CHEMICAL ENGINEERING TRANSACTIONS

VOL. 54, 2016

A publication of
AIDIC
The Italian Association
of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Selena Sironi, Laura Capelli
Copyright © 2016, AIDIC Servizi S.r.l.,
ISBN 978-88-95608-45-7; **ISSN** 2283-9216

Application of an Electronic Nose Coupled to a Gas Analyser for Measuring Ammonia

Fernando Campo[a], Andy Blanco-Rodríguez[b], Rodolfo Valiente[c], Bradies Lambert[d], Liliam Becherán[d], Henrique de Melo Lisboa[b], Alejandro Durán[d], Alejandro Garcia-Ramirez*[a]

[a]Center for Technological Earth and Sea Sciences (CTTMar), University of Vale do Itajaí (UNIVALI), 88302-202, Itajaí, Santa Catarina, Brasil.
[b]Laboratory of Air Quality Control (LCQAr), Department of Sanitary and Environmental Engineering (ENS), Federal University of Santa Catarina (UFSC), 88040-900, Florianópolis-SC, Brazil.
[c] Microelectronics Research Center (CIME), José Antonio Echeverría Higher Polytechnic Institute (CUJAE), 19390, Havana, Cuba.
[d] Institute of Science and Technology of Materials (IMRE), University of Havana (UH), 10400, Havana, Cuba.
garcia.ramirez@gmail.com

This work presents the application of an electronic nose (e-nose) based on MOS sensors, for quantitative analysis of ammonia, and its calibration regarding a commercial gas analyser. Features were extracted from the sensors responses in the time domain, the phase space and the frequency domain. Each feature was evaluated considering its repeatability discrimination capability, information content and redundancy, by using a signal to standard deviation ratio (S/σ) and Principal Components Analysis (PCA). Based on this assessment, two parameters were selected from the responses of two sensors: the slope of the desorption transient and the first coefficient of a Daubechies 1 Wavelet decomposition of 11 levels. Afterwards, the four features selected were used as predictor variables in a Partial Least Square (PLS) regression, for correlating the e-nose response with the ammonia concentration measured with a gas analyser. This regression model allows the e-nose to predict the ammonia concentration values with a $R^2 = 0.99$ and RMESCV of 1.40 ± 0.04 ppm.

1. Introduction

It is known that ammonia emissions affect the environment in several ways, such as over-fertilization and acidification of soils near ammonia sources, contamination of surface waters, fine particulate matter formation, and nitrous oxide formation (Hristov et al., 2011). Inside poultry barns, these emissions also have a negative impact in both humans and animals' health (Wathes et al., 2002). These and other related factors have made the measurement of ammonia emissions an issue of interest.

In this context, the so-called electronic noses (e-noses) have been taking an important role in recent years for monitoring these emissions (Lorwongtragool et al., 2010; Sohn et al., 2008; Abdullah et al., 2012). E-noses perform odour measurements based on the perception of the human nose, overcoming disadvantages derived from human psyche. E-noses can also operate continuously, performing real-time and *in-situ* monitoring for long time periods. In addition, their results can be correlated to other analytical instruments like olfactometers, gas chromatographs coupled to mass spectrometers (GC-MS) and gas analysers, providing information related to the concentration of a certain gas.

For correlating the responses of electronic noses with analytical instruments, linear and non-linear regression techniques have been reported (Lei et al., 2015; Michishita et al., 2010). These techniques use a group of variables, easier (or cheaper) to measure, to predict results that otherwise would be more difficult to obtain. Nevertheless, due to the highly redundant information provided by the sensors used in e-noses, some issues like data dimensionality and co-linearity must be treated in order to increase reliability and simplicity to the prediction model.

The above mentioned problems can be solved, in some extent, by using an appropriate pre-processing stage in e-noses, where features are extracted and the most discriminative and informative ones are selected, depending on the application. According to its nature, features could be classified as: extracted directly from the sensors responses in the time domain; extracted from the responses in phase space; extracted from the spectral analysis; and extracted from parametrical models (Yan et al., 2015). Feature selection, on the other hand, may be achieved by means of filters, wrappers or embedded solutions. Among these, filters provide a more general approach with suitable results in most applications, besides, they are computationally attractive (Scott et al., 2007). With this kind of algorithms, information is filtered according to certain criteria.

