Sensors & Actuators: B. Chemical 361 (2022) 131668

Contents lists available at ScienceDirect

Sensors and Actuators: B. Chemical

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

Drift in a popular metal oxide sensor dataset reveals limitations for gas classification benchmarks

Nik Dennler [a,b], Shavika Rastogi [a,b], Jordi Fonollosa [c,d], André van Schaik [b], Michael Schmuker [a,*]

[a] UH Biocomputation Research Group, Centre for Computer Science and Informatics Research, University of Hertfordshire, Hatfield, United Kingdom
[b] International Centre for Neuromorphic Systems, Western Sydney University, Sydney, Australia
[c] Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain
[d] Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain

ARTICLE INFO

Keywords:
Metal oxide gas sensors
Wind tunnel dataset
Sensor drift
Gas recognition

ABSTRACT

Metal oxide (MOx) gas sensors are a popular choice for many applications, due to their tunable sensitivity, space efficiency and low cost. Publicly available sensor datasets are particularly valuable for the research community as they accelerate the development and evaluation of novel algorithms for gas sensor data analysis. A dataset published in 2013 by Vergara and colleagues contains recordings from MOx gas sensor arrays in a wind tunnel. It has since become a standard benchmark in the field. Here we report a latent property of this dataset that limits its suitability for gas classification studies. Measurement timestamps show that gases were recorded in separate, temporally clustered batches. Sensor baseline response before gas exposure were strongly correlated with the recording batch, to the extent that baseline response was largely sufficient to infer the gas used in a given trial. Zero-offset baseline compensation did not resolve the issue, since residual short-term drift still contained enough information for gas/trial identification using a machine learning classifier. A subset of the data recorded within a short period of time was minimally affected by drift and suitable for gas classification benchmarking after offset-compensation, but with much reduced classification performance compared to the full dataset. We found 18 publications where this dataset was used without precautions against the circumstances we describe, thus potentially overestimating the accuracy of gas classification algorithms. These observations highlight potential pitfalls in using previously recorded gas sensor data, which may have distorted widely reported results.

1. Introduction

Over the last 50 years, artificial olfaction has evolved from an almost niche field of study into a thriving interdisciplinary research area. Many use cases have been addressed, for example the detection of hazardous gases or pollutants [1], spoilage localization [2], mobile olfactory robotics [3], health monitoring [4] and medical screening [5]; and artificial olfaction is expected to address many more use cases in the future [6]. A key challenge in artificial olfaction is to identify a range of odorants at high specificity. One way to achieve this is to use an array of multiple gas sensors, each with a rather large selectivity and low specificity, and extract the identity of the presented odor using pattern recognition. Metal oxide (MOx) gas sensors are widely used for such sensor arrays. Their sensing layer can be tuned to different analyte classes and they require little electronic periphery, which simplifies sensor design, reduces cost and saves space. One big drawback of MOx sensors is their susceptibility to sensor drift—the gradual and unpredictable variation of signal response over time when exposed to identical analytes under the same conditions [7]. Drift is mostly due to chemical and physical interactions on the sensor site, such as sensor ageing (reorganization of the sensor surface over time) and sensor poisoning (irreversible or slowly reversible binding of previously measured gases or other contamination). Environmental effects such as changes in humidity, temperature or pressure also affect the sensor response. The impact of sensor drift can be reduced by careful experimental design that avoids any correlation between gas identity and drift, for example by randomizing the order of analytes presented. Where this is not possible or not desired, it is essential to be aware of the presence of drift and design analysis algorithms accordingly.

Setting up an electronic olfaction system still requires custom design

* Corresponding author.
E-mail address: m.schmuker@herts.ac.uk (M. Schmuker).

https://doi.org/10.1016/j.snb.2022.131668
Received 20 August 2021; Received in revised form 12 January 2022; Accepted 4 March 2022
Available online 8 March 2022

of electronics and data analysis procedures. These designs and algorithms must be based on reliable data. There are many parameters that can affect MOx sensor recordings; e.g. environmental conditions like temperature and humidity, technical constraints like wind tunnel design/construction, flow control of analytes, turbulent dispersal, gas availability and associated safety requirements, among others. Previously recorded datasets from reputable sources are therefore popular in the Artificial Olfaction/Mobile Robot Olfaction (AO/MRO) community, since they reduce the need for recording data in the initial design stages. A number of datasets are publicly available, covering a range of tasks and use cases [8–16].

One of the most popular datasets contains MOx sensor data sampled in a wind tunnel, for different gases and different experimental parameters, over a time of 16 months [9] (downloadable at [17]). This publication has been cited more than 100 times.[1] The dataset has been used as a benchmark for gas classification algorithms in at least 18 publications [18–35]. It has also been used for gas source location estimation [36–39] and other applications [40–45].

Here, we reveal a fundamental limitation of the dataset published in [9]. First, we observed that gases were not presented in random order, but in distinct batches, sometimes recorded weeks or months apart. In consequence, the sensor recordings were contaminated by slow baseline drift effects that are characteristic for the time of recording. We show that since both gas identity and sensor baseline correlate with time, it is possible to identify trials using a specific gas only by looking at the baseline response, before any gas has been released. In addition, we show that even after correcting for slow drift by subtracting the average of the first few sensor readings of each experimental trial, residual short-term drift effects are characteristic enough to identify trials where specific gases have been used, using the baseline alone. Moreover, when further minimizing the impact of drift by selecting the least-affected subset of recordings and compensating for drift as much as possible, the gas classification performance is far inferior to the numbers we obtained when using the full dataset. Therefore we conclude that this dataset is only of limited use for gas classification benchmarking, and that previously reported classification results based on this dataset are likely overestimating the true accuracy of gas recognition. Finally, we give a perspective on how the measurement protocol could be improved to mitigate this problem, and elaborate on what tasks the dataset can be appropriately used for, i.e., tasks that are not affected by the drift contamination.

2. Dataset

The dataset in question [9] consists of 18000 time-series measurements recorded over a period of 16 months from a 72 MOx gas sensor array-based chemical detection platform exposed to 10 different analyte gases (Acetone, Acetaldehyde, Ammonia, Butanol, Ethylene, Methane, Methanol, Carbon monoxide, Benzene, and Toluene). The sensor platform consisted of nine modules, each equipped with eight MOx sensors (see Table 1 for sensor types). It was placed in a 2.5 m × 1.2 m × 0.4 m flat-bed wind tunnel, at six different distances from the gas inlet, perpendicular to the wind direction (see Fig. 1a for a schematic). Each sensor module was integrated with a sensor controller, which enabled data collection at 12-bit resolution and a sampling rate of 100 Hz. Gas flow was adjusted by computer-supervised mass flow controllers. A multiple-step motor-driven exhaust fan controlled the wind speed.

