Science of the Total Environment 1004 (2025) 180741

Contents lists available at ScienceDirect

Science of the Total Environment

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

Portable ion mobility spectrometry and partial least squares discriminant analysis for odour source discrimination in wastewater treatment plants

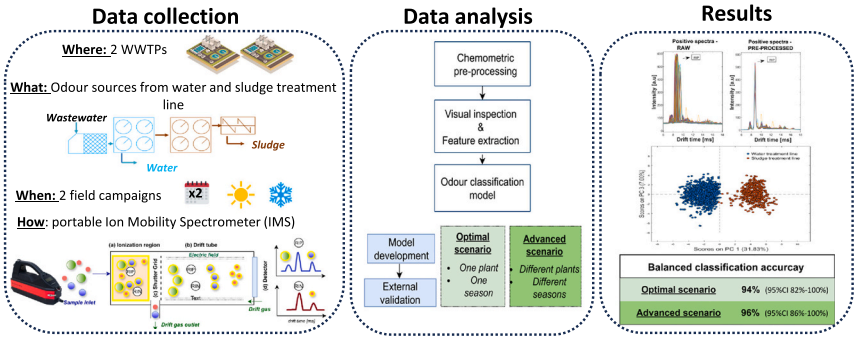
Veronica Villa [a], Luis Fernandez [b,c,*], Beatrice Julia Lotesoriere [a], Javier Alonso-Valdesueiro [b], Agustín Gutiérrez-Gálvez [b], Laura Capelli [a], Santiago Marco [b,c]

[a] Politecnico di Milano, Department of Chemistry, Materials and Chemical Engineering Giulio Natta (DCMC), Piazza Leonardo da Vinci 32, 20133, Milano, Italy
[b] Department of Electronics and Biomedical Engineering, Universitat de Barcelona, 08028, Barcelona, Spain
[c] Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, 08028, Barcelona, Spain

HIGHLIGHTS

- Portable Ion Mobility Spectrometry discriminates odour sources in wastewater plants.
- Achieved 94 % balanced accuracy within a single plant and season.
- Model transfer across two different plants with 96 % balanced accuracy.
- Ionic information is key for reliable odour source classification.
- Promising approach for improved real-time instrumental odour monitoring in the field.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Spectral analysis
Environmental monitoring
Pollutant
Volatile organic compounds
Process control
Odour nuisance

ABSTRACT

Odour emissions from Wastewater Treatment Plants are a relevant issue concerning environmental and social impact, regulatory compliance, and plant management. Instrumental Odour Monitoring Systems are widely used for real-time odour emissions monitoring, but seasonal and plant variability limit their long-term reliability. Therefore, new sensing technologies and approaches are being studied to improve their reliability and the transferability of predictions between different plants and seasons. In this context, this work investigates the suitability of portable Ion Mobility Spectrometers to discriminate the main odour sources in Wastewater Treatment Plants. Two measurement campaigns were carried out in different seasons, considering different odour sources in two independent plants. Through a proper data analysis approach, based on the importance of ionic information, portable Ion Mobility Spectrometry proved effective in discriminating odour sources from the two main process lines: water and sludge treatment. In the first phase, conducted in the same plant and season, a balanced classification rate of 94 % (95 %CI: 82 %–100 %) was achieved. Subsequently, including seasonal and plant variability, a model trained on one plant was applied to the second. The direct transfer of the calibration achieved a balanced classification accuracy of 96 % (95 %CI: 86 %–100 %), confirming the relevance of the selected ions for odour assessment across different plants. These results suggest that portable Ion Mobility Spectrometry is a technology that deserves further attention for instrumental odour monitoring. The consistent

* Corresponding author at: Department of Electronics and Biomedical Engineering, Universitat de Barcelona, 08028, Barcelona, Spain.
E-mail address: lfernandez@ub.edu (L. Fernandez).

https://doi.org/10.1016/j.scitotenv.2025.180741
Received 7 August 2025; Received in revised form 1 October 2025; Accepted 14 October 2025
Available online 24 October 2025

classification rates obtained both within a single plant and when transferring the model demonstrate that Ion Mobility Spectrometry, combined with feature selection, can reliably identify ions specifically relevant for odour emission assessment.

1. Introduction

In recent years, specific regulations and methodologies have been developed for monitoring odour emissions from various industrial activities, responding to the increasing public awareness of air quality (Bax et al., 2020). While odours are typically non-toxic and not directly harmful to human health, prolonged exposure has been linked to stress, headaches, nausea, and respiratory discomfort, affecting overall well-being (Piccardo et al., 2022). At the European Union level, several regulations have addressed the issue of odour exposure. EN 13725:2022 (CEN, 2022) introduced dynamic olfactometry to measure odour concentration, while EN 16841:2016 (CEN, 2016) standardized field inspections to assess odour exposure in surrounding areas. In particular, Italy has issued specific regulations on odour impact assessment, further developing the European regulatory framework (Bokowa et al., 2021). This regulatory development has opened a path for setting acceptability criteria for odour emissions ensured by standards in odour monitoring.

To ensure regulatory compliance, chemical analysis, dynamic olfactometry, and dispersion modelling are effective tools to characterize odour emissions and evaluate their impact (Muñoz et al., 2010). However, monitoring emissions is not exclusively about complying with acceptability criteria. From an industrial perspective, it also serves as a key management tool. Anomalous or excessive odour emissions often signal malfunctions or inefficiencies, requiring intervention and mitigation (Brinkmann et al., 2016). Therefore, early detection and characterization of odour release are fundamental also to reducing operational costs.

Among the most relevant pollution-emitting plants (Bax et al., 2020; Muñoz et al., 2010), Wastewater Treatment Plants (WWTPs) are significant contributors. The sludge treatment line is the most intense source due to anaerobic processes in sludge storage, digestion, and dewatering units, releasing high levels of H_2S, NH_3, and VOCs (González et al., 2022; Haider et al., 2022; Senatore et al., 2021). Water treatment units, such as grit chambers and primary sedimentation tanks generate less intense but more widespread odour emissions, due to their large surface areas (Czarnota et al., 2023; González et al., 2022; Haider et al., 2022; Lebrero et al., 2011; Muñoz et al., 2010; Senatore et al., 2021). The diversity of odour sources makes their control and monitoring challenging. Therefore, identifying whether odours originate from water or sludge treatment is important for targeted mitigation, industrial efficiency, and regulatory compliance (Barbu et al., 2018; Czarnota et al., 2023).

To tackle this, Instrumental Odour Monitoring Systems (IOMS) are widely employed in WWTPs for monitoring odour emissions directly over the emitting sources (Blanco-Rodríguez et al., 2018; Prudenza et al., 2022) or at the plant fence line boundary (Cangialosi et al., 2021). They have proved effective in classifying and quantifying odour emissions from different process stages (Blanco-Rodríguez et al., 2018; Burgués et al., 2022, 2021; Moufid et al., 2022; Prudenza et al., 2022; Wang et al., 2023).

IOMS are technologies designed to continuously monitor and analyse ambient air, correlating sensor outputs with odour presence and intensity (Oliva et al., 2021). Among these devices, electronic noses (EN) are the most used. They are instruments equipped with an array of gas sensors and a data processing unit that is able to detect, classify, and quantify odours in real time. EN can be equipped with partially specific or non-specific gas sensors, as industrial odour emissions usually do not depend on a single compound but rather on a complex fingerprint of VOCs. Examples include Metal Oxide Sensors (MOX) (Blanco-Rodríguez et al., 2018; Burgués et al., 2022, 2021; Moufid et al., 2022; Prudenza et al., 2022; Wang et al., 2023), Electrochemical sensors (EC) (González et al., 2022; Prudenza et al., 2022), and Photoionization Detectors (PID) (González et al., 2022; Prudenza et al., 2022). Therefore, data acquired by EN typically undergo visual inspection, for instance by means of Principal Component Analysis (PCA) (Blanco-Rodríguez et al., 2018; Burgués et al., 2021; Moufid et al., 2022; Prudenza et al., 2022).

By coupling EN responses with suitable machine learning approaches, it is possible to classify different odours using algorithms such as Linear Discriminant Analysis (Wang et al., 2023), Multilayer Perceptron, or Support Vector Machine (SVM) (Moufid et al., 2022; Prudenza et al., 2022; Wang et al., 2023) but also to estimate odour "quantity" using methods such as Partial Least Squares (PLS) Regression (Blanco-Rodríguez et al., 2018; Burgués et al., 2022, 2021; Moufid et al., 2022; Wang et al., 2023), Random Forest (RF) (Cangialosi et al., 2021), and Artificial Neural Networks (ANN) (Cangialosi et al., 2021). In general, EN can be employed directly over the emitting sources (Blanco-Rodríguez et al., 2018) for process control purposes (Prudenza et al., 2023, 2022), to enhance plant management and interventions to resolve odour emissions issues (Zarra et al., 2022) or to map emissions from different sources (Burgués et al., 2022). They also provide better understanding of different contributions in odour emissions, or insight into wastewater quality (Wang et al., 2023). The use of EN for real-time emission monitoring at plant fence lines is also becoming increasingly popular (Cangialosi et al., 2021). This type of deployment produces information about general anomalous malfunctioning, permits the implementation of alarm thresholds for odour emissions (Cangialosi et al., 2021), and provides information to identify the origin of odour nuisance (Cangialosi et al., 2021).

However, despite EN effectiveness and adaptability to different industrial applications and environments, EN are generally affected by several problems, including sensitivity to humidity and temperature fluctuations, as well as instrumental drift (Khorramifar et al., 2023; Robbiani et al., 2023). Lack of reproducibility and generalizability are also major challenges, requiring considerable time and effort for dedicated calibration of each instrument and facility, as models cannot be directly transferred between identical instruments or similar plants. These issues are worsened by seasonal variability and plant-specific conditions, making it difficult to establish generally applicable calibration models. Many of these issues are intrinsic to the sensor technologies employed in EN and, although several machine learning approaches attempt to mitigate them, they remain a major challenge in real-world applications. There is a significant amount of research on calibration transfer methods in laboratory settings (Fernandez et al., 2016; Fonollosa et al., 2016; Reimringer and Bur, 2023; Robin et al., 2023), but, as far as we are aware, there is still a lack of approaches for applying them to more complex, real-world situations.

