012. (Cited on page 37.)

[194] Gary Nelson. *Gas mixtures: preparation and control*. CRC Press: Boca Raton, FL, USA, 1992. (Cited on page 38.)

[195] Javier Burgués, Juan Manual Jimenez-Soto, and Santiago Marco. Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models. *Analytica Chimica Acta*, 1013:13–25, 2018. (Cited on page 39.)

[196] TimTaffner Yue Li Michael Bischoff,Bernd Niemeyer. Test Gas Generation fromPure Liquids_An Application_Oriented Overview of Methods in a Nutshell. *International Journal of Chemical Engineering*, 2012, 2011. (Cited on page 39.)

[197] Javier G. Monroy, Achim Lilienthal, Jose Luis Blanco, Javier Gonzalez-Jimenez, and Marco Trincavelli. Calibration of MOX gas sensors in open sampling systems based on Gaussian Processes. In *Proceedings of IEEE Sensors*, 2012. (Cited on page 39.)

[198] Ahmad Shakaff Ali Yeon, Kamarulzaman Kamarudin, Retnam Visvanathan, Syed Muhammad Mamduh Syed Zakaria, Ammar Zakaria, and Latifah Munirah Kamarudin. Gas Source Localization via Behaviour Based Mobile Robot and Weighted Arithmetic Mean. In *IOP Conference Series: Materials Science and Engineering*, 2018. (Cited on page 40.)

[199] Matteo Reggente and Achim J. Lilienthal. Three-dimensional statistical gas distribution mapping in an uncontrolled indoor environment. In *AIP Conference Proceedings*, volume 1137, pages 109–112, 2009. (Cited on page 40.)

[200] Sepideh Pashami, Achim J. Lilienthal, Erik Schaffernicht, and Marco Trincavelli. TREFEX: Trend estimation and change detection in the response of MOX gas sensors. *Sensors (Switzerland)*, 13:7323–7344, 2013. (Cited on page 40.)

[201] Hiroshi Ishida, Gouki Nakayama, Takamichi Nakamoto, and Toyosaka Moriizumi. Controlling a gas/odor plume-tracking robot based on transient responses of gas sensors. *IEEE Sensors Journal*, 5(3):537–545, 2005. (Cited on pages 40, 44, and 45.)

[202] G. Kowadlo and R. a. Russell. Using naïve physics for odor localization in a cluttered indoor environment. *Autonomous Robots*, 20(3):215–230, jun 2006. (Cited on pages 40 and 44.)

[203] J. Fonollosa, L. Fernández, A. Gutiérrez-Gálvez, R. Huerta, and S. Marco. Calibration transfer and drift counteraction in chemical sensor arrays using Direct Standardization. *Sensors and Actuators, B: Chemical*, 236:1044–1053, 2016. (Cited on page 40.)

[204] Javier Gonzalez-Jimenez, Javier G. Monroy, and Jose Luis Blanco. The Multi-Chamber Electronic Nose—An Improved Olfaction Sensor for Mobile Robotics. *Sensors*, 11(6):6145–6464, jan 2011. (Cited on pages 40, 53, 54, and 55.)

[205] Qing Hao Meng, Wei Xing Yang, Yang Wang, and Ming Zeng. Collective odor source estimation and search in time-variant airflow environments using mobile robots. *Sensors*, 11(11):10415–10443, 2011. (Cited on page 40.)

[206] Syed Muhammad Mamduh Syed Zakaria, Retnam Visvanathan, Kamarulzaman Kamarudin, Ahmad Shakaff Ali Yeon, Ali Yeon Md Shakaff, Ammar Zakaria, and Latifah Munirah Kamarudin. Development of a scalable testbed for mobile olfaction verification. *Sensors (Switzerland)*, 15(12):30894–30912, 2015. (Cited on pages 40, 60, and 61.)

[207] L. A. Currie. Detection and quantification limits: Origins and historical overview. *Analytica Chimica Acta*, 391(2):127–134, 1999. (Cited on page 41.)

[208] Paolo Montuschi, Sergei A. Kharitonov, and Peter J. Barnes. Exhaled carbon monoxide and nitric oxide in COPD. *Chest*, 120(2):496–501, 2001. (Cited on page 42.)

[209] Simon Davies, Patrik Spanel, and David Smith. Quantitative analysis of ammonia on the breath of patients in end-stage renal failure. *Kidney International*, 52(1):223–228, 1997. (Cited on page 42.)