This work presents the correlation of an e-nose output with a gas analyser for measuring ammonia concentrations. Correlation makes it possible to predict the response of the gas analyser with the e-nose, by means of a PLS regression model. In order to increase accuracy and simplicity of the desired model, features were extracted and selected, reducing dimensionality and redundancy. The features were extracted directly from the time responses, from the phase space and from the spectral analysis based on Fourier and Wavelet transforms. The extracted parameters were selected according to their repeatability, discrimination capabilities, information content and redundancy. For this selection, the signal to noise ratio (S/σ) of each feature was computed, and a principal components analysis was performed. The selected variables were used as predictors for the PLS regression.

2. Materials and methods

This section presents the experimental procedure and describes the feature extraction techniques applied to the responses of the e-nose sensors, as well as the choice criteria for the preprocessing stage.

2.1 Experimental procedure

Eight ammonia concentration values were prepared with a dilution system, and were measured with Dräger X-am® 7000 gas analyser and the e-nose, following the sequence described in the next paragraph. The dilution system was mainly composed by a cylinder containing ammonia at a concentration of 100 ppm. Ammonia was diluted with dry air, and the different levels of dilution were achieved by regulating the flow of both gases. All the experiments were performed under laboratory conditions of humidity and temperature.

For each experimental set, the gas with the desired concentration was first introduced inside a Tedlar bag using the dilution system. Then, the ammonia concentration was measured by the gas analyser and, right afterwards, the bag was placed in the e-nose gas inlet for analysis. This procedure was repeated for all 8 concentration values. Also, a reference bag containing pure air was prepared and measured with the gas analyser showing 0 ppm of ammonia.

For each bag, two sets of measurements were performed using the e-nose. Within each set, six replicates were acquired from the responses of each gas sensor. A blank response was acquired from the reference bag before each set of measurements. Considering that the corresponding electric signals were sampled at a frequency of 5 Hz during 350 s, the amount of data at the end of the measuring process was:

8 (concentration points) X 2 (sets of measurements) X 6 (replicates) X 6 (sensors) X 350 s X 5 Hz = 96 (gas samples) X 6 (sensors) X 1750 (digital samples)

2.2 Data preprocessing

Each sensor was processed by a moving average filter, and each corresponding baseline was treated using the differential method (Arshak et al., 2004). An exploratory analysis was performed for discarding none responsive sensors and outliers. For feature extraction, various techniques were applied in both the time and frequency domains, and in the phase space. The analysis performed in the frequency domain was based on Fourier and Wavelet transforms. Subsequently, the best features were selected according to their repeatability, discrimination and redundancy. The selection criterions were based on Principal Components Analysis (PCA) and the signal to standard deviation ratio reported by Eklov et al. (1997).

2.2.1 Signal to Standard Deviation Ratio (S/σ)

The signal to standard deviation ratio is a method reported by Eklov et al. (1997) for selecting features extracted from the responses of a gas sensor array according to their repeatability and discrimination. The signal (S), in this case, represents the variation of a feature’s mean value at all concentration points, and the standard deviation (σ) represents the variation around the mean value at each concentration point. Thus, the signal describes the spread between the feature’s mean values for the different concentrations (discrimination), while the standard deviation describes the feature’s variation within replicates (repeatability). A high value of S indicates that the feature discriminates well different concentrations; a low σ, on the other hand, denotes good repeatability for the replicates. So, the higher the ratio, the more discriminative and repeatable the feature is.

3. Results and discussion

After an exploratory data analysis, it was observed that only three of the sensors showed some level of response for different concentration values. These sensors were the TGS2602, TGS826 and TGS2610. Therefore, the responses of the other sensors were discarded. It was also observed during the experiments, that the responses corresponding to the first replicate of each set of measurements showed lower response values than the others. For that reason, these replicates were considered as outliers and removed from data. The information resulting from this analysis corresponded to the responses of 3 sensors to 80 gas samples. Since baselines had already been removed, each response vector was composed by 1250 digital samples. After this process of irrelevant and noisy data removal, the dimensionality was reduced from 96 X 6 X 1750 to 80 X 3 X 1250.

3.1 Feature extraction

Table 1 summarizes the parameters extracted, the dimensionality and signal to noise ratio (S/σ) of the corresponding data. The data dimensionality was reduced to 80 X 3 X 28, being 28 the number of features extracted. However, it was still necessary to reduce redundant and noisy information by using feature selection.

