



**HAL**  
open science

## Development of a normalized multi-sensors system for low cost on-line atmospheric pollution detection

Zaher Al Barakeh, Philippe Breuil, Nathalie Redon, Christophe Pijolat,  
Nadine Locoge, Jean-Paul Viricelle

### ► To cite this version:

Zaher Al Barakeh, Philippe Breuil, Nathalie Redon, Christophe Pijolat, Nadine Locoge, et al.. Development of a normalized multi-sensors system for low cost on-line atmospheric pollution detection. *Sensors and Actuators B: Chemical*, 2017, 241, pp.1235 à 1243. 10.1016/j.snb.2016.10.006 . hal-01379359

**HAL Id: hal-01379359**

**<https://hal.science/hal-01379359>**

Submitted on 17 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Development of a normalized multi-sensors system for low cost on-line atmospheric pollution detection.

Zaher Al Barakeh<sup>1,2</sup>, Philippe Breuil<sup>1</sup>, Nathalie Redon<sup>2,3</sup>, Christophe Pijolat<sup>1</sup>, Nadine Locoge<sup>2,3</sup>, Jean-Paul Viricelle<sup>1</sup>.

<sup>1</sup>*Ecole Nationale Supérieure des Mines de Saint Etienne, SPIN-EMSE, CNRS:UMR5307, LGF F-42027, Saint Etienne, France.*

<sup>2</sup>*Mines Douai, SAGE, F-59508 Douai, France.*

<sup>3</sup>*Université de Lille, F-59000, Lille, France.*

## Abstract:

Air Pollution monitoring and measurement are generally done using sampling techniques and analysis equipment often heavy, complex and expensive. Although these methods offer a high measurement precision which is essential to answer standards requirements, they are not adapted for quality oriented applications where simple information with low precision can be sufficient. The use of semiconductor gas sensors networks can provide the answer for a “low cost” system intended for such applications in air pollution detection fields. Three identical portable autonomous sensors arrays were built, each containing nine commercial semiconductor sensors especially chosen to detect a large range of pollutants usually encountered in ambient air and for a large part of them regulated. In order to overcome the temporal instability and the lack of reproducibility of these sensors, a calibration and normalisation procedure was developed. The obtained systems were used for on-site pollution monitoring in association with the French National Network of Accredited Associations for Air Quality Monitoring (AASQA). Gathered data from sensors systems and network data (NO, NO<sub>2</sub>, CO, PM<sub>2,5</sub>, ...) were treated using nonlinear regression algorithms like Neural Networks with an original “fuzzy logic” type pre-treatment in order to compute a model able to predict the membership degree for three predefined pollution categories: traffic, urban and photochemical pollution, along with a pollution index describing the severity of the predominant pollution. The prediction rate was estimated system per system, and site per site for six sites. It has been shown that it was possible to obtain a quasi-universal model with a success rate over 80%.

**Key words:** Air pollution detection, semi-conductor gas sensors, sensors array, calibration and normalization procedure, fuzzy logic, Neural Network calibration.

## 1. Introduction

Atmospheric pollution monitoring is usually done by a network of measuring stations. These stations are based on high performance industrial gas analyzers placed in fixed monitoring sites. Measurements from these stations are used to define air quality indexes which depend on some regulated gases (NO<sub>2</sub>, SO<sub>2</sub>, O<sub>3</sub> and/or particles concentrations (PM<sub>10</sub>).

For example, in France, nearly eight hundred permanent stations throughout the territory contribute to foresee short -and mid- term air quality forecast. The measurements done by these stations are based on high performance industrial gas analyzers, which offer an excellent selectivity and precision towards targeted gases. However their high price and heavy weight make their use complicated [1], especially for some application where selectivity and precision are not required [2] like pollution meshing [3] or source detection [4].

A complementary method for air pollution monitoring lies in the use of gas sensors arrays based on semiconductor gas sensors [3, 5-7]. Semiconductor gas sensors have been optimized for more than thirty years [8], giving them a high sensitivity for certain gas detection, a fast response time of up to several seconds, a small size, low cost (few euros) compared to other types of sensors, and ease of use in electronic systems which make them attractive cheap consumers devices [7] [9].

Even if semiconductor gas sensors show a low selectivity [10-14], reproducibility and repeatability [13, 15] to specific gases in a complex mixture, each type of sensor reacts differently to each family of gas species [14]. So, with adequate pattern recognition algorithms (multivariate analyses such as neural networks), this drawback is used as an advantage to identify an overall signature of pollution [16-20].

This point constitutes the sensors array approach. Such systems can be used to perform semi-quantitative measurements [21]. However, complex situations with large numbers of gas can limit their performance [22].

Most of time, pollution detection with commercial gas sensors is usually focused on identifying one single pollutant in the atmosphere [1] [7] [23] [24]. The presence of different interfering gases in the atmosphere renders these detections difficult and vulnerable for mistakes.

Therefore, in this work, we present three identical autonomous portable systems each containing 9 different commercial semi-conductor gas sensors, one humidity sensor and a temperature sensor especially developed to address ambient air pollution problematic. First, through many field experiments, we have evaluated the capability of these multi-sensors arrays to distinguish different types of air pollution like traffic, urban or photochemical pollution, and then we have associated them to an index level, according to French air quality standards. To reach this goal, the developed systems were placed in six different air network stations. Previously, one year of data history from each station was used to establish correlations and repeated patterns between pollutants and atmospheric pollution types. Neural networks models are built from sensors responses in order to predict membership of pollution to the categories previously defined (traffic, urban or photochemical pollution). Two field campaigns, during winter and summer in order to take into account the influence of atmospheric conditions, were conducted to establish the modelling and to validate it.

The portability and robustness of the models for different sensors arrays on different stations was also evaluated.

