

Drift Compensation on Electronic Nose Data Relevant to the Monitoring of Odorous Emissions from a Landfill by OPLS

Carmen Bax^{*a}, Stefano Prudenza^a, Agustin Gutierrez-Galvez^b, Laura Capelli^a

^aPolitecnico di Milano, Department of Chemistry, Materials and Chemical engineering "Giulio Natta", piazza Leonardo da Vinci 32, Milan 20133, Italy;

^bUniversity of Barcelona, Department of Biomedical Engineering, Martí i Franquès 1, 08028 Barcelona, Spain
 carmen.bax@polimi.it

Nowadays electronic noses (e-noses) are often prescribed in the permits of new or existing plants for a continuous monitoring of ambient air quality at receptors or plant fenceline. To do this, the e-nose has to guarantee a stable classification and quantification performance over time. However, the problem of drift (i.e., progressive deviation over time of sensor responses) becomes critical in the case of continuous odour monitoring in the field. Thus, specific strategies for drift compensation need to be defined. This paper focuses on the development a specific drift correction model based on OPLS to mitigate drift effects on e-nose data relevant to three olfactometric campaigns carried out at a landfill over three years. The paper aims to reduce costs associated to the recalibration of the decisional model to be performed before every seasonal e-nose monitoring campaign prescribed in the permit of the landfill. The OPLS model was built on the data collected in the first two campaigns, while data relevant to the most recent campaign were used to test its efficacy in compensating drift. The results achieved pointed out the potentialities of the OPLS model to mitigate drift effects, thereby allowing to extend the applicability of the model developed within an olfactometric campaign to the subsequent ones. The classification performance achieved involving the OPLS correction (i.e., 75%) was considerably higher than the one achieved on non-corrected data (i.e., about 55%).

1. Introduction

Since the first report of electronic noses by Persaud and Dodd in 1982 (Persaud and Dodd, 1982), the instruments have experienced a significant development both in terms of hardware and software. Several applications of e-noses in different fields have been proposed. The most interesting and promising fields for the application of electronic noses concern the food industry (Loutfi *et al.*, 2015; Di Rosa *et al.*, 2017), medical diagnosis (Simeng *et al.*, 2013; Wilson, 2015; Capelli *et al.*, 2016), and environmental monitoring (Capelli, L. *et al.*, 2014; Deshmukh *et al.*, 2015).

In recent years, electronic noses have become very popular air quality monitoring tools and a considerable growth of their applications has been registered. Indeed, being capable to perform a real time characterization of the ambient air directly where the odour presence is lamented, e-nose can be installed at receptor or at plant fenceline with the purpose of continuously detecting, identifying and quantifying odours from the plant under exam (Capelli, L. *et al.*, 2014). These instruments are more and more frequently prescribed in the permits of new or existing plants. For this reason, the data provided by e-noses are required to be reliable and stable since they are gaining legal value. New prescriptions foresee periodical monitoring in different seasons to assess the odour impact relevant to the plant under different meteorological conditions, or the permanent installation of one or more e-noses at plant fenceline to verify compliance of the odour concentration with prescribed limits. To cope with these scopes, the e-nose needs to guarantee a constant performance over time. In particular, the stability of classification and quantification models represents a fundamental requirement for long time monitoring. Nevertheless, the problem of drift (i.e., the deviation of sensor responses over time (Padilla *et al.*, 2010; Ziyatdinov *et al.*, 2010; Fonollosa *et al.*, 2016)) is emphasized in the case of odour monitoring in the field, where the exposure to the atmospheric conditions (humidity variation and temperature or pressure fluctuations) makes sensor responses even more sensible to every source of noise

(Vitali *et al.*, 2018). The progressive deviation of sensor responses occurring over the monitoring period results in the worsening of the instrument capability to correctly detect and classify odours. Thus, the calibration model should be updated periodically in the case of continuous monitoring or before each monitoring campaign in the case of repeated monitoring campaigns, thereby resulting in high operative costs. Many approaches to cope with this problem have been developed and proposed in the scientific literature over the years (Laref *et al.*, 2017; Rudnitskaya, 2018). All of them focus on the estimation of the drift trend and on its removal from the system. Some models require the use of calibrants (Artursson *et al.*, 2000), while others work directly on the data collected for training the instrument for the specific application (Padilla *et al.*, 2010). Orthogonal signal correction (OSC) is one of the most commonly applied methods for drift compensation. Compared to others, it has several advantages. Indeed, not requiring the use of calibrants, it allows to reduce time and costs associated to its implementation. Moreover, it can be efficiently applied also on small datasets as long as they contain useful information for estimating the drift direction.