Table 1: Extracted features.

Feature	Dimension	S/σ	Description
Features extracted directly from the time response			
VMax	***80 X 3***	***3,418***	Maximum response value (stationary feature).
TSub	80 X 3	0,366	Rise time.
IntSub	80 X 3	2,953	Integral of the adsorption transient.
IntDes	***80 X 3***	***3,450***	Integral of the desorption transient.
PenSub	80 X 3	2,867	Maximum slope during the adsorption transient.
NSub	80 X 3	1,536	Interception in the y-axis of the line with slope PenSub.
PenDes	***80 X 3***	***3,932***	Minimum slope during the desorption transient.
NDes	***80 X 3***	***3,749***	Interception in the y-axis of the line with slope PenDes.
Win(1)	80 X 3	2,550	Integral of the product of 4 window functions as in (GUTIERREZ-OSUNA and NAGLE, 1999), applied to the whole pulse.
Win(2)	***80 X 3***	***3,297***	
Win(3)	***80 X 3***	***3,393***	
Win(4)	***80 X 3***	***3,034***	
Features extracted from the phase space			
Int	80 X 3	2,048	Integral of the whole response in the phase state.
DA	80 X 3	1,238	Distance from the baseline to the maximum slope during adsorption (ZHANG et al., 2008).
DB	80 X 3	0,773	Distance from the maximum slope during adsorption to the stationary state of the response (ZHANG et al., 2008).
DC	***80 X 3***	***3,551***	Distance from the stationary state of the response to the minimum slope during desorption (ZHANG et al., 2008).
DD	***80 X 3***	***3,194***	Distance from the minimum slope during desorption to the baseline (ZHANG et al., 2008).
Features extracted from spectral analysis			
CoefFr	***80 X 3 X 7***	***3,076***	The seven lowest Fourier coefficients. They correspond to the DC component and the first three lowest frequencies. This was the set of components with the highest signal to noise ratio compared to the rest of the Fourier coefficients.
NormFr	***80 X 3***	***3,281***	Euclidian norm of the vector formed by the selected Fourier coefficients.
CoefWvt	***80 X 3***	***3,005***	Coefficient of the first level of decomposition of an 11 levels Daubechies 1 Wavelet decomposition. This was the level with the highest signal to noise ratio compared to the other levels.
AppWvt	***80 X 3***	***3,296***	Approximation coefficient of the Wavelet decomposition mentioned above.
NormWvt	***80 X 3***	***3,188***	Euclidian norm of the vector formed by the approximation and the coefficient of the first decomposition.

3.2 Feature selection

It was considered as a necessary condition for the selection of a feature, that its variation along all the concentration values were three times bigger than the variation around a single concentration. Thus, features with a S/σ greater than 3 were preselected, and the rest were discarded. As a result, for the subsequent analysis, 20 features were considered as having suitable repeatability and discrimination, and they appear highlighted in Table 1.

In the next phase of selection, PCAs were performed and a selection criterion based on redundancy and noise was used. In order to facilitate the analysis, features were divided in three groups: extracted from time and phase response, extracted from Fourier analysis and extracted from Wavelet analysis.

It was observed, in all three cases, that the first principal component (PC1) represented the variation from the lower to the higher concentrations, while PC2 contained information for discrimination between the lower concentrations, together with noise in higher concentrations. For that reason, a criterion for selection was that the feature provided most of the information in PC1. Features with too much information in PC2 were considered noisy and were not selected. Also, the features very close to each other were considered redundant and were not selected.

At the end of this procedure, features were selected from each group. In general, in all three analysis, the response of sensor TGS2610 showed to be source of noise and was therefore discarded. In the case of the time and phase space features, the selection was: PenDes and DC. The Fourier coefficients selected were three, which corresponded to the zero frequency component (CofFr4), and the first positive (CofFr5) and negative (CofFr3) frequency components. From the Wavelet analysis the feature selected was the coefficient of the first level of decomposition (CofWvt). Summarizing, the features selected after this process were: PenDes TGS826, PenDes TGS2602, DC TGS826, DC TGS2602, CofFr3 TGS826, CofFr4 TGS826, CofFr5 TGS826, CofFr3 TGS2602, CofFr4 TGS2602, CofFr5 TGS2602, CofWvt TGS826 and CofWvt TGS2602.

