Available online at www.sciencedirect.com
SCIENCE DIRECT®
Sensors and Actuators B 103 (2004) 122–128
SENSORS AND ACTUATORS B CHEMICAL
www.elsevier.com/locate/snb

Analysis of neural networks and analysis of feature selection with genetic algorithm to discriminate among pollutant gas

M. Aleixandre*, I. Sayago, M.C. Horrillo, M.J. Fernández, L. Arés, M. García, J.P. Santos, J. Gutiérrez

Laboratorio de Sensores, Instituto de Física Aplicada (IFA), Consejo Superior de Investigaciones Científicas (CSIC), Serrano 144, E-28006 Madrid, Spain

Available online 1 June 2004

Abstract

A multisensor based on tin and tin–titanium oxides has been utilised to detect pollutant gases (NO_2, CO, toluene and octane). The sensitive layers are deposited by r.f. reactive sputtering. Some tin oxide sensors are doped with Pt. Measurements are carried out with single gases and gas mixtures (two and three gases) in dry air at 250 °C.

An exhaustive analysis of several networks and feature extraction and selection is done to discriminate among four different pollutant gases. First the sensor responses are analysed with principal component analysis (PCA). The results are not good enough for mixtures. Then several pre-processing techniques and several artificial neural networks (ANN) are studied. Two models of neuronal networks are used: probabilistic neural network (PNN) and multilayer perceptrons (MLP). A selection of the sensors and of pre-processing techniques was made with a genetic algorithm (GA).

Keywords: Multisensor; Air pollutants; Neural networks; Genetic algorithm

1. Introduction

Atmospheric pollution may be defined as the presence in the atmosphere of harming substances to the man or the environment. Those pollutant can be the result of man-made activities or natural process. Typical atmospheric pollutants from man-made activities include nitrogen oxides, carbon monoxide, sulfur dioxide and hydrocarbons. All these pollutants are called primary pollutants because they are released directly into the atmosphere [1].

The control and monitoring of pollutants for ambient air quality is at the present limited by the techniques currently approved by the existing standards. Those techniques need the use of costly analytical equipment. Continuous monitoring of pollutant in air is expensive, so high-density networks of measurements are usually impracticable.

A potentially more cost-effective way of performing such measurements is using metal-oxide gas-sensor arrays. Sensor array systems use the global information formed by the responses of all sensors to discriminate among gases. The signals of the multisensor can be evaluated by techniques of principal component analysis (PCA) and artificial neural networks (ANN) with a selection of the possible features through a genetic algorithm that facilitates the classification task [2].

Four gases have been selected as pollutant gases: NO_2, CO, toluene and octane. CO and NO_2 are two of the main toxic gases resulting in the polluting processes in the cities. Toluene and octane are volatile organic compounds (VOCs). VOCs are important polluting agents because they can: (1) react with nitrogen oxides in the presence of sunlight to form ozone and photochemical smog; and (2) be toxic to humans, animals or vegetation.

In this work, mixtures of selected gases have been detected and a feature extraction followed by an analysis of the several artificial neural networks have been used for their discrimination. The impact of the different features are described. The performance and training speed of a back propagation neural network and a radial basis neural network has been compared. Also a genetic algorithm has been applied to select the features.

* Corresponding author.
E-mail address: manuel@ifa.cetef.csic.es (M. Aleixandre).
URL: http://www.ifa.csic.es.

0925-4005/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2004.04.044

2. Experimental

The multisensor consisted of eight sensing elements deposited by r.f. reactive sputtering from SnO_2 and TiO_2 targets. Four tin oxide sensors were doped with different amounts of Pt. The deposition conditions, dopants and thermal treatment were described in previous works [3]. The characteristics of each sensor are listed in Table 1, doping levels are expressed as sputtering times (seconds).

The multisensor device was placed in a stainless steel test chamber and it was characterised by dc electrical measurements.