Table 1
Metal oxide (MOx) sensors included in each 8-sensor array. All sensors were manufactured by Figaro USA, Inc [46].

Col. no.	Sensor model	Col. no.	Sensor model
1	TGS 2611	5	TGS 2600
2	TGS 2612	6	TGS 2600
3	TGS 2610	7	TGS 2620
4	TGS 2602	8	TGS 2620

Different experimental conditions were tested, namely three different wind speeds set by the fan (0.1 m s^{-1}, 0.21 m s^{-1}, 0.34 m s^{-1}) and five different sensor operating voltages (4.0 V, 4.5 V, 5.0 V, 5.5 V, 6.0 V). Before each measurement, a combination of the experimental parameters *gas*, *location*, *wind speed*, *operating voltage* was selected, until each combination was repeated 20 times. Each measurement lasted for 260 s, where gas was released between t = 20 s and t = 200 s. Before and after each experiment, the wind tunnel was ventilated at the maximum speed (0.34 m s^{-1}) for two minutes to assert the reestablishment of sensor response baseline.

The data is deposited as raw sensor data with one file per trial and parameter combination. The time of recording of individual measurements was encoded as part of the name of the file containing the time-series, alongside the parameters used in that recording and the trial sequence.

For our analysis, we interpolated and re-sampled the data for dealing with missing data points, and further converted the sensor voltage readings V_{sensor} given in the dataset to sensor resistance values R_{sensor}, according to Eq. (1),

$$R_{sensor} = 10\,\mathrm{k\Omega} \times \frac{3.11\mathrm{V} - V_{sensor}}{V_{sensor}}. \tag{1}$$

The readings of sensor 1 for all boards were discarded due to excessive sensor noise. Fig. 1b shows the responses of a sensor board to one gas in a typical trial. Unless stated otherwise, we used the wind tunnel location P4 B5 (wind-downstream from the gas source, see Fig. 1a for wind tunnel schematics), as we expected a high gas exposure at that location.

3. Results

3.1. Non-random order of gas measurements

We extracted the times of recording from the filenames to analyse the temporal order of measurements. Fig. 1c shows when each the 18000 measurements have been made, arranged by gas identity and sensor position. It is evident that gases have not been measured in random order, but in separate batches that cluster in time. Only rarely do measurements of different gases overlap in time (as for Ammonia, CO and Toluene); more often, measurement batches are several weeks apart (e.g., Toluene and Methane). In no case have gases been alternated on a per-trial basis. In addition, we observed that also other experimental parameters like distance-to-source, wind speed, sensor temperature were selected in a sequential fashion rather than in random order (not shown).

The batched arrangement of gas identity and parameter settings is not evident from the description of the dataset provided by the authors, neither in the original paper, nor in the documentation contained in the UCI repository [17]. Describing the experimental protocol, it was stated that (quote) *"This measurement procedure was reproduced exactly for each gas category exposure, landmark location in the wind tunnel, operating temperature, and airflow velocity in a random order and up until all pairs were covered."* [9]. This could be read as to imply that all experimental parameters that define an experiment were selected randomly before each trial, including which gas to release—which would mitigate, to a large extent, the detrimental effect of baseline drift on gas identification benchmarks—which is not the case, as we show here.

[1] According to Google Scholar as of December 2021.