In this context, other sensor technologies could be considered as complementary approaches to enhance odour monitoring capabilities. There are some laboratory applications of gas chromatography coupled with ion mobility spectrometry (GC-IMS) that have proved effective in identifying distinct VOC fingerprints from samples collected at different stages of a WWTP process, highlighting the chemical differences between process lines and their respective odour impacts. Gas chromatography separates mixture components based on their chemical properties, while Ion Mobility Spectrometry (IMS) distinguishes ions by their drift velocity in an electric field, adding a second dimension to the analysis. However, GC-IMS systems are typically fixed laboratory instruments, limiting their applicability for real-time field monitoring.

Therefore, our purpose is to explore IMS standalone technology as a tool for identifying odour sources in WWTPs. This technology allows us

to exploit ion mobility spectrometry in a portable system, making it well-suited for field application and onsite odour monitoring (Epping and Koch, 2023; Márquez-Sillero et al., 2011). However, IMS analysis in environmental conditions can be affected by factors such as slight variations in temperature, pressure, and relative humidity, as well as background interference, leading to issues like misalignment, noise, and baseline drift, making it challenging to compare different analyses when considering only raw spectra data. To address these challenges, specific signal and data processing methods were applied to ensure comparable analysis across different plants and seasons and data consistency.

To the best of our knowledge, this approach in real field applications for characterizing odour sources in WWTPs has not been implemented before. Therefore, this work aims to investigate the potential of a portable IMS to classify odour emissions from WWTPs, representing a novel approach and application. Furthermore, beyond the analytical performance, the cost-effectiveness of the proposed approach is particularly relevant, as portable IMS reduces infrastructure requirements and enables in-situ monitoring at lower operational costs compared to conventional laboratory-based techniques.

Samples were collected in two different WWTPs during two measurement campaigns directly over the emitting sources from the two main process lines, i.e., the water and sludge treatment. The potential of IMS to classify odours from the two process lines was investigated, along with a data analysis approach that ensured objective and repeatable results. This approach enabled the preliminary generalization of a classification model from one plant to another, reducing both calibration and computational demands and illustrating the feasibility of the method.

The article is organized as follows: in Section 2, Materials and Methods are discussed. Relevant information about the Instrumentation, Sampling sites in the plants, measurement campaigns, and data analysis is provided. In Section 3, Results and Discussion presents the most important features of the collected data and the results provided for the trained models are presented and discussed. Also, a proper discussion about the implications of this study on odour control and management technology is provided. In Section 4, conclusions based on the results presented in Section 3 and future work derived from the discussion provided in the same section are summarized. Finally, a Table of Acronyms is provided in the Supplementary Materials.

2. Materials and methods

2.1. Instrumentation and technology

In this study, the use of a handheld IMS device minimizes both equipment and maintenance costs compared to more complex GC–IMS setups. While still providing selective and sensitive measurements, this equipment is more suitable for field deployment. Data was collected directly at the WWTPs emission sources with an IMS: GDA2 from AIRSENSE Analytics GmbH. It is a handhewass detector array equipped with 4 gas sensors (1 Photo Ionization Detector (PID), 1 Electrochemical Cell (EC), and 2 Metal Oxide Sensors (MOX)) and one water chemistry Ion Mobility Spectrometer (IMS) able to alternate between positive and negative modes. The GDA2 is also equipped with an automatic dilution system: the sampled gas of interest can be diluted with filtered clean air at different ratios that can be set by the operator. In this study, we focused exclusively on the spectra signals from the IMS in the GDA2 with a scan rate of 1 s and a sample rate of 18.25 KHz.

IMS is an analytical technique used to detect volatile organic compounds (VOCs) through ionization of the sampled molecules. In the GDA2 device, ionization is initiated by a Ni-63–β radioactive source emitting β particles (17 keV) (Gabelica and Marklund, 2018), which initially ionize N_2 and O_2 molecules (Gabelica and Marklund, 2018). The atmospheric pressure chemical reaction with water molecules in air forms positively charged H_3O^+ ions (reactant positive ions, RIP), while interactions with oxygen molecules generate negatively charged O_2^- ions (reactant negative ions, RIN). Once formed, these reactant ions serve as intermediaries for the ionization of target molecules (M). Indeed, RIP and RIN act as "charge" reservoirs, transferring positive or negative charge to the target molecules (M) through collisions, ionizing them via proton transfer in positive mode (electron transfer), or proton abstraction and cluster formation in negative mode (Gabelica and Marklund, 2018). In other words, when the peak area corresponding to an ion increases in the spectrum, the area of the RIP (or RIN) decreases accordingly.

The ionized molecules are introduced into a drift region where they are accelerated by a homogeneous electric field (Fig. 1). Their terminal velocities, influenced by their mass, charge, and cross-section, cause them to reach the detector at different drift times, typically in the millisecond range (Fig. 1). The resulting drift times produce peaks in the spectra, each corresponding to a specific ion, providing a unique chemical fingerprint of the sampled gas (Novillo and Verónica, 2015). This fingerprint reflects the molecular composition of the sample, distinguishing various target compounds based on their interactions with RIP and RIN.

2.2. Sampling sites

The sampling sites were two WWTPs, from now on designated as Plant 1 (P1) and Plant 2 (P2) and schematically represented in Fig. 2. WWTPs are facilities designed to reduce wastewater pollution to acceptable limits by means of several unit operations. The two plants considered in the study are located adjacent to each other. While they share some similarities, they operate independently in most of their processes. They treat different amounts of civil residues, measured in daily flow of water in cubic meters (Q), with flowrates of QP1 = 130.000 m^3/day, and QP2 = 217.000 m^3/day from different sources, i.e., providing the service to different municipalities. WWTPs typically include two main process lines: the water treatment line (WTL) and the sludge treatment line (STL) (Fig. 2) (Barbu et al., 2018). In P1, WTL includes a water collector, a grid chamber for screening, and a desander, followed by four primary sedimentation tanks operating in parallel (Fig. 2). These are followed by two biological tanks and, finally, four secondary sedimentation tanks (Fig. 2). P2 is composed of the same process units as P1 including an additional physical-chemical treatment section after the wastewater collector and before the primary sedimentation (Fig. 2). However, during the measurement campaigns the physical-chemical treatment section of P2 was not operational, meaning that, in practice, the water treatment lines of both plants were identical in terms of process units. The water separated during the secondary sedimentation stage in both plants is subsequently treated together through a filtration and chlorination process, with the addition of UV treatments to reach potability standards (Fig. 2). The solid residues from the WTL, referred to as sludge, are processed in a second line, called the STL. In P1, this line consists of a sludge conveyor, three thickeners (two of which operate while one is under maintenance on a rotational basis), a sludge mixer, and a biodigester to produce biogas for energy generation (Fig. 2). Similarly, P2 includes two floaters (to remove suspended solids, grease and colloids from the water) and a biodigester also dedicated to biogas production for energy purposes.

The digested residual sludge from both plants' biodigesters is then treated together through mechanical dehydration to produce a final solid product that can be used as fertilizer or disposed of as waste.

The two plants can be treated as independent when considering the water treatment line, up to the secondary sedimentation, and for the sludge treatment line, up to the biodigesters (Fig. 2). Most of these processes are associated with relevant diffuse sources of odour and VOCs characterized by plant, seasonal, and temporal variability depending not only on the process units directly involved but also on the variable chemical composition of the treated wastewater from sewage (Beghi et al., 2012; Lebrero et al., 2011).

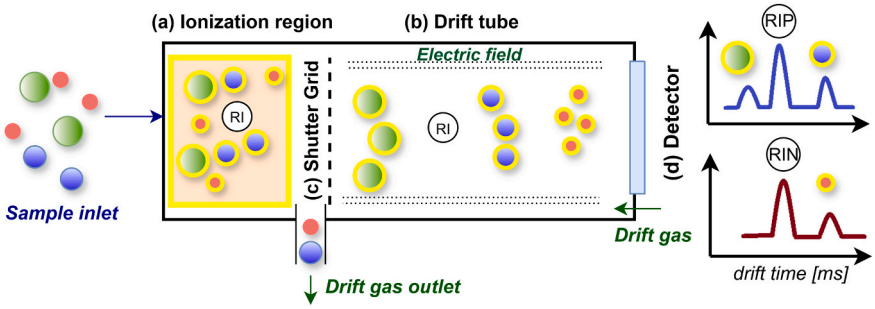


Fig. 1. Schematic representation of Ion Mobility Spectrometer (IMS). (a) Ionization source region in which the sample is ionized, (b) Drift tube where the ionized molecules are accelerated by an electric field, (c) shutter grid allows the ionized molecules go into drift tube and (d) detector where the charge of molecules is converted into a current output. RI: reference reactant ions; RIP: positive reactant ions; RIN: negative reactant ions.