[210] Jingying Zhang, Xin Yao, Rongbin Yu, Jianling Bai, Yun Sun, Mao Huang, Ian M Adcock, and Peter J Barnes. Exhaled carbon monoxide in asthmatics: a meta-analysis. *Respiratory research*, 11(1):50, 2010. (Cited on page 42.)

[211] Alejandro C Olivieri. Analytical figures of merit: from univariate to multiway calibration. *Chemical reviews*, 114(10):5358–5378, 2014. (Cited on page 43.)

[212] Avraham Lorber, Klaas Faber, and Bruce R Kowalski. Net Analyte Signal Calculation in Multivariate Calibration. *Analytical Chemistry*, 69(8):1620–1626, 1997. (Cited on page 43.)

[213] M C Ortiz, L A Sarabia, A Herrero, M S Sánchez, M B Sanz, M E Rueda, D Giménez, and M E Meléndez. Capability of detection of an analytical method evaluating false positive and false negative (ISO 11843) with partial least squares. *Chemometrics and Intelligent Laboratory Systems*, 69(1-2):21–33, nov 2003. (Cited on page 43.)

[214] Anita Singh. Multivariate decision and detection limits. *Analytica chimica acta*, 277(2):205–214, 1993. (Cited on page 43.)

[215] Alejandro C. Olivieri, Nicolaas M. Faber, Joan Ferré, Ricard Boqué, John H. Kalivas, and Howard Mark. Uncertainty estimation and figures of merit for multivariate calibration (IUPAC Technical Report). *Pure and Applied Chemistry*, 78(3):633–661, 2006. (Cited on page 43.)

[216] A Muñoz de la Peña, A Espinosa-Mansilla, M I Acedo Valenzuela, H C Goicoechea, and A C Olivieri. Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxycillin and clavulanic acid by stopped-flow kinetic analysis. *Analytica Chimica Acta*, 463(1):75–88, 2002. (Cited on page 43.)

[217] J. Hodgkinson, Q. Shan, and R. D. Pride. Detection of a simulated gas leak in a wind tunnel. *Measurement Science and Technology*, 17(6):1586–1593, 2006. (Cited on page 44.)

[218] Erik Schaffernicht, Marco Trincavelli, and Achim J Lilienthal. Bayesian Spatial Event Distribution Grid Maps for Modeling the Spatial Distribution of Gas Detection Events. *Sensor Letters*, 12(6):1142–1146, 2014. (Cited on page 44.)

[219] Ji-Gong Li Ji-Gong Li, Qing-Hao Meng Qing-Hao Meng, Fei Li Fei Li, Ming Zeng Ming Zeng, and D. Popescu. Mobile robot based odor source localization via particle filter. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 2984–2989, 2009. (Cited on page 44.)

[220] Anies Hannawati Purnamadjaja and R. Andrew Russell. Robotic pheromones: Using temperature modulation in tin oxide gas sensor to differentiate swarm's behaviours. *9th International Conference on Control, Automation, Robotics and Vision, 2006, ICARCV '06*, 2006. (Cited on pages 44 and 49.)

[221] Jacques Nicolas and Anne-Claude Romain. Establishing the limit of detection and the resolution limits of odorous sources in the environment for an array of metal oxide gas sensors. *Sensors and Actuators B: Chemical*, 99(2-3):384–392, may 2004. (Cited on page 45.)

[222] M. Kuske, M. Padilla, A. C. Romain, J. Nicolas, R. Rubio, and S. Marco. Detection of diverse mould species growing on building materials by gas sensor arrays and pattern recognition. *Sensors and Actuators, B: Chemical*, 119(1):33–40, 2006. (Cited on page 45.)

[223] George F Fine, Leon M Cavanagh, Ayo Afonja, and Russell Binions. Metal oxide semi-conductor gas sensors in environmental monitoring. *Sensors*, 10(6):5469–5502, 2010. (Cited on page 45.)

[224] C. Borrego, A. M. Costa, J. Ginja, M. Amorim, M. Coutinho, K. Karatzas, Th Sioumis, N. Katsifarakis, K. Konstantinidis, S. De Vito, E. Esposito, P. Smith, N. Andr??, P. G??rard, L. A. Francis, N. Castell, P. Schneider, M. Viana, M. C. Minguill??n, W. Reimringer, R. P. Otjes, O. von Sicard, R. Pohle, B. Elen, D. Suriano, V. Pfister, M. Prato, S. Dipinto, and M. Penza. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmospheric Environment*, 147(2):246–263, 2016. (Cited on page 45.)