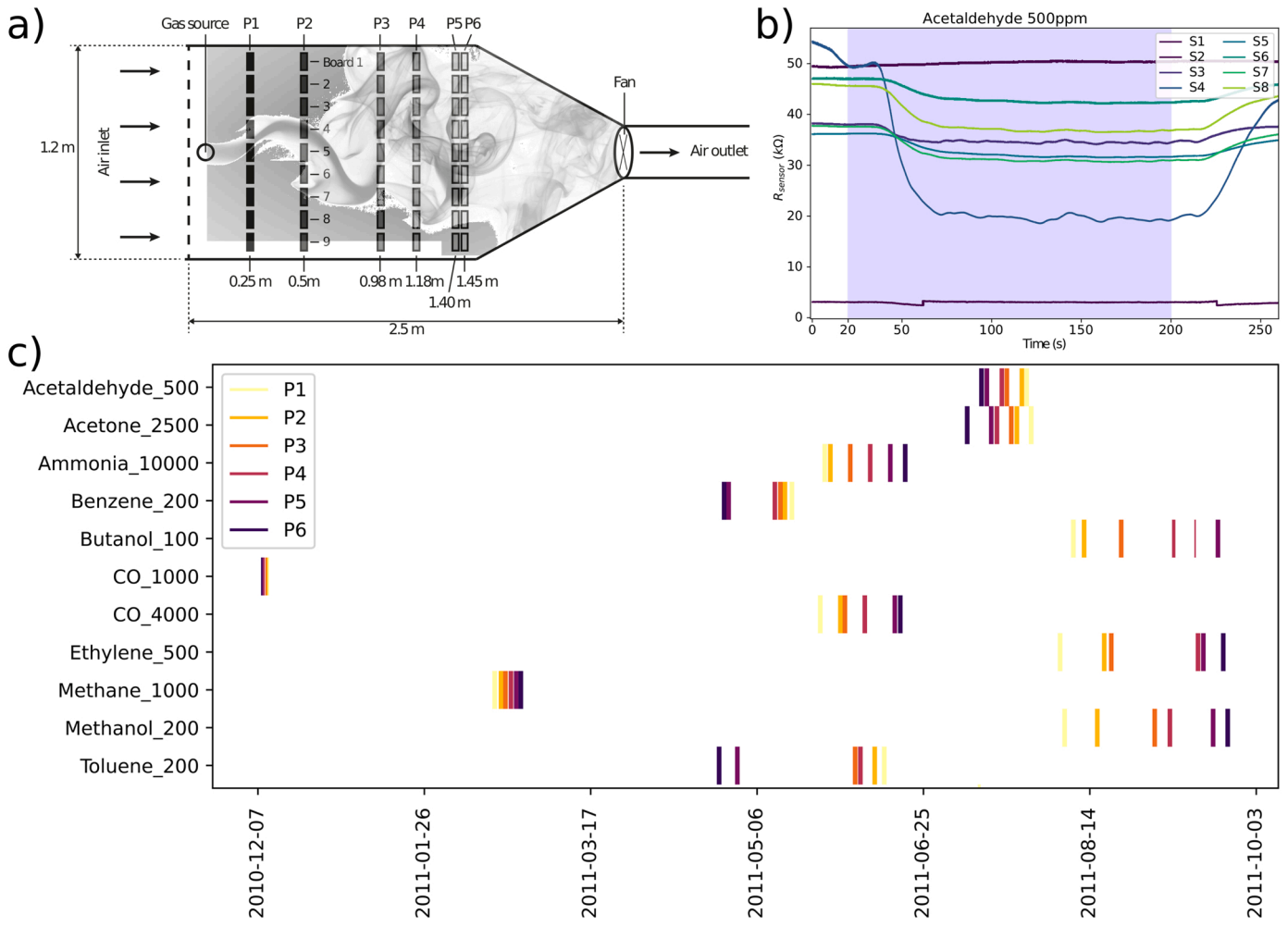

Fig. 1. Experimental procedure of Vergara et al. [9] **a)** Sensor boards with 8 MOx sensors each are placed at different locations in a wind tunnel. Air gets sucked in by a fan, and a gas source can be opened or closed. **b)** Measured sensor resistance for all sensors on one sensor board (location P4, module 5, Acetaldehyde, 0.21 m s^{-1} airflow velocity, 6 V operating voltage, trial 1). The shaded portion denotes the period during which the analyte was injected into the wind tunnel. **c)** Event-plot of timestamps for different gas trials. Each vertical line represents 300 trials, which were performed too close to each other for them to be visually distinguishable in this representation. The row name indicates the measured gas and its concentration in parts-per-million (*ppm*). CO at 1000 ppm was removed from further analysis since significantly fewer trials were performed than for the other analytes.
a) Adapted from [9].

3.2. Drift in baseline over time

We investigated the sensor baseline across trials, where here we defined baseline as the sensor readings measured before gas is released into the wind tunnel. Fig. 2a shows the trial-wise average of sensor baseline values at times $t < t_{release} = 20$ s, for a fixed sensor board location, operating temperature and airflow velocity, versus the date of recording. We observed that baseline varies significantly over time. Long-term drift can be observed as significant discontinuities between recording sessions. Since gas presentations were batched, the baseline pattern often correlated with gas identity. In addition, substantial baseline drift could be observed within some recording sessions.

3.3. Spatial distribution of baseline variations

By design, the gas plume does not distribute homogeneously across the wind tunnel, but disperses in a turbulent manner. Consequently, the total gas exposure at different sensor sites could vary, which may alter each sensor's response differently. Here we investigated how variations in the baseline response were distributed across the wind tunnel and the sensor board, as a proxy for sensor drift effects. To quantify the variations, we calculated the coefficient of variation (c_v) of the baseline, for each sensor and each board location. c_v is given by the fraction between the standard deviation σ and the mean μ (Eq. (2)),

$$c_v = \frac{\sigma}{\mu}. \tag{2}$$

We discriminated between *long-term* baseline variations over the whole duration of the experiment, and *short-term* drift within single trials. To quantify long-term variations, σ and μ were computed from the distribution of trial-wise averages of sensor baseline values, thus c_v described the variation of the baseline across the whole experimental duration. For short-term variations, c_v was taken as the average of the within-trial σ-to-μ ratios. We observed a distinct spatial pattern in the distribution of long-term baseline variations across the wind tunnel (Fig. 2b). The long-term drift effects were strongest in sensor boards close to the center line of the wind tunnel, where gas concentration was expected to be highest. This observation suggests that long-term drift could be caused by exposure to the sample gas. Long-term drift affected all sensors, although sensor 4 was affected most strongly (Fig. 2c. This sensor is a Figaro TGS 2602, which is targeted towards "Air pollutants (VOCs, ammonia, H2S)" according to Figaro's website.