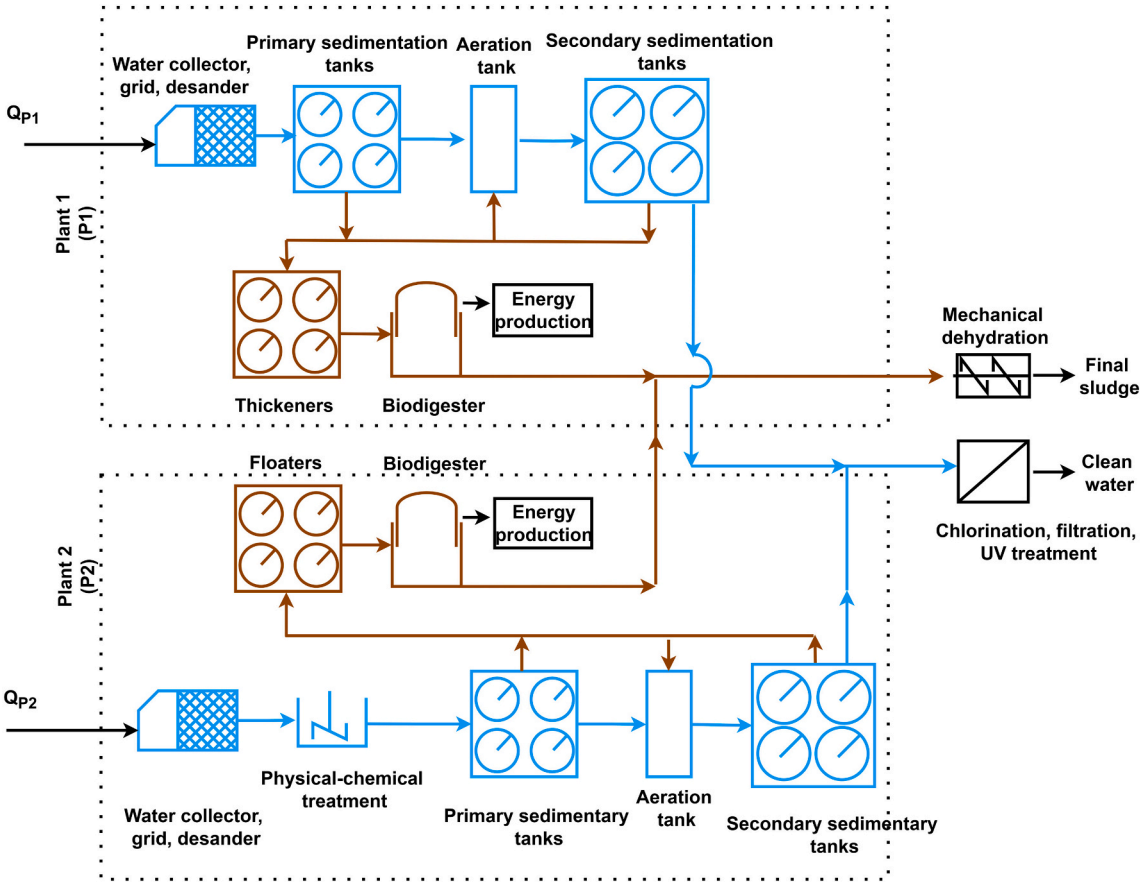


Fig. 2. Schematic representation of the process unit in Plant 1 (P1) and Plant 2 (P2), the main independent sections and the common final units of mechanical dehydration and water purification.

2.3. Measurement campaigns and samples collection

Two measurement campaigns were carried out, one in January and one in June 2023, to account for the odour-emitting sources variability due to seasonal and inlet flow rate wastewater variations. Each measurement campaign included both plant P1 and plant P2. Samples were collected directly over the sources in field conditions employing the GDA2 with its automatic dilution system to avoid sensors' saturation, set at 10 % gas sample and 90 % filtered air. Field measurements were organized to sample from both WTL and STL in both plants, focusing on the most emission-intensive sources. The January measurement campaign was carried out over two days, the first day sampling P1 sources and the second day sampling P2. The June measurement campaign was carried out over three consecutive days, sampling every day in both P1 and P2. In both campaigns and for each plant, samples were collected randomly and with repetitions from both classes, WTL and STL. A schematic report of the sampled sources in P1 and P2 during the days of measurement in January and June is reported, respectively, in Table 1 and Table 2.

P2 dataset presents less homogeneity in data distribution because of adverse weather conditions during June field measurements. P1 dataset, on the other hand, was better structured, more numerous, and with a more evenly distributed set of each class across the different measurement days, making it more suitable for calibration.

2.4. Data processing

Data processing included an initial pre-processing of the collected

Table 1
Number of samples collected in P1 with the GDA2 during January and June field campaigns.

P1							
	WTL				STL		
	Wastewater collector	Grid/ Screening	Primary sedimentation	Aeration tank	Thickener	Mixer	Total n° of independent samples per day
Day 1 (January)	1	–	2	1	2	1	7
Day 2 (June)	–	–	2	–	1	3	6
Day 3 (June)	1	1	–	–	1	1	4
Day 4 (June)	1	2	3	–	2	2	10
Total n° of independent samples per class	14				12		26

Table 2
Number of independent samples collected in P2 with GDA2 during January and June field campaigns.

P2						
	WTL			STL		
	Wastewater collector	Grid/ screening	Primary sedimentation	Floater	Sludge deposit	Total number of independent samples per day
Day 1 (January)	–	1	1	–	1	3
Day 2 (June)	–	–	–	2	1	3
Day 3 (June)	–	–	–	1	2	3
Day 4 (June)	3	2	5	2	2	14
Total n° of independent samples per class	12			11		23

spectra to prepare the data, followed by exploration of patterns using principal component analysis (PCA), extraction of the most important information, and, finally, the development and validation of a model to classify the different odours.

The pre-processing involved basic signal processing algorithms to enhance signal quality because spectra sampled under field conditions often contain interferences, background noise, baseline drift, and peak shifts. The proposed feature extraction relies on peak intensities in both the positive and negative polarity. Therefore, accurate comparison of peak heights requires initial pre-processing to align the peaks, ensuring they are mapped consistently across different samples, and removing the baseline and noise. After the feature extraction stage, the dimensionality of the feature vector is substantially reduced. Principal Component Analysis (PCA) was applied to visualize the reduced dataset and assess its potential for odour classification.

Two different scenarios were considered for the odour classification models: the first one focused on a single plant and a single field campaign (P1 June), while the second one considered both plants across two seasons (P1 and P2, in January and in June).

In scenario 1, two models were developed for odour source classification using the P1 dataset collected in June. The algorithm used for this purpose was Partial Least Squares Discriminant Analysis (PLS-DA). The first model was trained on data from the three days of measurements in June, excluding one sample at a time for external validation developing several train-test combinations. As described above, each sample is composed of several spectral points, which are highly correlated with each other. Consequently, for a fair comparison, internal cross-validation (ICV) on each training set determined the optimal number of latent variables (LVs) for the model. Each model was externally validated on the independent test set, and its performance was assessed in terms of a final Balanced Classification Accuracy based on the prediction of each external validation. In classification tasks where the number of samples per class is approximately balanced, the classification rate (overall accuracy) is a suitable measure of classifier performance. However, when the dataset is substantially imbalanced, the majority class tends to dominate, leading to an overestimation of the classification rate. To fairly evaluate classification quality under such conditions, it is more appropriate to use the Balanced Classification Accuracy (BCA) defined in Eq. (1), which equally weights the accuracy achieved in each class and thus provides a more reliable performance estimate across imbalanced datasets. In addition, since the BCA is a random variable (it depends on the sample set), it is important not only to report its point estimate but also to characterize its variability. Confidence intervals serve this purpose: they define the range within which we would expect the observed value of the statistic to for a given percentage of repeated experiments (typically 95 %, although other levels such as 99 % can also be used). Narrower confidence intervals indicate that the estimated performance is more stable and reproducible. As a general principle, the larger the sample size, the narrower the confidence intervals, thereby increasing our confidence in the robustness of the classifier's performance. An additional optimization step was performed by selecting the most relevant features based on Variable Importance in Projection (VIP) scores from the PLS-DA model. Raw predictions with the PLS-DA algorithm are made based on the single spectrum A second model was then built using these selected features, with the number of latent variables re-optimized. This refined model was subsequently validated on the independent test set. After confirming the classification of the two odour sources within a single plant and campaign, Scenario 2 expanded the analysis to cover a broader range of variability. Again, two classification models were created. One model was trained on the entire data from P1 (January and June) and validated on the independent data from P2 (January and June) exploring the capacity of the method to be applied across plants. Internal cross-validation and BCA on the external validation set were used to assess performance. The other model was also developed using VIP scores for feature selection, following the same procedure as in scenario 1. A detailed explanation of this procedure can be found in Section 2.9.

Data analysis was carried out in MATLAB (2024B version) with PLS_Toolbox 9.1 (Eigenvector Research) and a specific package for IMS signal processing SmartIMS (Oller Moreno, 2018), previously developed by the authors. Both traditional programming methods and the Graphical User Interface (GUI) were employed. The data processing steps are schematized in Fig. 3.

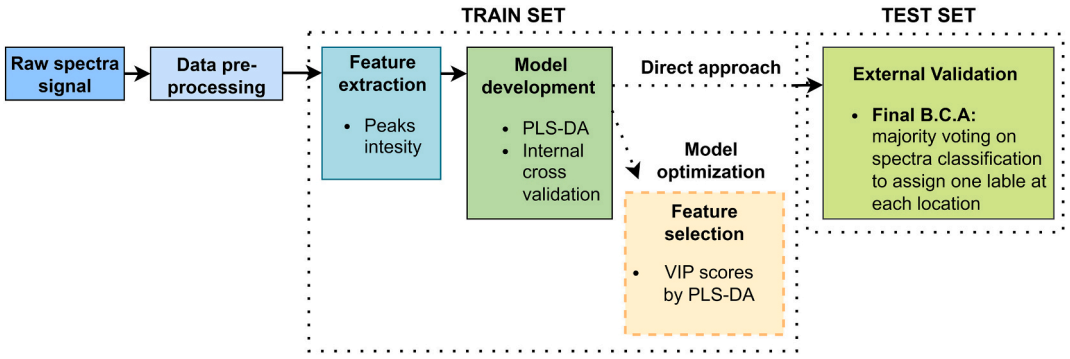


Fig. 3. Data processing steps.

2.5. IMS signals data structure

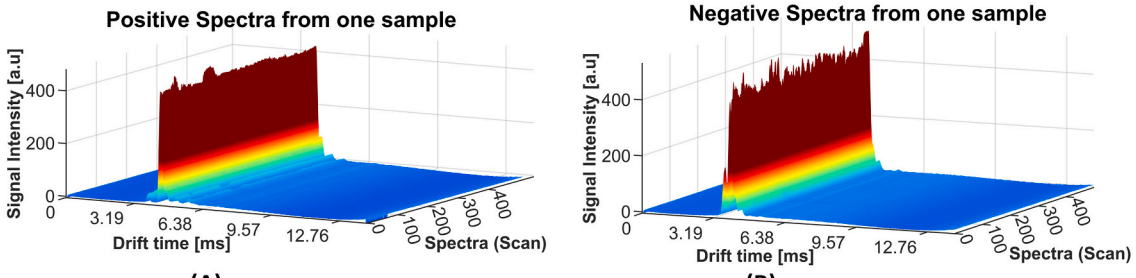
GDA2 acquires a time series of alternately positive and negative ion mobility spectra (Pomareda Sesé, 2013). Each scan contains *2 m* (m = 511) spectral samples resulting in *m* points from the positive spectra and m points for the negative spectra. For data processing purposes, the signals were organized in a 2D matrix, as shown in Fig. 4. The rows of the matrix were arranged by grouping all scans for each sample (k) together sequentially: the first n_1 rows correspond to the scans from sample k = 1, the next n_2 rows correspond to the scans from sample k = 2 and so on (Fig. 4). The spectral acquisition at each location lasted from 3 to 10 min per sample, according to the specific case (for practical reasons in the field). Consequently, the number of spectra for each measurement (n_k) ranged from 180 to 600.