## **2. Sensors arrays and responses corrections**

### 2.1.Measuring System development:

Three identical sensor arrays (Fig 1) were developed, each containing nine commercial semiconductor gas sensors (from Figaro Sensors and SGX Sensortech, Table 1) and also humidity (Measurement Specialities HS1101) and temperature (Texas Instruments LM35) sensors. The gas sensors were selected to detect the widest gas species usually present in urban atmosphere. We focused firstly on regulated pollutants ( $O_3$ , NO,  $NO_2$ , CO) in order to correlate them to databases provided by the network of air quality monitoring stations. Secondly, we targeted components sensitive to VOCs because those gases can also be specific to kind of contamination like traffic or urban pollution. Finally, we completed our choice, with sensors able to detect unregulated gas like  $NH_3$ ,  $H_2S$ , or alkanes because of their specificity to large types of pollution. Table 1 presents the sensors chosen with their commercial references, their target gases and the detection range as specified by the constructors.



Fig 1: Sensors array including 9 semiconductor sensors

Table 1: Chosen semiconductor based gas sensors with their target gases and their specified detection range, according to constructor specifications (\* TGS: Figaro Sensors, MICS: SGX Sensortech)

Sensor*	Target gases	Constructor detection range
MICS5525	CO	1-1000 ppm
MICS2710	$NO_2$	0,5-5 ppm
MICS2611	$O_3$	0,1-1 ppm
TGS2600	Large range of air contaminants	500-10000 ppm
TGS2602	VOC-Odors- $H_2S$	30-10000 ppm
TGS2611	Combustible gases - Methane	500-10000 ppm
TGS2620	Alcohol and solvent vapours	50- 10000 ppm
MICS5521	CO-VOC-Hydrocarbons	0,1 - 100 ppm

MICS5914	NH <sub>3</sub>	300-40000 ppm
----------	-----------------	---------------

All these gas sensors undergo a variation in conductance when the gas concentration in their environment is changed. A dimensionless quantity called "relative conductance variation" was used to characterize the sensor response (Eq.1).

$$X_R = (G - G_0) / G_0 \quad (1)$$

With:

- $X_R$ : relative conductance variation of a sensor
- $G$ : sensor's conductance
- $G_0$ : reference conductance (base-line) measured under "reference air" (synthetic air, nitrogen /oxygen 80/20 vol.% with H<sub>2</sub>O, 12g/m<sup>3</sup> (1.5% volume)).

It is clearly noticeable that these specified detection ranges are much higher than the usual pollutants concentrations in the atmosphere. Therefore testing these sensors under relatively low concentrations of pollutant gases was a necessity. To perform such experiments, a test bench allowing to characterize each array under defined gases was used. These gases were methane, toluene, nitrogen dioxide, carbon monoxide, ozone, ammonia and ethanol. For each gas, the concentrations were fixed by dilution in reference air described previously, thanks to massflow regulators. These concentrations were chosen according to levels that can be encountered in ambient air. Fig 2 presents sensors responses from one of the sensors arrays measured in such conditions.

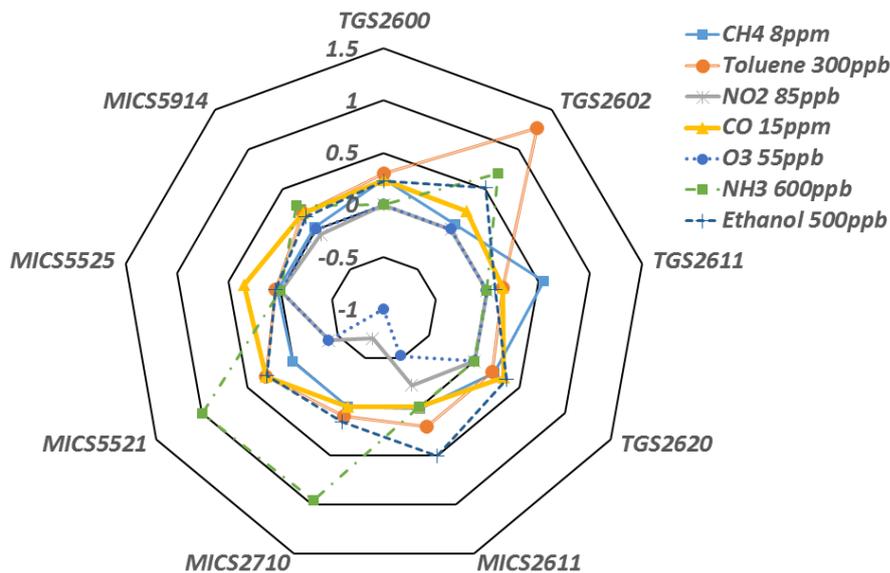


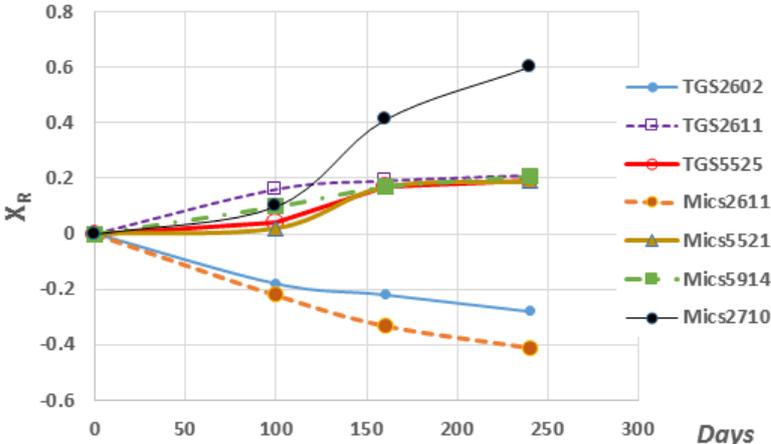
Fig 2 : Sensors relative conductance variation ( $X_R$ ) for low concentration of pollutants

Each signature is specific for a given gas at a given concentration. These sensors show a significant response toward different gases at concentrations much lower than the detection

ranges specified by their constructors, tested concentrations levels can be representative of what can be found in polluted sites.