The values for within-trial short-term coefficients of variation were

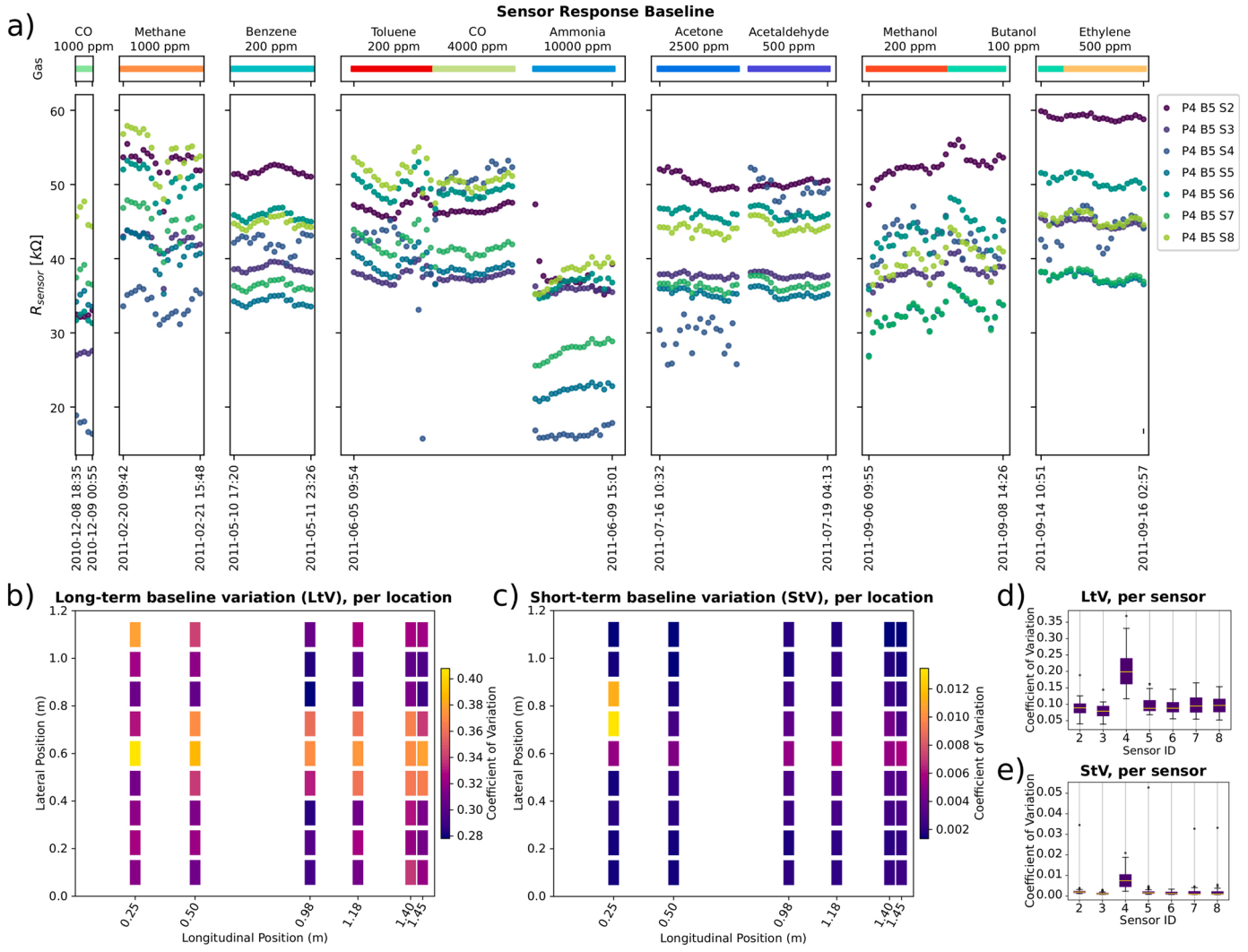

Fig. 2. Drift analysis of Vergara et al. dataset [9]. **a)** Baseline for each sensor and experimental trial. Dots represent the mean sensor resistance during the time before gas release (20 s). Top row indicates the gas and its concentration (in *ppm*) used in the corresponding sessions. **b)**–**e)** Local baseline variation analysis using the coefficient of variation, for spatial wind tunnel location (**b)** and **c)**) and sensor board (**d)** and **e)**). **b)** and **d)** display the long-term baseline variation across the whole experimental duration (16 months), where **c)** and **e)** display the averaged within-trial, short-term baseline variation. Data shown here was obtained with wind flow speed 0.21 m s^{-1} and hotplate voltage 6 V. All ten gases and sensors 2–8 were considered. For **a)**, only location 4 and board 5 were considered (see Fig. 1a for wind tunnel schematics).

naturally lower in magnitude but exhibited a similar pattern as observed for long-term variations, both spatially and per-sensor (Figs. 2d and 2e). This indicates that also within-trial drift was highest for those sensors that were exposed to the highest gas concentrations.

3.4. Gas clustering and classification

Since both the baseline drift and the identity of the gas used in a trial correlate with time, we tested how much information about the gas could be obtained from the baseline signal alone. Fig. 3a shows a Principal Components Analysis (PCA) plot of the raw baseline values for each gas, at a range of times after the start of trials. Each plot presents a snapshot of a 100 ms time window, within which the time-series data of the sensor responses was averaged and used for the PCA. The PCA was computed using all trials in all windows, thus each snapshot is a projection of the data into this shared PC space. Distinct gas-specific clusters could be observed already at t = 0 s, before any gas was released into the tunnel at $t_{release} = 20$ s. The clusters change slightly between 30 and 40 s, which we assume is when the gas had reached the sensors.

Next, we attempted to compensate for long-term drift effects by subtracting the average of the first 100 ms window, i.e. for t ∈ [0.0 s, 0.1 s). The data then only contains the difference of the sensor response relative to the start of a trial. This is a standard procedure when dealing with MOx-sensor data. For Fig. 3b we computed a PCA on the data compensated for long-term drift and used the same windows as before to visualize the evolution of sensor responses. By design, at t = 0 s there are no visible clusters. Interestingly, although the zero-baseline has been subtracted, we still observe the formation of clusters before the release of the gas. We interpret this observation as the manifestation of short-term drift within trials (cf. Figs. 2d and 2e). It indicates that short-term drift also changed over time, in a way that correlated with gas identity.

These observations were confirmed using a time-windowed supervised classification approach with a soft-margin Support Vector Machine (SVM) classifier. We used a linear kernel with regularization parameter $C = 1.0$. The classifier was trained and tested separately for each time window, using the same features as for the PCA (time-series data of the sensor responses). We used a 4-to-1 random training to test split, i.e. training on 80% of the trials in each window and testing using the remaining 20%, repeated 10 times with different 4-to-1 random splits.