2.6. Spectra pre-processing

To improve spectral quality noise was reduced by 2nd order Savitzky-Golay filters of optimized length. Baseline was estimated with a low order polynomial and peak alignment was based on a multiplicative factor with respect to the RIP or RIN position (Novillo and Verónica, 2015; Oller Moreno, 2018; Savitzky and Golay, 1964; Szymańska et al., 2016; Vu and Laukens, 2013). The specific parameters for each pre-processing algorithm (window width and polynomial order in baseline correction) were optimized by visual inspection of the pre-processed spectra (refer to Section 3.4). Across the dataset, there are minor variations in peak positions between spectra from different samples due to temperature and humidity differences. In this step, all scans were aligned with respect to chosen reference RIP and RIN values for mitigating this misalignment. This alignment stage allowed for consistent comparison of peak positions and intensities across samples and provided a basis for reliable feature extraction. The data pre-processing steps are schematized in Fig. 5.

2.7. Feature extraction

Peak heights were extracted as features, as they indicate the concentration of ions. Each peak represents the presence of a specific ion (monomer or dimer), providing a unique ion-fingerprint of the analyzed samples and valuable information about their composition (Gabelica and Marklund, 2018; Gao et al., 2021; Vera et al., 2016). Peak heights were extracted from the positive and negative spectra of WTL and STL of the training dataset.



		Drift time (positive spectra)			Drift time (negative spectra)		
		0,0318 [ms]	...	28,5 [ms]	0,0318 [ms]	...	28,5 [ms]
Sample 1	Scan 1						
	...						
	Scan n_1						
Sample k	Scan 1						
	...						
	Scan n_k						

(C)

Fig. 4. (A) positive spectra; (B) negative spectra; (C) Dataset structure with k samples, nk scans/spectra each sample, m positive spectral samples and m negative spectral samples.

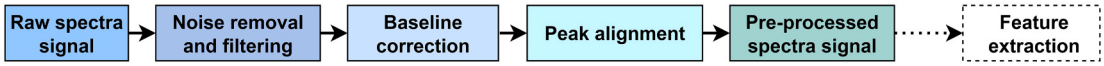


Fig. 5. Data pre-processing steps.

2.8. Exploratory analysis

After pre-processing and feature extraction, PCA was used for data visualization and interpretation to assess its ability to distinguish odour sources based on plant and seasonal variability (Bro and K. Smilde, 2014). Score plots of the principal components (PCs) were examined to identify patterns. At each plot, the amount of explained variability in percentage (the proportion of total variability explained by each principal component) has been added to the component label to aid the discussion presented in Section 3. Due to the variability of emissions across plants and over time in WWTPs, PCA was applied step by step in this study to explore differences between samples, campaigns, and plants, supporting the development of the classification models.

2.9. Model building, optimization and validation

The proposed preprocessing and modelling pipeline relies on standard algorithms (baseline correction, peak alignment, PLS-DA), which are computationally lightweight and easily transferable, thereby supporting both reproducibility and cost-efficiency in data analysis. The PLS-DA algorithm was used to develop the odour classification models. Two different scenarios were considered in the development and validation of classification models. The first scenario included data from one plant in one season. In this regard, P1 dataset from June proved most suitable in terms of homogeneity and data distribution. Then, the second scenario included both plants and both seasons, P1 and P2 in January and June, to assess the model's transferability. In both cases, the dataset was partitioned into a train set for model development with proper internal cross validation (ICV) and feature selection, and a test set for external validation. Data were auto-scaled with respect to the train set.

In the first scenario, specifically, a double leave-one-out scheme was performed to partition data in train and external validation and for internal cross validation. Iteratively, one sample from P1 June dataset was used for external validation. The remaining samples constituted the calibration set (Table 3). For each of the 19 Cal–Ext.Val combinations, ICV was performed using a leave-one-block-out (LOBO) approach. A LOBO approach was chosen, meaning that iteratively all spectra from one sample were excluded from the calibration set and used as part of the test set for internal validation, while the remaining samples formed the training set. The entire block of spectra from the same sample was excluded at each iteration from training set to avoid information leakage. Since they belong to the same measurement, they were highly correlated and not independent. Each of the 18 resulting training–validation combinations were tested with different numbers (q) of latent variables (LVs), ranging from 1 to N, where N is the number of extracted features. For every $LV_{i,q}$, a table that shows how well the model's predictions match the true labels, highlighting correct predictions and errors (Confusion Matrix or CM) was produced. Each prediction for the 18 samples resulted in a single confusion matrix, $CM_{i,q}$.

The Balanced Classification Accuracy (BCA) was calculated for each LV in each partition was calculated following Eq. (1):

$$BCA = 1 - BER \tag{1}$$

where BER is the Balanced Error derived from the confusion matrix following Eq. (2):

$$BER = 0.5\left(\frac{FN}{TP+FN} + \frac{FP}{FP+TN}\right) \tag{2}$$

Here, TP, TN, FP, and FN are respectively the True Positives, True Negatives, False Positives, and False Negatives from the sample prediction confusion matrix (CM). Uncertainty of BER was estimated by the bootstrap method.

For every partition i, the $BER_{i,q}$ can be plotted against the number of $LV_{i,q}$. The "knee-point" of this plot, where the error stops decreasing significantly, was identified as the optimal number of LVs for the partition i, balancing model complexity and classification performance.

Once the internal cross-validation was performed, the optimal number of latent variables (LVs) for each *i*-calibration set was retained to develop *i* classification models (Table 4). Each developed model was then tested on the corresponding test set for external validation (according to the partition Table 3). Final prediction was performed at the level of individual spectra with final measurement classifications determined by weighted majority (≥50 % agreement among scans), accounting for class imbalance during calibration through a corrective weighting factor. The prediction in external validation of each i Cal-Ext. Val combination was collected in a single and final CM to calculate a final BCA. BCA with respect to measurement classification was used as performance indicator. Additionally, a step of model optimization was considered before external validation: keeping the optimal number of LVs, a PLS-DA model was developed on the calibration set to compute the VIP scores and select the most relevant features. Internal cross validation was then repeated to re-optimize the number of LVs according to the same ICV scheme but employing only the selected variables. These optimized models were then externally validated on the test sets and performances were evaluated in terms of samples classification. Again, a final CM aggregating the results in external validation was built to calculate a final BCA for scenario 1.

After the performances in a reference scenario, within the same plant (P1) and within a single measurement campaign (i.e., one season, June), were investigated, a second model including the entire dataset of P1 (January and June) was developed and validated on P2 (January and June).

In this case, the calibration dataset (P1 January–June) is composed of four days of measurement. The same data analysis pathway was performed but for internal cross-validation, a leave-one-day-out (LODO) approach was applied, considering each measurement day as an independent block (Table 1). In each iteration, data from three blocks were used for model training, while the fourth block was kept for validation

Table 3
Data partition of P1 June dataset for i Cal-Ext.Val combinations where C are the data for calibration while E the data for external validation.

Scenario 1: P1 June												
Data partition i = 1, …,19												
	WTL				STL						Combination Cal-Ext.Val	
Partition 1	E	C	C	C	C	C	C	C	C	C	à	Model 1
Partition 2	C	E	C	C		C	C	C	C	C	à	Model 2
Partition **i**	…	…	…	…	…	…	…	…	…	…	à	Model **i**
Partition 19	C	C	C	C	C	C	C	C	C	E	à	Model 19

Table 4
Data splitting for internal cross validation: Train dataset (T) and internal validation (V) with P1 June dataset for each i Ext.Val-Train partition. Data partition i = 1 is reported as an example.

Cal-Ext.Val partition $i = 1$ Train-test combinations $j = 1, \ldots, 18$										Calibration set SET: P1 June (Scenario 1) *Internal cross validation*	
										Output for each LV à $CM_{i,q}$	Final output à $BER_{i,q}$ and $BCa_{i,q}$
Model i,j	**WTL**			**STL**						$\sum_{j=1}^{18}\left(Samples\ predictions_{i,q}\right)_j$ à $\boldsymbol{CM_{1,q}}$	**$BER_{1,q}$** and **$BCA_{1,q}$** per q = 1,2, …,N for **optimal number of LVs**
Model 1,1	V	T	T	T	T	T	T	T	T		
Model 1,2	T	V	T	T	T	T	T	T	T		
Model 1,j	…	…	…	…	…	…	…	…	…		
Model 1,18	T	T	T	T	T	T	T	T	V		

Table 5
Data splitting for internal cross validation: Train dataset (T) and internal validation (V) with P1 January and June as calibration dataset.

					TRAIN SET: P1, January and June (Scenario 2) *Internal cross validation*	
	Day 1	Day 2	Day 3	Day 4	Output for each LV à CM_q	Final output à BER_q and BCA_q
Model 1	V	T	T	T	$\sum_{j=1}^{18}\left(Samples\ predictions_{i,4}\right)_j$ à $\boldsymbol{CM_{1,q}}$	**BER_q** and **BCA_q** per q = 1,2, …,N for **optimal number of LVs**
Model 2	T	V	T	T		
Model 3	T	T	V	T		
Model 4	T	T	T	V		

(Table 5). As described above, the number of LVs was optimized based on samples predictions, before and after performing the features selection with VIP scores. Finally, the entire P1 dataset (i.e., all four blocks) was used to develop the final calibration model, which was subsequently tested on the P2 dataset. Model performance on P2 samples was evaluated using BER and BCA with respect to sample classifications based on spectra-majority voting as described for scenario 1.

