Electronic noses: a review of signal processing techniques

E.L.Hines, E.Llobet and J.W.Gardner

Abstract: The field of electronic noses, electronic instruments capable of mimicking the human olfactory system, has developed rapidly in the past ten years. There are now at least 25 research groups working in this area and more than ten companies have developed commercial instruments, which are mainly employed in the food and cosmetics industries. Most of the work published to date, and commercial applications, relate to the use of well established static pattern analysis techniques, such as principal components analysis, discriminant function analysis, cluster analysis and multilayer perceptron based neural networks. The authors first review static techniques that have been applied to the steady-state response of different odour sensors, e.g. resistive, acoustic and FET-based. Then they review the emerging field of the dynamic analysis of the sensor array response. Dynamic signal processing techniques reported so far include traditional parametric and nonparametric ones borrowed from the traditional field of system identification as well as linear filters, time series neural networks and others. Finally the authors emphasise the need for a systems approach to solve specific electronic nose applications, with associated problems of sensor drift and interference.

1 Introduction

An electronic nose is an instrument comprising an array of electronic chemical sensors with partial specificity, and an appropriate pattern recognition system that is capable of recognising both simple and complex odours. This definition restricts the term electronic nose to those types of intelligent sensor systems that are used to sense odorant molecules in an analogous manner to the human nose [1]. However, an array of nonspecific gas sensors can also be used to detect single components or mixtures of gases/vapours.

Each individual sensor i within the electronic nose produces a time-dependent electrical signal $x_{ij}(t)$ in response to an odour j. The dynamic sensor signal depends on several physical parameters, such as the speed of the flow that carries the odour from the source to the sensor array, the nature of the odour, the diffusion and reaction of the odour within the active sensing material, and ambient conditions (e.g. pressure, temperature and humidity). It is common practice to use only the static (i.e. steady-state) values of the sensor signals rather than the dynamic (i.e. transient) response when applying electronic noses. The sensor outputs are expected to reach constant asymptotic values when the input signal is fixed at a constant value. The response generated by the sensor array can then be represented by a time-independent vector $\boldsymbol{r}_j = (r_{1j}, r_{2j}, ..., r_{ij}, ..., r_{nj})$ where n is the number of odour sensors in the array. Then the response to a set of m odours can be regarded as a set of m column vectors, which are best represented by a response matrix $\boldsymbol{R}$:

$$\boldsymbol{R} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & r_{ij} & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix} \quad (1)$$

Odour sensors do not behave as independent sensors; instead, an individual sensor will respond to a variety of odours but with varying cross-sensitivity. So the off-diagonal terms of the response matrix $\boldsymbol{R}$ are usually nonzero. It is under these conditions that pattern recognition techniques are required to process the signals generated by an array of sensors and thereby extract the pertinent information in a nontrivial manner (i.e. when $m > n$) [2].

It is only in the last few years that the use of dynamic signals from a multisensor system has received any significant attention. There are several reasons why dynamic signal processing techniques are important to the field of electronic noses:

- Recent reports suggest that under certain test conditions the dynamic response of solid-state gas sensors contains useful information. The dynamic response holds information about the chemical kinetics of the sensor and these vary with both sensor type and analyte. This additional information can be extracted from the transient response of a sensor to a controlled change in the analyte concentration (e.g. concentration modulation) or to a change in the temperature of operation of the sensor (e.g. temperature modulation). In some applications an enhancement of the sensor array selectivity from these techniques has been reported [3–6].
- Some sensors respond very slowly to weakly interacting odours. Non-steady-state measurements are required when the environmental changes are on the same time-scale as the sensor response. This may help to broaden the field of

IEE Proceedings online no. 19990670

DOI: 10.1049/ip-cds:19990670

Paper first received 23rd September 1998 and in revised form 29th June 1999

E.L. Hines and J.W. Gardner are with the Electrical and Electronic Engineering Division, School of Engineering, University of Warwick, Coventry CV4 7AL, UK

E. Lobet is with the Departament d'Enginyeria Electrònica, Universitat Rovira i Virgili, Autovia de Salou s/n, 43006 Tarragona, Spain

application of intelligent sensor systems (e.g. continuous pollution monitoring).

• The sample delivery system and the sensor array are both parts of a dynamic system. The time taken by the system to reach steady state depends on parameters such as flow rate, volume of the sensor chamber, odour diffusion rate and sensor reaction rate. When the sensors are modelled using steady-state values, the calibration time can be very long, especially when a multicomponent calibration is performed. Therefore, dynamic modelling may significantly reduce the time of each calibration experiment [7, 8].

• Even when sensors are exposed to identical gas mixtures, they do not give stable responses over a long period of time. In other words, sensor signals tend to show significant temporal variation, typically referred to as 'long-term drift'. This variation may be due to unknown processes in the sensor system, like poisoning, ageing or gradual changes in the environment (i.e. temperature and humidity). Drift may seriously affect calibration. Therefore, when an intelligent sensor system is to be operated for a long period of time, long-term drift should be addressed by the pattern recognition algorithms [9, 10].