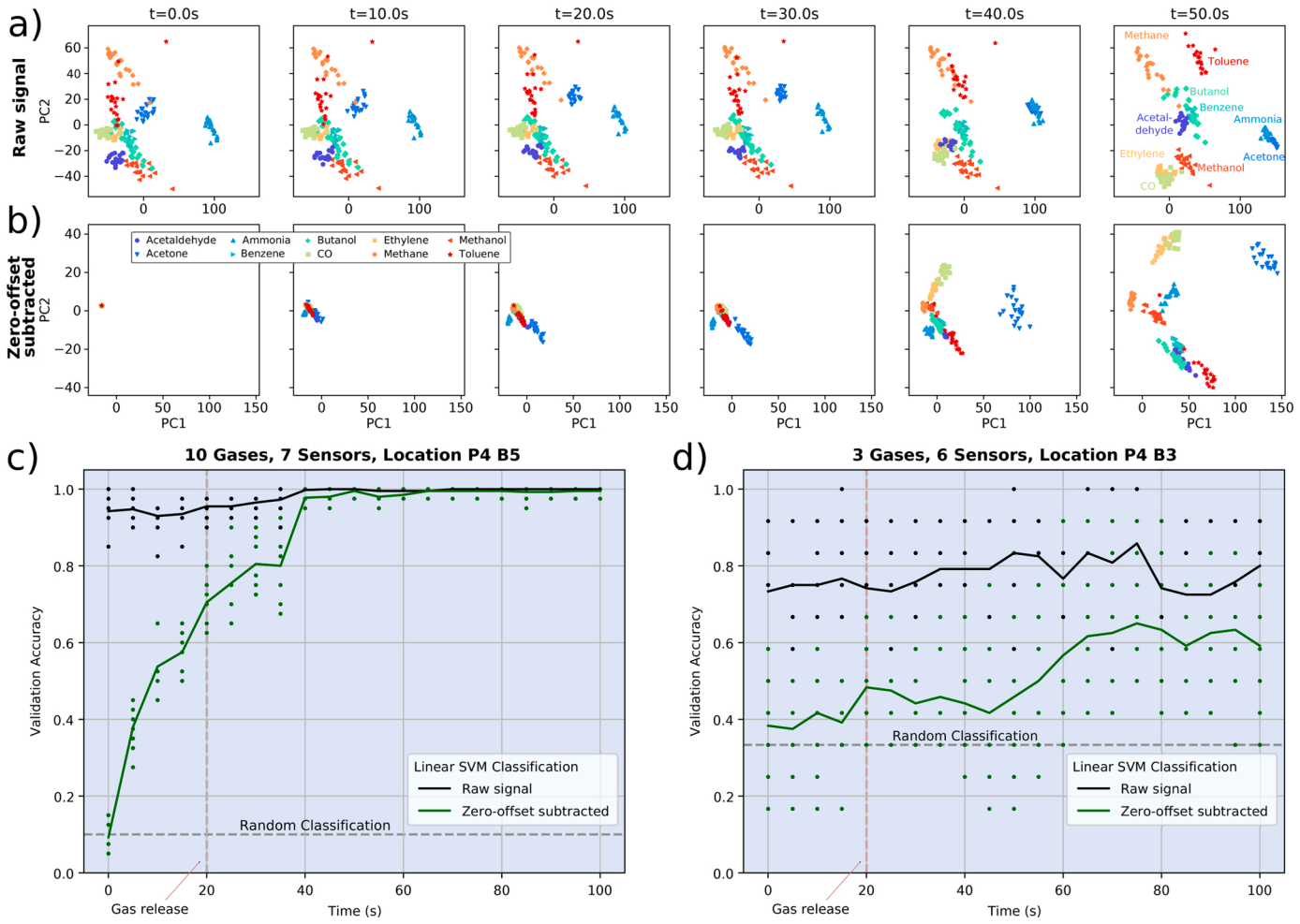

Fig. 3. PCA analysis and SVM gas classification of Vergara et al. dataset [9]. **a) and b)** Principal Component Analysis (PCA) of samples within a 100 ms time window, at different starting times. Each color and shape corresponds to a different gas. For **a)**, the raw resistance signal was considered, where for **b)**, the zero-offset was removed by subtracting the mean resistance in the first 100 ms for each sensor. **c) and d)** Classification results using a Linear Support Vector Machine classifier. The trials for each gas were randomly split in training and validation datasets with a ratio of 80–20. Black corresponds to the raw resistance signal, whereas green corresponds to the zero-offset subtracted signal. For all experiments shown here, the wind flow speed was fixed at 0.21 m s^{-1}, and the hotplate voltage was set to 6 V. For **a)–c)**, all ten gases and sensors 2–8 have been considered, at location 4 and board 5 (see Fig. 1a for wind tunnel schematics). For **d)**, the gases Methanol, Ethylene and Butanol have been considered, measured with sensors 2–3 and 5–8, at location 4 and board 3.

Fig. 3c shows the classifier performance. As expected, the classifier yielded near-perfect gas recognition performance on the raw data (i.e., without compensation for long-term drift), with an average accuracy of 94.3% already on the first time window of a trial, for $t \in [0.0\text{ s}, 0.1\text{ s})$. Test accuracy increased slightly for later time windows. From $t = 40$ s on it converged at 100%. We assume that this is when the sensor board was maximally exposed to the gas.

Compensating for baseline offset did not rectify these classification artefacts. While test accuracy was random for the time window at $t = 0$ s, it was clearly above random already at $t = 5$ s (Fig. 3c, green line). It increased further to around 80% at $t = 35$ s, before making a step to near 100% at $t = 40$ s.

Taken together, we observed that the time window before gas exposure contained enough information to identify the gas used in a particular trial, even before the gas has been released into the wind tunnel. Baseline compensation for long-term drift reduced the extent of the problem, but there was still sufficient information contained in the short-term drift dynamics that allowed identification of the gas used in a given trial far above chance level. Noteworthy here is that even gases that are measured in close temporal proximity (such as CO at 4000 ppm and Ammonia) separated well in PC space. We suspect that not only time-related sensor ageing, but also the slow recovery phase after gas exposure and permanent sensor poisoning could play a role in causing the baseline drift effects we observed here.

3.5. *Restricted data subset*

Based on our findings in Section 3.1–3.4, we selected a subset of the data that would be least affected by the drift effects we observed. We selected this subset by three constraints. First, only Methanol, Ethylene and Butanol were considered, since they have been measured within close temporal proximity (see Figs. 1c and 2a). Second, we removed Sensor 4 from the analysis, as it appears to be particularly affected by drift (see Figs. 2c & 2e). Third, we used data from sensor board 3 rather than sensor board 5, since our analysis suggested that it was, on average, less affected by drift (see Figs. 2b & 2d).

We repeated the SVM classification task in this, according to our analysis, less compromised subset. The results are displayed in Fig. 3d. The classification accuracy for the raw signal is initially still well above chance level at around 75%, without changing significantly after gas release. This indicates that long-term drift effects are pronounced enough to enable trial identification even in the restricted dataset. Moreover, classification accuracy increased only very slightly after gas onset. This indicates, paradoxically, that actual gas exposure made little

difference for "gas" recognition in the restricted dataset.

The picture changed after compensating for long-term drift by subtracting the baseline offset at t = 0 s. Classification accuracy was only slightly above the chance level of 33.3% until gas release. After gas release, accuracy slowly increased to slightly above 60%. Therefore, we conclude that the restricted dataset is suitable as a gas classification benchmark when compensating for baseline offset. It should be noted though that a gas recognition accuracy of 60% is much lower than what we and others have reported for the original dataset. On the other hand, the sensor board we selected was located slightly lateral to the downwind axis from the source, therefore likely not as strongly exposed to the gas plume, which potentially affects classification performance negatively (but also apparently reduces sensor drift).

4. Discussion and conclusion

In our analysis, we have shown that the different gases have been measured in time-separated batches and not in random order, which makes the data susceptible to sensor drift effects. We have shown that the sensor response baseline correlates with the time and order of measurement, consistent with long-term drift behaviour. We have also shown that the sensor response baseline alone is enough for 'accurate' gas classification, even after compensating for long-term drift by removing the offset at t = 0 s. This means that the dataset cannot be used for gas classification benchmarks without further precautions.

In an attempt to alleviate this limitation, we identified a subset of the dataset, which, under certain conditions, could be used for gas classification benchmarking. The subset contains three gases that have been measured in close temporal proximity, at a location that appears to be less affected by drift, while disregarding one most affected sensor. After applying long-term drift compensation to this subset, we observed what would be expected from a clean experiment: Gas identification accuracy was near chance level at the beginning of a trial and rose only after the gas has reached the sensor. However, gas classification performance under those conditions was much lower than when using the full dataset, in spire of the reduced complexity of the task due to the smaller number of gases.

We therefore conclude that there is substantial information about the gas identity in the baseline. It must be assumed that this information will also interfere with actual gas sensor response. Therefore, the classification accuracy of a benchmark will overestimate the accuracy that could be obtained without the drift effect. These findings suggest that many, if not all, of the previous studies using this dataset overestimated the performance of their gas recognition algorithms. In fact, previous studies have acknowledged the exceptionally high classification accuracy obtained on this dataset, compared to others [23,25].

We identified 18 publications that are potentially affected [18–35]. At least one study reported a classification accuracy of 100% [29]. None of those 18 studies described subtracting the baseline or other attempts that could address the detrimental effects of drift. Only four studies described data normalisation efforts [18,19,28,30] that would remove the absolute scaling of the data, which should be a standard procedure when dealing with MOx sensor data. Since the different studies use different subsets of the dataset, and only a small fraction of the teams provided their analysis code [20,25], a thorough numerical comparison of their algorithmic performance is beyond the scope of this work.