[225] Naomi Funazaki, Akihide Hemmi, Satoshi Ito, Yasukazu Asano, Yukio Yano, Norio Miura, and Noboru Yamazoe. Application of semiconductor gas sensor to quality control of meat freshness in food industry. *Sensors and Actuators: B. Chemical*, 25(1-3):797–800, 1995. (Cited on page 45.)

[226] Noureddine El Barbri, Eduard Llobet, Nezha El Bari, Xavier Correig, and Benachir Bouchikhi. Electronic nose based on metal oxide semiconductor sensors as an alternative technique for the spoilage classification of red meat. *Sensors*, 2008. (Cited on page 45.)

[227] A Perera, A Pardo, D Barrettino, A Hierlermann, and S Marco. Evaluation of fish spoilage by means of a single metal oxide sensor under temperature modulation. *Sensors and Actuators B: Chemical*, 146(2):477–482, 2010. (Cited on page 45.)

[228] Amalia Berna. Metal oxide sensors for electronic noses and their application to food analysis, 2010. (Cited on page 45.)

[229] Amy Loutfi, Silvia Coradeschi, Ganesh Kumar Mani, Prabakaran Shankar, and John Bosco Balaguru Rayappan. Electronic noses for food quality: A review. *Journal of Food Engineering*, 144:103–111, 2015. (Cited on page 45.)

[230] Piet Bergveld. Sensors for biomedical applications. *Sensors and Actuators*, 10(3-4):165–179, 1986. (Cited on page 45.)

[231] Alphus D Wilson and Manuela Baietto. Advances in electronic-nose technologies developed for biomedical applications. *Sensors*, 11(1):1105–1176, 2011. (Cited on page 45.)

[232] Marco Righettoni, Anton Amann, and Sotiris E Pratsinis. Breath analysis by nanostructured metal oxides as chemo-resistive gas sensors. *Materials Today*, 18(3):163–171, 2015. (Cited on page 45.)

[233] Teodor Šundić, Santiago Marco, Alex Perera, Antonio Pardo, Simone Hahn, Nicolae Bârsan, and Udo Weimar. Fuzzy inference system for sensor array calibration: prediction of CO and CH 4 levels in variable humidity conditions. *Chemometrics and Intelligent Laboratory Systems*, 64(2):103–122, 2002. (Cited on pages 45 and 49.)

[234] Alexey A. Tomchenko, Gregory P. Harmer, and Brent T. Marquis. Detection of chemical warfare agents using nanostructured metal oxide sensors. *Sensors and Actuators B: Chemical*, 108(1):41–55, 2005. (Cited on page 45.)

[235] K Brudzewski, S Osowski, and W Pawlowski. Metal oxide sensor arrays for detection of explosives at sub-parts-per million concentration levels by the differential electronic nose. *Sensors and Actuators B: Chemical*, 161(1):528–533, 2012. (Cited on page 45.)

[236] Thomas Hübert, L Boon-Brett, V Palmisano, and Mark Andreas Bader. Developments in gas sensor technology for hydrogen safety. *International Journal of Hydrogen Energy*, 39(35):20474–20483, 2014. (Cited on page 45.)

[237] Jose Prado, Goncalo Cabrita, and Lino Marques. Bayesian sensor fusion for landmine detection using a dual-sensor hand-held device. In *Industrial Electronics Society, IECON 2013-39th Annual Conference of the IEEE*, pages 3887–3892. IEEE, 2013. (Cited on page 45.)

[238] Praveen K. Sekhar, Eric L. Brosha, Rangachary Mukundan, Wenxia Li, Mark A. Nelson, Ponnusamy Palanisamy, and Fernando H. Garzon. Application of commercial automotive sensor manufacturing methods for NOx/NH3 mixed potential sensors for on-board emissions control. *Sensors and Actuators, B: Chemical*, 144(1):112–119, 2010. (Cited on page 45.)

[239] L Boon-Brett, J Bousek, and P Moretto. Reliability of commercially available hydrogen sensors for detection of hydrogen at critical concentrations: part II–selected sensor test results. *International Journal of Hydrogen Energy*, 34(1):562–571, 2009. (Cited on page 45.)