2.10. PLS-DA modelling

The PLS-DA algorithm was implemented to build the classification models capable of discriminating odour sources by combining dimensionality reduction with discriminant analysis. PLS-DA is a supervised technique particularly suitable for datasets with more features than observations (2 m > k) and where multicollinearity exists among the predictors. It uses the features X (n x N) extracted from sensor signals, where n represents the number of samples and N the number of features. The method works by extracting latent variables (LVs), which are linear

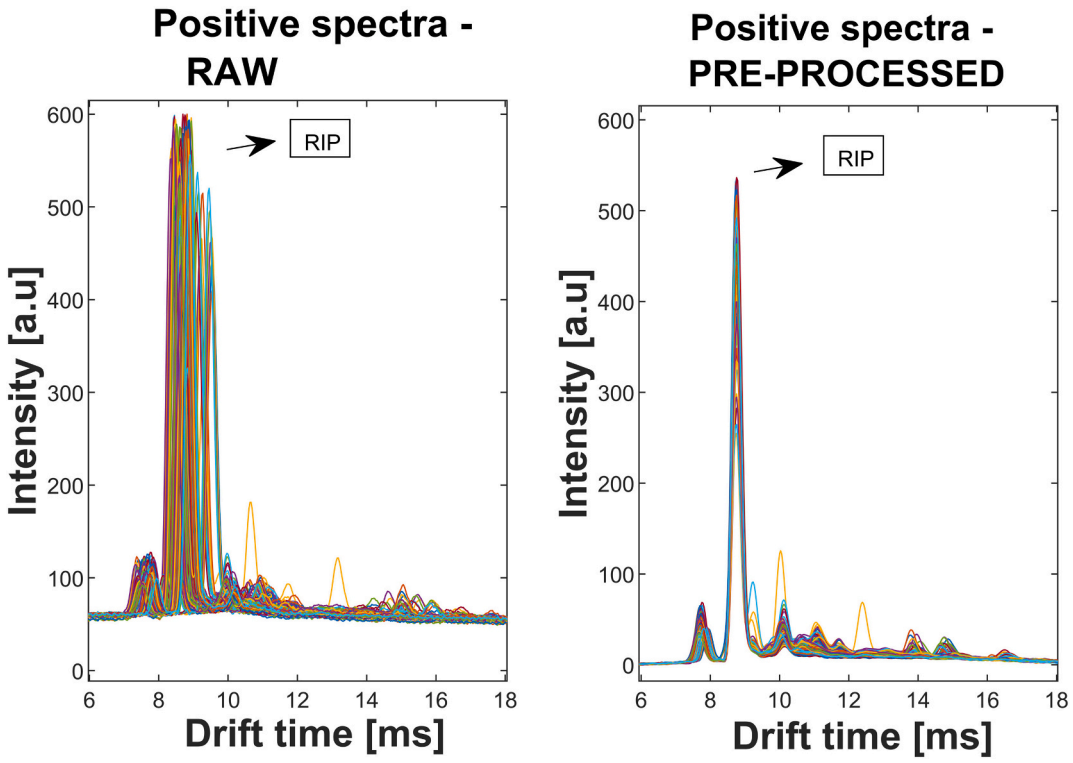


Fig. 6. Positive spectra from field samples collected during June measurements from both plants P1 and P2: (A) Raw spectra signals, (B) Pre-processed spectra signals after noise removal, baseline correction and spectra alignment respect to the same reference RIP and RIN position.

combinations of the original features, to reduce the dataset's dimensionality while maximizing the covariance between the predictors and the class labels. This approach enables the identification of key features influencing the classification and provides a robust model for separating observations into distinct groups based on their characteristics (Lee et al., 2018; Marco and Gutierrez-Galvez, 2012).

2.11. Feature selection

Feature importance was evaluated on each training set using VIP scores, which quantify each variable's contribution to the LVs. Features with VIP scores greater than 1 were considered significant and selected, while those below 1 were excluded. This process reduced the number of features to an optimal subset for model development (Mehmood et al., 2012).

3. Results and discussions

3.1. Raw sensor signal pre-processing

Fig. 6 shows an example of positive spectra points from all the samples collected in both P1 and P2 during field measurements, before and after pre-processing. The samples were collected from different odour sources, at different times, and from different plants (as described in Section 2.3 *Measurement campaigns and samples collection*). In the raw signals extracted from IMS (Fig. 6(A)), noise, baseline and peaks misalignments were present. The 2nd order Savitzky-Golay (SVG) filters were optimized in length, 10 was identified as the optimal width, avoiding excessive noise without distorting the shape of the peaks. Multiplicative spectral alignment based on RIP/RIN position was applied to standardize peak positions and facilitate ion matching within and across samples. Spectral alignment is a key step; omitting this step impairs feature extraction and reduces the model's ability to discriminate samples' odour class. Baseline removal also allowed a fair extraction of peak intensities. In this case, the baseline removal resulted optimal with a fourth order polynomial fitted using the first 16 % and the last 33 % of each individual spectrum where no peaks are present.

3.2. Feature extraction

Because of different chemical composition, samples from different origins resulted in different spectra. The differences are evident in terms of ionic presence, i.e., in terms of peaks in the spectra. Differences are both qualitative, meaning that peaks are in different positions because they are derived from different chemical compounds, and quantitative, meaning that same nominal peaks in different samples can have different heights due to varying quantities of the same chemical compound. PCA was first applied to the full positive and negative spectra. In this case, neither classes nor macro-classes were distinguishable; redundancy was evident in the data with general overlapping in the PCs space. Most of the variance in the definition the PCs was associated to the RIP and RIN peaks. Since these reference ions are not chemically informative for the discrimination of the odour sources they were removed from the spectra. Additionally, feature extraction was based on the intensity of the ions peaks, resulting in a significant reduction of dimensionality. The extraction of the features was performed on the calibration dataset for each scenario investigated.

Feature extraction in both scenario (P1 June and P1 January and June) resulted in the identification of 17 different peaks in the positive spectra signals and 17 different peaks in the negative spectra signals. A total of 34 peaks were identified across all the spectra. The RIP and the RIN were discarded as suggested above.

3.3. Exploratory analysis

The PCA score plots of the pre-processed signals, and the 32 extracted features were computed to understand the relationship among data and the instrument's potential in differentiating between two odour classes of interest, WTL and STL. In general, in the next plots and graphics, data concerning the WTL are coloured in blue and light-blue shades while data from STL are in orange and brown shades. First, data from the same measurement campaign (January) and the same plant (P1) were examined (Fig. 7). From visual inspection, it is visible that the information to distinguish WTL and STL classes is contained in PC1. In Fig. 7, the results are reported in the PC1-PC3 plane for visualization purposes. Indeed, along the PC1 scores space (Fig. 7(A)) the samples belonging to the same process lines, WTL or STL, cluster together, forming two clear distinguishable groups (Fig. 7(B)). This statement is also reinforced by the amount of explained variability corresponding to each PC. While PC3 represents only 7 % of the total variability of the dataset, PC1 represents almost 32 % of the variability, leading to a visual separation of the dataset. These two principal components represent almost 40 % of the variability of the dataset which explains the scattering observed in Fig. 7 and the need of better models to achieve clear source discrimination. Similar results are observed in Figs. 8 and 9.

This outcome highlights the instrument's ability to differentiate between the two odour classes analyzed under controlled conditions (i.e., the same measurement campaign and the same plant). The differentiation between the two odour classes would allow the identification of the odour source from one process line to another distinguishing the causes and enabling targeted mitigation interventions.

Next, the study considers the effects of seasonal and plant variability that characterize the WWTPs' emissions. Seasonal variability was evident, since the same emitting sources, e.g., WTL samples from P1, behave differently in January compared to June, resulting in distinct clusters for each season (Fig. 8(A)) in the PC score plots. Analogously, plant variability was observed, as samples from the same emitting sources collected during the same season at different plants (e.g., WTL samples in January from P1 and P2) grouped into distinct regions in the PC score plot (Fig. 8(B)). STL samples from different seasons and different plants behave similarly to the WTL samples shown, clustering differently according to season and plants.

Finally, when combining data from both plants across the two seasons (January and June) the variability is much higher, making it harder to distinguish between WTL and STL (Fig. 9). The samples showed more dispersion into the score plots without clear clustering of the two odour (Fig. 9). This highlights how seasonal and plant variability of WWTPs' emissions sources adds complexity to differentiating between the two odour classes. These characteristics are driven by the intrinsic nature of odour emissions from WWTPs, which are strongly influenced by wastewater composition, environmental factors and process parameters (Pasciucco et al., 2023). It is important to account for these aspects during model calibration and to manage this complexity in real-field scenarios. For these reasons, we considered developing at first a model in a reference case considering one plant (P1) in one season (June) and then including seasonal variability with a calibration on the entire P1 (January and June).

3.4. PLS-DA model building, optimization and validation

The models were built as described in Section 2.9. In the first scenario, when only P1 in June was considered, the evaluation of the BER vs. LVs (Balanced Error vs number of Latent Variables) curve, derived from internal cross-validation across the different combinations, indicated a number of LVs ranging from 1 to 3, with 2 being the most frequently selected value. Among these 14 peaks, 11 were in the positive spectra at positions 7.3 ms, 8.8 ms, 9.2 ms, 9.6 ms, 10.2 ms, 10.7 ms, 11.3 ms, 13.3 ms, 14.4 ms, 14.6 ms, 16.3 ms, and 3 in the negative spectra in position 6 ms, 7.9 ms, 9.6 ms respectively. The same features were extracted from the samples for external validation. The first model, calibrated and validated within the same plant (P1) and same season

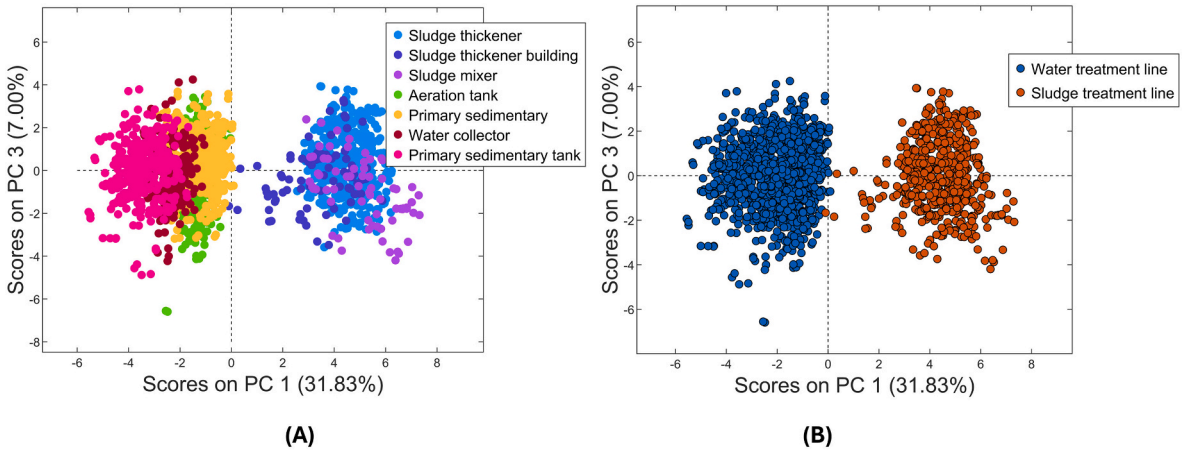


Fig. 7. PC1 and PC3 scores plot of P1 samples from January field campaign coloured by: (A) The independent samples (Sludge thickener, sludge thickener building, Sludge mixer, Aeration tank, Primary sedimentary tank, Water collector, Primary sedimentary); (B) The two odour classes WTL and STL.

