CHEMICAL ENGINEERING TRANSACTIONS
VOL. 95, 2022
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Selena Sironi, Laura Capelli

ISBN 978-88-95608-94-5; **ISSN** 2283-9216
DOI: 10.3303/CET2295030

Preliminary Study for the Implementation of a Software Method for Humidity Compensation in E-Noses for Outdoor Applications

Stefano Prudenza[a], Alejandro Pinzolas Rubio[b], Carmen Bax[a]*, Marco Marzocchi[c], Marco Casadio[c], Laura Capelli[a]

[a] Politecnico di Milano, Departement of Chemistry, Materials and Chemical Engineering, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy.
[b] School of Engineering and Architecture, University of Zaragoza. Departement of Chemical Engineering and Environmental Technologies. C/María de Luna, 3. 50018 Zaragoza, Spain
[c] Sacmi s.c., via Provinciale Selice 17/a, 40026 Imola, Italy.
carmen.bax@polimi.it

The competitive adsorption of water over Volatile Organic Compound (VOC) in Metal Oxide Sensor (MOS) is well known in literature and is one of the main disturbing factors that contributes to worsen electronic nose (e-nose) performance in classifying and quantifying odours. This constitutes one of the most important limitations for the large-scale diffusion of e-noses since it hinders their use for several applications. In this paper, we investigate the possibility to implement the Orthogonal Signal Correction (OSC) method for compensating the relative humidity effect on MOS sensors. Two different e-noses have been used for this study, each one equipped with a different sensor array comprising only MOS sensors. In order to investigate the relative humidity (RH) effect on the sensors response for different compounds, four calibrants (acetone, butanol dimethyl-disulphide and toluene) have been analyzed at differente levels of RH (20%, 50% and 80%, for a fixed temperature of 20°C) for training the instruments to compensate its intereference. The concentration of the sample analyzed has been set equal to 2.5 ppm. The results achieved proved that the OSC implementation can be a suitable method for the mitigation of the RH interference on MOS sensors, since its application significantly improved the performances of classification, increasing the global accuracy above 70% for both e-noses considered in this study.

1. Introduction

Gas sensing system have proven the capability to recognize and discriminate different gases and odours, an ability that is gaining much interest because of its possible wide range of uses and application. They are currently being used for several purposes, ranging from ambient air monitoring, food quality control, medical diagnosis, process control, etc... (Capelli, Sironi, and Rosso 2014; Loutfi et al. 2015; Collins and Moy 1995). However, their functioning is influenced by various ambient parameters which act as disturbing factors, worsening the system performances. Among them, humidity has a major impact since water molecules absorb on the MOS sensors in a competitive way compared to the target gas to be analyzed (Nenova and Dimchev 2013). To enhance the measurement accuracy, compensating the relative humidity (RH) impact on gas sensors is of prime importance. To face this problem, there are two possible approaches, involving either hardware or software modifications. In the first case, the sampling system can be modified introducing a section for the selective removal and/or control of the humidity in the gas stream to be sampled. Even though several options are available for this purpose, a general drawback of such hardware solutions is that they generally increase the instrument complexity, thus resulting in an increase of the final costs both for purchase and maintenance of the e-nose. On the other hand, considering a software modification, it is possible to introduce a signal correction on the data acquired by the instrument that compensates the raw sensor responses based on the relative humidity content of the gas stream. In this case, no hardware modifications of

Paper Received: 1 April 2022; Revised: 1 June 2022; Accepted: 8 June 2022
Please cite this article as: Prudenza S., Pinzolas Rubio A., Bax C., Marzocchi M., Casadio M., Capelli L., 2022, Preliminary Study for the Implementation of a Software Method for Humidity Compensation in E-noses for Outdoor Applications, Chemical Engineering Transactions, 95, 175-180 DOI:10.3303/CET2295030

the system are required, but it is necessary to develop an appropriate model for training the instrument for such purpose. Indeed, this solution allows to reduce the complexity of the sampling system of the e-nose, thus minimizing its final dimensions and costs. In this paper, we investigate the possibility of implementing the Orthogonal Signal Correction (OSC) (Wold et al. 1998) method for compensating the RH interference on two different e-noses, each equipped with a different MOS sensor array. To appositely train the system to recognize how the different RH levels influence the sensors signal responses, four calibrants have been selected for the analysis: acetone, butanol dimethyl-disulphide and toluene. The concentration used for the tests has been set equal to 2.5 ppm to ensure a good response of the sensor array to the target gases despite the RH masking effect. Three different levels of humidity have been considered for training the compensation model: 20%, 50% and 80%, both evaluated at a temperature of 20°C. These levels have been chosen to almost cover all the spectrum of the RH for a fixed temperature, except for the extreme conditions of severe dry and wet environments that are rarely encountered in real scenarios. Once the models have been developed and tuned with an internal validation, an external validation has been carried out for evaluating the performance of the systems in an independent way.