[240] Yoon-Sung Kim, In-Sung Hwang, Sun-Jung Kim, Choong-Yong Lee, and Jong-Heun Lee. CuO nanowire gas sensors for air quality control in automotive cabin. *Sensors and Actuators B: Chemical*, 135(1):298–303, 2008. (Cited on page 45.)

[241] Richard Kwor. Carbon monoxide detectors. *Carbon Monoxide Toxicity. CRC; Boca Raton, FL*, pages 61–82, 2000. (Cited on page 45.)

[242] Jia Zhou, Po Li, Song Zhang, Yingcai Long, Feng Zhou, Yiping Huang, Pengyuan Yang, and Minhang Bao. Zeolite-modified microcantilever gas sensor for indoor air quality control. *Sensors and Actuators, B: Chemical*, 94(3):337–342, 2003. (Cited on page 45.)

[243] Gilles Adam, Sébastien Lemaigre, Xavier Goux, Philippe Delfosse, and Anne-Claude Romain. Upscaling of an electronic nose for completely stirred tank reactor stability monitoring from pilot-scale to real-scale agricultural co-digestion biogas plant. *Bioresource technology*, 178:285–296, 2015. (Cited on page 45.)

[244] V. Sberveglieri, E. Núñez, D. Zappa, E. Comini, and A. Pulvirenti. Classification of different roasting processes by MOX nanowire. In *Procedia Engineering*, volume 87, pages 572–575, 2014. (Cited on page 45.)

[245] J. Fonollosa, A. Solórzano, and S. Marco. Chemical sensor systems and associated algorithms for fire detection: A review. *Sensors (Switzerland)*, 18(2), 2018. (Cited on page 45.)

[246] Sunandan Baruah and Joydeep Dutta. Nanotechnology applications in pollution sensing and degradation in agriculture, 2009. (Cited on page 45.)

[247] Jehuda Yinon. Field detection and monitoring of explosives. *TrAC - Trends in Analytical Chemistry*, 21(4):292–301, 2002. (Cited on page 45.)

[248] K.C. Persaud, A.M. Pisanelli, Stefan Szyszko, Marcus Reichl, Gerhard Horner, Willi Rakow, H.J. Keding, and Heinrijk Wessels. A smart gas sensor for monitoring environmental changes in closed systems: results from the MIR space station. *Sensors and Actuators B: Chemical*, 55(2-3):118–126, 1999. (Cited on page 45.)

[249] Michael V. Storey, Bram van der Gaag, and Brendan P. Burns. Advances in on-line drinking water quality monitoring and early warning systems. *Water Research*, 45(2):741–747, 2011. (Cited on page 45.)

[250] C. M. McEntegart, W. R. Penrose, S. Strathmann, and J. R. Stetter. Detection and discrimination of coliform bacteria with gas sensor arrays. *Sensors and Actuators, B: Chemical*, 70(1-3):170–176, 2000. (Cited on page 45.)

[251] Stoyan Nihtianov and Antonio Luque. *Smart Sensors and MEMS: Intelligent Sensing Devices and Microsystems for Industrial Applications*. Woodhead Publishing, 2018. (Cited on page 47.)

[252] Daniel Rüffer, Felix Hoehne, and Johannes Bühler. New Digital Metal-Oxide (MOx) Sensor Platform. *Sensors (Basel, Switzerland)*, 18(4), 2018. (Cited on pages 47, 50, 52, 56, and 57.)

[253] P. K. Clifford and D. T. Tuma. Characteristics of semiconductor gas sensors I. Steady state gas response. *Sensors and Actuators*, 3(C):233–254, 1982. (Cited on pages 47 and 48.)

[254] Fabrizio Formisano, Ettore Massera, Saverio De Vito, Antonio Buonanno, Girolamo Di Francia, and Paola Delli Veneri. Tinynose, an auxiliary smart gas sensor for RFID tag in vegetables ripening monitoring during refrigerated cargo transport. In *Lecture Notes in Electrical Engineering*, volume 319, pages 217–221, 2015. (Cited on page 47.)

[255] Jianwei Gong, Quanfang Chen, Ming Ren Lian, Nen Chin Liu, Roberts G. Stevenson, and Fatos Adami. Micromachined nanocrystalline silver doped SnO2H2S sensor. *Sensors and Actuators, B: Chemical*, 114(1):32–39, 2006. (Cited on page 48.)

[256] Zhi Chen and Chi Lu. Humidity Sensors: A Review of Materials and Mechanisms. *Sensor Letters*, 3(4):274–295, 2005. (Cited on page 48.)