A final PCA was performed with these features and its corresponding biplot graph is shown in Figure 1.

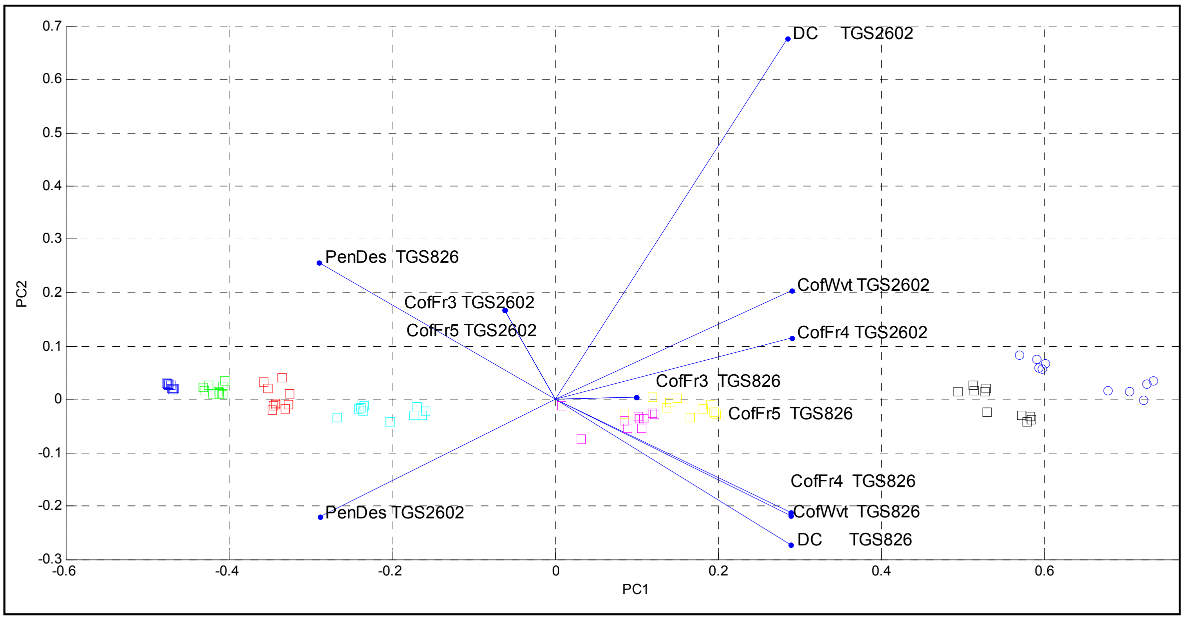

Figure 1. Biplot graph depicting the features selected from each group.

Following the same analysis as before, the features DC, extracted from the phase space of both sensors were discarded for having too much information in PC2. Fourier low frequency coefficients (CofFr3 and CofFr5) were redundant in both sensors' responses, and, specifically, those corresponding to TGS2602 contained noise. On the other hand, the CofFr4 coefficient, corresponding to the zero frequency component, provided almost the same information as the Wavelet coefficient. Since Wavelet components contain both frequency and time information, the feature CofFr4 was also discarded and the final selection was: the minimum slope during desorption transient of the sensors TGS2602 and TGS826, and the Wavelet coefficient of the sensors TGS2602 and TGS826.

Figure 2 shows the PCA plot using these final features. It can be observed suitable discrimination between groups and good repeatability, although higher dispersion was observed at high concentrations. This result proves that dynamic features also provide information of concentration and therefore they could be used in quantitative analyses like the one performed in this work.

