Chemometrics methods for the identification and the monitoring of an odour in the environment with an electronic nose.

Jacques NICOLAS, Anne-Claude ROMAIN, Jindriska MATERNOVA

Fondation Universitaire Luxembourgeoise
"Environmental Monitoring"
Avenue de Longwy, 185
B 6700 - ARLON
BELGIUM
e-mail : nicolas@ful.ac.be

Running title : Chemometrics applied to electronic nose in the environment.

ABSTRACT

The purpose of the paper is to briefly review some researches regarding the adaptation of the electronic nose principle to recognise some malodour sources in the environment, if possible directly in the field, and to monitor the odour intensity continuously.
Research aims at improving the portability and the user-friendliness of the instrument, together with testing what kind of signal may be used to monitor the odour.

A laboratory-made electronic nose, constituted of an array of tin-oxide sensors, is used in different configurations. The ambient air is either sampled around environmental sources (landfill, urban waste composting facilities, …), or directly transferred into the sensor chamber in the field.

Two main options are considered : firstly, identifying the source of odour in the background and among interfering odours and, secondly, when the malodour is recognised, trying to monitor it continuously in order, for example, to assess the nuisance or to control an odour abatement system.

Chemometrics methods are generally used for both purposes. They provide quick answers and allow to evaluate the relationships between variables and between observations at a glance.
They are applied on the sensor signals, eventually preprocessed by a suitable algorithm.

Non-supervised analyses, such as Principal Component Analysis (PCA), provide basically a performance evaluation of the system during the development phase. On the contrary, supervised analyses, such as Discriminant Analysis (DA), or some Neural Networks algorithms are quite appropriate to make a reliable recognition in real time, when the system is developed.

To predict the odour intensity, different techniques are tested : either using only one of the sensor elements, or applying different chemometrics techniques, such as Multilinear Regression (MLR) on the original measured sensor signals, Principal Component Regression (PCR), or Partial Least Squares regression (PLS). The latter seems to be the most adapted model for the intensity prediction.

INTRODUCTION

Numerous activities are responsible for various annoyances in the surroundings. Among these environmental pollutions, malodours are the less accepted by the complainants. The malodours emanating from some industrial, agricultural or service activities greatly impair the comfort status requested in our civilised countries.

The increasing public concern about such source of environment pollution has given rise to the implementation of laws and directives. The usual method of odour measurement is based on dilutions of odoriferous samples assessed by human panel and on statistical calculations. Although this procedure gives the right odour human sense evaluation, it is strongly affected by subjectivity. A second way of measurement is the chemical analysis of the odorous mixture by gas chromatography/mass spectrometry (GC-MS) that provides accurate concentration of each compound in the sample but not the total olfactory perception.

Furthermore, those two methods are not designed for real-time and continuous operation on site. Indeed, the environmental malodours being characterised by a puff dispersal, spot samplings of the polluted air are not sufficient to observe the annoyance perceived by the population.
Consequently, for several years, there is a need for a portable air pollution monitor.
The development of sensor array technology so called "electronic noses" for odour classification may offer an objective and on-line instrument for assessing environment odours.
However, among the potential application areas for electronic noses, the monitoring of our environment constitutes a real challenge. The rigorous experimental conditions which apply in the laboratory are no longer usable in the field. Particularly, the input odour can't be considered as a well-defined variation of the gaseous ambience with respect to clean air. In the field of environmental monitoring, the background is an ever-changing chemical mixture against which we want to detect the rise of a particular odour - although the exact profile of that rise is both unknown and variable [1].

In fact, few studies are devoted to environmental applications in the field [2,3,4,5,6]. All of them are restricted to the identification of very specific odours, chiefly at the emission, just near the source. The majority of them apply the electronic nose to the detection of hazardous compounds or of olfactory nuisance in the agricultural and the breeding sectors.

The scientific articles generally show the graphical results of the model calibration step, which consist commonly in 2-dimensional scatterplots identifying some groups of different odours. Using the calibrated classification model to recognise unknown odour sources is less usual and, anyhow, such prediction is seldom reported for field application.

For 5 years, the department "Environmental Monitoring" of F.U.L. has been trying to assess the ability of a detector based on the electronic nose principle to discriminate between some environmental odours, if possible directly in the field, and to monitor them continuously. Such an approach should aim at a better understanding of the odour release, by relating it to the process which caused the emission. But the most interesting challenge of the continuous monitoring of malodour in the field is the real time control of odour abatement techniques.

ELECTRONIC NOSE PRINCIPLE

The electronic nose principle seems well suited to such objectives.
It is made of an array of non specific gas sensors. When all the sensor responses are put together, they form a pattern which is typical of the odour presented to the array, like a signature.
The method is characterised by a learning phase, during which a great number of gaseous mixtures are presented to the sensor array.

Thanks to a pattern recognition engine, using chemometrics techniques, a library of typical patterns for the sensor signals is constituted, each of them corresponding to a specific environmental odour.

For our work, we have chosen so far to use tin oxide sensors, based on the increase of electrical conductivity when a gas compound interacts with their surface. Their use is thus very simple. Their major drawback, for applications in the field, is that the sensing element must be maintained at an elevated temperature, typically above 300°C, in order to improve the kinetics of adsorption. And so, the sensor array consumes about 1 Ampere.
But, generally, tin oxide sensors are quite robust, the influence of water vapour on their response is acceptable, and they are commercially available : almost all the results presented in this paper are obtained with sensors from the Japanese company Figaro.
The option of working with a home-made instrument is justified by the fact that commercial electronic noses are generally not suitable for environmental applications in the field : they are rather expensive and, usually, they are intended for laboratory applications for which the odour is sampled from the head space above a liquid or a solid.
In the field, the air is directly sampled in the gaseous ambience, with its humidity, and its temperature.
Also, the usual way of using the sensors is to work by cycling between a reference air, free from any contaminant, and the odorous sample. The response of the sensor is generally the difference between the signal, after equilibrium in the odorous ambience, and the base line generated in pure air.
But the use of a pure air cylinder is not convenient for field applications : that is heavy and cumbersome. Alternatively, filtering the ambient air through charcoal or molecular sieves needs very restricting maintenance conditions.
We decided not to work with perfectly pure air, but with the ambient air, coarsely filtered and not dried. The goal is not to create a base line for the measurement reference, but simply to purge the sensor vessel, so as to regenerate the sensors.
Obviously, the recorded signal is less pretty than the one obtained with a commercial electronic nose, under the rigorous operating conditions of the laboratory, and with perfectly controlled and stable head space. But anyhow, the results are promising and seems sufficiently accurate for field applications. If the results are actually worse, the requirements are less restricting. For example, in food analysis, one have to class very accurately three or four coffee beans, or edible oils, in order to grant a quality label to the product. The monitoring of a given odour in the environment aims just at detecting the rise of the odour in the background, in order to supply a warning signal when the odour increases beyond a given threshold : that is only an "on-off" signal, and not an accurate measurement.