Since the first algorithm proposed by Wold *et al.* (Wold *et al.*, 1998), several models, based on their idea, have been implemented (Svensson *et al.*, 2002). Among them, the orthogonal projection to latent structures (OPLS) (Trygg and Wold, 2002) proved easier to apply and faster in finding the solution, achieving the same performances of other approaches (Svensson *et al.*, 2002).

This paper investigates the possibility to apply the OPLS model to data relevant to the monitoring of odorous emissions from a landfill, representing a very common application for e-noses for the environmental monitoring. In fact, diffuse odorous emissions from landfills are hardly quantifiable and characterizable by dispersion modeling (Capelli *et al.*, 2013; Shen *et al.*, 2020), and the e-nose monitoring has become a scientifically recognized method for their continuous monitoring (Capelli *et al.*, 2013; Capelli, Laura *et al.*, 2014; Bax *et al.*, 2020; Karakaya *et al.*, 2020).

In particular, this paper focuses on the development of a specific drift correction model for training data relevant to three monitoring campaigns carried out at the same landfill in 2017, 2018 and 2019, to extend the validity of the calibration model over time. Thusly, avoiding periodical recalibrations of the training model before e-nose installation at the sensible receptor for the monitoring of the ambient air.

To do this, the oldest campaigns (i.e., 2017 and 2018) were used to train the OSC model, identify and remove the drift direction related to sensor aging or external factors. Conversely, the newer campaign (i.e., 2019) was used to test the efficacy of the correction, which was assessed by the classification rate improvement obtained after the OSC compensation.

2. Materials and methods

2.1 The electronic nose

The electronic nose EOS507F used for this study, commercialized by Sacmi s.c., has been developed in collaboration with the Olfactometric Laboratory of the Politecnico di Milano specifically for environmental odour monitoring in open field and at far distance from the emission source, i.e. at receptors (Dentoni *et al.*, 2012). The EOS507F is equipped with 6 Metal Oxide Semiconductor (MOS) gas sensors, characterized by high sensitivity, which enables its effective use for the detection of odors at very low concentrations, typical of the receptor level (Eusebio *et al.*, 2016). Moreover, the electronic nose is equipped with specific systems for humidity regulation and realization of reference air, allowing outdoor use, even in presence of variable weather conditions (Dentoni *et al.*, 2012). The instrument has also an automatic calibration system that periodically checks instrument performance and performs a preliminary compensation for drift by pre-processing raw sensor signals. Nevertheless, this method proved to be effective in removing drift only for odours chemically similar to the reference calibrant (i.e., n-butanol). Thus, in case of real environmental samples, this approach should be integrated with specific correction models to completely mitigate the aging effects on sensor responses.

2.2 Dataset

Data involved for the development of the drift correction model were collected at the landfill in different olfactometric campaigns carried out over three years (i.e., 2017, 2018 and 2019) to train the e-nose for the continuous monitoring at the receptor located at about 2 km south from the landfill, where the presence of odours attributable to the plant was lamented. The odour sources considered for instrument training were fresh waste disposal and pretreatment units, landfill gas leakages from landfill surface, off-gases from landfill gas combustor, and leachate tanks (Davoli *et al.*, 2003; Romain *et al.*, 2008). Accordingly, 4 odour classes were defined for futher data processing: Fresh Waste (FW), Landfill Gas (LG), Lecheate (L) and Off-gases Combustor (OC). Samples collected at emission sources were analysed by dynamic olfactometry to determine their odour concentration and dilution factors to be applied before presenting them to the e-nose to build the

Training Set (TS). Indeed, the TS odour concentration range must be representative of the conditions at the receptor, accounting for the fact that odour concentrations measured at the emission sources are generally higher than the concentration levels that are typically found at far distance. For the specific application, odour samples with an odour concentration ranging from 15 ou_E/m³ to 350 ou_E/m³ were considered for the three different campaigns.