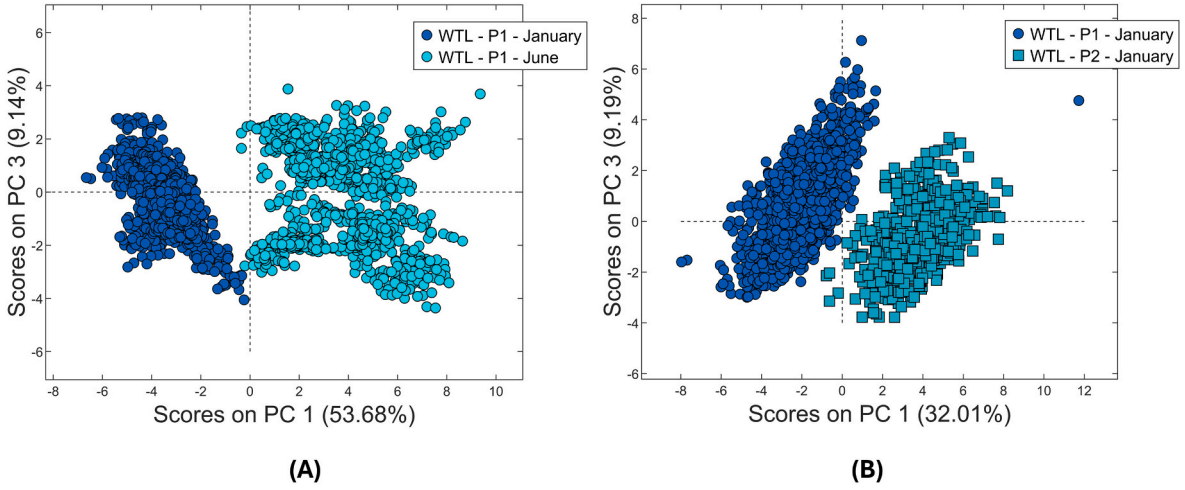


Fig. 8. PC1 and PC3 scores plot of: (A) WTL samples from P1 collected in January and the ones collected in June; (B) WTL samples from P1 and P2 collected in January. Data from P1 are represented by circle markers while P2 as squared markers.

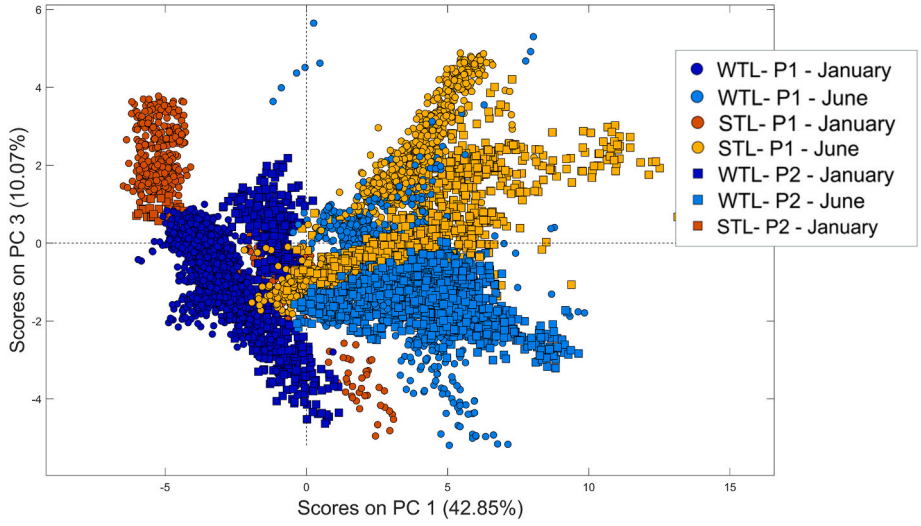


Fig. 9. PC1 and PC3 scores plot of the whole dataset: samples collected from P1 and P2 in January and June, both WTL and STL.

(June), achieved an 89 % Balanced Classification (*95 %CI* 75 %–100 %) when externally validated without the step of model optimization (i.e., calibrated with all the 32 extracted features). The same model developed with the thirteen VIP-selected features resulted in a balanced classification of 94 % (*95 %CI* 82 %–100 %) (Table 6).

In the second scenario, when the entire P1 dataset (January and June) was considered for calibration, two LVs were identified as the optimal number. Among the 32 extracted peaks, 7 were selected: at positions 7.3 ms, 8.8 ms, 9.6 ms, 10.2 ms from the positive spectra, and the ones at positions 6.5 ms, 7.9 ms, 8.9 ms from the negative spectra. The same features were extracted from the validation dataset from P2.

This second model achieved 82 % balanced classification accuracy (*95 %CI* 64 %–97 %) in external validation on data from P2 January and June considering all the features, and 96 % (*95 %CI.* 86 %–100 %) when only the 7 selected features were used for model development (Table 6).

This demonstrates not only the potential of classifying samples in controlled conditions, such as the same plant and the same seasons, but also the model's strong performance in classifying new samples from a different plant, even in the presence of seasonal and plant variability. The classification model's ability to handle the complexity introduced by these variabilities, as observed in Section 3.3, is particularly significant. This indicates that the supervised PLS-DA model can partially mitigate the effects of seasonal and plant variability by focusing on the classification of target odour sources.

3.5. Discussion

Despite the challenges posed by seasonal and plant variability, as evident from PC scores plot analysis (Figs. 7, 8 and 9), the PLS-DA classification model effectively used the information contained in the spectral signals features, i.e., the peak heights, addressing this variability, at least for the differences between plants. In both cases, scenario 1 and scenario 2, the optimization step in model development improved classification performance. When different seasons are included in model calibration, a lower number of features was selected as the most important suggesting that the most essential information for the classification of two sources is retained in a few peaks. The development of approaches the investigation of new techniques to monitor odour emissions in WWTPs with generalized and reliable systems are important under different perspectives. Distinguishing between odour sources from WTL and STL is valuable for understanding the specific contributions of these process lines to overall emissions in WWTPs. This differentiation is particularly relevant for implementing targeted preventive and mitigation strategies, optimizing plant operations and control, and ensuring compliance with environmental regulations. Prevention and mitigation of odour emissions from a WWTP may vary between WTL and the STL, and to implement targeted and cost-effective interventions, it is crucial to identify the primary sources and address them correctly (Muñoz et al., 2010). When odours mainly originate from the WTL, mitigation and prevention efforts should focus on the early stages of wastewater collection. At these stages, covering systems can be installed to contain odour emissions, coupled with localized ventilation (Brinkmann et al., 2016; Muñoz et al., 2010; Senatore et al., 2021). If significant emissions are detected and covering is not sufficient, air collection systems with forced ventilation can be implemented to convey odorous air to treatment units. Common treatment methods include biofilters, which degrade odour compounds using bacteria; these typically use organic material like compost or wood chips, as they handle lower odour loads (Brinkmann et al., 2016; De Sanctis et al., 2022; Senatore et al., 2021). Chemical scrubbers and thermal or catalytic oxidation may also be applied (Senatore et al., 2021).

Additionally, wastewater additives, such as oxidizing agents (chlorine, potassium permanganate) or bacteria-based products, help prevent hydrogen sulphide formation and the associated malodours (Brinkmann et al., 2016). On the other hand, when the odour mainly originates from STL, different targeted interventions are applicable (Muñoz et al., 2010; Senatore et al., 2021). In the STL, where odours mainly result from anaerobic fermentation of organic material, prevention strategies involve covering and isolating sludge thickening tanks and dewatering stations, with dedicated air extraction and conveyance systems (Brinkmann et al., 2016; Haider et al., 2022; Senatore et al., 2021). The recovered air is treated using activated carbon filters, bio-scrubbers, or biofilters. Due to higher odour loads, biofilters in this line often incorporate zeolites or activated carbons to enhance adsorption capacity (Márquez et al., 2021). Chemical additives to the sludge, such as iron salts or oxidants, help reduce hydrogen sulphide formation.

Moreover, optimizing sludge retention times, as well as ensuring proper transport and storage under controlled conditions, can minimize uncontrolled fermentation and reduce unwanted emissions (Haider et al., 2022; McNevin and Barford, 2000; Muñoz et al., 2010; Senatore et al., 2021). Therefore, identifying the main odour sources among the WTL and the STL is a critical step to implement suitable process control strategies aimed at reducing odour impacts. While these lines are interdependent, different control strategies are typically used to ensure optimal and efficient operations (Barbu et al., 2018). Commonly monitored variables include ammonia, nitrogen, chemical oxygen demand (COD), biochemical oxygen demand (BOD), and oxygen content, which help assess the quality of the effluent water, ensure compliance with effluent limits, and control operational costs (Barbu et al., 2018, 2016). Summarizing, an efficient IOMS capable of discriminating odour sources and quantifying odour emissions from WTL or STL would contribute to improve process control and further optimizing plant management.

Accordingly, the final performance of IOMS can be improved by accurately classifying the odour sources in a plant and then appropriately adjusting the prediction models. In this contribution, the relevance of having equipment capable of analyzing the chemistry of different source emissions in source classification has been highlighted. Furthermore, the associated problems with portable IMS technology (plant-to-plant and temporal variabilities) have been studied by using data coming from two different plants. These problems have been addressed by designing a clever data processing pipeline. Indeed, the potential of the IMS instrument in classifying odour sources in WWTPs is clearly demonstrated through the application of the PLS-DA algorithm. However, the verification of the robustness of the model and the

Table 6
Classification results.

Model	Train	Test	N° LVs	Approach	Features	BCA (95 % CI)
First scenario	P1 June 1st-2nd day	P1 June 3rd day	1÷3 (ν0 2)	Direct	All 32 features	89 % (75 %–100 %)
				Optimized	14 selected features *(Positive: 7,3 ms, 8,8 ms, 9,2 ms, 9,6 ms, 10,2 ms, 10,7 ms, 11,3 ms, 13,3 ms, 14,4 ms, 14,6 ms, 16,3 ms; Negative: 6,0 ms, 7,9 ms, 9,6 ms)*	94 % (82 %–100 %)
Second scenario	P1 January and June	P2 January and June	2	Direct	All 32 features	82 % (64 %–97 %)
				Optimized	7 selected features *(Positive: 7,3 ms, 8,8 ms, 9,6 ms, 10,2 ms; Negative: 6,5 ms, 7,9 ms, 8,9 ms*	96 % (86 %–100 %)

repeatability of these results will require more measurement campaigns covering at least one-year variability. Moreover, the direct transferability of models across plants should be assessed, including more facilities of the same category in future studies.