SAMPLING, MEASUREMENT AND DATA PROCESSING

Firstly, we have tested the recognition of synthetic odours, made of chemical compounds, typical of environmental odours : that is alcohols, esters, amines, aldehydes, ketones and sulfides.

A second set of experiments concerns several real environmental odour sources. As the study aims at evaluating the ability of the array to identify each odour, the samples are collected near the source and not in the surrounding. The malodours are drawn into Tedlar® bags by evacuating a pressure vessel containing the bag. Spot samples of approximately 60 and 80 l volume are collected during about twenty minutes.
Five odorous sources are sampled, they cover a range of typical environmental odours.
Those are :
- a rendering plant, in the vicinity of the oven,
- a paint shop in a coachbuilding, the sample is collected either during or after the primer spray painting work of a car door, inside the workshop,

– a waste water treatment plant, near the fresh sludge aerobic treatment work,
– urban waste composting facilities, near the compost deposit area, which is under a shelter,
– and a printing house (in fact there were two different printing houses for that study).

A total of 59 samples are collected during a period of 7 months, between March and October, in various climate conditions, and sometimes at various operating conditions.

All those samples are then analysed in the lab, by an array consisting in 12 individual commercial tin oxide gas sensors (Figaro Engineering Inc.), placed in a cubic chamber. A constant power voltage is supplied to the sensor heaters. The sensor resistance is measured by a computer controlled data acquisition unit (HP 3421A). Additionally, the chamber temperature and humidity are measured. Data acquisition and real time graphic display are controlled by a computer program written in Labwindows language.
A complete measurement cycle consists in first drawing across the sensor chamber dry odourless air, bubbling into saturated salt solution (KCl in melting ice), in order to reactivate the initial semiconductor properties while keeping them at suitable humidity, and then pumping the sampled odour across the sensor array. The operator controls the signal stabilisation thanks to a real-time algorithm, based on a low-pass filter. Afterwards, data is off-line processed by two commercial sofware packages (Statistica and Matlab).

In order to test the feasibility of a field electronic nose, a "mobile detector" is also used. It is simply made of an electronic board on which 8 tin oxide sensors are soldered. The sensors are heated by a battery-powered power supply. They are simply in static contact with the ambient odorous air. The measurements are made with a portable data logger in the field after stabilisation, and several times during the test period.

But, in this case, the main problem is the influence of the air movement around the sensor, which modifies the conditions of heat convection, and thus the temperature of the sensor.

So, a portable instrument has been developed, made of 6 TGS sensors in an aluminium vessel of about 100 cm^3. A constant gas flow rate of 150 ml/min is provided by a small pump, and the system operates by a series of cycles, alternating 5 minutes of "air" (sampled in a Tedlar bag from ambient air, far from the source), and 5 minutes of odorous gas transferred directly into the sensor chamber. The whole system is powered by a 12 Volts battery.

That instrument is chiefly applied for the assessment of the odour generated by a landfill of urban waste. Two kinds of odours are perceived by the neighbouring population : either the one of the fresh refuse (esters, sulphur organic compounds, solvents, …) or the one of the biogas generated by the decomposition of the organic matter under anaerobic conditions (trace elements, such as H_2S, NH_3 and some VOC's in a mixture essentially composed of odourless compounds : methane and carbon dioxide).
During the sampling time, the operator tries to identify the odour with his nose and, later on, his feeling is compared with the sensor signals.

The last application concerns the monitoring of indoor air pollution.
The department "Environmental Monitoring" at FUL collaborates with a service of investigation of indoor pollution. They measure some pollutants in buildings, and for some of them, the only analysis method is the GC-MS at the laboratory level from samples of the ambient air. However, the sample is not always representative of the typical ambience : may be a window was just opened before the sampling, or somebody entered in the room with a perfume, or cooking vapours interfere with the ambience.
So, the service would be interested in an instrument to monitor continuously the

ambience in a room, so as to point out a warning signal when a given pollutant rises above a threshold. That signal may serve to switch on the sampling device.
For that application, we were specially aware of the ambience of solvents in the buildings (benzene, toluene, xylene, and other ones), and we have used a multisensor from the Swiss manufacturer Microsens. Those sensors are based on thin-film, metal oxide technology. They are packaged in standard TO8 metal can (12 pins) and each package contains either 4 or 6 sensors.

CLASSIFICATION RESULTS

The purpose of the first experiments made on synthetic odours is essentially to study the influence of the water content of the sample on the sensor signals and on the pattern recognition. Indeed, the odorous mixture generated by any industrial source may exhibit a water content ranging from zero to about saturation. Consequently, the semiconductor resistance variation is modified or even reversed when humidity changes [7].

Each sample is prepared in a Tedlar bag, but under uncontrolled external conditions, and thus under various humidity levels. Data is further processed by a neural network with 18 log-sigmoïd neurons and a backpropagation algorithm. The network is firstly trained with a set of signals generated by the odour at any humidity level. In such conditions, the network is able to recognise 6 new samples. New samples means : 6 validation samples, not used for the calibration procedure.
In a second step, a similar operation is performed, based on the same samples, but this time, the training is performed only with samples at low water content. Then, the model so calibrated is tested for the recognition of more humid samples, but it is no more able to recognise all the validation samples.
Thus, those experiments show that, as long as sampling and learning are carried out under many different humidity conditions, and not under particular ones for a given source, the classification remains relevant, and the humidity may be considered as a "neutral" variable.