2. Material and methods

2.1 Electronic noses

The e-noses used in this project are two “EOS507c” commercialized by Sacmi s.c, named EOS02 and EOS03. The use of two different instruments to conduct the test is related to the purpose of increasing results robustness and thus the validity of the method proposed. These instruments mount two different sensors arrays: the EOS02 is equipped with 4 commercial sensors produced by Figaro (i.e. TGS 2600, 2602, 2603, 2444) and two non-commercial tin-oxide-based MOS sensors produced by Sacmi, while the EOS03 comprises 1 Figaro sensor (TGS 2610) and 5 non-commercial MOS produced by Sacmi made of tungsten and zinc oxide. The sampling system of the “EOS507c” operates mixing the incoming gas sample to be analysed with reference air in a ratio 1:1, obtaining a total flux of 50 cm^3/s. This reference air is generated through a filtration of the ambient air on activated carbon, to remove any traces of odorous compounds. Those instruments are already equipped with a cutting-edge system for RH control: the system allows to tailor the incoming gas stream dew point before entering the sensor chamber. Despite being very useful for outdoor odour monitoring, this system makes the instrument extremely complex and delicate. Indeed, most of the times that environmental analyses are interrupted, the problem is related to a malfunctioning of the humidity regulation system. For this reason, we decided to study alternative software control systems, and thus bypassed the RH control system present on the e-noses for our tests.

2.2 Sample preparation and analysis

The calibrant used for the tests have been prepared with a concentration of 5 ppm, but since the e-noses used are equipped with a dilution system that mixes the sample with filtered air in a ratio 1:1, the final concentration that reached the sensor chamber was equal to 2.5 ppm. The samples preparation has been made using the method reported by Sartore et al. (2022) consisting in a syringe-pump injecting a certain amount of liquid calibrant in a stream of filtered air to obtain the desired gas concentration that is then used to fill a Nalophan bag. Once the sample was prepared, the second phase consisted in tailoring its RH. For this purpose, a climatic chamber was used. Because of the Nalophan permeability to humidity, by leaving the samples at a fixed temperature of 20°C and at the desired level of RH inside the climatic chamber, they reach the external RH after ca. 2 hours. The e-nose analysis lasted a total of 40 minutes, comprising 20 minutes of “during”, in which the gas sample is sent to the sensors chamber, and other 20 minutes of “after”, when reference air is sucked for cleaning the sensors and restore its response to the original value. Globally 24 samples have been prepared for training the instruments, 6 for each calibrant, and 2 for each RH level considered. For the validation set 12 samples have been used, 3 for each calibrant, and one for each RH level.

2.3 Data Processing

The data processing pathway used in this study is reported in Figure 1 and will be described more in detail in the next paragraphs.

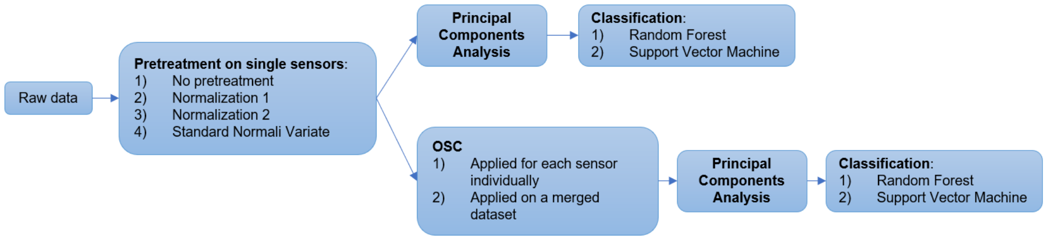

Figure 1. Data processing pathway used for the raw signal elaboration.

2.3.1 Pretreatment

Once the data are collected from the e-noses they are first pretreated, to compensate for differences in the magnitudes of the signals. In this study we have considered 4 different scenarios as reported in Table 1

Table 1. The four different scenarios considered for the initial pretreatment of the raw data.