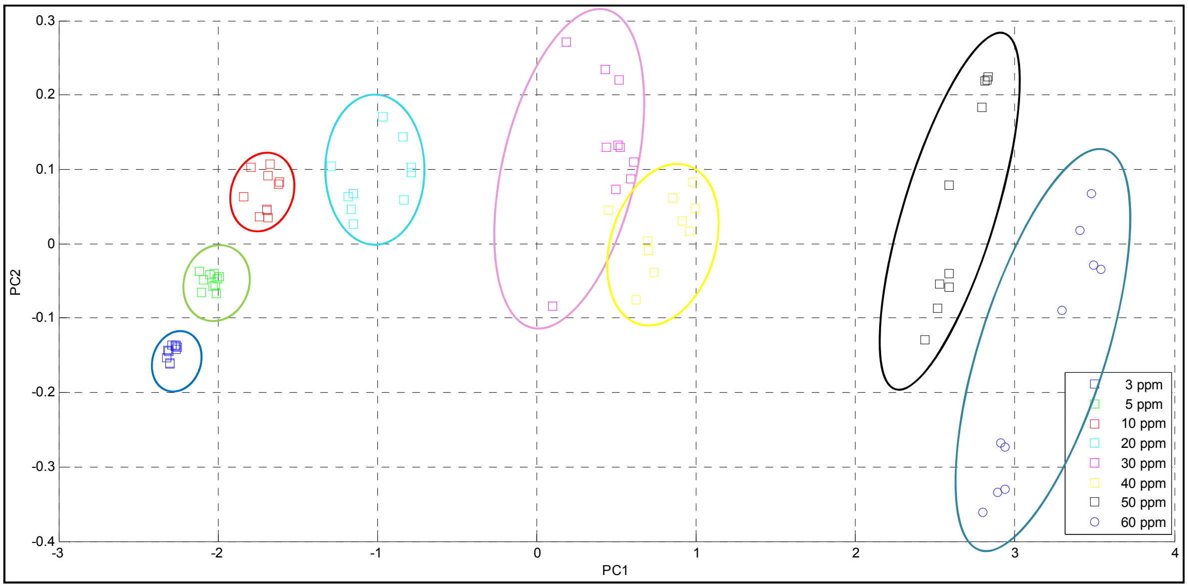

Figure 2. PCA graph using the minimum slope during desorption transient of the sensors TGS2602 and TGS826, and the Wavelet coefficient of the sensors TGS2602 and TGS826.

3.3 E-nose calibration

A PLS regression was performed using the finally selected features as predictors, see Figure 3. After ten regressions, the model showed a root mean square error by cross-validation (RMSECV) between 1.36 and 1.43 ppm. According to the model, the real ammonia concentration is related to the e-nose responses by a line of slope 0.99 ± 0.02, a *y*-intercept in 0.15±0.47 ppm and a R^2 = 0.9914. Since both intervals of slope and *y*-intercept contain the values 1 and 0 respectively, it could be affirmed that the relation can be represented by a line with unitary slope and no intercept. This means that the selected variables are able to predict the concentration of ammonia.