A 200 ml/s flow of dry air was used to achieve a line base. Then an automatic system controlled via PC switched the flow of gas with pollutants for 10 min. Afterwards dry air was fluxed until achieving a base line stable again. The resistance of the sensors is recorded during all the experiment every 60 s.

Measurements were carried out with single gases and gas mixtures (two and three gases) in dry air at 250 °C.

The concentrations go from 50 to 200 ppm for reducing gases and from 0.5 to 2 ppm for NO_2. Every experiment is repeated three times. The number of experiments carried out for binary and ternary mixtures are 72 and 96 respectively.

Table 1
Multisensor distribution

Sensor	Semiconductor material
S1	SnO_2 (300 nm)
S2	SnO_2 (400 nm)
S3	SnO_2 (150 nm)–Pt (8 s)–SnO_2 (150 nm)
S4	SnO_2 (150 nm)–Pt (16 s)–SnO_2 (150 nm)
S5	SnO_2 (200 nm)–Pt (8 s)–SnO_2 (200 nm)
S6	SnO_2 (200 nm)–Pt (16 s)–SnO_2 (200 nm)
S7	SnO_2 (300 nm)–TiO_2 (150 nm)
S8	SnO_2 (400 nm)–TiO_2 (150 nm)

3. Data analysis

Feature extraction methods, pattern recognition algorithms and the genetic algorithm were implemented using Matlab [4].

3.1. Principal component analysis

Principal component analysis is a linear method to reduce the dimensionality of the data. Once this analysis is done we can plot all data in only two axes.

The PCA results for individual gases are shown in Fig. 1. As can be observed the cluster for NO_2 is well separated from the rest. Nevertheless, for reductor gases (CO, toluene and octane) the separation is not so good and in some cases they are overlapped.

For binary and ternary mixtures the PCA plot shows a great overlapping among the measurements with and without NO_2 (Fig. 2). The overlapping becomes worse if we plot the other gases.

3.2. Feature extraction

In many cases the sensor response can also be improved considering the transient states (drift, abrupt changes). These states can be a very important part of the sensor signal

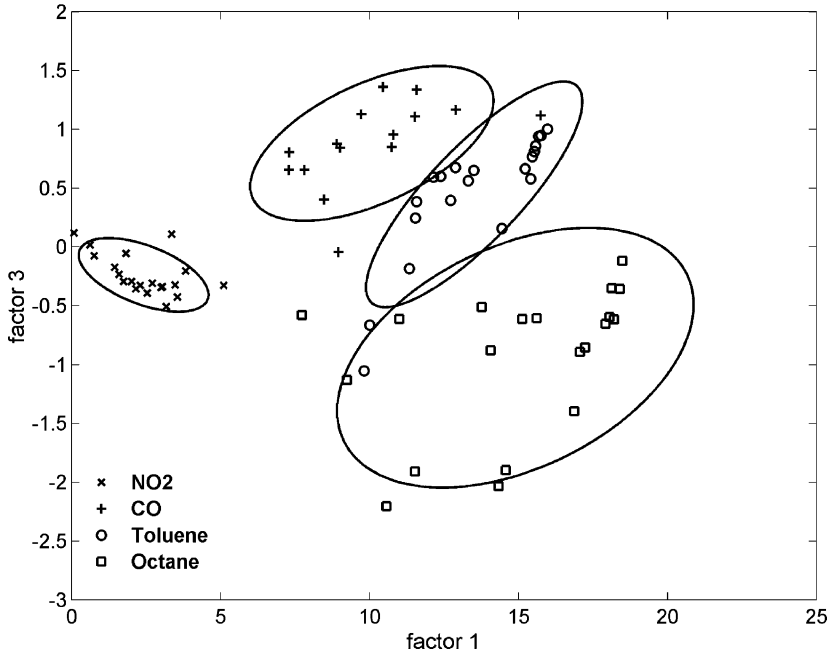

Fig. 1. PCA plot for single gases.