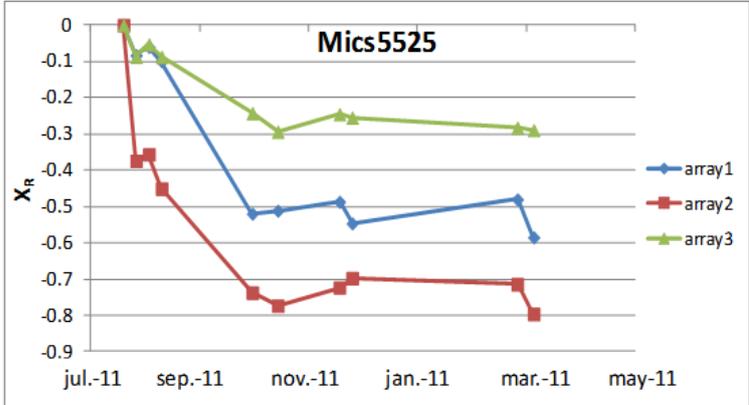
*2.2. Baseline drifts correction:*

The drift of the sensors signal can cause significant problems of temporal validity for the produced models. So a periodic measurement of this change is necessary to ensure proper monitoring of the baseline. Sensors baseline drift was studied by measuring and giving each month a reference conductance ( $G_0$ ) to each sensor calculated under reference air for specific conditions of temperature and relative humidity (50% at 25°C). Fig 3 shows the evolution of sensors relative conductance variation  $X_R$  (defined by eq 1) for zero air flow during 7 months of tests for the first prototype array.



**Fig 3 : Relative conductance variation ( $X_R$ ) under zero air variation during 7 months of measurements**

Baseline drift value is specific to each sensor. It is imperative that this test is always repeated under the same conditions. Drifts can take different shapes depending on the sensor’s type, but sensors of same type can also have different drifts. Fig 4 shows the evolution of the baseline of Mics 5525 sensors from the 3 different arrays for another period of 8 months.



**Fig 4: Baseline drift registered for different Mics5525 sensors from three arrays during 8 months of observation.**

This figure shows a decreasing trend similar between the three different Mics 5525 sensors. However, the amplitude of the variation varies from one sensor to another. We note that responses have sharply decreased in the first months (July-November), and after this period, they were relatively stable.

This confirms that a periodic measurement of these changes is necessary to ensure proper monitoring of the baseline. Therefore a calibration correction was done on a weekly basis to detect this drift in order to have new measurements of the reference  $G_0$ , although it can probably be reduced to a monthly basis with no significant changes if sensors are aged at least 3 months before use.

### 2.3. Sensitivity change compensation:

Besides the drift of the baseline, semi-conductor gas sensors also tend to undergo changes in their sensitivity through time. Moreover, same sensors type can have different sensitivity to the same gases, inducing a low reproducibility between them. To homogenise a relative repeatable and reproducible responses between the 3 arrays, a normalisation procedure was elaborated. The principle is based on normalization of the relative conductance of each same type sensors under one reference gas as shown in Table 2. Four gases were chosen as representatives of pollutants in the atmosphere (CO, CH<sub>4</sub>, NO<sub>2</sub>, NH<sub>3</sub>), and each sensor was normalized toward the gas for which the sensor shows highest sensitivity.

Sensor	Reference gases N	C <sub>N</sub>
MICS5525	CO	20 ppm
MICS2710	NO <sub>2</sub>	0.4 ppm
MICS2611	NO <sub>2</sub>	0.4 ppm
TGS2600	NH <sub>3</sub>	4 ppm
TGS2602	NH <sub>3</sub>	4 ppm
TGS2620	NH <sub>3</sub>	4 ppm
TGS2611	CH <sub>4</sub>	20 ppm
MICS5521	CO	20 ppm
MICS5914	NH <sub>3</sub>	4 ppm

**Table 2: Reference gas and concentration chosen for each sensor for normalization**

Normalization procedure for a given sensor “i” was made by dividing its relative conductance variation  $X_{R,i}$  by its relative conductance variation  $X_{N,i}$  under the chosen normalization gas “N” with a concentration  $C_N$ :

$$S_i = X_{R,i} / X_{N,i} \quad (2)$$

This test can also be periodically used as a diagnostic tool to detect any sensor failure that implies its replacement. Figure 5 shows the sensors response to reference gases and the effect of this sensitivity normalization on the 3 multi-sensors arrays:

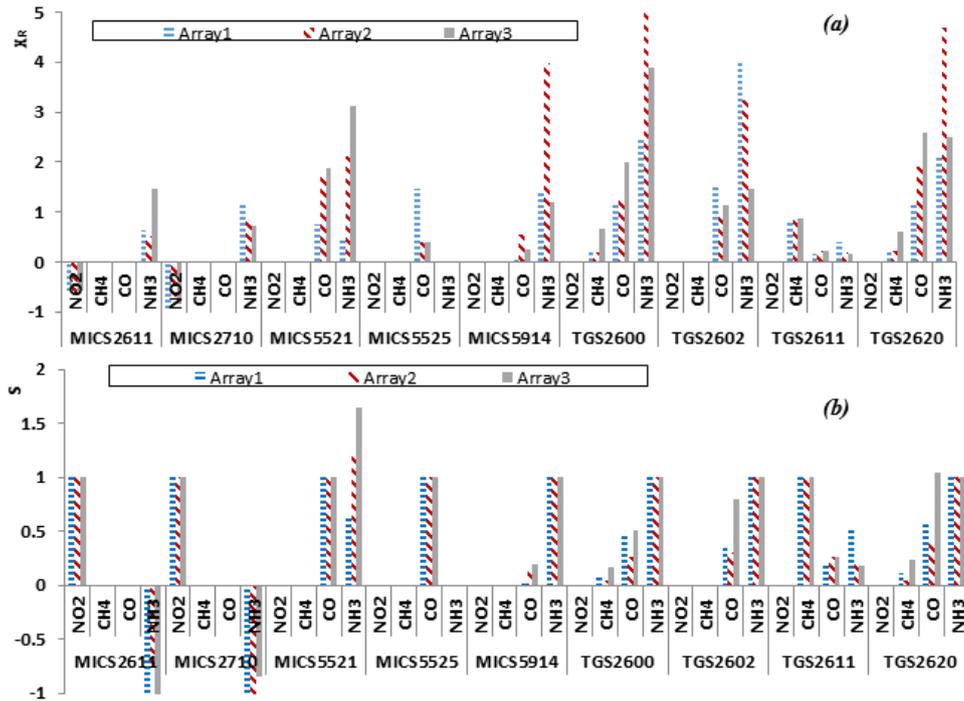


Fig 5: Sensors responses to normalization gases: (a)  $X_R$  before normalization, (b)  $S$  after normalization .