The second set of experiments concerns real environmental odour sources.
Beyond the graphical appearance of classification results, showing nice ellipses around well separated groups, the end user would like to have at his disposal a simple function allowing him to clearly evaluate the membership to a given group, in order to make a decision.

In this spirit, non-supervised analyses, such as Principal Component Analysis (PCA), provide basically a performance evaluation of the system during the development phase, but they do not really represent learning tools aiming at the calibration of a recognition model.
Fig. 1 shows the plot of the two first principal components (factors 1 and 2, containing 93% of the total variance) for the 59 responses of the sensor array to the five sets of odours. For that analysis, the used variable is ($1-R_{nor}$), where R_{nor} is the normalised resistance of each sensor, i.e. the resistance of the sensor, divided by the quadratic mean of all sensor resistance values. The net effect of that normalisation is to reduce the dependency of the array response to the odour concentration, and also to reduce a little bit the effect of the sensor drift. That is a very classical way to pre-process the signal. But, what is less classical is the fact that the best classification is always obtained with the raw resistance, and not with the difference between the signal and the base line (i.e. : the response to the reference air). The reason is that, in this case, the "reference" is not a very pure dry air and the small variations of the concentration of trace elements in the reference air influence the classification [8].

In figure 1, four clusters are clearly observed even if some points stand out of their group. Only the "compost" and the "rendering" data sets are not well separated one from the other. The similarity of some chemical compounds (especially ammonia, aldehydes, alcohols and fatty acids), identified in those two odours by

GC-MS and specific colorimetric cartridges, can explain the similarity of the two corresponding patterns.

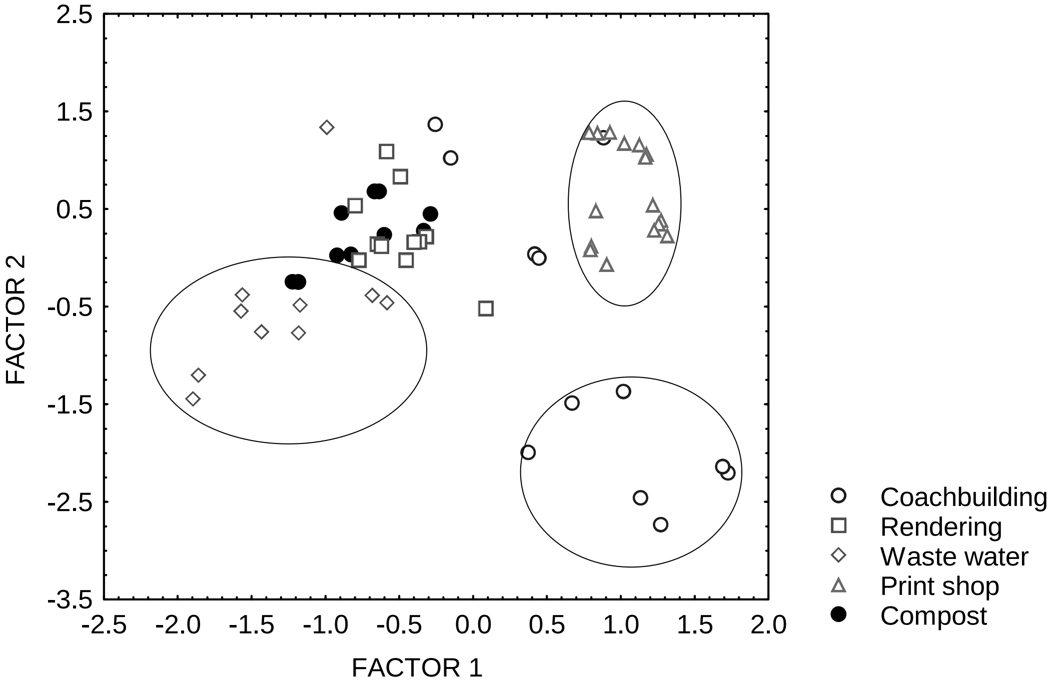

Figure 1 : Results of a PCA on 59 sensor responses (1-R_{nor}) to five odours

Moreover, the "loading plot", i.e. the plot of the 12 sensor signals again in the plane of the two first factors, brings information about the role of each sensor. Here, the left part is mostly influenced by the sensors more sensible to water vapour (TGS2180, TGS883), which confirms the localisation of the observations related to "waste water" and to "composting" in figure 1. Such results are thus particularly useful to design the instrument, to select the sensors in the array and to evaluate the performance of the whole system for the odour sources classification. However, PCA doesn't provide classification functions and some groups remain difficult to separate, such as "compost" and "rendering".

On the contrary, supervised analyses, such as Discriminant Analysis (DA), or some Neural Networks algorithms, should not be regarded as evaluation tools. The first reason is that they always give satisfactory results (in particular, a total of 50 % of cases correctly classified seems a rather good result, but it is simply normal : such result is already reached with groups randomly created). Secondly, the model that they build is too specific to the operating conditions : it is valid for a given sensor array, a given gas flow rate and a given test protocol.
Now, these fixed conditions are prevailing during the final utilisation phase of the apparatus in the field. There, the e-nose is not any more in the development phase, and it must provide a function of the type "yes or no", based on a preliminary training.
The various experiments conducted until now in the field show that the classification functions provided by the DA procedures are quite appropriate to make a reliable recognition in real time, when the system is developed. Those functions are linear combinations of signals, providing as many classification scores as identified groups. A particular case is assigned to the group for which it has the highest classification score. Their use is very simple, very convenient, and leads to unambiguous classification results.
Alternatively, artificial neural network with the Radial Basis Function (RBF) architecture may

be used to handle also some "unknown" class [9].