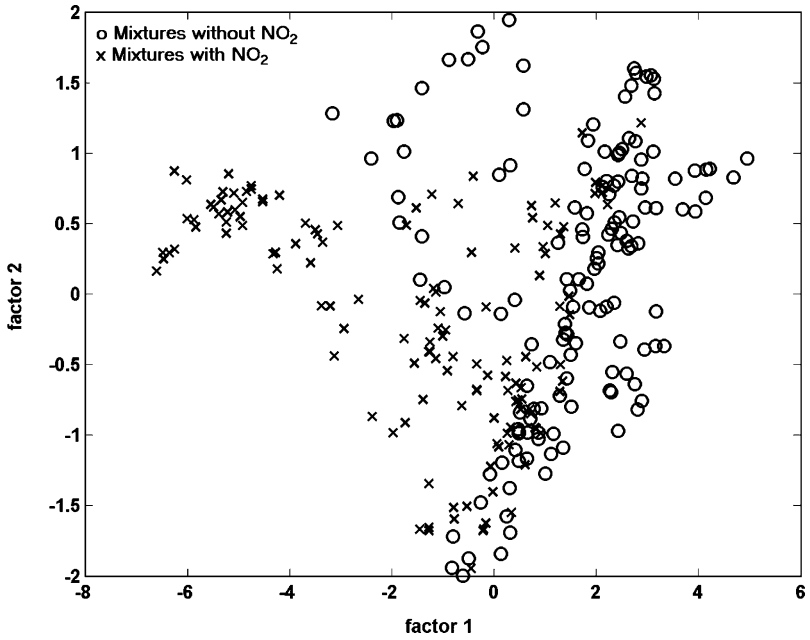

Fig. 2. PCA plot of mixtures with and without NO_2.

and the analysis techniques should analyse them. Different feature extractions can be considered to obtain more information and to facilitate therefore the separation and identification of compounds. These coefficients were selected because they were significantly different for the analysed gases.

In Fig. 3 is shown the sensor signal during a detection cycle. The measurement procedure is carried out in three steps. In each one of these a constant flow (200 ml/m) is supplied to the multisensor. The first step determines the sensor baseline. The second step is the detection process (10 min). In the last step, the recovery process re-establishes the baseline (30 min).

Fig. 4 shows the response of sensor S3 to different ternary mixtures (M1 and M2). The concentrations of each mixture are shown in Table 2. The M1 mixture is formed by nitrogen dioxide, toluene and octane. On the other hand, the mixture M2 is formed by carbon monoxide, toluene and octane (reducing gases). The sensor response is specific for each concentration and for each gas. Each sensor has a different rate of desorption and adsorption that is conditioned by the reactions on the sensing surface.

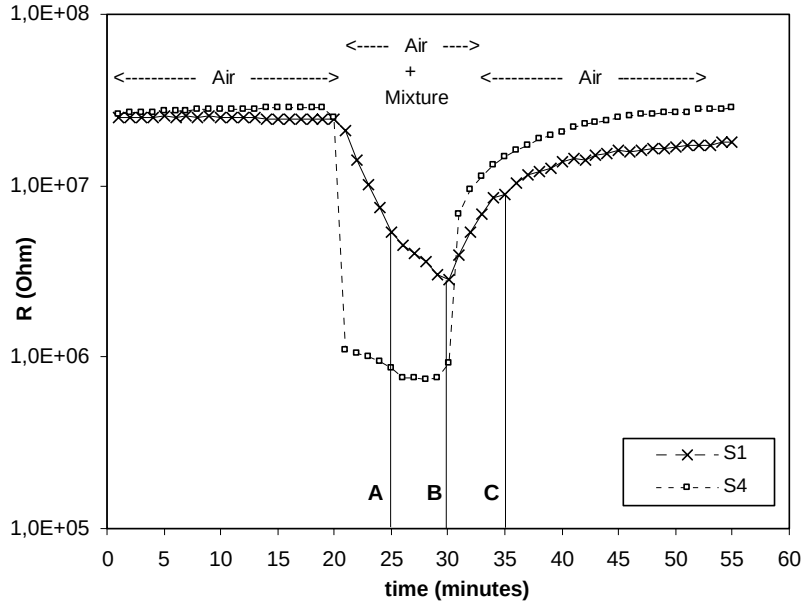

Fig. 3. Variation of the sensor resistance during a typical detection process.