Fig 5a represents relative conductance of sensors under target gases before normalization, and Fig 5b shows these variations after normalization. Although differences of normalized response to a given target gas is still observable between the three sensors of same type for each array, normalization has decreased in a remarkable way the amplitude of these variations.

### 3. On site measurements and monitoring method

#### 3.1. Field measurement sites and classes of pollution:

Two measurement campaigns (one in summer and one in the end of autumn) conducted on site during this work were done in six stations of the National Network of Accredited Associations for Air Quality Monitoring of France. AASQA networks are based on different types of monitoring stations placed in specifically determined position according to the EU directive (2008/50/EU).

These six monitoring stations, distributed in the north of France, are representative of three different categories of majority pollutions, named urban, suburban and traffic as shown in Table 3. For each pollutant, mean and maximum value as well as standard deviation over one year history's data are reported. During these measuring campaigns, measurements were done during 5 to 14 consecutive days to ensure the registration of special events, like traffic spikes for example, and also to get data for the calculation of the neural network modelling.

**Table 3 : Statistical analysis of pollutant concentrations in six AASQA stations**

Station		NO ( $\mu\text{g}/\text{m}^3$ )			NO <sub>2</sub> ( $\mu\text{g}/\text{m}^3$ )			O <sub>3</sub> ( $\mu\text{g}/\text{m}^3$ )		
Station	Category	Mean	Max	Std dev	Mean	Max	Std dev	Mean	Max	Std dev
Creil	Urban	9	243	16	24	122	17	42	235	29
Douai	Urban	9	296	18	25	100	15	39	182	27
Bethune	Urban	6	191	13	22	150	16	44	182	27
Nogent	Suburban	12	254	21	26	139	17	43	236	30
Roubaix	Traffic	32	801	46	46	301	25	---	---	---
Valenciennes	Traffic	20	405	29	37	152	18	---	---	---

Previously, one year data history issued from analyzers and collected from these stations were studied in order to define a credible pollution categorization according to its source. The analysis was oriented to find a repeatable patterns or specific characteristics that can define existing pollution in a station. Thus repeated patterns in stations from similar category (urban, suburban or traffic) were studied. Then three pollution categories were defined using a fuzzy logic approach from the measurements of O<sub>3</sub>, NO, and NO<sub>2</sub> concentration in the atmosphere:

- Traffic pollution: sites near emission of pollutants, high correlation between NO and NO<sub>2</sub>, ozone absent (reaction with other primary pollutants),
- Urban pollution: more far from emission, NO is partially destructed.
- Photochemical pollution: high concentration of ozone, low concentrations of NO and NO<sub>2</sub>, typical of suburban sites,

However, if the traffic, urban and photochemical pollution types are most probable on respectively traffic, urban and suburban site types, each type of pollution can still be met on each type of site.

As boundaries between categories are not so easy to define, which may be detrimental to the modelling, we have chosen a “fuzzy logic” approach. Fuzzy logic was introduced in 1965 by Lotfi Zadeh [25] based on his mathematical theory of fuzzy sets. By introducing the notion of a membership function (MF) allowing a condition to be in a state other than true or false unlike boolean logic, it provides flexibility and makes possible to take into account inaccuracies and uncertainties. Indeed, a boolean logic uses a step function defined with one threshold while fuzzy logic uses a membership function, which is a sigmoid function that can be defined with two thresholds  $l_1$  and  $l_2$ , to calculate a “membership degree” (MD) (Fig 6 and Table 4). For a given assertion, this membership degree indicates its probability, from 0 (not probable) to 1 (very probable). The two thresholds are calculated from the median and the average values of the concentrations for one year of measurements and for the six stations.

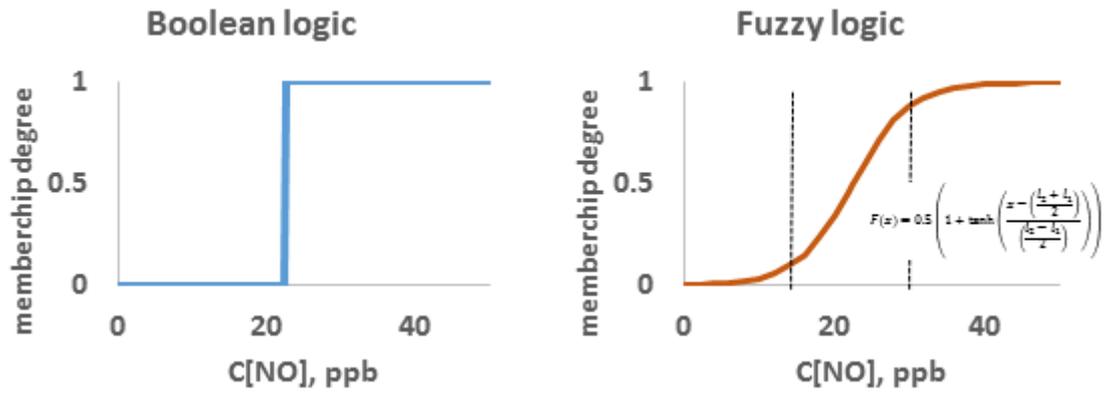


Fig 6 : Membership function of the assertion “pollution with NO”, Boolean logic and Fuzzy logic

Table 4 : Gas concentration thresholds for each pollution category

Thresholds	NO	NO <sub>2</sub>	O <sub>3</sub>
l <sub>1</sub>	15 µg/m <sup>3</sup>	20 µg/m <sup>3</sup>	30 µg/m <sup>3</sup>
l <sub>2</sub>	30 µg/m <sup>3</sup>	40 µg/m <sup>3</sup>	50 µg/m <sup>3</sup>

Then the membership degree of the different classes can be calculated from the membership degree of the pollutants and their equations (Table 5), following the fuzzy logic theory laws [25].