With such supervised techniques, the 59 observations are perfectly classified within the 5 groups, so as new validation cases, unknown to the system [8].
But for practical purpose, it is not needed to distinguish between 5 sources. A more plausible situation should be to detect the rising of an odour in the background. The figure 2 shows the result of a discriminant analysis carried out with the 59 observations in the surrounding of the sources and with an equivalent number of observations made in odourless air. Such model is able to detect the presence of any odour with respect to odourless air : that seems an obvious result, but it is sometimes sufficient for many applications [10].

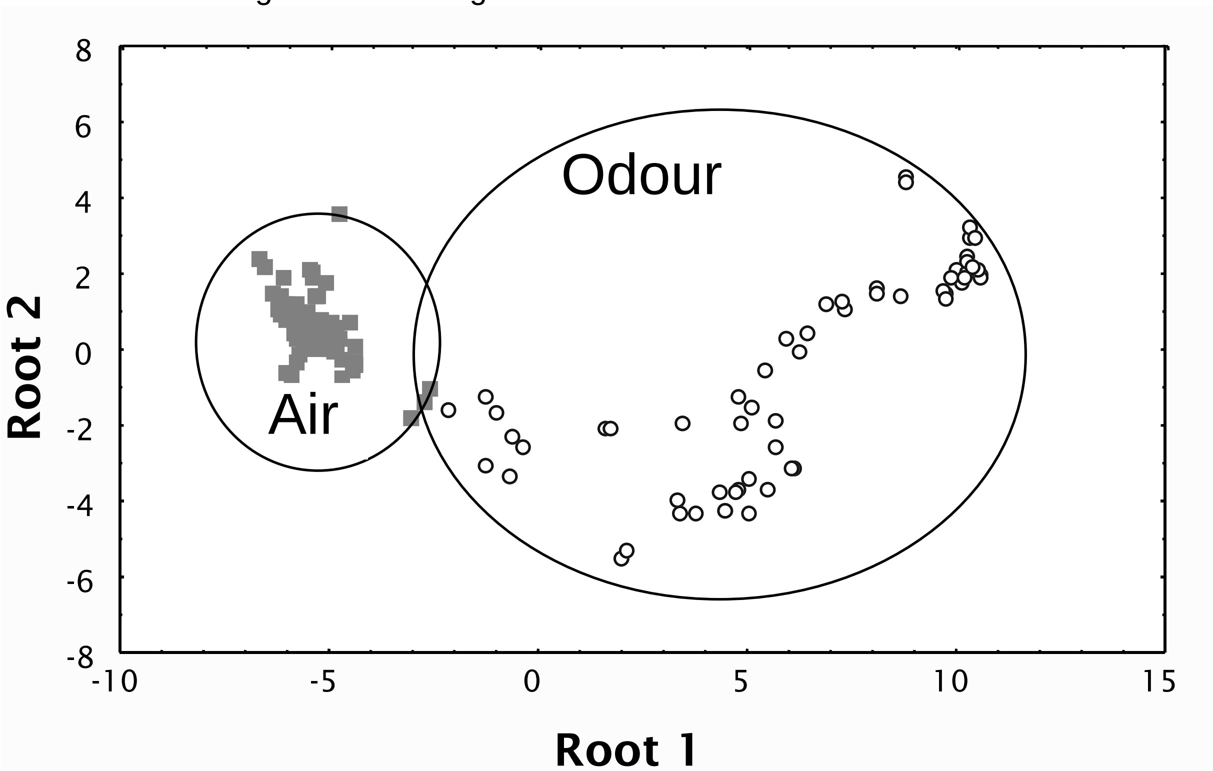

Fig. 2 : Discriminant model calibrated by DA for 59 observations in the surrounding of odorous sources and 59 observations in odourless air.

Moreover, such result can already be obtained with more simple device. Even with the "mobile detector", simply made of 8 tin oxide sensors soldered on an electronic board, it is possible to recognise the 5 sources with a relatively good accuracy .

The "portable instrument", with a dynamic gas flux, gives rise to the same type of results for well differentiated sources of odour.
It has been also used around a landfill site, characterised by more mixed odours generated either by the fresh refuse or by the biogas, or by other sources, such as truck exhaust gas. The instrument is moved at different locations on the landfill area, either in the vicinity of the fresh refuse, sometimes when the trucks pour out the refuse, sometimes when the waste is at rest, or at various distances from a landfill gas extraction well. During the sampling time, the technician tries to identify the odour with his nose. He has recognised the biogas odour 72 times and the fresh refuse odour 69 times (these observations correspond to the same number of measurements with the electronic nose). The 141 groups of responses to the 6

sensors constitute the input data set of a Principal Component Analysis.

Figure 3 shows the plot of the 141 scores in the plane of the two first principal components. Despite an obvious trend, bringing together the "biogas" points to the left part of the diagram and the "fresh refuse" ones to the right, the plot doesn't show a clear separation between the two groups. Such result is however quite normal and corresponds to the reality of the odorous perception : both odours are generally present at the same time on the site.