Method	Equation
No pretreatment	-
Normalization 1	$y_i = (x_i - x_i^{min})/(x_i^{max} - x_i^{min})$
Normalization 2	$y_i = x_i/x_i^{max}$
Standard Normal Variate (SNV)	$y_i = (x_i - \bar{x})/\sigma$

Here *x* represents the raw sensor signal and *y* represent the new sensor value after pretreatment. The two normalization techniques set the sensor values in a range comprised between 0 and 1 for “Normalization 1”, and between -1 and 1 for “Normalization 2”, whereas the SNV sets the mean and the variance value of the new curves to 0 and 1, respectively. Each of these techniques has been applied on each sensor individually.

2.3.2 OSC

The OSC algorithm has been applied following the method proposed by Tom Fearn (Fearn 2000; Bax et al. 2021). In this study, two different scenarios have been investigated, one consisting in the application of the OSC on each MOS sensor individually, to develop a specific compensation model for each sensor. The other one foresees the union of the different sensors signals in one single curve, by attaching one to the other the resistance value registered during the analysis as suggested by Padilla et al. (2010), on which the OSC is then applied. This way, the compensation model developed is “generalized” and applied simultaneously to all the sensors. One degree of freedom of the OSC is the number of orthogonal components to be removed from the system, for which it is necessary to find an optimum. To identify this optimum the classification performances, obtained by the subsequent application of the PCA and RF/SVM, for the different orthogonal components have been compared, and the best scenario with the maximum classification accuracy has been chosen as the optimal model.

2.3.3 Principal Component Analysis (PCA)

PCA was applied for reducing the dataset dimensionality to an acceptable number of principal components (PC) for the subsequent classification task. In the scenario in which the OSC was not applied, before the PCA implementation, the different sensor’s curves after pretreatment have been merged in one single curve, in the same way as it has been done with the OSC.

2.3.4 Classification

After dimensionality reduction, the first 4 PC scores obtained from the PCA are passed to two different machine learning algorithms: Random Forest (RF) (Breiman 2001) and Support Vector Machine (SVM) (Sun 2014). The choice of considering only the first 4 PC has been decided by looking at the variance explained, which resulted always >80%. The use of two different algorithms was done for the purpose of comparing of the impact of the OSC implementation on two different machine learning techniques. For tuning the internal parameters required by the two algorithms, a 5-fold Cross validation was applied on the training data.

2.3.5 External Validation

Finally, for testing the classification performances in an independent way, an external validation test was performed. The data belonging to this test underwent the same preteatment previously illustrated; they have been processed using the models already developed on the training data.

3. Results

3.1 Relative humidity interference

To verify the effect of the RH interference on MOS, some of the resistances registered by the sensors during the analysis have been plotted. Depending on the RH level of the sample, the plateau of the curves shifts, proving that the water present in the gas significantly affects the sensors response.

3.2 OSC correction

3.2.1 OSC optimization

Due to space limitations, as an example, here only the optimization procedure carried out on the EOS02 using the RF is illustrated, with the purpose to provide the logic behind the choices made. The same procedure has been applied for the SVM and for the EOS03 data, resulting in the same conclusion that will be reported in this section. After data pretreatment, the OSC implementation followed. As stated in paragraph 2.3.2, for this technique it is necessary to select an optimal number of Orthogonal Components (OC) to be removed from the original dataset. For this purpose, a comparison between the classification performances, on both the train and the test set were carried out by varying the number of orthogonal components removed. Figure 2 and Figure 3 report the accuracies obtained for both the cases where the OSC has been implemented on the single sensors and on the “merged” sensors curves. For both scenarios considered there is a maximum of classification performances on the train set by removing 1 OC. By increasing the number of components removed the classification accuracy worsens, meaning that too much information’s are being removed from the original data.