Table 5 : Determination of membership degrees for the 3 categories of pollution from the concentrations of pollutants according to Boolean or Fuzzy logic equations

pollution class	Boolean logic equation	Fuzzy logic equation
Traffic	$[NO] * [O_3]$	$MD(NO) * (1 - MD(O_3))$
Urban	$[NO_2] * [NO] * [O_3]$	$MD(NO_2) * (1 - MD(NO)) * (1 -$
Photochemical	$[O_3]$	$MD(O_3)$

### 3.2. Pollution index:

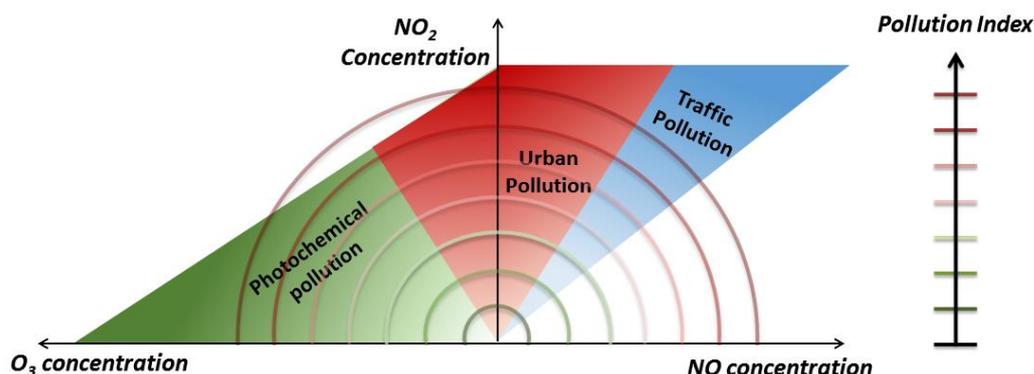
To complete this categorization with a quantitative information, a pollution index (PI) was created in order to describe the pollution intensity. This index is inspired from the index defined by AASQA [26], named ATMO, which is defined as the maximum of 4 minor indexes calculated out of the concentration of 4 pollutants (SO<sub>2</sub>, O<sub>3</sub>, NO<sub>2</sub>, PM10), and is used to inform the public about the ambient air quality in France. In our case, the chosen pollution categories are based on only 3 minor index calculated from pollutants concentrations (NO, NO<sub>2</sub> and O<sub>3</sub>). Each index is integer and varies from 1 to 10, 1 being the case of a no or low concentration of a pollutant and 10 being the maximum defined threshold in the European directives. The highest value among them conveys the value of the overall index. O<sub>3</sub> and NO<sub>2</sub>

indexes can be calculated in the same way defined by ATMO, but a new index for NO had to be calculated (Table 6) as this gas is not included in the AASQA measurements. It was estimated in order to keep the same frequency distribution for pollution indexes of NO and NO<sub>2</sub> encountered in the year 2010.

**Table 6 : Minor pollution indexes (PI) as a function of pollutants concentrations for the three pollutants considered in this study (NO, NO<sub>2</sub> and O<sub>3</sub>)**

<b>Pollution</b>	<b>NO (<math>\mu\text{g}/\text{m}^3</math>)</b>	<b>NO<sub>2</sub> (<math>\mu\text{g}/\text{m}^3</math>)</b>	<b>O<sub>3</sub> (<math>\mu\text{g}/\text{m}^3</math>)</b>
<b>1</b>	<b>0-9</b>	0-29	0-29
<b>2</b>	<b>10-34</b>	30-54	30-54
<b>3</b>	<b>35-74</b>	55-84	55-79
<b>4</b>	<b>75-154</b>	85-109	80-104
<b>5</b>	<b>160-239</b>	110-134	105-129
<b>6</b>	<b>239-314</b>	135-164	130-149
<b>7</b>	<b>315-379</b>	165-199	150-179
<b>8</b>	<b>380-439</b>	200-274	180-209
<b>9</b>	<b>440-499</b>	275-399	210-239
<b>10</b>	<b><math>\geq 500</math></b>	$\geq 400$	$\geq 240$

The three categories of pollution and the pollution index are illustrated as a function of the concentrations of pollutants on Fig 7.



**Fig 7 : Illustration of categories of pollution and pollution index**

Then our goal is to predict, from the gas sensors signals, this pollution index and the class of pollution which is the one with the highest membership degree.

### 3.3. Artificial Neural Network Modelling:

In order to perform model of the array sensor response, we have opted for an artificial neural network (ANN, Fig 8) for its ability to model non-linear phenomena. The input data of

the ANN are the normalized responses  $S_i$  from the 9 sensors forming each multi-sensor system, to which are added relative humidity, temperature, time of day and if it is a weekday or not. Output data “Y” correspond to the membership degree of the 3 categories of pollution to which we add the pollution index. The output layer is connected to the input layer via one hidden layer each consisting of 4 to 12 neural nodes. The relations between the different layers are either linear combinations of the preceding layer for input and output layers, either non-linear combination for the hidden layer. A hyperbolic tangent function was used as the non-linear function in our case.

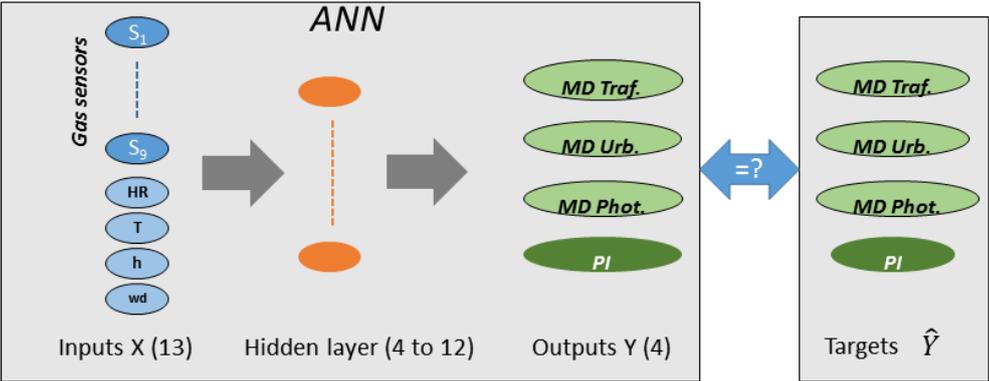


Fig 8 : Structure of the artificial neural network used for the pollution categorization

The building of the target data “ $\hat{Y}$ ”, to which the output data Y resulting from the ANN is compared during the learning step, is summarized on Fig 9. Starting from gas concentrations measured by gas analysers, membership degree and pollution index for each three gas are calculated. Then, the overall index and the membership degree of the three pollution category constitute the target data “ $\hat{Y}$ ”.