[257] R. Ionescu, A. Vancu, and A. Tomescu. Time-dependent humidity calibration for drift corrections in electronic noses equipped with SnO2gas sensors. *Sensors and Actuators, B: Chemical*, 2000. (Cited on pages 48 and 50.)

[258] G. Korotcenkov and B.K. Cho. Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey). *Sensors and Actuators B: Chemical*, 156(2):527–538, aug 2011. (Cited on pages 48 and 51.)

[259] K. Kamarudin, V.H. Bennetts, S.M. Mamduh, R. Visvanathan, A.S.A. Yeon, A.Y.M. Shakaff, A. Zakaria, A.H. Abdullah, and L.M. Kamarudin. Cross-sensitivity of metal oxide gas sensor to ambient temperature and humidity: Effects on gas distribution mapping. In *AIP Conference Proceedings*, volume 1808, 2017. (Cited on pages 48 and 49.)

[260] G Korotcenkov and B K Cho. Engineering approaches to improvement of conductometric gas sensor parameters. Part 2: Decrease of dissipated (consumable) power and improvement stability and reliability. *Sensors and Actuators B: Chemical*, 198:316–341, 2014. (Cited on pages 48 and 53.)

[261] N. Yamazoe, Y. Kurokawa, and T. Seiyama. Effects of additives on semiconductor gas sensors. *Sensors and Actuators*, 4(C):283–289, 1983. (Cited on page 48.)

[262] Jordi Fonollosa, Luis Fernández, Ramón Huerta, Agustín Gutiérrez-Gálvez, and Santiago Marco. Temperature optimization of metal oxide sensor arrays using Mutual Information. *Sensors and Actuators, B: Chemical*, 187:331–339, 2013. (Cited on pages 48 and 50.)

[263] Andrew P. Lee and Brian J. Reedy. Temperature modulation in semiconductor gas sensing. *Sensors and Actuators, B: Chemical*, 60(1):35–42, 1999. (Cited on pages 49 and 51.)

[264] Arturo Ortega, Santiago Marco, Alex Perera, Teodor Šundic, Antonio Pardo, and Josep Samitier. An intelligent detector based on temperature modulation of a gas sensor with a digital signal processor. *Sensors and Actuators B: Chemical*, 78(1-3):32–39, aug 2001. (Cited on page 49.)

[265] Alexander Vergara, Eduard Llobet, Jesús Brezmes, Xavier Vilanova, Peter Ivanov, Isabel Gràcia, Carles Cané, and Xavier Correig. Optimized temperature modulation of micro-hotplate gas sensors through pseudorandom binary sequences. *IEEE Sensors Journal*, 5(6):1369–1377, 2005. (Cited on page 49.)

[266] Eugenio Martinelli, Davide Polese, Alexandro Catini, Arnaldo D'Amico, and Corrado Di Natale. Self-adapted temperature modulation in metal-oxide semiconductor gas sensors. *Sensors and Actuators, B: Chemical*, 161(1):534–541, 2012. (Cited on pages 49 and 51.)

[267] I T Jolliffe. *Principal Component Analysis. Second Edition*. Springer, 2002. (Cited on page 49.)

[268] Svante Wold, Arnold Ruhe, Herman Wold, and W J Dunn III. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, 5(3):735–743, 1984. (Cited on page 49.)

[269] Kieu An Ngo, Pascal Lauque, and Khalifa Aguir. High performance of a gas identification system using sensor array and temperature modulation. *Sensors and Actuators, B: Chemical*, 124(1):209–216, 2007. (Cited on page 49.)

[270] Xin Yin, Lei Zhang, Fengchun Tian, and David Zhang. Temperature Modulated Gas Sensing E-Nose System for Low-Cost and Fast Detection. *IEEE Sensors Journal*, 16(2):464–474, 2016. (Cited on page 49.)

[271] Jae Ho Sohn, Michael Atzeni, Les Zeller, and Giovanni Pioggia. Characterisation of humidity dependence of a metal oxide semiconductor sensor array using partial least squares. *Sensors and Actuators, B: Chemical*, 131(1):230–235, 2008. (Cited on page 49.)

[272] F. Hossein-Babaei and V. Ghafarinia. Compensation for the drift-like terms caused by environmental fluctuations in the responses of chemoresistive gas sensors. *Sensors and Actuators, B: Chemical*, 143(2):641–648, 2010. (Cited on pages 49 and 52.)

