Quantitative gas detection with semiconductor micro-sensors and chemometrics

P. Breuil, N. Perdreau and C. Pijolat

École Nationale Supérieure des Mines, Centre SPIN, 158, cours Fauriel, 42023 Saint-Étienne Cedex 2, France

The use of the chemometric techniques to exploit the selectivity of gas sensors, although current in the case of multisensors systems, is tested here with only one gas sensor. The method consists in measuring the electric conductance of the sensitive element, a tin dioxide bar, at various temperatures. Although these gas sensors follow non linear laws, the modelization algorithm Partial Least Square (PLS) is preferred to neural networks. A comparison of the two algorithms is made, as well as a study of the portability of the model and stability of the prediction. Lastly, the concept is tested on an autonomous prototype in real conditions.

Introduction

Tin dioxide (SnO_2) is a N type semiconductor. The adsorption of gases on its surface can involve great variations of electric conductivity. Thanks to this significant sensitivity, this material is very used in the realization of gas sensors. Currently, a lot of applications use this type of components [1]. The principal disadvantage of tin dioxide is its low selectivity. In order to improve it, many studies have been done: doping, addition of metals, thin layers or thick layers technologies… [2,3] Generally, these modifications involve only a weak improvement of the selectivity.

Another solution, very studied too, is the chemometric approach. In addition to the development of micro sensors, the originality of our study lies in the exploitation of the information of only one sensor cycled in temperature whereas the current tendency is the multisensors system [4] which one of the most known applications is the "electronic nose".

The aim of our study is the quantitative detection of a binary or ternary mixture of ethanol vapour (0-80 ppm), carbon monoxide (0-300 ppm) and methane (0-1000 ppm) in the air. These gases and their range of concentration are representative of the majority of the expected applications: so, in domestic use, these three gases may coexist.

Sensors and selectivity

The sensitive element that we developed is a sintered tin dioxide bar (2.5 × 0.5 × 0.5 mm) [3]. It is stuck on an alumina substrate (Fig. 1) with an integrated heating system (20 to 500 °C). The variations of electric conductivity are due partly to a modification of the point defects of the semiconductor in the presence of oxidizing or reducing gases. But they are also related to the catalytic reactions of surface, in particular in the presence of organic gases or vapours [2,5].

The curves conductance-temperature G = *f*(T) of some sensors may be characteristic of the gases present in air (Fig. 2). It is then necessary that these data be acquired in decreasing temperature in order to avoid memory effects.

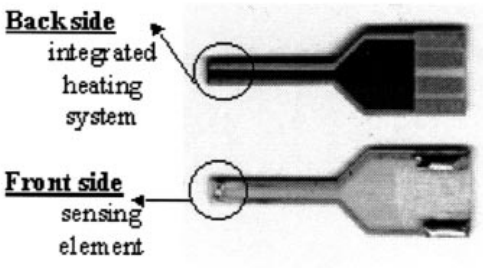


Figure 1. Substrate + sensitive element.

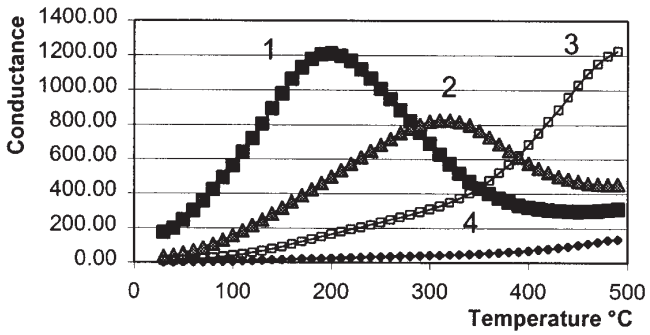


Figure 2. Conductance in decreasing temperature: (100 °C / ') (1): Ethanol (80 vpm), (2): CO (300 vpm), (3): CH4 (1000 vpm), (4): air.