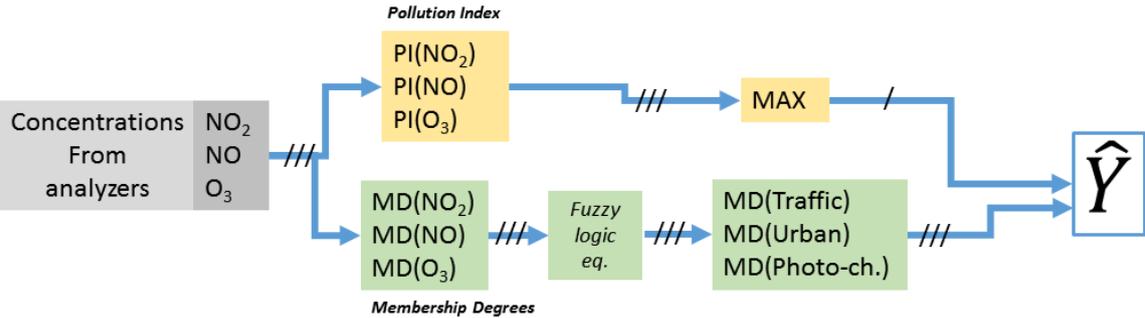


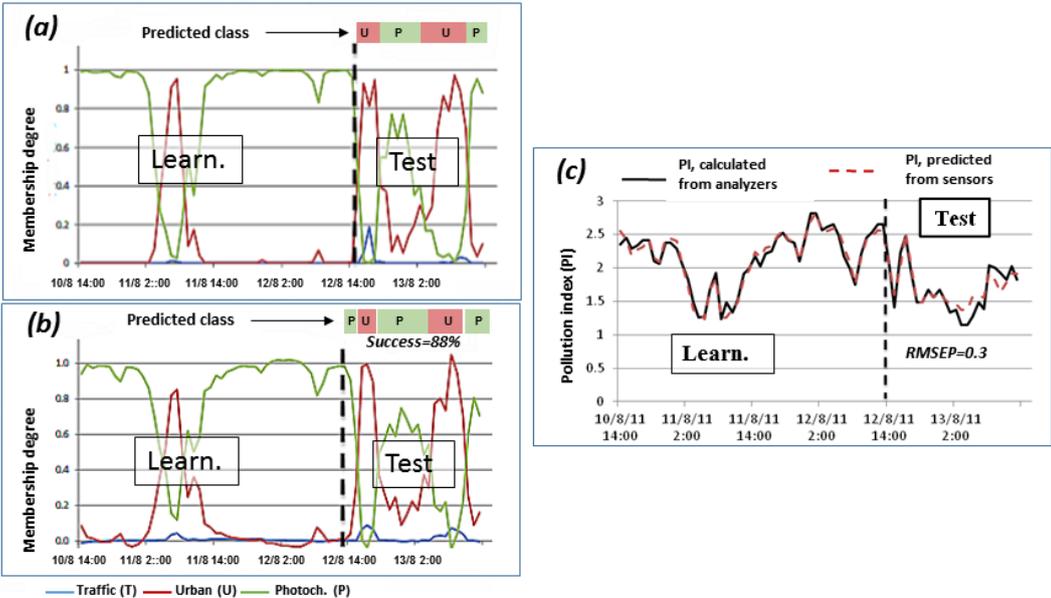
Fig 9 : Building of the target data  $\hat{Y}$  to be compared with output data “Y” from the artificial neural network

For a typical ANN-modelling, three sets of data are used: the two to five first measurement days were used for the learning, followed by one day for the validation and one day for the test. The modelling is iteratively optimised during the learning step by adjusting the coefficients of the neural networks so that the Root Mean Error of Prediction (RMSEP)

comparing the Y variables calculated from the conductance of the sensors, and  $\hat{Y}$  calculated from measurements of gases analysers, is as low as possible for the learning set. This iterative optimisation is stopped when the RMSEP for the validation set no longer decreases. Once the coefficients of the model are optimized, they are applied to the test set, and a classification rate is calculated: it represents the percentage of right classifications compared to the categorization obtained by measurements done with AASQA analysers. Similarly, the PI prediction quality is quantified with the RMSEP.

**4. Results of classification and prediction:**

Fig 10 shows a short example of learning/classification/prediction. Results are shown over a three days period. The two first correspond to the learning period while the last one is the test period. At any time, the predicted classes corresponds to the highest membership degree of the 3 classes. Agreement between predicted class from sensor array and from analysers is observed for more than 88% of the test duration. There is also a very good correlation between predicted and calculated pollution index. The RMSEP of the pollution index is less than 0.3 during the test period.



**Fig 10 : Short example of classification and prediction from array n° 3, Béthune, period from 10/8 to 13/8; 2 days for artificial neural network learning, 1 day for test. (a): classification from analyzers data, (b): classification from gas sensors data, (c): pollution index (PI): calculated from analyzers and predicted from sensors array**

Table 7 shows the overall classification rates and RMSEP for the pollution index for the three multi-sensors arrays deployed on the six stations. Data are issued from more than 200 hours of test and 1200 hours of learning and validation acquired during the summer and

autumn campaigns. All performances are close to each other and especially satisfactory since all classification rates are higher than 85%.