[273] R Piedrahita, Y Xiang, N Masson, J Ortega, A Collier, Y Jiang, K Li, R P Dick, Q Lv, M Hannigan, and Others. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques*, 7(10):3325, 2014. (Cited on pages 49 and 50.)

[274] N. Masson, R. Piedrahita, and M. Hannigan. Approach for quantification of metal oxide type semiconductor gas sensors used for ambient air quality monitoring. *Sensors and Actuators, B: Chemical*, 208:339–345, 2015. (Cited on pages 49, 50, 52, and 53.)

[275] Ramon Huerta, Thiago Mosqueiro, Jordi Fonollosa, Nikolai F. Rulkov, and Irene Rodriguez-Lujan. Online decorrelation of humidity and temperature in chemical sensors for continuous monitoring. *Chemometrics and Intelligent Laboratory Systems*, 157:169–176, 2016. (Cited on pages 49 and 50.)

[276] R. Andrew Russell. Survey of robotic applications for odor-sensing technology. *International Journal of Robotics Research*, 20(2):144–162, 2001. (Cited on page 51.)

[277] Thorsten Conrad, Pascal Hiry, and A Schutze. PuMaH-a temperature control and resistance read-out system for microstructured gas sensors based on PWM signals. In *Sensors, 2005 IEEE*, pages 8—-pp. IEEE, 2005. (Cited on page 52.)

[278] Santiago Marco, Arturo Ortega, Antonio Pardo, and Josep Samitier. Gas identification with tin oxide sensor array and self-organizing maps: Adaptive correction of sensor drifts. *IEEE Transactions on Instrumentation and Measurement*, 47(1):316–321, 1998. (Cited on page 53.)

[279] Alexandre Perera, Niko Papamichail, Nicolae Bârsan, Udo Weimar, and Santiago Marco. On-line novelty detection by recursive dynamic principal component analysis and gas sensor arrays under drift conditions. *IEEE Sensors Journal*, 6(3):770–782, 2006. (Cited on page 53.)

[280] M. Padilla, A. Perera, I. Montoliu, A. Chaudry, K. Persaud, and S. Marco. Drift compensation of gas sensor array data by Orthogonal Signal Correction. *Chemometrics and Intelligent Laboratory Systems*, 100(1):28–35, jan 2010. (Cited on page 53.)

[281] A. Ziyatdinov, S. Marco, A. Chaudry, K. Persaud, P. Caminal, and A. Perera. Drift compensation of gas sensor array data by common principal component analysis. *Sensors and Actuators, B: Chemical*, 146(2):460–465, 2010. (Cited on page 53.)

[282] Santiago Marco and Agustín Gutierrez-Galvez. Signal and data processing for machine olfaction and chemical sensing: A review, 2012. (Cited on page 53.)

[283] Marta Padilla, Jordi Fonollosa, and Santiago Marco. Improving the Robustness of Odor Sensing Systems by Multivariate Signal Processing. In *Human Olfactory Displays and Interfaces: Odor Sensing and Presentation*, pages 296–316. IGI Global, 2013. (Cited on page 53.)

[284] A.C. Romain and J. Nicolas. Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview. *Sensors and Actuators B: Chemical*, 146(2):502–506, apr 2010. (Cited on page 53.)

[285] Dominique Martinez, Lotfi Arhidi, Elodie Demondion, Jean-Baptiste Masson, and Philippe Lucas. Using insect electroantennogram sensors on autonomous robots for olfactory searches. *Journal of visualized experiments: JoVE*, (90), 2014. (Cited on pages 53 and 54.)

[286] Piotr Batog and Andrzej Wołczowski. Odor markers detection system for mobile robot navigation. In *Procedia Engineering*, volume 47, pages 1442–1445, 2012. (Cited on pages 53 and 54.)

[287] Javier G. Monroy, Javier Gonźalez-Jiḿenez, and Jose Luis Blanco. Overcoming the slow recovery of MOX gas sensors through a system modeling approach. *Sensors (Switzerland)*, 12(10):13664–13680, 2012. (Cited on pages 53, 54, and 55.)

[288] Jordi Fonollosa, Sadique Sheik, Ramón Huerta, and Santiago Marco. Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring. *Sensors and Actuators B: Chemical*, 215:618–629, 2015. (Cited on pages 53 and 55.)