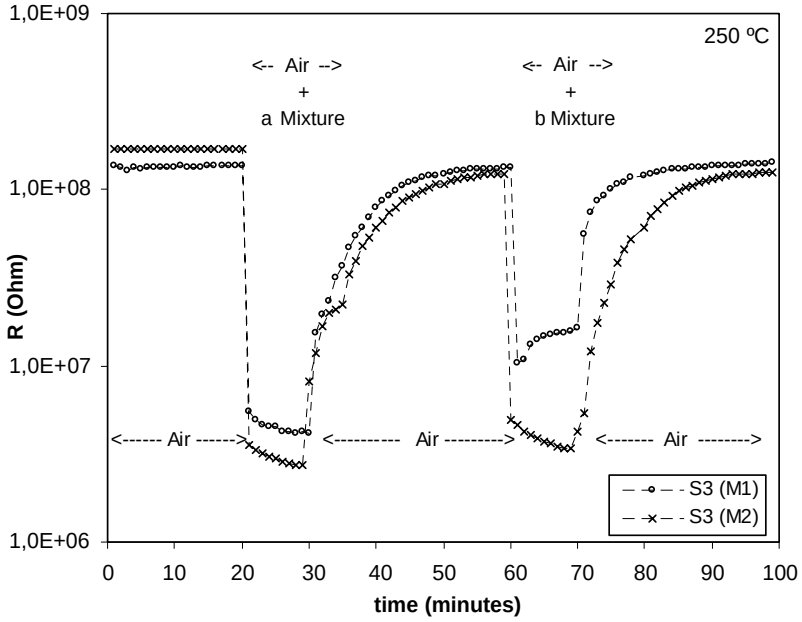

Fig. 4. Dynamic response to different mixtures of sensor S3.

A number of different features of the data were used to train the networks. Maximum response was measured when the flow of pollutant gas is switched off/turn off (10 min). Responses at 5, 10 and 15 min correspond to the middle of the duration of the pollutant gas flow (5 min, data at point A of Fig. 3), the final state of the sensor when the pollutant gas is turned off (10 min, data at point B of Fig. 3), and the response at 5 min of the desorption phase (15 min, data at point C of Fig. 3). Other extraction was made with the response curve made every minute during 20 min starting when pollutant gas flow is switched on. The last one was the wavelet expansion of the response curve. To do this analysis a Bior 2.2 wavelet was used because the difference between the original signal and the signal reconstructed with that wavelet was very low [5].

In all cases the responses have been normalised so their mean was 0 and their standard deviation was 1.

3.3. Neuronal networks

Two kind of neural networks were utilised to classify the different gases, a multilayer perceptron (ANN) and a probabilistic neural network (PNN).

Table 2
Mixture concentration

	NO_2 (ppm)	CO (ppm)	Toluene (ppm)	Octane (ppm)
M1				
(a) Mixture	0.5	150	150	0
(b) Mixture	2	50	50	0
M2				
(a) Mixture	0	150	150	150
(b) Mixture	0	150	50	50

The neural networks were trained with different sets of features. The performance was evaluated with leave-one-out cross validation and the results are shown in Tables 4–6 whose parameters are defined in Table 3:

- The letters (A, B, C and D) represent the numbers of occurrences that satisfy the row and column specifications.
- Accuracy as the percentage of correct predictions.
- Sensitivity as the percentage correct predictions when the gas is present.
- Specificity as the percentage of correct predictions when the gas is not present.

The multilayer perceptron used presents the following characteristics:

- Each gas has an independent network.
- Each of them have a hidden layer of five neurons.
- The output layer is composed by only one neuron.
- The neural network is trained with a back propagation algorithm.