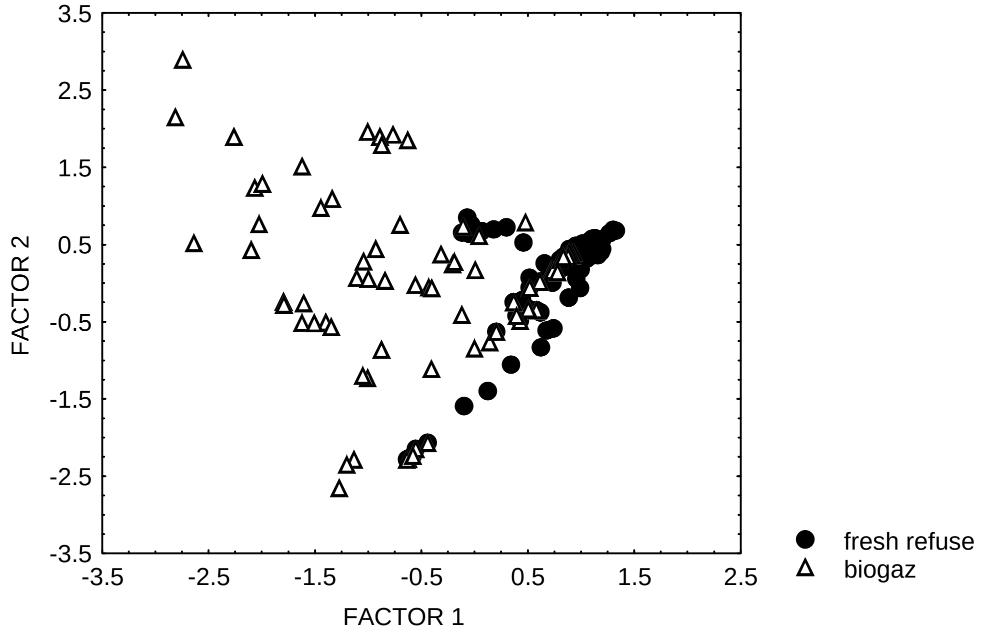

Fig. 3 : PCA on 141 observations around a landfill area - Plot of scores in the plane of the two first factors (94.7 % of the total variance)

Nevertheless, such results are sometimes sufficient for the final user, e.g., the manager of the landfill area. For him, a first type of information should be a warning signal when the odour level increases beyond a given threshold (figure 2), then, when he is warned of the presence of an odour, he may want to identify its origin : so, PCA or DA results, such as those of figure 3, even if they show only general trends may be used as decision making support.

However, what is essential for him is to identify the right odour in the background, without any confusion with odourless air or with other odorous sources. The research aims thus at a good separation between the observations related to a specific source and the other ones.

Figure 4 shows for example the results of a DA in the plane of the two first roots for 388 observations made with the mobile detector around different odorous sources in the environment, and particularly in the vicinity of the urban waste composting area. It distinguishes rather clearly the compost from the air and also from "other" odours. "Air" means either pure odourless air, or ambient air, far from any odorous source, and "other", in this case, means essentially solvent odours.

Such findings are promising for further development of odour detector operating in real time in the field.

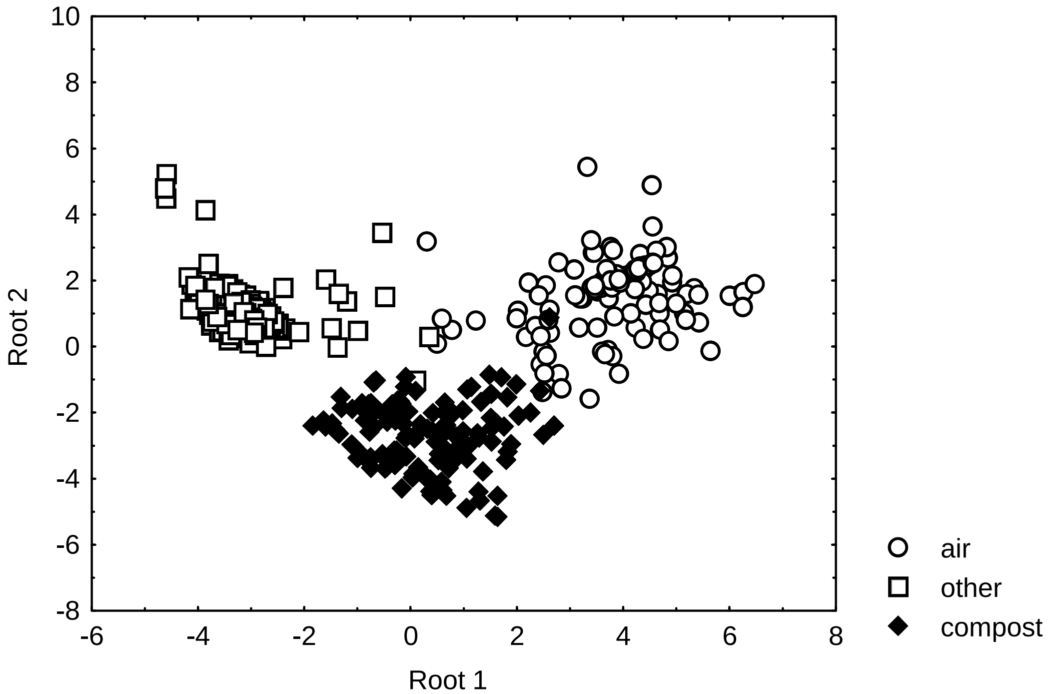

Figure 4 : Observations in the plane of the two first roots of a DA for different sources in the environment.

MONITORING THE ODOUR INTENSITY

As long as only the classification of odour sources is concerned, the models calibrated during the learning phase, either by statistical methods (PCA, DA, …), or by Artificial Neural Networks, could be included in the mobile detector for the on-line identification of unknown odours. When the identification is achieved, the estimation of the odour intensity by the system can be done by appropriate techniques. Whatever the final application of the monitoring may be -either assessing the nuisance or controlling an abatement system- the useful signal should be related to the monitored odour, and shouldn't be influenced by disturbing odours or ambient parameters.
For more than ten years, several techniques are proposed in the literature [11]. In short, they use either the signal of one sensor element, or a weighted signal of all elements, obtained by a suitable regression technique. Are those proposed techniques suited for the on-line monitoring of environmental odours?

Field applications

Some techniques have been tried out in the field, either for the assessment of the odour intensity around a landfill area, or for the monitoring of odour "events" in the vicinity of settling ponds in a sugar factory, or to track a specific malodour around a given source in the environment.

– On the landfill site, at each measurement location, the operator notes his feeling of odour intensity on a 4 level scale. Among the 141 observations in total, 69 refer to "fresh refuse", including 21 observations with intensity 0 and 72 refer to "biogas", including 24 zero-intensity observations.
  Together with the intensity assessment, the 6 sensor signals of the portable instrument are recorded in a data logger and then off-line processed by Statistica or Matlab procedures.

– The odour generated by the settling pond of the sugar factory is investigated with a 4-sensor array in a continuous way during 3 months (Spring 1998). The observations made by the factory staff may be used to validate the detected odorous events.