Article available at http://analusis.edpsciences.org or http://dx.doi.org/10.1051/analusis:2000140

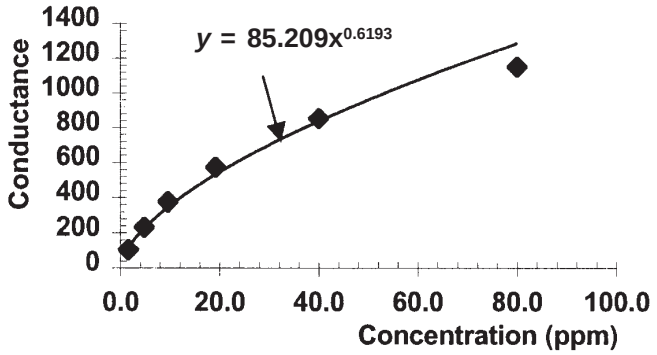


Figure 3. Conductance for different ethanol concentrations at 220 °C.

The relation conductance-concentration G = *f*(C) for each gas is generally non-linear (Fig. 3), and moreover non additive if there is a mixture of several gases.

Methods

The application of a model based on physicochemical knowledge of the phenomena is often limited by many assumptions. This is why we preferred to use the chemometrics techniques for the modelization of the sensor and the prediction of the concentrations.

In this paper, we will call "sample" a complete experiment of measurement in decreasing temperature, *i.e.* a set of 27 conductance variables and 3 concentration variables. The models are carried out from the calibration data (conductances (X variables) at 27 different temperatures, from 500 to 200 °C, concentrations (Y variables) of the 3 gases). These data (about 40 samples) are acquired with the help of an automated test bench which can generate known concentrations of several gases or vapor mixtures (ethanol 0-80 ppm, carbon monoxide 0-300 ppm, methane 0-1000 ppm and humidity 0-90 %). The test bench is also used to control the temperature of four sensors and to measure their conductance.

The algorithms which are used for the modelization are:

– PLS (Partial Least Square) [6,7]: in fact "PLS2", which makes the same factor decomposition for all the Y variables. The soft was Unscrambler 6.1 (Camo).

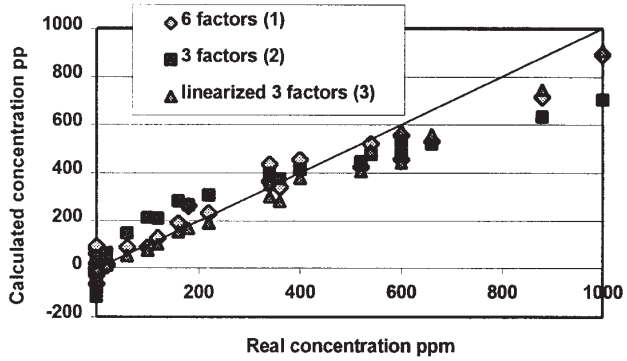
– Neural Networks (NN): We used one intermediate layer, the learning algorithm was backpropagation , the soft was Matlab with a modified "Neural Network" toolbox.

As modelization are carried out using a relatively low number of samples (about 40), a very close attention was paid to the problem of the overmodelization. For PLS, the solution is classical, it consists in validating the choice of the number of factors (latent independent variables) with a second set of samples, known as "validation set". With neural networks the solution consists in testing, during the training and after each iteration, the network on the "validation set" and to optimise the iteration count thus.

The goal of this study being to compare the performances of prediction between the different methods and also between various gases, we use as "performance criterion", the relative RMSEP (Root Mean Square Error of Prediction = square root of the mean of the errors of prediction) normalized in percentage of the range of concentration of the gas, and we call it RRMSEP %. The gas concentrations of the mixtures of the three gases are selected randomly with a uniform distribution on all the ranges.

Justification of the choice of PLS

The neural networks are generally preferred with methods such as PLS to modelize non linear phenomena. However PLS, under certain conditions, may rigorously modelize non linear laws, in particular polynomials, so it will be able to modelize any few non linear laws approximately. It can thus modelize in a satisfactory way conductivity of our sensors.



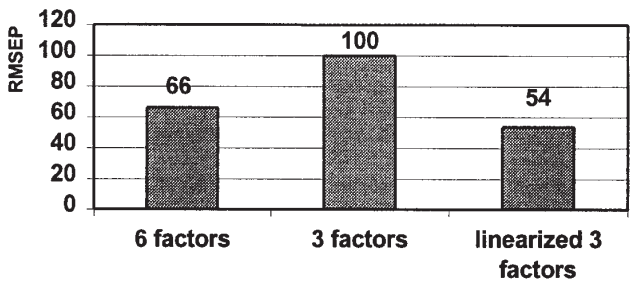


Figure 4. Modelization of non-linearities (Prediction of the concentration of methane in ternary mixture).