Due to the popularity of this dataset as a benchmarks for electronic olfaction, the overestimation of accuracy that could be obtained may have distorted the state-of-the-art in gas recognition. The uncritical use of this dataset may have impeded progress in the field to a considerable extent, for example by casting unjustified doubt on other datasets where classification scores were lower (but possibly more relevant). We hope that the analysis presented here may enable a more realistic assessment of gas classification algorithms, and further encourage the collection and sharing of novel gas sensor datasets.

It should be noted that this dataset, despite its limitations, is an excellent example of how datasets should be shared. It contains the raw measurement data and all timestamps of the recordings. This is unfortunately not common practice in the field—often, only derived features are shared. We expressly acknowledge the effort Vergara et al. have made to share the data as accurately as possible. Only through their diligence and attention to detail was it possible to identify the underlying limitations. The dataset still has unique features which make it a tremendous resource for machine olfaction research. It is one of the very few available datasets which have been recorded with a very high temporal resolution in a wind tunnel. Therefore, it includes temporal dynamics of odor concentration which are due to turbulent dispersal. This feature of the dataset has given rise to a study demonstrating that information about source proximity can reliably be extracted from turbulent plumes using metal oxide sensors [36], which has been replicated independently [39] and confirmed using newly recorded data [38]. Such studies are not affected by the adverse effects discussed in the present study, since they do not attempt to identify odorants, but focus only on the temporal dynamics of odorant concentration induced by turbulence, which is largely independent of odorant identity.

Our study highlights that it is still difficult to obtain clean and reliable data for gas recognition benchmarks. Besides the challenges in designing and manufacturing a gas sensor setup, planning a recording campaign robust against drift could hold its own pitfalls that may not be evident from the outset, and even go undetected for years after publication, even for highly cited datasets. Our findings highlight once more the importance of thoroughly checking the validity of third-party datasets before using it as a basis to develop algorithms for gas sensing.

A few recommendations emerge from our analysis towards best practices for designing MOx gas sensor datasets and sharing them. First and foremost, it is imperative to use a reference gas at short time intervals that will allow the identification and quantification of deviations in sensor response. Second, individual gases or mixtures should ideally be presented in a pseudo-randomized order, as should any varied parameter (e.g. wind speed, hotplate voltage). If a randomized presentation order is not feasible, one should record multiple batches for the same set of parameters at separate points in time. Training and testing data splits should then be selected from batches that were time-separated (as in [47]), which would allow for a more realistic performance evaluation. Finally, external parameters that could affect sensor behavior should be measured and reported, e.g. ambient temperature and humidity, and the exact time of the recording. Some MOx gas sensor datasets that implemented such principles are openly available [12,13,15].

Reliable data is the foundation for progress in the development of algorithms for gas sensing. The large number of citations of the original publication of the dataset analyzed here indicates that such data is much sought after and of high value for the community. It underlines the requirement for future efforts to record and publicly share gas sensing data for the progress of the field as a whole.

CRediT authorship contribution statement

Nik Dennler: Conceptualization, Investigation, Formal analysis, Software, Visualization, Writing – original draft, Writing – review & editing. **Shavika Rastogi:** Conceptualization, Investigation, Validation, Writing – review & editing. **Jordi Fonollosa:** Writing – review & editing. **André van Schaik:** Conceptualization, Writing – review & editing, Supervision. **Michael Schmuker:** Conceptualization, Writing – review & editing, Supervision.

Declaration of Competing Interest

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

Code Availability Statement

Code for the analysis and for reproducing the figures is available[48] under the Apache 2.0 license (http://www.apache.org/licenses/LICENSE-2.0).

Acknowledgements

We thank A. J. Lilienthal, M. Psarrou and S. Sutton for fruitful discussions on multiple occasions, which led to valuable insights. MS was funded by the NSF/CIHR/DFG/FRQ/UKRI-MRC Next Generation Networks for Neuroscience Program (NSF award no. 2014217, MRC award no. MR/T046759/1), and the EU Flagship Human Brain Project SGA3 (H2020 award no. 945539). JF acknowledges the Spanish Ministry of Economy and Competitiveness DPI2017-89827-R, Networking Biomedical Research Centre in the subject area of Bioengineering, Biomaterials and Nanomedicine, initiatives of Instituto de Investigación Carlos III, Share4Rare Project (Grant agreement 780262), and ACCIÓ (Innotec ACE014/20/000018). JF also acknowledges the CERCA Programme/Generalitat de Catalunya and the Serra Húnter Program. B2SLab is certified as 2017 SGR 952.