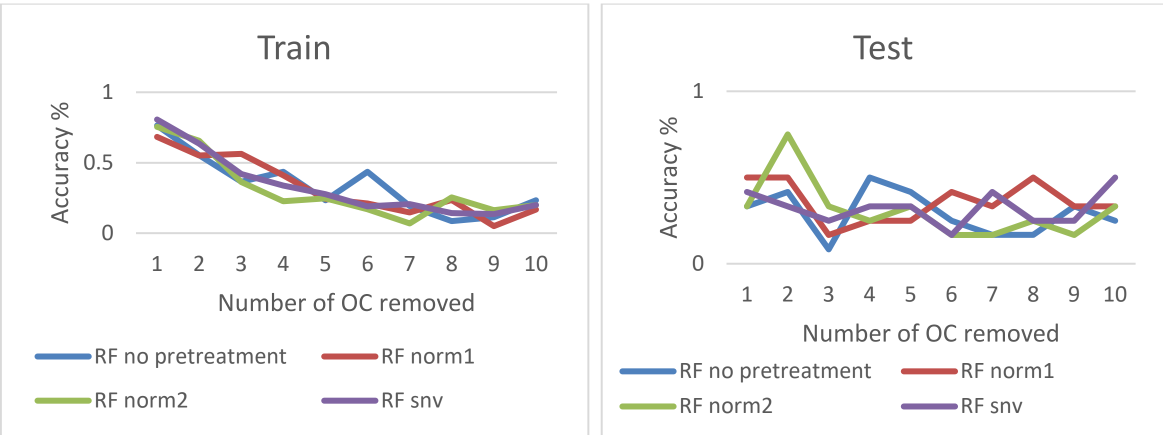

Figure 2. OSC classification performances for its application on every sensor individually obtained with the RF algorithm, as function of the Orthogonal Components (OC) removed from the original data.

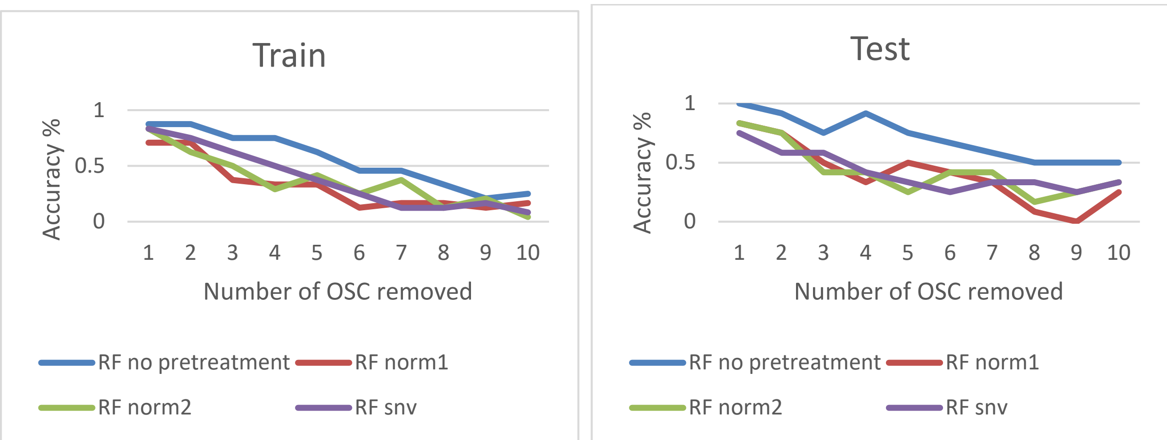

Figure 3. OSC classification performances for its application on the merged sensor curves obtained with the RF algorithm, as function of the Orthogonal Components (OC) removed from the original data.

By looking at the validation test performances a difference among the two approaches emerges. The OSC applied on the single sensors it is not able to transfer the train performances on the test set, since here the accuracy randomly fluctuates around 50%. Conversely, if the OSC is applied on the "merged" sensors curves, the test set follows the same trend obtained on the train, with a peak of the performances after the removal of 1 OC. This is a clear sign that among the two options the only suitable is the application of the OSC on all the sensors merged in a single curve. For this reason, in the subsequent section only the results obtained with this scenario will be presented. Based on these evaluations, the optimal number of OC to be removed resulted to be 1. Similar results have been obtained by implementing the SVM, and they have been confirmed also on the EOS03 data. The next paragraph only show the results obtained by removing 1 OC.

3.2.2 Results

The performances obtained with the EOS02 before and after the OSC implementation, considering the removal of only 1 OC, for both RF and SVM are reported in Figure 4 and Figure 5, whereas the performances obtained with the EOS03 before and after the OSC implementation, considering the removal of only 1 orthogonal component are reported in Figure 7 and Figure 8.