[289] Santiago Marco, Antonio Pardo, Fabrizio A.M. Davide, Corrado Di Natale, Arnaldo D'Amico, Andreas Hierlemann, Jan Mitrovics, Markus Schweizer, Udo Weimar, and Wolfgang Göpel. Different strategies for the identification of gas sensing systems. *Sensors and Actuators, B: Chemical*, 34(1-3):213–223, 1996. (Cited on pages 53 and 54.)

[290] A Pardo, S Marco, and J Samitier. Nonlinear inverse dynamic models of gas sensing systems based on chemical sensor arrays for quantitative measurements. *Ieee Transactions on Instrumentation and Measurement*, 47(3):644–651, 1998. (Cited on pages 53 and 54.)

[291] Mehmet K. Muezzinoglu, Alexander Vergara, Ramon Huerta, Nikolai Rulkov, Mikhail I. Rabinovich, Al Selverston, and Henry D.I. Abarbanel. Acceleration of chemo-sensory information processing using transient features. *Sensors and Actuators, B: Chemical*, 137(2):507–512, 2009. (Cited on page 55.)

[292] Drager - Risk Management Program. Gas Dispersion. Technical report, Drägerwerk AG & Co. KGaA, 2013. (Cited on page 57.)

[293] C. Lytridis, E. E. Kadar, and G. S. Virk. A systematic approach to the problem of odour source localisation. *Autonomous Robots*, 20(3):261–276, 2006. (Cited on page 57.)

[294] Achim Lilienthal, Holger Ulmer, H Frohlich, Felix Werner, and A ZeIl. Learning to detect proximity to a gas source with a mobile robot. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, volume 2, pages 1444–1449. IEEE, 2004. (Cited on page 57.)

[295] T. Yamanaka, H. Ishida, T. Nakamoto, and T. Moriizumi. Analysis of gas sensor transient response by visualizing instantaneous gas concentration using smoke. *Sensors and Actuators, A: Physical*, 69(1):77–81, 1998. (Cited on page 58.)

[296] H. Ishida, T. Yamanaka, N. Kushida, T. Nakamoto, and T. Moriizumi. Study of real-time visualization of gas/odor flow image using gas sensor array. *Sensors and Actuators, B: Chemical*, 65(1-3):14–16, 2000. (Cited on page 58.)

[297] Ronald J Adrian. Twenty years of particle image velocimetry. *Experiments in fluids*, 39(2):159–169, 2005. (Cited on page 58.)

[298] H. G. Maas, A. Gruen, and D. Papantoniou. Particle tracking velocimetry in three-dimensional flows - Part 1. Photogrammetric determination of particle coordinates. *Experiments in Fluids*, 1993. (Cited on page 58.)

[299] B Hiller and R K Hanson. Simultaneous planar measurements of velocity and pressure fields in gas flows using laser-induced fluorescence. *Applied optics*, 27(1):33–48, 1988. (Cited on page 58.)

[300] J P Crimaldi and J R Koseff. High-resolution measurements of the spatial and temporal scalar structure of a turbulent plume. *Experiments in Fluids*, 31(1):90–102, 2001. (Cited on page 58.)

[301] Wei Yan. *DEVELOPMENT OF HYBRID PARTICLE TRACKING ALGORITHMS AND THEIR APPLICATIONS IN AIRFLOW MEASUREMENT WITHIN AN AIRCRAFT CABIN MOCK-UP*. PhD thesis, University of Illinois at Urbana-Champaign, 2010. (Cited on page 58.)

[302] Marcel Loomans. *The Measurement and Simulation of Indoor Air Flow*. University of Eindhoven, 1998. (Cited on page 58.)

[303] Maureen Simonds, Hongkui Xiao, and Steven P Levine. Optical remote sensing for air pollutants-review. *American Industrial Hygiene Association Journal*, 55(10):953–965, 1994. (Cited on page 58.)

[304] Rodney A Brooks and Giovanni Di Chiro. Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging. *Physics in Medicine & Biology*, 21(5):689, 1976. (Cited on page 58.)

[305] A. C. Drescher, D. Y. Park, M. G. Yost, A. J. Gadgil, S. P. Levine, and W. W. Nazaroff. Stationary and time-dependent indoor tracer-gas concentration profiles measured by OP-FTIR remote sensing and SBFM-computed tomography. *Atmospheric Environment*, 31(5):727–740, 1997. (Cited on page 58.)