**Table 7 : Multi-sensors arrays classification success rates and pollution index prediction error (RMSEP)**

Site	Type	Multi-sensors array n°1		Multi-sensors array n°2		Multi-sensors array n°3	
		Classification rate	RMSEP PI	Classification rate	RMSEP PI	Classification rate	RMSEP PI
Creil	Urban	87%	0.04	90%	0.03	89%	0.05
Nogent	Suburban	90%	0.03	91%	0.01	89%	0.01
Saint-Amand	Suburban	89%	0.03	87%	0.03	88%	0.03
Valenciennes	Traffic	91%	0.02	90%	0.01	89%	0.02
Béthune	Urban	91%	0.03	98%	0.02	98%	0.03
Douai	Urban	91%	0.01	92%	0.01	93%	0.02

These performances show the ability of a multi-sensors system to obtain relevant information about atmospheric pollution, using a model based on neural networks. But it is also important to study the domain of validity of the calculated models, assessing their portability from one multi-sensors array to another and also from one measuring site to another. The aim is to build a robust model. To check the first criteria, inter-array portability, various combinations of arrays were made for both learning and predictions steps on one site. Table 8 reports the results obtained in the case of measurements done on the station of Nogent.

**Table 8 : Example of inter-arrays model portability for the urban site of Nogent**

Array use for model learning...	Array use for Prediction done	Classification Success Rate	Pollution index RMSEP
array 1	array 1	90%	0.03
array 2	array 2	91%	0.01
array 3	array 3	89%	0.01
array 1	array 2	25%	0.2
arrays 1+2	arrays 1+2	83%	0.06
arrays 1+ 2	array 3	72%	0.1

As it can be seen, the model calculated on a multi-sensors array (n°1) and applied on itself has very good performances but is useless because when this same model is applied to another multi-sensors array (n°2), it has a very low success rate, around 25%. In contrast, a model calculated on a set of two multi-sensors arrays (n°1+n°2) and applied to a multi-sensors array which data were not used in the calculation of the model parameters (n°3), has correct success rate over 70%. These results tend to demonstrate that if the number of multi-sensors arrays is large enough for learning phase, the model becomes more general, but slightly less efficient.

Anyway, with only 3 multi-sensors arrays, the statistical size seems to be good enough to make a pertinent modelling.

In order to study the inter-site portability, same types of experiments were carried out by changing the site for learning and prediction with a given array (Table 9).

**Table 9 : Example of inter-sites model portability**

Model learning site	Prediction site	Array 1	
		Classification success rate	Pollution index RMSEP
Bethune – Urban	Bethune - Urban	91%	0.03
Bethune - Urban	Nogent - Suburban	47%	0.15
All stations out of Bethune (all types)	Bethune - Urban	86%	0.09

Like for inter-array portability, a model calculated from data collected on one station (Bethune) and applied to another one (Nogent) presents a bad success rate, around 50%. However, if the parameters are calculated from a set of different stations (including urban types), and applied to a station that was not taking into account in the calculation, then the results are satisfying, always upper than 80%. So, the more sites representative of each pollution categorization are involved in the learning and validation phases, the more robust is the model.

## 5. Conclusions

The concept of qualitative and semi-quantitative pollution detection using a gas sensor array was presented in this work. The method is based on modelling temporal responses of conductances in multi-sensors systems. Three identical sensors arrays with nine different commercial semiconductor gas sensors were constructed. Calibration and standardization techniques were elaborated to correct drifts of baseline and of sensitivity and assure an inter-array homogeneity in sensors response. The study was carried out by measuring existing pollution in different polluted sites. Six sites were chosen, and the pollution type definition was made based on pollutant concentrations measured by air quality network stations. A pollution index was also introduced to describe pollution's severity. This combination of predicted pollution category and pollution index can be used to inform the public on the pollution in real time, and with a high spatial accuracy. Such multi-sensors arrays can be multiplied and used in a way to mesh a sensible zone to be monitored. Up to now, and according to our knowledge, no equivalent instrument combining multi-sensors arrays with

models based on neural networks using the proposed semi-quantitative approach exists. The classification success rate for pollution categories prediction had a success rate higher than 80% with a low error in the prediction of the pollution index. Presented results show that the model can be applied to a large number of uncharacterised arrays of sensors (thanks to calibration / normalization procedure) and unknown sites if a sufficient number of reference sites and arrays for the learning and validation phases of the neural network are used. Further modelling approach could be applied in order to build equivalent categorizations for inside air quality problematic.