The structure of the ANN is described in detail elsewhere [6]. Results of the classification are shown in Table 4, the

Table 3
Parameter to evaluate the network

	Predicted	
	Gas present	Gas not present
Real		
Gas present	A (true-positive)	C (false-negative)
Gas not present	B (false-positive)	D (true-negative)

Accuracy = A + D/(A + B + C + D); Sensitivity = A/(A + C); Specificity = D/(D + B).

Table 4
Radial basis neural network results

	NO_2	CO	Toluene	Octane
Responses at 10 min				
Accuracy (%)	79	83	87	85
Sensitivity (%)	81	82	82	85
Specificity (%)	78	84	91	85
Responses at 5, 10 and 15 min				
Accuracy (%)	95	89	99	94
Sensitivity (%)	95	86	100	93
Specificity (%)	96	91	98	94
All points				
Accuracy (%)	88	73	92	79
Sensitivity (%)	90	74	93	80
Specificity (%)	86	71	90	78
Wavelet				
Accuracy (%)	86	74	90	75
Sensitivity (%)	92	75	95	73
Specificity (%)	80	74	83	76

best results were obtained with the response to 5, 10 and 15 min.

The probabilistic neural network has the following characteristics:

- Each gas also has an independent network.
- The transfer function of the hidden layer is a radial basis one.
- The output layer has two competitive neurons, one for the presence of gas and another for its absence.
- The weights of this network are calculated directly from the training set so its training times are much shorter than the training times of the former neural network [7].

In the Fig. 5 a schematic diagram of one layer of the PNN network is shown. The new data is presented to the network. Then the distance with each data computed, multiplied by a bias is passed to the transfer function. Then all the amounts from the same class are added together and the class with the highest value is the winner. In this case the classes are the presence or absence of a gas in the mixture. Each gas have one of this layer. The results of the PNN are shown in Table 5. Also for this type of network the best results are obtained with the answer to 5, 10 and 15 min. Although the differences among the different features are smaller than obtained with the MLP networks.

The network performance trained only with the response at 10 min are the worst. As we add more features to the training set the performance improves. However there is a limit, too many features not only requires more calculation time and space memory but also the performance does not improves. This is due to the fact that the network has to generalise a much more complex dependance and sometimes get stuck in a local minima.

Table 5
Back propagation neural network results

	NO_2	CO	Toluene	Octane
Responses at 10 min				
Accuracy (%)	92	87	96	86
Sensitivity (%)	93	90	97	87
Specificity (%)	91	84	95	86
Responses at 5, 10 and 15 min				
Accuracy (%)	96	91	97	95
Sensitivity (%)	96	91	97	95
Specificity (%)	96	91	96	94
Wavelet				
Accuracy (%)	92	88	96	90
Sensitivity (%)	96	91	96	93
Specificity (%)	95	86	96	88

The back propagation with the data of the measurements made every minute during 20 min was not trained due to the long training times and memory needed.

3.4. Genetic algorithm

The best feature set of the multisensor measurements had 24 components. In order to reduce this large number a genetic algorithm was implemented [7].

Genetic algorithm (GA) is a computer technique that imitates the evolutionary process of the living organisms. While evolving, living things acquire not only inherent parental traits in the process of evolution, but are also influenced by changes in the environment. Then, through many generations, gene information makes many changes and adapt to the environment [8].

The GA has been used extensively as a method of search in optimization problems [9]. In this work the GA is used to search the best subset of features with the following considerations:

- Each individual is composed by six combinations of a sensor and a method of feature extraction done above whose characteristics have been commented previously (Section 3.2). The initial population is large enough to have twice each combination of sensor and feature.
- The evolution of the individuals is done by one-point-crossover with a rate of 90%. Also mutation of the genotype is allowed with a mutation rate of 5%.