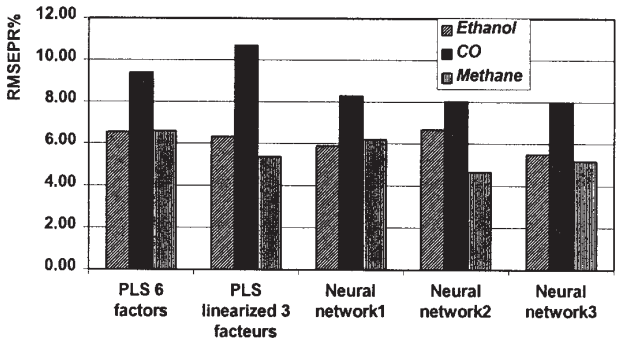


Figure 5. RRMSEP % for different methods of modelization.

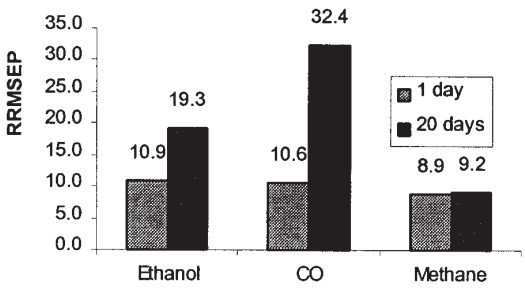


Figure 6. RRMSEP % after 1 day and 20 days of operation.

But this modelization can be done only with a number of factors more significant than for the linear cases: thus, if there are 3 components, a model with 3 factors is not precise enough: in figure 4 (curve 1), the distribution of the points around the bisecting line shows a problem of non linearity, and 6 factors are then required in order to have an optimal prediction (curve 2). Then the model is less robust, *i.e.* it is more sensitive to the measurement fluctuations. But we know that the curves conductance-concentration can be approximated by a power type law (Fig. 3) where exponent *n* lies between 0.25 (methane) and 0.6 (ethanol and carbon monoxide). It can thus also be planned to carry out a preprocessing of linearization on the concentrations. We will thus do the calibration and the prediction not with C but with C^n and will obtain a good model for 3 components with only 3 factors (Fig. 4, curve 3).

With the neural networks, we encountered an obstructing problem of bad repeatability: the initialization of the neurons being made more or less randomly, the same calibration data set will lead to different networks and so will not have exactly the same performances (Fig. 5). The results of prediction of the various methods on the same set of samples are summarised in figure 5.

It will be noticed that the performances are quite identical. The neural networks (with a 3 neurons intermediate layer) seem slightly more powerful but suffer from bad repeatability. So they cannot be easily retained within the scope of a comparative study. Moreover, PLS has an advantage, which we wish to use. This algorithm allows, by examination of the conductance residuals, to invalidate the prediction if there happen events, which did not occur during the calibration phase. The results presented in the following paragraphs were carried out with PLS, but neural networks would give quite the same.

Stability tests

Results of prediction (Fig. 6) with mixtures show that stability after 20 days of operation is good for methane, average for ethanol and bad for carbon monoxide. Let us note here that the prediction with carbon monoxide is generally less powerful, undoubtedly because it has an intermediate characteristic curve between the two other gases.

Portability tests

Within the aim of an industrial application, it is not very interesting to make a calibration for each sensor. The study of portability consists in checking if a model build from one or more sensors can be used on a whole batch of other sensors. Here, starting from 3 sensors A, B and C, we carried out 3 experiments of prediction on the sensor A (Fig. 7).

The model built with the sensor used for the prediction (1: A/A) gives obviously the best results. But a "portable" model must be able to be used on various sensors. Under these conditions, the model carried out starting from 2 sensors (3: B+C/A) is more powerful than the one carried out from only one (2: B/A). In fact we integrate into a model with several sensors only common information and we thus avoid the overmodelization which consists in modelling the informations characteristic to each sensor.