## References:

- [1] M. Kamionka, P. Breuil, C. Pijolat, Calibration of a multivariate gas sensing device for atmospheric pollution measurement. *Sensors and Actuators B, chemical* 118 (2006) 323-327, <http://dx.doi.org/10.1016/j.snb.2006.04.058>.
- [2] N. Yamazoe, K. Shimano, New perspectives of gas sensor technology. *Sensors and Actuators B, Chemical* 138 (2009) 100-107, <http://dx.doi.org/10.1016/j.snb.2009.01.023>.
- [3] Y. Ma, M. Richards, M. Ghanem, Y. Guo, J. Hassard, Air Pollution Monitoring and Mining Based on Sensor Grid in London, *Sensors* (2008) 3601-3623, <http://dx.doi.org/10.3390/s80603601>.
- [4] P. Ködderitzsch, R. Bischofa, P. Veitenhansl, W. Lorenz, G. Bischoff, Sensor array based measurement technique for fast-responding cigarette smoke analysis, *Sensors and Actuators B, Chemical* 107, (2005) 479-489, <http://dx.doi.org/10.1016/j.snb.2004.11.007>.
- [5] C. Pijolat, C. Pupier, M. Sauvan, G. Tournier, R. Laulauze, Gas detection for automotive pollution control. *Sensors and Actuators B, Chemical* 59 (1999) 195-202, [http://dx.doi.org/10.1016/S0925-4005\(99\)00220-8](http://dx.doi.org/10.1016/S0925-4005(99)00220-8).
- [6] B. Rivière, J.P. Viricelle, C. Pijolat, Development of tin oxide material by screen-printing technology for micro-machined gas sensors *Sensors and Actuators B, Chemical* 93 (2003) 531-537, [http://dx.doi.org/10.1016/S0925-4005\(03\)00173-4](http://dx.doi.org/10.1016/S0925-4005(03)00173-4).
- [7] M. Kamionka, P. Breuil, C. Pijolat, Atmospheric pollution measurement with a multi-materials sensing device, *Materials Science and Engineering: C* 26 (2006) 290-296, <http://dx.doi.org/10.1016/j.msec.2005.10.047>.
- [8] J. Watson, The tin oxide gas sensor and its applications, *Sensors and Actuators. B, Chemical* 5 (1984) 29, [http://dx.doi.org/10.1016/0250-6874\(84\)87004-3](http://dx.doi.org/10.1016/0250-6874(84)87004-3).
- [9] S. Zampolli, I. Elmi, F. Ahmed, M. Passini, G. Cardinali, S. Nicoletti et L. Dori, An electronic nose based on solid state sensor arrays for low-cost indoor air quality monitoring applications, *Sensors and Actuators B, Chemical* 101(2004) 101, 39-46, <http://dx.doi.org/10.1016/j.snb.2004.02.024>.
- [10] M. Padilla, A. Perera, I. Montoliu, A. Chaudry, K. Persaud, S. Marco, Drift compensation of gas sensor array data by Orthogonal Signal Correction. *Chemometrics and Intelligent Laboratory Systems*, 100 (2010) 28-35, <http://dx.doi.org/10.1016/j.chemolab.2009.10.002>.
- [11] O. Tomic, H. Ulmer, J.E. Haugen, Standardization methods for handling instrument related signal shift in gas-sensor array measurement data, *Analytica Chimica Acta* 472, (2002). 99-111, [http://dx.doi.org/10.1016/S0003-2670\(02\)00936-4](http://dx.doi.org/10.1016/S0003-2670(02)00936-4).
- [12] G. Korotcenkov, B. Cho, Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey), *Sensors and Actuator B, Chemical* 156 (2011) 527-538, <http://dx.doi.org/10.1016/j.snb.2011.02.024>.

- [13] A.C. Romain, J. Nicolas, Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview, *Sensors and Actuators B, Chemical* 146 (2010) 502-506, <http://dx.doi.org/10.1016/j.snb.2009.12.027>.
- [14] S. De Vito, M. Piga, L. Martinotto, G. Di Francia, CO, NO<sub>2</sub> and NO<sub>x</sub> urban pollution monitoring with on field calibrated electronic nose by automatic bayesian regulation *Sensors and Actuators B, Chemical* 143 (2009) 182-191, <http://dx.doi.org/10.1016/j.snb.2009.08.041>.
- [15] O. Tomic, T. Eklöv, K. Kvaal, J.-E. Haugen, Recalibration of a gas-sensor array system related to sensor replacement *Analytica Chimica Acta* 512, (2004) 199-206, <http://dx.doi.org/10.1016/j.aca.2004.03.001>.
- [16] R. Shaffer, S. Rose-Pehrsson, R. McGill, A comparison study of chemical sensor array pattern recognition algorithms *Analytica Chimica Acta* 384, (1999) 305-317, [http://dx.doi.org/10.1016/S0003-2670\(98\)00780-6](http://dx.doi.org/10.1016/S0003-2670(98)00780-6).
- [17] V. Sommer, P. Tobias, D. Kohl, H. Sundgren, I. Lundstriim, Neural networks and abductice networks for chemical sensor signals: a case comparision *Sensors and Actuators B, Chemical* 28 (1995) 217-222, [http://dx.doi.org/10.1016/0925-4005\(95\)01721-6](http://dx.doi.org/10.1016/0925-4005(95)01721-6).
- [18] C. D. Natale, F. Davide, A. D'Amico, Pattern recognition in gas sensing: well-stated techniques and advances *Sensors and Actuators B, Chemical* 23,(1995) 111-118, [http://dx.doi.org/10.1016/0925-4005\(94\)01279-2](http://dx.doi.org/10.1016/0925-4005(94)01279-2).
- [19] E. Hines, E. Llobet, J. Gardner, Electronic noses: a review of signal processing techniques *IEEE Proceedings – Circuits Devices and Systemes* 146, (1999) 297-310, <http://dx.doi.org/10.1049/ip-cds:19990670>.
- [20] P. Szecowka, A. Szuczurek, B. Licznanski, On reliability of neural network sensitivity analysis applied for sensors array optimisation. *Sensors and Actuators B, Chemical* 157, (2011) 298-303, <http://dx.doi.org/10.1016/j.snb.2011.03.066>.
- [21] R.M. Negri, S. Reich S. Identification of pollutant gases and its concentrations with a multisensors array. *Sensors and Actuators B, Chemical* 75, (2001) 172-178.
- [22] W. Bourgeois, The use of sensor arrays for environmental monitoring: interests and limitations. *The Royal Society of Chemistry* 5 (2003) 852-860, <http://dx.doi.org/10.1039/B307905H>
- [23] W. Tsujita, A. Yoshino, H. Ishida, T. Moriizumi. Gas sensor network for air-pollution monitoring. *Sensors and Actuators B, Chemical* 110 (2005) 304-311, <http://dx.doi.org/10.1016/j.snb.2005.02.008>.
- [24] H. Shurmer, P. Corcoran, M. James, Sensitivity enhancement for gas sensing and electronic nose applications *Sensors and Actuators B, Chemical* 16 (1993) Issues 1-3, 256-259. [http://dx.doi.org/10.1016/0925-4005\(93\)85191-C](http://dx.doi.org/10.1016/0925-4005(93)85191-C).
- [25] L.A. Zadeh, Fuzzy sets. *Information and Control* 8 (1965) 338–353. [http://dx.doi.org/10.1016/S0019-9958\(65\)90241-X](http://dx.doi.org/10.1016/S0019-9958(65)90241-X).
- [26] “Arrêté du 22 juillet 2004 relatif aux indices de la qualité de l'air », NOR: DEVP0430272A (<http://www.legifrance.gouv.fr>), 2004 (accessed 27.04.16).