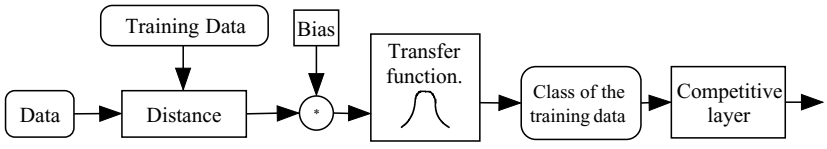

Fig. 5. Schematic diagram of one layer of the PNN network.

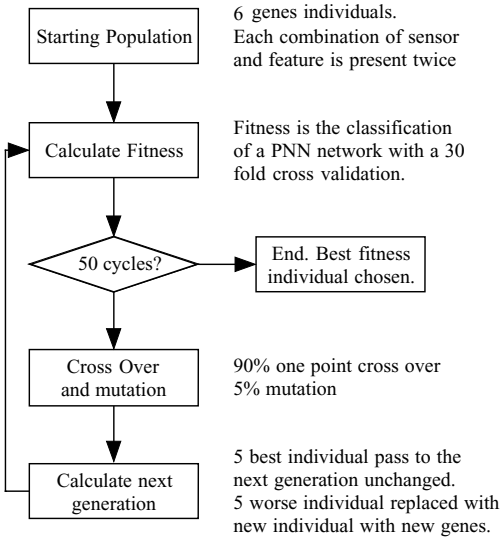

Fig. 6. Scheme of algorithm diagram.

- The five best individuals of the population pass without changes to the following generation. Also in each generation, the five worse individuals are eliminated and replaced by new individuals composed by the combination of the sensor and the feature not present in the genetic population. This is in order to ensure that in each combination a sensor and a feature are present in the evolution of the next generation.

As the fitness function of the genetic algorithm we train a PNN and evaluate his performance with a 30-fold-cross validation. The details of this network have already been commented in Section 3.3.

When the algorithms reach 50 generations the gene configuration is very stable and do not change. So the algorithm is terminated at 50 generations. Fig. 6 shows the algorithm diagram.

The combination of sensors and features selected by the genetic algorithm is applied to the PNN and MLP networks. The results are in Table 6, they are very similar to those of the latter networks, but the feature space has been reduced to only six factors. The GA selects four different sensors (S2, S3, S4 and S6). Also the features selected are different, two are from the adsorption process (at 5 min), one from the final state (at 10 min), one from the desorption (at 15 min) and two from the wavelets coefficients.

The performance of both networks (see Table 6) shows that the networks do not lose classification power with the reduction of the factor space.

Table 6
Back propagation neural network results (with GA)

	NO_2	CO	Toluene	Octane
Radial basis				
Accuracy (%)	95	91	96	93
Sensitivity (%)	93	91	97	94
Specificity (%)	96	91	95	92
Back propagation				
Accuracy (%)	92	95	96	96
Sensitivity (%)	93	94	96	96
Specificity (%)	90	96	96	96

4. Conclusions

PCA allows a good discrimination between NO_2 and the reductor gases (CO, toluene and octane). For reductor gases the separation is not so good and in some cases they are overlapped.

The pre-processing techniques and artificial neural networks are necessary to detect and classify the gases in the mixtures. We use different networks and feature extraction to discriminate among pollutant gases.

The best performance was achieved with the responses at 5, 10 and 15 min in both networks. We can conclude that it is better to carry out a feature extraction based on the properties of the signal than a general method of feature extraction.

The overall performance of the back propagation neural network is better than the performance of the radial basis. But the training times are much longer. When the optimal selection of features is used to train the network the performance difference between both networks is reduced.

The genetic algorithm was used to reduce the features to only six. The performance of this six features is similar to the performances of the others feature selections, so the feature space has been reduced without loss of discrimination power.

Also the good result obtained with the classification make us think about quantification of the gases in the mixtures. But the networks trained up to date keep a lot of quantification errors. Probably only a semi-quantitative network could give a good performance. This semi-quantitative analysis will be very useful to several applications as alarms or pollutant level detectors. Next work will address this line.