References

[1] J.R. Stetter, P.C. Jurs, S.L. Rose, Detection of hazardous gases and vapors: pattern recognition analysis of data from an electrochemical sensor array, Anal. Chem. 58 (4) (1986) 860–866.
[2] D. Maier, R. Hulasare, B. Qian, P. Armstrong, et al., Monitoring carbon dioxide levels for early detection of spoilage and pests in stored grain, in: Proceedings of the 9th International Working Conference on Stored Product Protection, vol. 1, 2006, 117.
[3] A.J. Lilienthal, A. Loutfi, T. Duckett, Airborne chemical sensing with mobile robots, Sensors 6 (11) (2006) 1616–1678, https://doi.org/10.3390/s6111616.
[4] N. Alizadeh, H. Jamalabadi, F. Tavoli, Breath acetone sensors as non-invasive health monitoring systems: a review, IEEE Sens. J. 20 (1) (2020) 5–31, https://doi.org/10.1109/JSEN.2019.2942693.
[5] F. Loizeau, H.P. Lang, T. Akiyama, S. Gautsch, P. Vettiger, A. Tonin, G. Yoshikawa, C. Gerber, N. de Rooij, Piezoresistive membrane-type surface stress sensor arranged in arrays for cancer diagnosis through breath analysis, in: 2013 IEEE 26th International Conference on Micro Electro Mechanical Systems (MEMS), IEEE, 2013, pp. 621–4.
[6] J.A. Covington, S. Marco, K.C. Persaud, S.S. Schiffman, H.T. Nagle, Artificial olfaction in the 21st century, IEEE Sens. J. 21 (11) (2021) 12969–12990, https://doi.org/10.1109/jsen.2021.3076412.
[7] A. Ziyatdinov, S. Marco, A. Chaudry, K. Persaud, P. Caminal, A. Perera, Drift compensation of gas sensor array data by common principal component analysis, Sens. Actuators B: Chem., vol. 146(2), 2010, pp. 460–5, selected Papers from the 13th International Symposium on Olfaction and Electronic Nose. ⟨https://doi.org/10.1016/j.snb.2009.11.034⟩.
[8] J. Rodríguez, C. Durán, A. Reyes, Electronic nose for quality control of colombian coffee through the detection of defects in cup tests, Sensors 10 (1) (2010) 36–46, https://doi.org/10.3390/s100100036. ⟨https://www.mdpi.com/1424-8220/10/1/36⟩.
[9] A. Vergara, J. Fonollosa, J. Mahiques, M. Trincavelli, N. Rulkov, R. Huerta, On the performance of gas sensor arrays in open sampling systems using inhibitory support vector machines, Sens. Actuators B: Chem. 185 (2013) 462–477, https://doi.org/10.1016/j.snb.2013.05.027.
[10] A. Vergara, S. Vembu, T. Ayhan, M.A. Ryan, M.L. Homer, R. Huerta, Chemical gas sensor drift compensation using classifier ensembles, Sens. Actuators B: Chem. 166 (167) (2012) 320–329, https://doi.org/10.1016/j.snb.2012.01.074.
[11] J. Fonollosa, I. Rodríguez-Luján, M. Trincavelli, A. Vergara, R. Huerta, Chemical discrimination in turbulent gas mixtures with mox sensors validated by gas chromatography-mass spectrometry, Sensors 14 (10) (2014) 19336–19353, https://doi.org/10.3390/s141019336.
[12] 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, Sens. Actuators B: Chem. 215 (2015) 618–629, https://doi.org/10.1016/j.snb.2015.03.028.
[13] A. Ziyatdinov, J. Fonollosa, L. Fernández, A. Gutierrez-Gálvez, S. Marco, A. Perera, Bioinspired early detection through gas flow modulation in chemo-sensory systems, Sens. Actuators B: Chem. 206 (2015) 538–547, https://doi.org/10.1016/j.snb.2014.09.001.
[14] J. Fonollosa, L. Fernández, A. Gutiérrez-Gálvez, R. Huerta, S. Marco, Calibration transfer and drift counteraction in chemical sensor arrays using direct standardization, Sens. Actuators B: Chem. 236 (2016) 1044–1053, https://doi.org/10.1016/j.snb.2016.05.089.
[15] J. Burgués, J.M. Jiménez-Soto, S. Marco, Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Anal. Chim. Acta 1013 (2018) 13–25, https://doi.org/10.1016/j.aca.2018.01.062.
[16] J.C. Rodriguez Gamboa, E.S. Albarracin E, A.J. da Silva, T.A.E. Ferreira, Electronic nose dataset for detection of wine spoilage thresholds, Data Brief 25 (2019), 104202, https://doi.org/10.1016/j.dib.2019.104202.
[17] A. Vergara, J. Fonollosa, J. Mahiques, M. Trincavelli, N. Rulkov, R. Huerta, Gas sensor arrays in open sampling settings data set, ⟨http://archive.ics.uci.edu/ml/datasets/Gas.sensor.arrays.in.open.sampling.settings⟩, (Accessed 09 August 2021).
[18] C. Battaglino, G. Ballard, T.G. Kolda, A practical randomized CP tensor decomposition, SIAM J. Matrix Anal. Appl. 39 (2) (2018) 876–901, https://doi.org/10.1137/17m1112303.
[19] N. Vervliet, L.D. Lathauwer, A randomized block sampling approach to canonical polyadic decomposition of large-scale tensors, IEEE J. Sel. Top. Signal Process. 10 (2) (2016) 284–295, https://doi.org/10.1109/jstsp.2015.2503260.
[20] N. Imam, T.A. Cleland, Rapid online learning and robust recall in a neuromorphic olfactory circuit, Nat. Mach. Intell. 2 (3) (2020) 181–191, https://doi.org/10.1038/s42256-020-0159-4.
[21] J.-H. Choi, J.-S. Lee, EmbraceNet: a robust deep learning architecture for multimodal classification, Inf. Fusion 51 (2019) 259–270, https://doi.org/10.1016/j.inffus.2019.02.010.
[22] P. Zhou, Y.-D. Shen, L. Du, F. Ye, X. Li, Incremental multi-view spectral clustering, Knowl.-Based Syst. 174 (2019) 73–86, https://doi.org/10.1016/j.knosys.2019.02.036.
[23] J.G. Monroy, E.J. Palomo, E. López-Rubio, J. Gonzalez-Jimenez, Continuous chemical classification in uncontrolled environments with sliding windows, Chemom. Intell. Lab. Syst. 158 (2016) 117–129, https://doi.org/10.1016/j.chemolab.2016.08.011.
[24] H. Fan, V.H. Bennetts, E. Schaffernicht, A.J. Lilienthal, A cluster analysis approach based on exploiting density peaks for gas discrimination with electronic noses in open environments, Sens. Actuators B: Chem. 259 (2018) 183–203, https://doi.org/10.1016/j.snb.2017.10.063.
[25] J.C.R. Gamboa, A.J. daSilva, I.C.S. Araujo, ESAE, CMDA, Validation of the rapid detection approach for enhancing the electronic nose systems performance, using different deep learning models and support vector machines, Sens. Actuators B: Chem. 327 (2021), 128921, https://doi.org/10.1016/j.snb.2020.128921.
[26] P. Zhou, Y.-D. Shen, L. Du, F. Ye, Incremental multi-view support vector machine, in: Proceedings of the 2019 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, 2019, pp. 1–9. ⟨https://doi.org/10.1137/1.9781611975673.1⟩.
[27] T.G. Kolda, D. Hong, Stochastic gradients for large-scale tensor decomposition, SIAM J. Math. Data Sci. 2 (4) (2020) 1066–1095, https://doi.org/10.1137/19m1266265.
[28] A. Mishra, N.S. Rajput, D. Singh, Performance evaluation of normalized difference based classifier for efficient discrimination of volatile organic compounds, Mater. Res. Express 5 (9) (2018), 095901, https://doi.org/10.1088/2053-1591/aad3dd.
[29] I. Araujo, J. Gamboa, A. Silva, Deep learning models for classification of gases detected by sensor arrays of artificial nose, in: Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional, SBC, Porto Alegre, RS, Brasil, 2019, pp. 844–55. ⟨https://doi.org/10.5753/eniac.2019.9339⟩.
[30] N. Vervliet, Compressed Sensing Approaches to Large-scale Tensor Decompositions (Ph.D. thesis), KU Leuven, 2018. ⟨https://lirias.kuleuven.be/1741494?limo=0⟩.
[31] J. Monroy, J. González-Jiménez, Real-time odor classification through sequential Bayesian filtering, in: Proceedings of the 16th International Symposium on Olfaction and Electronic Noses, 2015.
[32] L. Gugel, Y. Shkolnisky, S. Dekel, Machine olfaction using time scattering of sensor multiresolution graphs, 2016, arXiv:1602.04358.
[33] I.-S. Chang, H.-j. Choi, G.-m. Park, Uci sensor data analysis based on data visualization, in: Proceedings of the Korean Society of Broadcast Engineers Conference, The Korean Institute of Broadcast and Media Engineers, 2020, pp. 21–4.
[34] S.H. Wang, T.I. Chou, S.W. Chiu, K.T. Tang, Using a hybrid deep neural network for gas classification, IEEE Sens. J. 21 (5) (2021) 6401–6407, https://doi.org/10.1109/JSEN.2020.3038304.
[35] J. Li, Z. Tao, Y. Wu, B. Zhong, Y. Fu, Large-scale subspace clustering by independent distributed and parallel coding, IEEE Trans. Cybern. 1 (c) (2021) 1–11, https://doi.org/10.1109/TCYB.2021.3052056.
[36] M. Schmuker, V. Bahr, R. Huerta, Exploiting plume structure to decode gas source distance using metal-oxide gas sensors, Sens. Actuators B: Chem. 235 (2016) 636–646, https://doi.org/10.1016/j.snb.2016.05.098.
[37] J. Burgues, Signal Processing and Machine Learning for Gas Sensors: Gas Source Localization with a Nano-drone (Ph.D. thesis), Universitat de Barcelona, 2019.
[38] J. Burgues, S. Marco, Wind-independent estimation of gas source distance from transient features of metal oxide sensor signals, IEEE Access 7 (2019) 140460–140469, https://doi.org/10.1109/access.2019.2940936.
[39] J. Burgués, S. Marco, Feature extraction for transient chemical sensor signals in response to turbulent plumes: application to chemical source distance prediction, Sens. Actuators B: Chem. 320 (2020), 128235, https://doi.org/10.1016/j.snb.2020.128235.
[40] K.-S. Lee, S.-R. Lee, Y. Kim, C.-G. Lee, Deep learning-based real-time query processing for wireless sensor network, Int. J. Distrib. Sens. Netw. 13 (5) (2017), 155014771770789, https://doi.org/10.1177/1550147717707896.
[41] J. Monroy, V. Hernandez-Bennetts, H. Fan, A. Lilienthal, J. Gonzalez-Jimenez, GADEN: a 3d gas dispersion simulator for mobile robot olfaction in realistic environments, Sensors 17 (7) (2017) 1479, https://doi.org/10.3390/s17071479.
[42] T. Schneider, N. Helwig, A. Schütze, Automatic feature extraction and selection for condition monitoring and related datasets, in: Proceedings of the 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2018, pp. 1–6. ⟨https://doi.org/10.1109/I2MTC.2018.8409763⟩.