2.3 Drift correction model

2.3.1 The OPLS

The OPLS (Trygg and Wold, 2002) is a technique for OSC implementation based on PLS algorithm. The main goal of OSC is to remove from a matrix X all the information non-correlated to a vector (or matrix) Y (i.e., concentration, toxicity, class belonging, etc...). This is done by imposing the condition of orthogonality between X and Y that ensures that only a minimum amount of information is removed from the system, thereby preserving relevant information for odour detection and classification. The OPLS algorithm starts by computing the first PLS component on X and Y, obtaining weights (w) and loadings (p). These are used to obtain orthogonalized weights (w_⊥), loadings (p_⊥) and scores (t_⊥) that are then subtracted from the system to obtain the corrected matrix:

$$X_{osc} = X - t_{\perp} p'_{\perp}$$

If one desires to remove more than one component, the algorithm can be simply repeated using the X_{osc} matrix obtained. Once the model is developed to correct new samples, the orthogonalized weights (w_⊥) and loadings (p_⊥) obtained from the training are applied to the new data to obtain new orthogonalized scores (t_⊥) to be subtracted from the data.

2.3.2 Implementation

We have implemented the OPLS algorithm for drift reduction using R as follows. The X matrix, defined in Section 2.4.1, consisted of the analyses of samples collected at the odour sources (i.e., fresh waste, leachate and landfill gas and off-gases combustor) at different concentrations. Data were processed in terms of Eos Unit, defined in equation (1), where R_i is the instantaneous resistance value for the ith sensor, R_{std,i} is the signal related to the analysis of the reference standard for the ith sensor, a_i and b_i are characteristic coefficients that depend on the nature of the sensor.

$$EU_i = a_i \cdot \left(\frac{R_i}{R_{std,i}} \right)^{b_i} \quad (1)$$

On the contrary, the Y matrix contains the information of the belonging class of each analysis. It is a binary matrix of four columns, each one representing one odour class, whose rows report 1 at the corresponding column that designate the class membership, while the other terms of the row are set to 0.

2.3.3 Validation

To validate the drift correction model on the landfill monitoring data, the dataset was divided into a train and a test set. The OPLS model was built on the training set, comprising the first two olfactometric campaigns (i.e., 2017 and 2018), which was used to estimate the drift direction to be removed from the dataset. Within this train set, 3-fold cross validation was used to select the number of components to be removed. Conversely, the test set, comprising data of the third campaign carried out in 2019, was used to test the performance of the OPLS model implemented. To evaluate this performance, we compared the classification rate of a 3-NN classifier before and after drift compensation.

3. Results

3.1 Preliminary considerations

Drift on the dataset was first explored with Principal Component Analysis (PCA). Figure 1 points out that drift occurs mainly on the direction of the second principal component (PC2). By visual inspection we can observe that there is a clear shift of the sensor responses across campaigns. The PCA projection of the sensor responses move from the lower part of the graph, corresponding to the samples collected in the 2017 campaign, to the higher part, where more recent campaigns are represented. Figure 2 shows the projection of all campaigns with a PCA model build using only data from 2017 and 2018 campaigns. It highlights that the Fresh Waste source is more prone to drift, since data of different campaigns spread through the PCA score plot. This shift associated to sensor aging makes the classification model developed on training data usable only within the same monitoring campaign. Indeed, the classification performance of a 3-NN classifier built on data of the 2017 campaign is 55% when applied to classify data collected within the 2019 campaign. This performance is clearly lower than that obtained within the same campaign (i.e., 86%) (Table 1).

Using a 3-NN classifier built on the 2017 campaign, all samples corresponding to leachate and fresh waste odour classes were misclassified as fresh waste or leachate samples, respectively.

Table 1. Classification performance of the classifier built on data relevant to the 2017 campaign

	Classification performance within the same campaign (2017)	Classification performance on more recent data (2019)
Accuracy (CI _{95%})	0.86 (0.62 – 0.95)	0.55 (0.43 – 0.65)

This result confirms the need for recalibration or shift compensation if we are to use a classification or regression model trained on data of previous measuring campaigns.