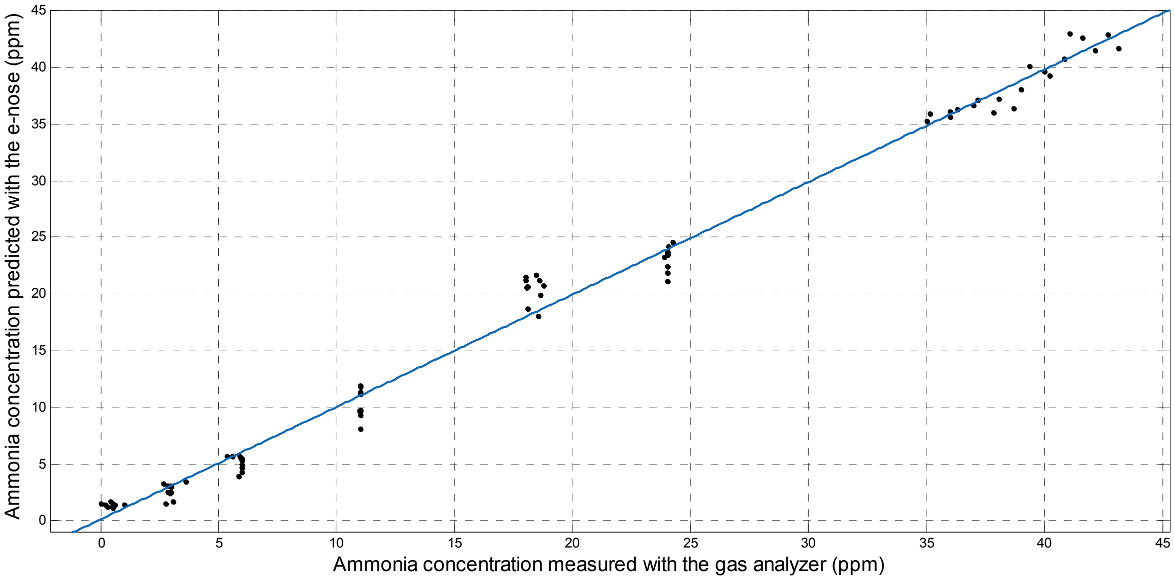

Figure 3. PLS regression with the selected features

4. Conclusions

In the present work, an e-nose was calibrated for measuring ammonia concentrations under laboratory conditions. It was observed that from an array of six sensors, only two of them provided useful information. In the same way, only two features were able to discriminate between the concentration samples, reducing noise and redundancy of the original responses. Thus, data dimensionality was reduced to 80 X 4 from the original amount of 96 X 6 X 1750. It was also observed that dynamic features are suitable for quantitative approaches providing accurate results by a PLS regression. Through the obtained regression model, the e-nose was able to predict the concentration of ammonia previously measured with a gas analyser, showing potentiality for monitoring ammonia emissions.

Acknowledgments

The authors thank Project CAPES-MES 139/11 “Desenvolvimento de narizes eletrônicos para a detecção de substâncias gasosas no meio ambiente: contribuição para a avaliação do impacto de odorantes”, and Programa de Estudantes-Convênio de Pós-Graduação (PEC-PG) from CAPES/CNPq–Brasil.

Reference

Abdullah, A., Shakaff, A., Adom, A., Zakaria, A., Saad, F., Kamarudin, L., 2012, Chicken Farm Malodour Monitoring Using Portable Electronic Nose System, Chemical Engineering Transactions, 30, 55 – 60.

Arshak, K. Moore, E., Lyons, G., Harris, J., Clifford, S., 2004, A review of gas sensors employed in electronic nose applications, Sensor Review, 24 (2), 181-198.

Eklov T., Martensson, P., Lundstrom, I, 1997, Enhanced selectivity of MOSFET gas sensors by systematical analysis of transient parameters, Analytica Chimica Acta, 353, 291 – 300.

Gutierrez-Osuna, R. and Nagle, H. T., 1999, A Method for Evaluating Data-Preprocessing Techniques for Odor Classification with an Array of Gas Sensors, IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, 29 (5), 626 – 632.

Hristov, A. N., Hanigan, M., Cole, A., Todd, R., Mcallister, T. A., Ndegwa, P. M., Rotz, A., 2011, Review: ammonia emissions from dairy farms and beef feedlots 1, Canadian journal of animal science, 91 (1), 1 – 35.

Lei, J., Hou, C., Huo, D., Li, Y., Luo, X., Yang, M., Fa, H., Bao, M., Li, J., Deng, B., 2015, Detection of ammonia based on a novel fluorescent artificial nose and pattern recognition, Atmospheric Pollution Research.

Lorwongtragool, P., Wongchoosuk, C., Kerdcharoen, T, 2010, Portable Artificial Nose System for Assessing Air Quality in Swine Buildings, 2010 International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 532 – 535.

Michishita, T., Akiyama, M., Hirano, Y., Ikeda M., Sagara, Y., Araki, T, 2010, Gas Chromatography/Olfactometry and Electronic Nose Analyses of Retronasal Aroma of Espresso and Correlation with Sensory Evaluation by an Artificial Neural Network, Journal of Food Science, 75 (9).

Scott, S. M.; James, D.; Ali, Z., 2007, Data analysis for electronic nose systems. Microchimica Acta, 156 (3-4), p. 183-207.

Sohn, J. H., Hudson, N., Gallagher, E., Dunlop, M., Zeller, L., Atzeni, M, 2008, Implementation of an electronic nose for continuous odour monitoring in a poultry shed, Sensors and Actuators B: Chemical, 133 (1), 60 – 69.

Wathes, C., Jones, J., Kristensen, H., Jones, E., Webster, A., 2002, Aversion of Pigs and Domestic Fowl to Atmospheric Ammonia, Transactions of the ASAE, 45 (5), 1605 – 1610.

Yan, J., Guo, X., Duan, S., Jia, P., Wang, L., Peng, C., Zhang, S., 2015, Electronic Nose Feature Extraction Methods: A Review, Sensors, 15, 27804 – 27831.

Zhang, S., Xie, C., Hu, M., Li, H., Bai, Z., Zeng, D., 2008, An entire feature extraction method of metal oxide gas sensors, Sensors and Actuators B: Chemical, 132 (1), 81 – 89.