- Tracking a specific odour is possible with the "mobile detector" (8 sensors soldered on an electronic board), which is moved in various spots around a given source (print shop, waste water treatment plant or composting area).
  The measurements are made with a portable data logger in the field after 30 minutes of stabilisation and several times during the test period.

With one sensor signal

Using one of the sensor elements, preferably that with the highest sensitivity towards the identified substance, is a rather easy solution. On the landfill area, we have chosen to monitor the signal of the Figaro sensor TGS2610 (more sensitive to a wide variety of combustible gas). The Kendall correlation coefficient between the TGS2610 resistance and the odour intensity class, as assessed by the field operator, is -0.87 and shows a rather strong relationship between the two variables. The model obtained by linear regression allows to predict the intensity class from the recorded sensor signal. The measured intensity class is correctly predicted in 65 % of the 141 observations made on the landfill area.
However, the TGS2610 sensor is sensitive to both sources (fresh waste and biogas), and probably also to many other ones, so, its signal cannot be used to detect the rise of a particular odour among other ones. The procedure will thus always include two steps : a first identification of the odour by a classification technique, followed by the monitoring of the intensity of the global odour.
But the sensor varies also with ambient air temperature and humidity, and that is more awkward, since that kind of variation is unpredictable.
Such solution is thus applicable for "pure" substance, in the laboratory, but is not really suited to field applications, where background odours and variable ambient conditions affect the sensor response.

With an "expert system"

In the absence of other odours in the field, the sole influence of ambient temperature and humidity could be taken into account (if they are measured) by a kind of "expert system". For example, the following rule can be applied :

> *If (the sensor resistance drops) AND (the humidity is stable or decreases) AND (the temperature is stable or decreases) THEN (there is probably an "odour event") OTHERWISE (any conclusion can't be drawn from the sensor resistance variation).*

In this case, the system is able to detect the emergence of the odour only when the temperature and the humidity don't influence the sensor response. Such rule has been experimented in Spring 1998 around the settling pond of a sugar factory and pointed out some "odour events", validated by the observation [8]. Figure 5 shows the evolution of a function made from such rule. An empirical algorithm, based on the linear combination of the 4 sensor signals, corrected by the values of the temperature and the humidity of the ambient air, is able to pinpoint some “odour events” during the day. One of them, around 9 o’clock a.m. (identified by the arrow), is validated by the observation : it corresponds to the daily partial emptying of the pond, involving the stirring of the water and the release of malodours. However, other such “odour events” couldn’t be further validated by the comparison with the human nose. Future works must go in that direction.

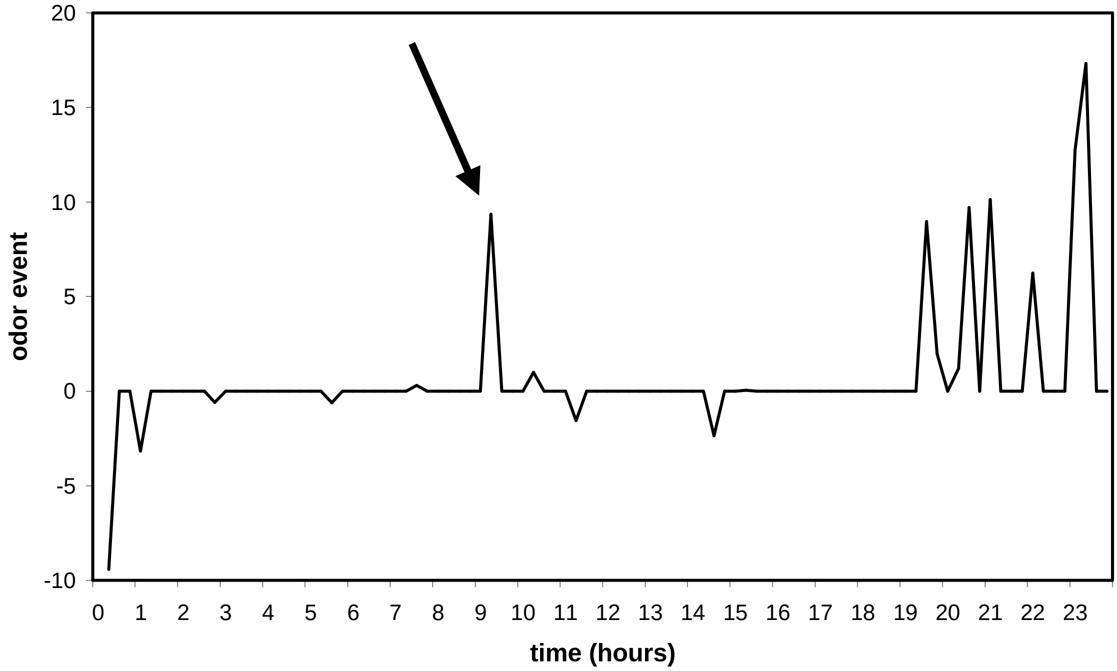

Figure 5 : Odour event pointed out around the settling pond of a sugar factory

With the classification functions of the discriminant analysis

Despite the good results obtained with only one sensor, using a weighted signal of all elements is more accurate and more "selective" to the monitored odour, i.e. minimises the cross-sensitivity.

If the classification model was previously calibrated by a supervised method, such as Discriminant Analysis (DA), the classification functions supplied as a standard result of such method can be used as "odour signal". Though they show very bad correlation with the intensity measured in the field, their selectivity to a given odour can be exploited for the continuous monitoring. Figure 6 shows such result obtained for 5 sources (print shop, coachbuilding, compost, waste water and background air, far from any odorous source) and with the mobile detector. Five classification functions are computed by the DA method, they are linear combinations of the 8 sensor signals and they are usually used to assign a new case into one of the known groups. Here, they are applied for the monitoring of the odour in the field [12].

During the validation phase, the same mobile detector is moved in various spots around a given source and, at each sampling time of the data logger, the data from the sensor signals are inserted into each previously calibrated classification function to develop a classification score for each group. The figure shows the graphical evolution of the 5 calibrated classification functions when the detector is moved around in the print shop. The scaling of the horizontal time axis is unessential : it shows only that the mobile detector is continuously moved away from the source.