To further enhance the performance of the IOMS, Calibration Transfer techniques may be considered as well (Fernandez et al., 2016; Fonollosa et al., 2016; Reimringer and Bur, 2023; Robin et al., 2023). The transferability property, if further confirmed or developed, would significantly enhance the field application of odour monitoring using this approach, drastically reducing calibration work, as well as field and computational effort. In the data processing pipeline presented here, PCA and PLS-DA algorithms proved effective for visualization and processing of IMS data, but these are not the only possible approaches.

In the scenarios considered (measurements within the same plant and campaign, and measurements across two plants and campaigns), consistent results in classification were obtained with both full feature models and those built after feature selection. Models based on selected features achieved better performance, which can be explained by the removal of variables that contributed little to class separation or that were strongly affected by noise. When comparing the confidence intervals of the BCA for both approaches, a large overlap was observed. This is a direct consequence of the limited number of independent samples used for model building. Increasing the number of samples would be necessary to confirm whether the models based on selected features are significantly better. For instance, we observed narrower CIs in the second scenario due to the larger number of samples. Alternative techniques could potentially improve classification accuracy, and they might provide valuable insights into the effectiveness and versatility of various methods in similar scenarios. Overall, these preliminary results on real field samples demonstrates the potential of IMS coupled with robust data analysis to classify odour emissions from WWTPs, laying the foundation for further research into its application for environmental odour monitoring.

It should be highlighted that, while there are is an increasing attention towards standardizing the way IOMS performances are evaluated, as proven by the recent Italian standard UNI 11761:2023 and the CEN Working Group WG 41, to date, there are no established regulatory requirements for IOMS minimum classification accuracy. In order to define some benchmark accuracy, it is possible to refer to previous recently published work discussing the attempt to develop IOMS for environmental odour monitoring at plant fencelines, reporting classification accuracies above 95 % on independent test sets after optimization of the model for compensation of humidity effects on sensors (Ratti et al., 2024), which are indeed comparable to the accuracy levels obtained in this study. As an additional consideration, portable IMS may allow both the identification of relevant chemical species as a subsequent step to the classification presented here. At present, only the spectral features (ion peaks) associated with discrimination are known. Assigning these features to ions corresponding to chemical species would require complementary analyses by gas chromatography coupled with mass spectrometry (GC–MS) to provide information on chemical composition.

Taken together, these aspects highlight that our approach is not only technically feasible but also cost-effective, offering a portable and relatively low-maintenance solution for odour source classification in WWTPs and the prediction of odour concentrations when proper calibration is performed.

4. Conclusions

As odour monitoring in WWTPs has become more necessary than ever, IOMS have emerged as convenient tools for emission monitoring, source detection, and managing compliance with regulatory standards. Due to the limitations of EN used as IOMS, it is important to explore the possibilities offered by other technologies which might improve odour source classification in the field. In this context, IMS may represent a valuable solution to be considered as a replacement for EN in field monitoring and classification. The results of this study demonstrate that the combined approach of IMS data, spectral analysis and classification models, such as PLS-DA, is effective in achieving a high discrimination capability between the two primary odour sources in a WWTP (i.e., WTL and STL). The optimized model, calibrated and validated within the same plant and the same measurement campaign, performed with an 89 % (*95 %CI* 75 %–100 %) balanced accuracy in predicting new samples. In controlled conditions with low seasonal variability, the potential of IMS-PLSDA is well demonstrated in discriminating the WTL from the STL and the most important information are retained in few peaks. This capability might contribute to evaluating the impact of each process line on the overall odour impact of the plant, and thus, to implementing targeted odour control strategies in a cost-effective manner.

The impact of seasonal and plant variability, considering datasets from different seasons (winter and summer) and from different WWTPs, has also been demonstrated in this contribution. Accounting for seasonal variability is particularly important, especially for real-world field monitoring in WWTPs, which requires continuous and reliable information about the odour emitted from the facility over a long period of time. Plant variability is also an issue when aiming to transfer a classification model between different plants, thereby reducing operational and computational effort needed to calibrate the monitoring systems. Here, we have presented a modelling approach that tackles this issue. A second model was calibrated using data from one plant across different seasons and successfully validated on data from a different plant, achieving a balanced classification accuracy of 96 % (*95 %CI.* 86 %–100 %). It was also shown that, when seasonal and plant variability are present, fewer chemical features (i.e., fewer peaks) can be retained to address the complexity of transferring a model from one plant to another.

Considering the performance of the presented models, this study highlights with this study the potential of IMS-PLSDA combination as a promising tool for odour source classification and source monitoring, not only in reference and controlled conditions but also across different facilities and with different environmental conditions. However, the limited number of sources and the variety of plants included does not allow generalization of the presented results.

To confirm the robustness of the model and the reproducibility of the results, further studies are required. Future work should include additional measurements across different seasons and from a larger number of WWTPs. This would allow for assessing the model's stability in real-world scenarios and introducing potential adjustments to enhance accuracy. Confidence intervals are expected to improve with a larger dataset for external validation. Additionally, exploring alternative data pre-treatment methods and classification models could offer valuable insights and potentially improve system performance. Such investigations would also contribute to a better understanding of how this technology can be deployed in a standardized and reproducible manner. Furthermore, extending the use of portable IMS as IOMS beyond this study could offer reliable tools to address issues of interest to the industrial community, such as the quantification of odour emission rates. Integrating odour source classification (via portable IMS), odour concentration quantification (complementing IMS with GC–MS and dynamic olfactometries), and airflow rate estimation techniques, could provide a very useful tool for odour monitoring (odour source control) and managing (malodour event forecasting) at an industrial level.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2025.180741.

CRediT authorship contribution statement

Veronica Villa: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation. **Luis Fernandez:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Data curation. **Beatrice Julia**

Lotesoriere: Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Javier Alonso-Valdesueiro:** Validation, Resources, Formal analysis, Data curation. **Agustín Gutiérrez-Gálvez:** Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Laura Capelli:** Writing – review & editing, Validation, Supervision, Resources, Investigation. **Santiago Marco:** Writing – review & editing, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Santiago Marco has patent issued to Universitat de Barcelona. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We would like to acknowledge, the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2021 SGR 01393); the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya; and the European Social Fund (ESF). This project has received funding from ATTRACT, a European Union Horizon 2020 research and innovation project, under grant agreement No. 101004462. 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. Authors of this report gratefully acknowledge the cooperation Silvia Doñate, Lidia Sauco, Lara Terren, and Mª Deseada Esclapez, from Depuración de Aguas del Mediterraneo.

Data availability

The authors do not have permission to share data.

References

Barbu, M., Vilanova, R., Santin, I., 2016. Control strategies for the sludge line in wastewater treatment plants. In: 2016 IEEE Int. Conf. Autom. Qual. Test. Robot. AQTR, pp. 1–6. https://doi.org/10.1109/AQTR.2016.7501332.

Barbu, M., Santin, I., Vilanova, R., 2018. Applying control actions for water line and sludge line to increase wastewater treatment plant performance. Ind. Eng. Chem. Res. 57, 5630–5638. https://doi.org/10.1021/acs.iecr.7b05298.

Bax, C., Sironi, S., Capelli, L., 2020. How can odors be measured? An overview of methods and their applications. Atmosphere 11, 92. https://doi.org/10.3390/atmos11010092.

Beghi, S.P., Santos, J.M., Reis Jr., N.C., de Sá, L.M., Goulart, E.V., de Abreu Costa, E., 2012. Impact assessment of odours emitted by a wastewater treatment plant. Water Sci. Technol. 66, 2223–2228. https://doi.org/10.2166/wst.2012.409.

Blanco-Rodríguez, A., Camara, V.F., Campo, F., Becherán, L., Durán, A., Vieira, V.D., de Melo, H., Garcia-Ramirez, A.R., 2018. Development of an electronic nose to characterize odours emitted from different stages in a wastewater treatment plant. Water Res. 134, 92–100. https://doi.org/10.1016/j.watres.2018.01.067.

Bokowa, A., Diaz, C., Koziel, J.A., McGinley, M., Barclay, J., Schauberger, G., Guillot, J.-M., Sneath, R., Capelli, L., Zorich, V., Izquierdo, C., Bilsen, I., Romain, A.-C., del Carmen Cabeza, M., Liu, D., Both, R., Van Belois, H., Higuchi, T., Wahe, L., 2021. Summary and overview of the odour regulations worldwide. Atmosphere 12, 206. https://doi.org/10.3390/atmos12020206.

Brinkmann, T., Giner, S.G., Yukseler, H., Roudier, S., Delgado, S.L., 2016. Best available techniques (BAT) reference document for common waste water and waste gas treatment/management systems in the chemical sector. In: Industrial Emissions Directive 2010/75/EU (Integrated Pollution Prevention and Control) [WWW Document]. JRC Publ. Repos. https://doi.org/10.2791/37535.

Bro, R., K. Smilde, A., 2014. Principal component analysis. Anal. Methods 6, 2812–2831. https://doi.org/10.1039/C3AY41907J.

Burgués, J., Esclapez, M.D., Doñate, S., Marco, S., 2021. RHINOS: a lightweight portable electronic nose for real-time odor quantification in wastewater treatment plants. iScience 24, 103371. https://doi.org/10.1016/j.isci.2021.103371.

Burgués, J., Doñate, S., Esclapez, M.D., Saúco, L., Marco, S., 2022. Characterization of odour emissions in a wastewater treatment plant using a drone-based chemical sensor system. Sci. Total Environ. 846, 157290. https://doi.org/10.1016/j.scitotenv.2022.157290.

Cangialosi, F., Bruno, E., De Santis, G., 2021. Application of machine learning for fenceline monitoring of odor classes and concentrations at a wastewater treatment plant. Sensors 21, 4716. https://doi.org/10.3390/s21144716.