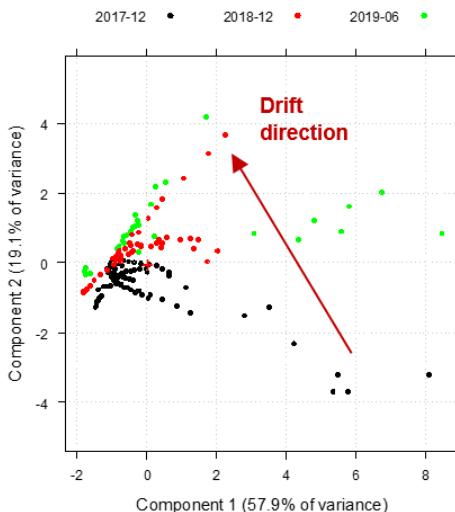


Figure 1. PCA score plot on three campaigns

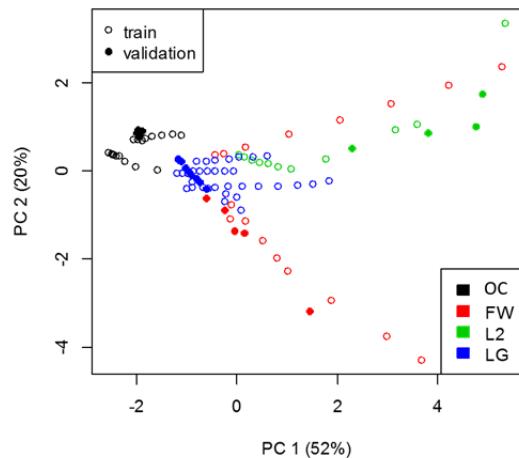


Figure 2. Projection of 2019 data on the PCA model relevant to previous campaigns

3.2 OPLS drift correction

An OPLS drift correction model was built using the olfactometric campaigns of 2017 and 2018 as training set, with the purpose of extending the stability of the e-nose calibration model and its applicability over time. A degree of freedom in the implementation of the drift correction model concerns the selection of the optimal number of OPLS components to be removed from the system. To determine the optimal number, 4 OPLS models were built differing in the number of OPLS components removed.

Figure 3 shows the PCA score plots performed on training data before and after the removal of 1- to 4-OPLS components. From a visual inspection, the removal of 1- and 2-OPLS component resulted in the improvement of sample clusterization, especially for the Fresh Waste class, which, as already mentioned, is the one most subjected to drift. Indeed, Fresh Waste samples relevant to different campaigns group in the same portion of the plot after drift compensation by 1-OPLS and 2-OPLS models. Conversely, the PCA score plot corresponding to models involving 3- or 4-OPLS components pointed out that samples from Fresh Waste and Leachate become indistinguishable: clusters relevant to different classes overlap.

The choice of the optimal number of OPLS components for the specific application was based on the comparison of the classification performance achieved by 1- to 4-OPLS models, assessed by 3-Fold cross validation (3-Fold CV). The 1-OPLS model achieved the highest classification accuracy compared to the others (i.e., about 91%) (Table 2). Thus, it was selected for processing data relevant to the more recent campaign (i.e., 2019).

Table 2. Classification performance assessed by 3-Fold CV achieved by 1-OPLS and 2-OPLS model compared with the one relevant to raw data

	No drift correction	1-OPLS model	2-OPLS model	3-OPLS model	4-OPLS model
Classification Accuracy (3-Fold CV)	0.655	0.910	0.805	0.802	0.767

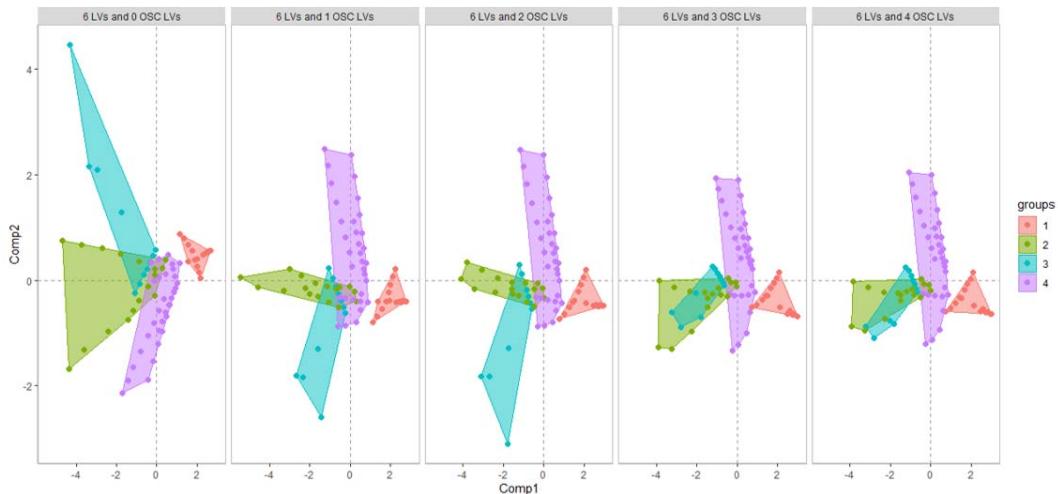


Figure 3. PCA on training data obtained by increasing the number of OSC removed. Classes are labelled as: 1) Off-gases landfill gas combustor, 2) Fresh Waste, 3) Leachate, 4) Landfill Gas.