References

[1] J. Colls, Air Pollution: An Introduction, Paperback, 1996.
[2] J.W. Gardner, P.N. Bartlett, Electronic Noses Principles and Applications, Oxford University Press, 1999.
[3] M.C. Horrillo, J. Getino, L. Arés, J.I. Robla, I. Sayago, F.J. Gutiérrez, Measurements of VOCs with a semiconductor electronic nose, J. Electrochem. Soc. 145 (7) (1998) 2486–2489.
[4] Wavelet Toolbox for Matlab®, Documentation, The Math Works Inc., 2000.
[5] A. Teolis, Computational signal processing with wavelets, Birkhauser, 1998.
[6] I. Sayago, M.C. Horrillo, S. Baluk, M. Aleixandre, M.J. Fernández, L. Arés, M. García, J.P. Santos, J. Gutiérrez, Detection of toxic gases by a tin oxide multisensor, IEEE Sens. J. 2 (5) (2002) 387–393.

[7] P. Boilot, E.L. Hines, M.A. Gongora, R.S. Folland, Electronic noses inter-comparison, data fusion and sensor selection in discrimination of standard fruit solutions, Sens. Actuators B 88 (2003) 80–88.
[8] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Canada, 1989.
[9] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, Handbook of Chemometrics and Qualimetrics. Part A, Elsevier, Netherlands, 1997.

Biographies

Manuel Aleixandre received his BSc in Physics from the Universidad Autonoma of Madrid, Spain, in 1999. At present he is a PhD student at the Laboratorio de Sensores (IFA-CSIC). His research interests include neural networks, fuzzy logic, statistics and multivariate data analysis.

Isabel Sayago received the doctorate degree from the Universidad Complutense of Madrid in 1993. Since then, she has been working on chemical sensor at the Laboratorio de Sensores (IFA-CSIC). She has worked on microsensor at the CNR-LAMEL Institute (Bologna, Italy). Her main research interests are on the preparation and characterisation of gas sensors (chemoresistive, SAW and cantilevers).

María del Carmen Horrillo received her PhD in Chemistry from the University Complutense of Madrid in 1992. From 1993 to 1995, she was working at the Institute for Advanced Materials of the European Commission's Joint Research Centre (Ispra, Italy). Since then she has been working at the Instituto de Física Aplicada (CSIC) on I + D of chemical microsensors and electronic noses for environmental protection and quality control of foods. Since 1999, she is the Head of the Department of Tecnología de Gases y Superficies.

María Jesús Fernández has graduated from the University Complutense of Madrid of Spain and received her PhD degree in 1992, both in Physics. She has worked in RF and microwave devices. Her current research interests involve electronic noses and chemical sensors.

Luis Arés received his degree in Physics from the University Complutense of Madrid. He is working at the on numerical calculus, cybernetics and chemical sensors.

Maria García received the degree in chemistry from the Universidad Autónoma de Madrid, Madrid, Spain in 1998. She is currently pursuing the PhD degree in chemical sensors at the Instituto de Física Aplicada (IFA), Consejo Superior de Investigaciones Cientificas (CSIC), Madrid. Her research activities are focused on chemical sensors for the detection of VOCs in food.

José Pedro Santos received his BSc (1987) and PhD (1995) in Physics from the University Complutense of Madrid. He has worked at the University of Milan (Italy), at the Institute of Advanced Materials of the European Commission's Joint Research Centre (Ispra, Italy) and at the Electronics Department of the University Complutense of Madrid. Presently he works at the Laboratorio de Sensores (IFA-CSIC) on several projects related to the development of sensors for volatile-compound detection.

Javier Gutiérrez is Director of the Instituto de Física Aplicada (IFA-CSIC), received the PhD in Physics from the University Complutense of Madrid in 1977. His research interest is chemical sensors: semiconductor sensors, gravimetric sensors and optoelectronic devices applied to artificial olfactometrics for environmental monitoring, consumer protection and safety of aliments.