CEN, 2016. Standard EN 16481-1:2016. Ambient Air-Determination of Odour in Ambient Air by Using Field Inspection-Part 1: Grid Method. https://standards.iteh.ai/catalog/standards/cen/8ddf0231-5dcb-438f-b9f8-f13329f2eb2f/en-16841-1-2016?srsltid=AfmBOorIXw6ke7S0ph7Jt2cVdA9eLZoqb2E7Het9MdawjpdznnXdPqCM.

CEN, 2022. Standard EN 13725:2022 Stationary source emissions - Determination of odour concentration by dynamic olfactometry and odour emission rate. https://store.uni.com/en/en-13725-2022.

Czarnota, J., Masłoń, A., Pajura, R., 2023. Wastewater treatment plants as a source of malodorous substances hazardous to health, including a case study from Poland. Int. J. Environ. Res. Public Health 20, 5379. https://doi.org/10.3390/ijerph20075379.

De Sanctis, M., Murgolo, S., Altieri, V.G., De Gennaro, L., Amodio, M., Mascolo, G., Di Iaconi, C., 2022. An innovative biofilter technology for reducing environmental spreading of emerging pollutants and odour emissions during municipal sewage treatment. Sci. Total Environ. 803, 149966. https://doi.org/10.1016/j.scitotenv.2021.149966.

Epping, R., Koch, M., 2023. On-site detection of volatile organic compounds (VOCs). Molecules 28, 1598. https://doi.org/10.3390/molecules28041598.

Fernandez, L., Guney, S., Gutierrez-Galvez, A., Marco, S., 2016. Calibration transfer in temperature modulated gas sensor arrays. Sens. Actuators B Chem. 231, 276–284. https://doi.org/10.1016/j.snb.2016.02.131.

Fonollosa, J., Fernández, L., Gutiérrez-Gálvez, A., Huerta, R., Marco, S., 2016. Calibration transfer and drift counteraction in chemical sensor arrays using direct standardization. Sens. Actuators B Chem. 236, 1044–1053. https://doi.org/10.1016/j.snb.2016.05.089.

Gabelica, V., Marklund, E., 2018. Fundamentals of ion mobility spectrometry. Curr. Opin. Chem. Biol. 42, 51–59. https://doi.org/10.1016/j.cbpa.2017.10.022.

Gao, W., Yang, X., Zhu, X., Jiao, R., Zhao, S., Yu, J., Wang, D., 2021. Limitations of GC-QTOF-MS technique in identification of odorous compounds from wastewater: the application of GC-IMS as supplement for odor profiling. Atmosphere 12, 265. https://doi.org/10.3390/atmos12020265.

González, D., Colón, J., Sánchez, A., Gabriel, D., 2022. Multipoint characterization of the emission of odour, volatile organic compounds and greenhouse gases from a full-scale membrane-based municipal WWTP. J. Environ. Manage. 313, 115002. https://doi.org/10.1016/j.jenvman.2022.115002.

Haider, K.M., Lafouge, F., Carpentier, Y., Houot, S., Petitprez, D., Loubet, B., Focsa, C., Ciuraru, R., 2022. Chemical identification and quantification of volatile organic compounds emitted by sewage sludge. Sci. Total Environ. 838, 155948. https://doi.org/10.1016/j.scitotenv.2022.155948.

Khorramifar, A., Karami, H., Lvova, L., Kolouri, A., Łazuka, E., Piłat-Rożek, M., Łagód, G., Ramos, J., Lozano, J., Kaveh, M., Darvishi, Y., 2023. Environmental engineering applications of electronic nose systems based on MOX gas sensors. Sensors 23, 5716. https://doi.org/10.3390/s23125716.

Lebrero, R., Bouchy, L., Stuetz, R., Muñoz, R., 2011. Odor assessment and management in wastewater treatment plants: a review. Crit. Rev. Environ. Sci. Technol. 41, 915–950. https://doi.org/10.1080/10643380903300000.

Lee, L.C., Liong, C.-Y., Jemain, A.A., 2018. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. Analyst 143, 3526–3539. https://doi.org/10.1039/C8AN00599K.

Marco, S., Gutierrez-Galvez, A., 2012. Signal and data processing for machine olfaction and chemical sensing: a review. IEEE Sensors J. 12, 3189–3214. https://doi.org/10.1109/JSEN.2012.2192920.

Márquez, P., Benítez, A., Caballero, Á., Siles, J.A., Martín, M.A., 2021. Integral evaluation of granular activated carbon at four stages of a full-scale WWTP deodorization system. Sci. Total Environ. 754, 142237. https://doi.org/10.1016/j.scitotenv.2020.142237.

Márquez-Sillero, I., Aguilera-Herrador, E., Cárdenas, S., Valcárcel, M., 2011. Ion-mobility spectrometry for environmental analysis. TrAC Trends Anal. Chem. 30, 677–690. https://doi.org/10.1016/j.trac.2010.12.007.

McNevin, D., Barford, J., 2000. Biofiltration as an odour abatement strategy. Biochem. Eng. J. 5, 231–242. https://doi.org/10.1016/S1369-703X(00)00064-4.

Mehmood, T., Liland, K.H., Snipen, L., Sæbø, S., 2012. A review of variable selection methods in partial least squares regression. Chemom. Intel. Lab. Syst. 118, 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010.

Moufid, M., Tiebe, C., El Bari, N., Hamada Fakra, D.A., Bartholmai, M., Bouchikhi, B., 2022. Pollution parameters evaluation of wastewater collected at different treatment stages from wastewater treatment plant based on E-nose and E-tongue systems combined with chemometric techniques. Chemom. Intel. Lab. Syst. 227, 104593. https://doi.org/10.1016/j.chemolab.2022.104593.

Muñoz, R., Sivret, E.C., Parcsi, G., Lebrero, R., Wang, X., Suffet, I.H. (Mel), Stuetz, R.M., 2010. Monitoring techniques for odour abatement assessment. Water Res. 44, 5129–5149. https://doi.org/10.1016/j.watres.2010.06.013.

Novillo, Guamán, Verónica, Ana, 2015. Multivariate Signal Processing for Quantitative and Qualitative Analysis of Ion Mobility Spectrometry Data, Applied to Biomedical Applications and Food Related Applications (Ph.D. Thesis). Universitat de Barcelona.

Oliva, G., Zarra, T., Pittoni, G., Senatore, V., Galang, M.G., Castellani, M., Belgiorno, V., Naddeo, V., 2021. Next-generation of instrumental odour monitoring system (IOMS) for the gaseous emissions control in complex industrial plants. Chemosphere 271, 129768. https://doi.org/10.1016/j.chemosphere.2021.129768.

Oller Moreno, S., 2018. Data Processing for Life Sciences Measurements with Hyphenated Gas Chromatography-ion Mobility Spectrometry.

Pasciucco, F., Pecorini, I., Iannelli, R., 2023. A comparative LCA of three WWTPs in a tourist area: effects of seasonal loading rate variations. Sci. Total Environ. 863, 160841. https://doi.org/10.1016/j.scitotenv.2022.160841.

Piccardo, M.T., Geretto, M., Pulliero, A., Izzotti, A., 2022. Odor emissions: a public health concern for health risk perception. Environ. Res. 204, 112121. https://doi.org/10.1016/j.envres.2021.112121.

Pomareda Sesé, V., 2013. Signal Processing Approaches to the Detection and Localization of Gas Chemical Sources Using Partially Selective Sensors (Ph.D. Thesis). TDX Tesis Dr. En Xarxa. Universitat de Barcelona.

Prudenza, S., Panzitta, A., Bax, C., Capelli, L., 2022. Electronic nose for real-time monitoring of odour emissions at a wastewater treatment plant. Chem. Eng. Trans. 95, 169–174. https://doi.org/10.3303/CET2295029.

Prudenza, S., Bax, C., Capelli, L., 2023. Implementation of an electronic nose for real -time identification of odour emission peaks at a wastewater treatment plant. Heliyon 9, e20437. https://doi.org/10.1016/j.heliyon.2023.e20437.

Ratti, C., Bax, C., Lotesoriere, B.J., Capelli, L., 2024. Real-time monitoring of odour emissions at the Fenceline of a waste treatment plant by instrumental odour monitoring systems: focus on training methods. Sensors 24, 3506. https://doi.org/10.3390/s24113506.

Reimringer, W., Bur, C., 2023. Promoting quality in low-cost gas sensor devices for real-world applications. Front. Sens. 4. https://doi.org/10.3389/fsens.2023.1317533.

Robbiani, S., Lotesoriere, B.J., Dellacà, R.L., Capelli, L., 2023. Physical confounding factors affecting gas sensors response: a review on effects and compensation strategies for electronic nose applications. Chemosensors 11, 514. https://doi.org/10.3390/chemosensors11100514.

Robin, Y., Amann, J., Schneider, T., Schütze, A., Bur, C., 2023. Comparison of transfer learning and established calibration transfer methods for metal oxide semiconductor gas sensors. Atmosphere 14, 1123. https://doi.org/10.3390/atmos14071123.

Savitzky, Abraham, Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047.

Senatore, V., Zarra, T., Galang, M.G., Oliva, G., Buonerba, A., Li, C.-W., Belgiorno, V., Naddeo, V., 2021. Full-scale odor abatement technologies in wastewater treatment plants (WWTPs): a review. Water 13, 3503. https://doi.org/10.3390/w13243503.

Szymańska, E., N. Davies, A., C. Buydens, L.M., 2016. Chemometrics for ion mobility spectrometry data: recent advances and future prospects. Analyst. 141, 5689–5708. https://doi.org/10.1039/C6AN01008C.

Vera, L., Companioni, E., Meacham, A., Gygax, H., 2016. Real time monitoring of VOC and odours based on GC-IMS at wastewater treatment plants. Chem. Eng. Trans. 54, 79–84. https://doi.org/10.3303/CET1654014.

Vu, T.N., Laukens, K., 2013. Getting your peaks in line: a review of alignment methods for NMR spectral data. Metabolites 3, 259–276. https://doi.org/10.3390/metabo3020259.

Wang, B., Li, X., Chen, D., Weng, X., Chang, Z., 2023. Development of an electronic nose to characterize water quality parameters and odor concentration of wastewater emitted from different phases in a wastewater treatment plant. Water Res. 235, 119878. https://doi.org/10.1016/j.watres.2023.119878.

Zarra, T., Galang, M.G., Oliva, G., Belgiorno, V., 2022. Smart instrumental Odour Monitoring Station for the efficient odour emission management and control in wastewater treatment plants. Chemosphere 309, 136665. https://doi.org/10.1016/j.chemosphere.2022.136665.