The external validation on the test set (i.e., 2019 campaign) proved the efficacy of the developed OPLS model in compensating for drift. Indeed, the 1-OPLS model allowed to improve the e-nose classification performance on recent data from 55% ($CI_{95\%}$ 43–65%) to 75% ($CI_{95\%}$ 65–84%), which is closest to the one achieved for data acquired few months after training in 2017 (i.e., about 86%).

Despite based on few data, this preliminary result confirms the potentialities and the applicability of the OPLS models not only to calibrant datasets, as already reported in the scientific literature, but also to real environmental samples.

4. Conclusions

This paper proposes the adoption of the OPLS algorithm to create a specific drift correction model for e-nose data relevant to the monitoring of odorous emissions from a landfill collected over three monitoring campaigns carried out in 2017, 2018 and 2019. The research aimed to develop a classification model stable over time, thereby reducing the burden related to periodical recalibrations of the instrument for the real-time monitoring of the ambient air at a receptor. To do this, the research involved the development of a specific drift compensation model on data from the first two olfactometric campaigns (i.e., 2017 and 2018) and the validation of its performance on more recent data collected in 2019. Indeed, the problem of sensor drift is one of the main limitation factors in the long-term and continuous use of e-noses since it causes a progressive worsening of the instrument classification performance.

The results achieved prove the capability of the OPLS model developed to compensate for drift effects not only on calibrant dataset, but also on real environmental samples. Indeed, the classification performance achieved involving the OPLS correction (i.e., 75%) was considerably higher than the one achieved on non-corrected data (i.e., about 55%). Future work should focus on the optimization of this preliminary model by including in the model the intrinsic variability associated to real odour sources, thereby allowing to achieve higher classification performances. Although this procedure does not prevent from periodical recalibration of the system, it allows to extend the period between recalibration and, thus, to reduce time and costs associated with environmental odour monitoring by e-nose.

References

- Artursson, T.; Eklöv, T.; Lundström, I.; Mårtensson, P.; Sjöström, M.; Holmberg, M. Drift correction for gas sensors using multivariate methods. *Journal of Chemometrics*, v. 14, n. 5-6, p. 711-723, 2000.
- Bax, C.; Sironi, S.; Capelli, L. Definition and Application of a Protocol for Electronic Nose Field Performance Testing: Example of Odor Monitoring from a Tire Storage Area. *Atmosphere*, v. 11, n. 4, p. 426, 2020.
- Capelli, L.; Dentoni, L.; Sironi, S.; Del Rosso, R. The need for electronic noses for environmental odour exposure assessment. *Water Science and Technology*, v. 69, n. 1, p. 135-141, 2013.
- Capelli, L.; Sironi, S.; Del Rosso, R. Electronic noses for environmental monitoring applications. *Sensors (Basel)*, n. 14(1), p. 19979-20007, 2014.