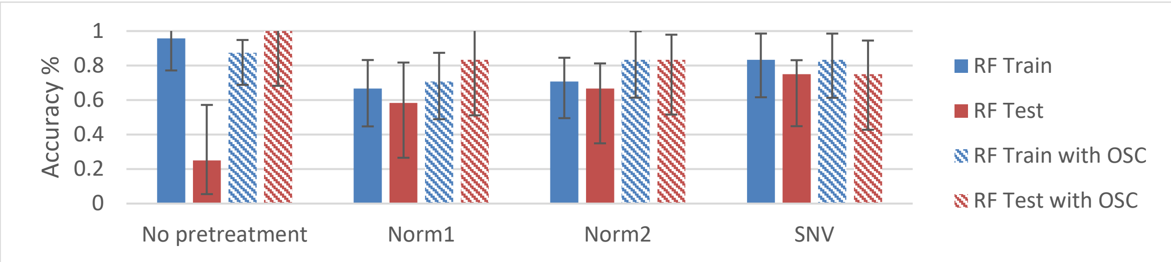

Figure 4. Performances obtained with the RF algorithm before and after the OSC implementation on EOS02.

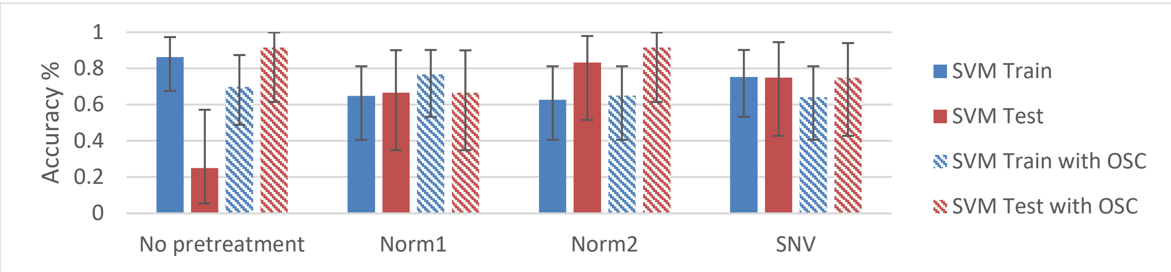

Figure 5. Performances obtained with the SVM algorithm before and after the OSC implementation on EOS02.

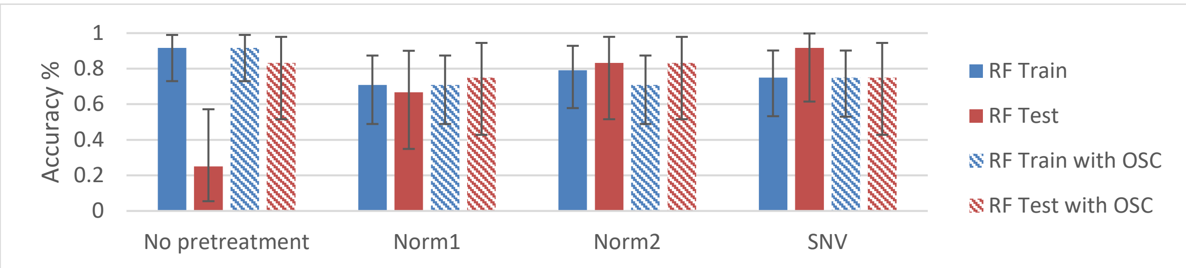

Figure 6. Performances obtained with the RF algorithm before and after the OSC implementation on EOS03.

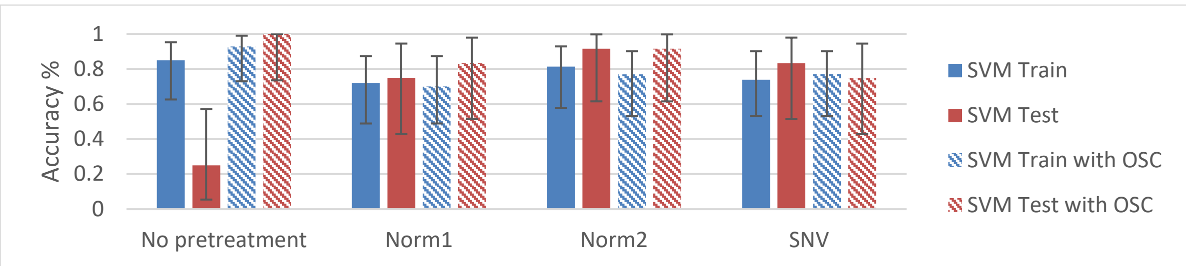

Figure 7. Performances obtained with the RF algorithm before and after the OSC implementation on EOS03.