The classification function corresponding to the printing shop has effectively the highest value when the detector is in the printing shop, but it suddenly drops when the detector is moved outside the shop, and increases again when the technician moves again inside. As expected, the classification function characterising the outside atmosphere is the “background” one.

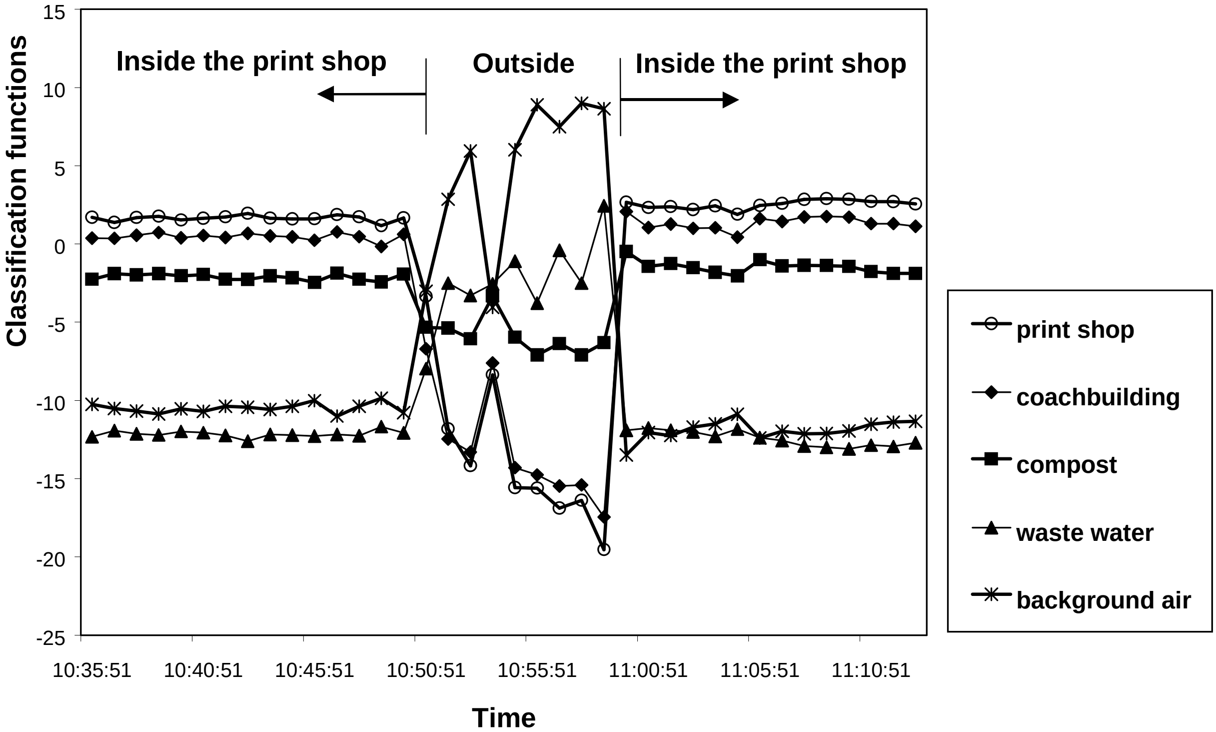

Figure 6 : Evolution of the DA classification functions, resulting from the learning phase with five sources, when the mobile detector is moved around in the print shop.

With regression methods

Different regression techniques may also be applied to predict the intensity of the odour. Multilinear Regression (MLR) on the original measured sensor signals (resistances) provides a rather good model. In the case of the odour generated by the landfill area, it is able to predict an intensity value in agreement with the measured one in 67 % of the cases. The resulting model, however, is a pure mathematical construction, which is convenient to predict intensity values inside the training sample, but which is less adapted to the prediction of new data.

Using the results of an unsupervised classification method, such as the factors supplied by a Principal Component Analysis (PCA), has a good chance to produce a more physical model, making more "sense" from a physical standpoint [13]. Unlike quantitation methods attempting to calculate the model coefficients from a direct regression of the odour intensity onto the sensor responses, the PCR method regresses the intensities on the PCA scores.

Still for the landfill, the principal component regression (PCR) includes in the model the first principal component, which is well correlated with the odour intensity, and the second one, which separates "biogas" from "fresh refuse". Including the third one in the regression provides a model which predicts the measured intensity in 69 % of the cases. Of course, the model converges towards the MLR one when the 3 remaining principal components are added. As that MLR model is worse than the model based on 3 principal components, it seems that some of the initial variables were not relevant for the prediction of the odour intensity.

PCR gives also excellent results for the prediction of the benzene concentration levels inside the buildings. The concentration measured by gas chromatography is compared with the signals of 6 elements of a Microsens multisensor. The MLR model, calibrated on the

6 sensor signals, explains 90 % of the variability in the benzene concentration. For chemometrics users, that MLR method is sometimes known as Inverse Least Square (ILS) method and is widely used for spectroscopic quantitation, as it tries to predict the concentration level from the spectral responses (or here, from the sensor signals), and not the opposite.
But, again, the PCR model, including 5 of the 6 principal components computed from the sensor signals, gives better prediction performances, since it exhibits a R-squared of 95.5 %. Those are still preliminary results, which are only based on 11 measurements in different buildings.
But that research is now going on, and it seems to be a very promising application of the electronic nose.

With Partial Least Square model

Finally, Partial Least Squares regression (PLS) captures the greatest amount of variance, like PCA, and also achieves correlation with the predictor variable (here, the odour intensity around the landfill site), like MLR. Instead of first decomposing the matrix of sensor responses into a set of eigenvectors and scores, and regressing them against the intensity as a separate step, PLS actually uses the intensity information during the decomposition process. This causes signals containing higher intensity information to be weighted more heavily than those corresponding to low intensity. Thus, the eigenvectors and scores calculated using PLS are quite different from those of PCR. The main idea of PLS is to get as much intensity information as possible into the first few loading vectors.
By combining the advantages of several other chemometrics methods, PLS should probably provide the most adapted model for the intensity prediction.