- Capelli, L.; Sironi, S.; Del Rosso, R. Electronic Noses for Environmental Monitoring Applications. *Sensors*, v. 14, n. 11, p. 19979-20007, 2014.
- Capelli, L.; Taverna, G.; Bellini, A.; Eusebio, L.; Buffi, N.; Lazzeri, M.; Guazzoni, G.; Bozzini, G.; Seveso, M.; Mandressi, A.; Tidu, L.; Grizzi, F.; Sardella, P.; Latorre, G.; Hurle, R.; Lughezzani, G.; Casale, P.; Meregalli, S.; Sironi, S. Application and Uses of Electronic Noses for Clinical Diagnosis on Urine Samples: A Review. *Sensors (Basel)*, v. 16, n. 10, 2016.
- Davoli, E.; Gangai, M.L.; Morselli, L.; Tonelli, D. Characterisation of odorants emissions from landfills by SPME and GC/MS. *Chemosphere*, v. 51, n. 5, p. 357-68, 2003.
- Denton, L.; Capelli, L.; Sironi, S.; Del Rosso, R.; Zanetti, S.; Della Torre, M. Development of an electronic nose for environmental odour monitoring. *Sensors (Basel)*, v. 12, n. 11, p. 14363-14381, 2012.
- Deshmukh, S.; Bandyopadhyay, R.; Bhattacharyya, N.; Pandey, R.A.; Jana, A. Application of electronic nose for industrial odors and gaseous emissions measurement and monitoring – An overview. *Talanta*, v. 144, p. 329-340, 2015.
- Di Rosa, A.R.; Leone, F.; Cheli, F.; Chiofalo, V. Fusion of electronic nose, electronic tongue and computer vision for animal source food authentication and quality assessment – A review. *Journal of Food Engineering*, v. 210, p. 62-75, 2017.
- Eusebio, L.; Capelli, L.; Sironi, S. Electronic Nose Testing Procedure for the Definition of Minimum Performance Requirements for Environmental Odor Monitoring. *Sensors (Basel)*, n. 16(9), p. 1424-8220, 2016.
- Fonollosa, J.; Fernández, L.; Gutiérrez-Gálvez, A.; Huerta, R.; Marco, S. Calibration transfer and drift counteraction in chemical sensor arrays using Direct Standardization. *Sensors and Actuators B: Chemical*, v. 236, p. 1044-1053, 2016.
- Karakaya, D.; Ulucan, O.; Turkan, M. Electronic Nose and Its Applications: A Survey. *International Journal of Automation and Computing*, v. 17, n. 2, p. 179-209, 2020.
- Laref, R.; Ahmadou, D.; Losson, E.; Siadat, M. Orthogonal Signal Correction to Improve Stability Regression Model in Gas Sensor Systems. *Journal of Sensors*, v. 2017, p. 9851406, 2017.
- Loutfi, A.; Coradeschi, S.; Mani, G.K.; Shankar, P.; Rayappan, J.B.B. Electronic noses for food quality: A review. *Journal of Food Engineering*, v. 144, p. 103-111, 2015.
- Padilla, M.; Perera, A.; Montoliu, I.; Chaudry, A.; Persaud, K.; Marco, S. Drift compensation of gas sensor array data by Orthogonal Signal Correction. *Chemometrics and Intelligent Laboratory Systems*, v. 100, n. 1, p. 28-35, 2010.
- Persaud, K.; Dodd, G. Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. *Nature*, v. 299, n. 5881, p. 352-355, 1982.
- Romain, A.; Delva, J.; Nicolas, J. Complementary Approaches to Measure Environmental Odours Emitted by Landfill Areas. *Sensors and Actuators B: Chemical*, v. 131, p. 18-23, 2008.
- Rudnitskaya, A. Calibration Update and Drift Correction for Electronic Noses and Tongues. *Frontiers in Chemistry*, v. 6, n. 433, 2018.
- Shen, S.; Wu, B.; Xu, H.; Zhang, Z. Assessment of Landfill Odorous Gas Effect on Surrounding Environment. *Advances in Civil Engineering*, v. 2020, p. 8875393, 2020.
- Simeng, C.; Yuchao, W.; Seokheun, C. Applications and Technology of Electronic Nose for Clinical Diagnosis. *Open Journal of Applied Biosensor*, v. Vol.02No.02, p. 12, 2013.
- Svensson, O.; Kourtzi, T.; MacGregor, J. An investigation of orthogonal signal correction algorithms and their characteristics. *Journal of Chemometrics*, v. 16, p. 176-188, 2002.
- Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, v. 16, n. 3, p. 119-128, 2002.
- Vitali, S.; Bax, C.; Sironi, S.; Capelli, L. Investigation of Electronic Nose Sensor Drift Correction Methods and Their Application to Environmental Samples. *Chemical Engineering Transactions*, v. 68, p. 289-294, 2018.
- Wilson, A.D. Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. *Metabolites*, n. 5(1), p. 140-63, 2015.
- Wold, S.; Antti, H.; Lindgren, F.; Öhman, J. Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, v. 44, n. 1, p. 175-185, 1998.
- Ziyatdinov, A.; Marco, S.; Chaudry, A.; Persaud, K.; Caminal, P.; Perera, A. Drift compensation of gas sensor array data by common principal component analysis. *Sensors and Actuators B: Chemical*, v. 146, n. 2, p. 460-465, 2010.