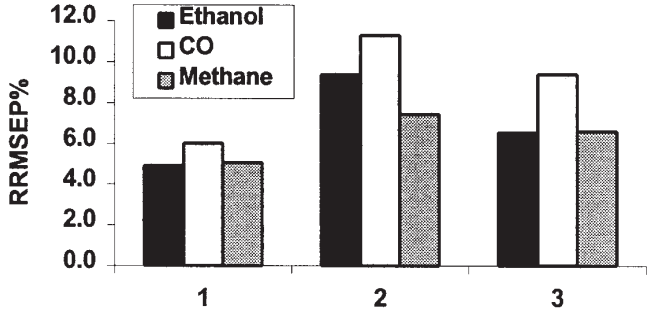


Figure 7. (1): Calibration with A; (2): Calibration with B; (3): Calibration with B and C.

Tests on prototype

In order to validate the experiments undertaken on the test bench, we developed autonomous demonstration equipment controlled by a PIC (Microchip) type microcontroller. We used this prototype with a concrete case: ethanol and methane detection in a room. The model used is a model with 2 components and 6 conductance variables at 6 different temperatures, it was calculated using PLS and the concentrations were linearized. The water vapour being a major interfering, various relative humidities were generated during the acquisition of the calibration samples.

We can see in figure 8 the good correlation between the concentrations calculated by the microcontroller and the concentrations measured with traditional analyzers (Photo Ionization Detector for ethanol and catalytic bead sensor for methane). We could check that the variations of relative humidity (20 to 80 %) which has a strong influence on the conductance, had a weak influence on the calculated concentrations.

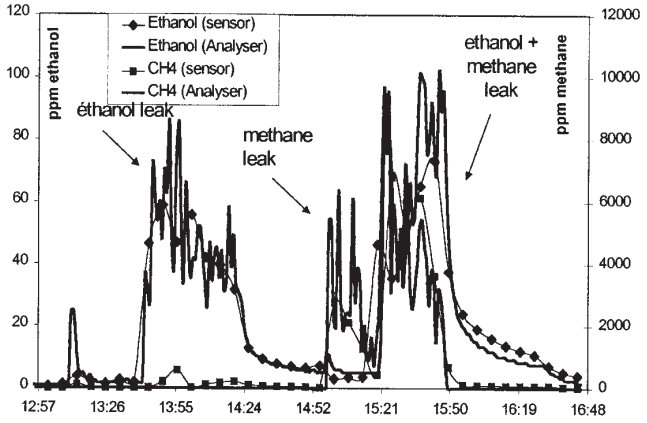


Figure 8. Prototype comparison – analyzer with gas leakages during 4 hours.

Conclusion

The possibility of using a gas sensor cycled in temperature associated with multivariate analysis to predict the concentrations of a mixture of 2 or 3 gases was shown. Although the neural networks can be slightly more powerful, it appears that within the framework of our study, PLS remains an interesting algorithm by its facility of use, its repeatability and the facility with which one can implement it in a microcontroller software.

In fact, when they are suitably used, the most powerful methods of multivariate analysis give close results. This shows that if we want to improve the system we will have to try to improve rather the sensor or the measurement procedure. The bad stability of the results, in particular for carbon monoxide, remains the principal problem and it could be interesting to try to modelize the conductance drifts.

References

1. Perdreau, N.; Pijolat, C.; Breuil P., *3rd IFAC symposium SICICA' 97,* Annecy, 1997, pp 411-416.
2. Breuil, P. Thesis, École des Mines de Saint-Étienne, France, Nov. 1989.
3. Sauvan, M. École des Mines de Saint-Étienne, France, 1999.
4. Göpel, W. *Sensors and Actuators B* **1998**, *52*, 125-142.
5. Heiland, G.; Kohl, D. in *Chemical sensors technology, Physical and chemical aspects of oxidic semiconductor gas sensors*; Lausanne: T Seiyama–Kodansha Ltd, Elsevier, 1988, pp15-38.
6. Martens, H.; Naes, T. *Multivariate Calibration;* Chichester: John Wiley and Sons, 1989.
7. Camo Anonymous Society, *The Unscrambler ® 6.0-6.1 User' S Guides;* Camo HAVE, Norway, 1996.