3.2.3 Discussion

The results obtained proved that the implementation of the OSC increases the classification performances on both the train and test data for almost all the scenarios considered and for both the e-noses involved in the study, resulting in a suitable method for the mitigation of the RH interference on the MOS sensors. In particular, the most significant increase is obtained in the case where no preteatments are applied on the raw sensor curves. In fact, in this case the train performances remain almost unchanged, but in the test set there is an average increase of the accuracy of almost 70%, passing from 20% accuracy without the OSC implementation to a 90-100% after its application. This means that without the correction, the classification model performances are not able to be replicated on new independent samples because of the RH interference, while after the OSC application on the new data this effect is compensated and the samples are correctly classified. In the other cases, where the normalizations and SNV are applied without the OSC implementation, there is a general increase in the performances obtained, meaning that also these techniques can somehow mitigate the RH effect in the first place. Still, if the OSC is then applied on the pretreated data, the global performance slightly increases, meaning that the RH compensation could be even pushed further and optimized. Finally, it is important to highlight that very similar results have been obtained on both the e-noses involved in this study, confirming the goodness of the OSC in compensating the RH interference on the MOS sensors.

4. Conclusions

In this preliminary study, the implementation of the OSC for compensating the RH interference has been tested on two different e-noses equipped with two different MOS sensors arrays. The tests involved the use of 4 different calibrants, analysed at a fixed concentration of 2.5 ppm and at three different levels of RH, i.e. 20%, 50% and 80%. The OSC implementation provided a general increase of the classification performances for both the e-noses. In particular, in the case where no pretreatments are applied on the raw signals, the performances obtained on the test set increase from a 20% of accuracy obtained without the OSC application to an average of 90% after the OSC implementation. This proved the potentiality of this technique in the mitigation of the RH competitive adsorption on the MOS sensors, despite the limited number of samples considered in this preliminary study. Further studies should in the first place involve a larger number of samples with the purpose of confirming the goodness of this method, and secondly try to transfer this correction technique from a laboratory application to real environmental scenarios.

References

Bax, Carmen, Stefano Prudenza, Agustin Gutierrez-Galvez, and Laura Capelli. 2021. "Drift compensation on electronic nose data relevant to the monitoring of odorous emissions from a landfill by opls." In *NOSE 2021*, 13-18. Italian Association of Chemical Engineering-AIDIC.

Breiman, Leo. 2001. 'Random forests', *Machine learning*, 45: 5-32.

Capelli, Laura, Selena Sironi, and Renato Del Rosso. 2014. 'Electronic noses for environmental monitoring applications', *Sensors*, 14: 19979-20007.

Collins, Mark A., and Laurent Moy. 1995. "The Electronic Nose for Process Control." In *Neural Networks: Artificial Intelligence and Industrial Applications*, edited by Bert Kappen and Stan Gielen, 297-302. London: Springer London.

Fearn, Tom. 2000. 'On orthogonal signal correction', *Chemometrics and intelligent laboratory systems*, 50: 47-52.

Loutfi, Amy, Silvia Coradeschi, Ganesh Kumar Mani, Prabakaran Shankar, and John Bosco Balaguru Rayappan. 2015. 'Electronic noses for food quality: A review', *Journal of Food Engineering*, 144: 103-11.

Nenova, Zvezditza, and Georgi Dimchev. 2013. 'Compensation of the impact of disturbing factors on gas sensor characteristics', *Acta Polytech. Hung*, 10: 97-111.

Padilla, Marta, A Perera, I Montoliu, A Chaudry, K Persaud, and S Marco. 2010. 'Drift compensation of gas sensor array data by orthogonal signal correction', *Chemometrics and intelligent laboratory systems*, 100: 28-35.

Sartore, Lorenzo, Elisa Polvara, Marzio Invernizzi, and Selena Sironi. 2022. 'Determination of Air Pollutants: Application of a Low-Cost Method for Preparation of VOC Mixtures at Known Concentration', *Sustainability*, 14: 9149.

Sun, Minghe. 2014. 'Support Vector Machine Models for Classification.' in John Wang (ed.), *Encyclopedia of Business Analytics and Optimization* (IGI Global: Hershey, PA, USA).

Wold, Svante, Henrik Antti, Fredrik Lindgren, and Jerker Öhman. 1998. 'Orthogonal signal correction of near-infrared spectra', *Chemometrics and intelligent laboratory systems*, 44: 175-85.