[306] Marc L. Fischer, Phillip N. Price, Tracy L. Thatcher, Carrie A. Schwalbe, Mathias J. Craig, Emily E. Wood, Richard G. Sextro, and Ashok J. Gadgil. Rapid measurements and mapping of tracer gas concentrations in a large indoor space. *Atmospheric Environment*, 35(16):2837–2844, 2001. (Cited on page 58.)

[307] Lori A Todd, Mallika Ramanathan, Kathleen Mottus, Robert Katz, Ashley Dodson, and Gary Mihlan. Measuring chemical emissions using open-path Fourier transform infrared (OP-FTIR) spectroscopy and computer-assisted tomography. *Atmospheric Environment*, 35(11):1937–1947, 2001. (Cited on page 58.)

[308] G. Tochon, J. Chanussot, J. Gilles, M. Dalla Mura, J. M. Chang, and A. L. Bertozzi. Gas plume detection and tracking in hyperspectral video sequences using Binary Partition Trees. In *Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote Sensing*, volume 2014-June, 2017. (Cited on page 59.)

[309] Christopher M Gittins and William J Marinelli. LWIR multispectral imaging chemical sensor. In *Air Monitoring and Detection of Chemical and Biological Agents*, volume 3533, pages 93–103. International Society for Optics and Photonics, 1999. (Cited on page 59.)

[310] V Farley, A Vallies, A Villemaire, M Chamberland, P Lagueux, and J Giroux. Chemical agent detection and identification with a hyperspectral imaging infrared sensor - art. no. 673918. *Electro-Optical Remote Sensing, Detection, and Photonic Technologies and Their Applications*, 6739:73918–73918\r259, 2007. (Cited on page 59.)

[311] D. Manolakis, L. G. Jairam, D. Zhang, and M. Rossacci. Statistical Models for LWIR Hyperspectral Backgrounds and their Applications in Chemical Agent Detection. In *Proceedings of SPIE Vol. 6565*, 2007. (Cited on page 59.)

[312] Alexander Haverkamp, Julia Bing, Elisa Badeke, Bill S. Hansson, and Markus Knaden. Innate olfactory preferences for flowers matching proboscis length ensure optimal energy gain in a hawkmoth. *Nature Communications*, 7, 2016. (Cited on pages 59 and 60.)

[313] A. Vergara, J. Fonollosa, J. Mahiques, M. Trincavelli, N. Rulkov, and R. Huerta. On the performance of gas sensor arrays in open sampling systems using Inhibitory Support Vector Machines. *Sensors & Actuators B: Chemical*, 185:462–477, 2013. (Cited on pages 59 and 60.)

[314] Roderick Shepherd, Stephen Beirne, King Tong Lau, Brian Corcoran, and Dermot Diamond. Monitoring chemical plumes in an environmental sensing chamber with a wireless chemical sensor network. *Sensors and Actuators B: chemical*, 121(1):142–149, 2007. (Cited on page 60.)

[315] Ayano Murai, Kamon Yoshimoto, Ryuichi Takemura, Haruka Matsukura, and Hiroshi Ishida. Robotic gas source localization assisted by active airflow generation. In *2015 IEEE SENSORS - Proceedings*, 2015. (Cited on page 60.)

[316] Shinzo Yamakawa and Akihumi Yamaguchi. Visualization of Methanol Gas Concentration Distribution Using a Fiber-Optic Sensor Array with Dye Coating. *Transactions of the Society of Instrument and Control Engineers*, 31(9):1273–1278, 1995. (Cited on page 60.)

[317] Yukio Hiranaka and Hiro Yamasaki. Locating gas sources using a gas sensor array. *Analytical sciences*, 7(Supple):1565–1568, 1991. (Cited on page 60.)

[318] Saverio De Vito, Grazia Fattoruso, Raffaele Liguoro, Antonio Oliviero, Ettore Massera, Carlo Sansone, Valentina Casola, and Girolamo Di Francia. Cooperative 3D air quality assessment with wireless chemical sensing networks. *Procedia Engineering*, 25:84–87, 2011. (Cited on page 60.)

[319] Paul Moore and John Crimaldi. Odor landscapes and animal behavior: Tracking odor plumes in different physical worlds. *Journal of Marine Systems*, 49(1-4):55–64, 2004. (Cited on page 62.)

[320] Shreejoy Tripathy. Odors pulsed at wing beat frequencies are tracked by primary olfactory networks and enhance odor detection. *Frontiers in Cellular Neuroscience*, 2010. (Cited on page 62.)