Indeed, testing PLS regression on the 141 observations on the landfill shows that the model provides 71 % of intensity prevision in agreement with the measured one. That is a very good result, knowing that, in the 29 % left, it remains probably a lot of errors of estimation due to the operator in the field. Moreover, like PCA, the PLS provides the classification of observations in two groups. Consequently, it should be used as sole tool, both to identify the source of the odour and to predicting its intensity.
As an example of such advantage, figure 7 shows the odour intensity on the landfill site as predicted by the PLS model as a function of the two first "latent variables" (the equivalent of the "factors" of the factorial analysis).

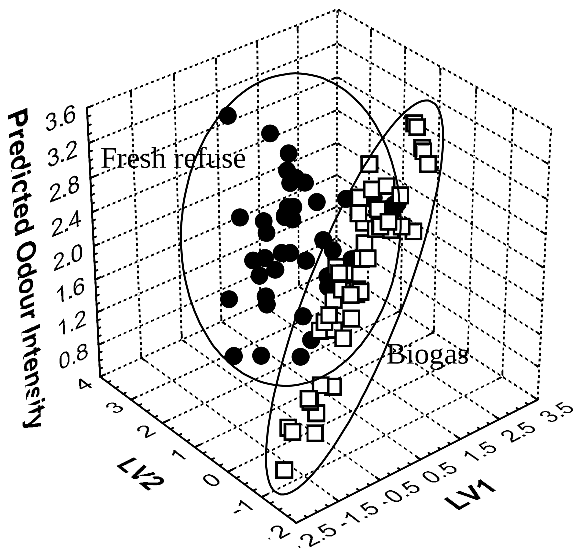

Figure 7 : Predicted odour intensity around a landfill site from a PLS model

CONCLUSIONS

The use of laboratory-based sensor array systems and of field electronic noses for measuring environmental odours have demonstrated that their responses can be correlated to the assessment of odour annoyance. However, research is now needed to translate these experiences into the assessment of environmental odours under variable conditions [14].

To become a reality, the routine use of electronic nose in the environment has first to overcome some difficulties, such as :

- understanding and controlling the influence of ambient parameters (temperature, humidity),
- detecting the very low concentration levels of odorous compounds in the atmosphere,
- identifying main environmental odours by typical signatures, in various operating conditions and non-constant odour intensity,
- taking into account the sensor drift and the non-linearity of the sensor response, which are general problems, but which are more pronounced for the monitoring of the highly variable odours in the environment

Such objectives can be achieved only by the improvement of sensors sensitivity and noise reduction.

But, if preliminary results are very encouraging and entitle us to be trustful for future developments, that is essentially due to the happy marriage between the sensor technology and the chemometrics.

One in vain would improve in a spectacular way the quality of the sensors, the identification and the continuous monitoring of patterns typical of environmental odours necessarily passes by the judicious choice of chemometrics techniques.

REFERENCES

[1] Gardner, J.W., and Bartlett, P.N. 1999, *Electronic Noses - Principles and applications,* Oxford University Press, Oxford, 245.

[2] Stuetz, R. M., Fenner, R. A. and Engin, G. 1999, *Assessment of odours from sewage treatment works by an electronic nose, H2S analysis and olfactometry*, Wat. Res., 33, 453-461.

[3] Rella, R. and Siciliano, P. 1999, *Air quality monitoring by beans of sol-gel integrated tin oxide thin films*, Sensors & Actuators B, 58, 283-288

[4] Carotta, M.C., Martinelli, G., Crema, L., Gallana, M., Merli, M., Ghiotti, G. and Traversa, E. 2000, *Array of thick film sensors for atmospheric pollutant monitoring*, Sensors & Actuators B, 68, 1-8

[5] Menzel, R. and Goschnick, J. 2000, *Gradient gas sensor microarrays for on-line process control - a new dynamic classification model for fast and reliable air quality assessment*, Sensors & Actuators B, 68, 115-122

[6] Becker, Th., Mühlberger, S., Bosch-von Braunmühl, C. Müller, G., Ziemann, T. and Hechtenbergla, K.V. 2000, *Air pollution monitoring using tin-oxide-based microreactor systems,* Sensors & Actuators B, 69, 108-119

[7] Romain, A.C., Nicolas, J. and André, Ph. 1997, *In situ measurement of olfactive pollution with inorganic semiconductors : limitations due to the influence of humidity and temperature,* Seminars in Food Analysis, 2, 283-296

[8] Romain, A.C., Nicolas, J., Wiertz, V., Maternova, J. and André, Ph. 2000, *Use of a simple tin oxide sensor array to identify five malodours collected in the field*, Sensors & actuators B, 62, 73-79

[9] Horner G. 1999, *Qualitative and Quantitative Evaluation Methods for Sensor arrays,* Proceedings of ISOEN99, Tübingen, September 20-22, 171-174

[10] Nicolas, J., Romain, A.C., Wiertz, V., Maternova, J. and André, Ph. 1999, *Utilisation*

élémentaire d'un réseau de capteurs "SnO_2" pour la reconnaissance d'odeurs environnementales, Proceedings Eurodeur 99, Paris, 15-18 Juin 1999

[11] Göpel, W., Hesse, J. and Zemel, J.N. (editors) 1989, *Sensors, A comprehensive survey (Volume I)*, VCH Verlagsgesellschaft mbH, Weinheim, 641

[12] Nicolas, J., Romain, A.C., Wiertz, V., Maternova, J. and André, Ph. 2000, *Using the classification model of an electronic nose to assign unknown malodours to environmental sources and to monitor them continuously,* Sensors & actuators B, 69, 366-371

[13] Wise, B.M. and Gallagher, N.B. 1998, *PLS_Toolbox 2.0 for use with MATLAB^{TM},* Eigen Research, Inc., Manson, 320.

[14] Stuetz, R.M. and Nicolas, J. 2001, *Sensor arrays : an inspired idea or an objective measurement of environmental odours ?,* 1st IWA International Conference on Odour and VOCs, Sydney, 25-28 March