[43] D. Mitchell, N. Ye, H.D. Sterck, Nesterov acceleration of alternating least squares for canonical tensor decomposition: momentum step size selection and restart mechanisms, Numer. Linear Algebra Appl. 27 (4) (2020), https://doi.org/10.1002/nla.2297.
[44] A. Cardellicchio, A. Lombardi, C. Guaragnella, Iterative complex network approach for chemical gas sensor array characterisation, J. Eng. 2019 (6) (2019) 4612–4616, https://doi.org/10.1049/joe.2018.5125.
[45] K. Gilman, L. Balzano, Grassmannian optimization for online tensor completion and tracking with the t-svd, 2021. arXiv:2001.11419.
[46] Figaro USA, Inc., ⟨http://www.figarosensor.com/⟩.
[47] S. Asadi, H. Fan, V.H. Bennetts, A.J. Lilienthal, Time-dependent gas distribution modelling, Robot. Auton. Syst. 96 (2017) 157–170, https://doi.org/10.1016/j.robot.2017.05.012.
[48] N. Dennler, Code: drift in a popular metal oxide sensor dataset reveals limitations for gas classification benchmarks, 2022. ⟨https://doi.org/10.5281/zenodo.6338624⟩.

Nik Dennler received the B. Sc. degree in Physics in 2018 and the M. Sc. degree in Neural Systems and Computation in 2020, both at the University of Zurich and ETH Zurich. During and after his studies, he worked in various research labs, such as Sensirion AG, the Sony Aerial Robotics Lab, IBM Research - Zurich and the Cantilever Sensor Lab of the University of Basel. Currently, he is a dual PhD candidate at the University of Hertfordshire (UK), and the Western Sydney University (AU), where he investigates on Neuromorphic Olfaction. His main interests are biologically inspired sensing mechanisms, event-based signal processing and neuromorphic engineering.

Shavika Rastogi received the B.Tech degree in Electronics and Communication Engineering from Mody University, Lakshmangarh (Raj.), India in 2017 and M.Sc. degree in Neuroscience from Albert-Ludwigs-Universität Freiburg, Germany in 2020 with Computational Neuroscience as her focus area. She is currently a PhD student in the field of Neuromorphic Olfaction/Event-based gas sensing, at the International Centre of Neuromorphic Systems at Western Sydney University, Australia in collaboration with the Biocomputation Research group at University of Hertfordshire, UK. Her research interests include olfaction, brain inspired computing and she wants to utilise her skills in making present day gas sensing as efficient as biological olfactory system.

Jordi Fonollosa received the Ph.D. degree in electronic engineering from the University of Barcelona in 2009. He joined the Universitat Politècnica de Catalunya in 2017, where he is Associate Professor under the Serra Húnter Program. His research efforts are focused on the development of algorithmic solutions for chemical detection systems. He has applied chemical sensing to a variety of applications, such as food quality control, fire detection, non-invasive human activity monitoring, and air quality control. He has also applied information theory to chemical sensing systems. His other strong interests include biologically inspired algorithms, signal recovery systems, and infrared sensing technologies. More information at http://b2slab.upc.edu/.

André van Schaik received the M.Sc. degree in electrical engineering from the University of Twente, Enschede, The Netherlands, in 1990 and the Ph.D. degree in electrical engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1998. He is the Director of the International Centre of Neuromorphic Systems at Western Sydney University. His research focuses on neuromorphic engineering and computational neuroscience.

Michael Schmuker is Professor of Neural Computation at the University of Hertfordshire. He obtained his Ph.D. in Chemistry from Goethe-University Frankfurt am Main in 2007, and his M.Sc. in Biology from Albert-Ludwigs University Freiburg in 2003. Michael's areas of expertise include Olfaction in animals and machines, Neuroscience, Neuromorphic Computing and Chemical Informatics. His research aims at translating computational processes observed in the brain into algorithms for sensory computing and machine learning.