• Finally, the baseline signal (in air) and response of a sensor can depend on its immediate prior chemical history. These changes can be considered to be like a 'short-term' drift. For example, a dynamic model that uses the knowledge of present and past inputs and outputs of the sensor would be able to predict its true baseline.

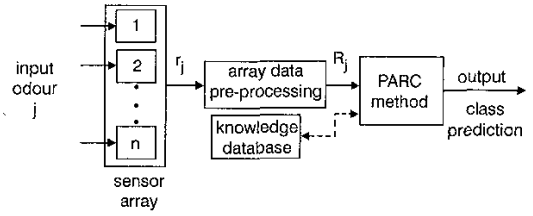

Fig. 1 *Basic structure of a data processing system for an electronic nose*

2 Static pattern analysis techniques

The static responses generated by an array of odour sensors may be processed using a variety of techniques. Fig. 1 illustrates the basic data-processing structure of an electronic nose. When a set of m odours is introduced to an array of n odour sensors, the sensor outputs are pre-processed so that the modified response matrix $\tilde{\boldsymbol{R}}$ can be fed into a pattern recognition (PARC) engine. According to eqn. 1, column j in $\tilde{\boldsymbol{R}}$ is the pre-processed response vector generated by the sensor array in the presence of odour j. The nature of a PARC engine is usually classified by the terms parametric or nonparametric and supervised or nonsupervised. A parametric technique is based on the assumption that the sensor data can be described by a probability density function (PDF) that *a posteriori* defines its spread of values (in most cases the assumption made is that data are normally distributed with a known mean and variance). Nonparametric methods do not assume any PDF for the sensor data and thus apply more generally. In a supervised PARC method, a set of known odours are systematically introduced to the electronic nose, which then classifies them according to known descriptors (classes) held in a knowledge base. Then, in a second stage, an unknown odour is tested against the knowledge base and the predicted membership class is given. Other PARC methods do not need a separate training stage but learn the different classes from the response vectors automatically (unsupervised learning). These methods are closer to the way that our brain is understood to work.

2.1 Pre-processing techniques

It is important to examine the data generated by an array of odour sensors so that the most informed choice of sensor, pre-processing and PARC method is made. Pre-processing of the response vectors should be designed to help analyse data from a specific problem, such as to linearise the output from the sensors or to compensate for concentration fluctuations in the response vectors by using a normalisation procedure. There is evidence that when analysing data from metal oxide odour sensors, the fractional change in conductance is generally a good choice with PARC methods [11]. The general expression is

$$r_{ij} = \frac{X_{ij}^{odour} - X_i^0}{X_i^0} \tag{2}$$

where X_i^0 is the conductance of the sensor i in air (i.e. baseline signal) and X_{ij}^{odour} is the conductance of the sensor i in the presence of odour j. This procedure helps compensate for temperature sensitivity of the sensors and linearises the mechanism that generates their concentration dependence. Moreover, when sample concentration is of no interest, but fine discrimination between odours is required, it may be useful to normalise the length of the response vectors to unity [11]. Then eqn. 2 becomes

$$\tilde{r}_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^{n} r_{ij}^2}} \tag{3}$$

The effect of this is to place the end of all response vectors on to the surface of a unit hypersphere and hence reduce the effect of concentration fluctuations. However, caution is needed because the method both assumes concentration linearity and enhances the noise in the case of small signals.

2.2 Parametric analysis techniques

2.2.1 Linear calibration methods: Linear multivariate calibration methods are often used to process sensor array data and obtain the concentrations within a multicomponent mixture [12]. Two common methods are partial least squares (PLS) [13] and principal components regression (PCR) [14], which assume that a linear inverse model can be applied to the data. In the model, the concentration vector $\boldsymbol{c}$ is related to the response matrix by

$$\boldsymbol{c} = \tilde{\boldsymbol{R}}\boldsymbol{M} + \boldsymbol{e} \tag{4}$$

where $\boldsymbol{M}$ is a regression vector containing all the model parameters and $\boldsymbol{e}$ is an error vector containing the concentration residuals. The main difference between PLS and PCR is that PLS includes information about the concentration vector in the model while PCR does not.

2.2.2 Discriminant function analysis: In discriminant function analysis (DFA) it is assumed that the data are multinormal-distributed and then the discriminant functions Z_j are determined. Each discriminant function is calculated by maximising the F-ratio on the analysis of the subject to Z_p being uncorrelated with Z_p ... Z_{p-1} within groups. The discriminant functions are linearly related to the sensor responses by the following equation:

$$Z_p = a_{1p}r_{1j} + a_{2p}r_{2j} + \ldots + a_{ip}r_{ij} + \ldots + a_{np}r_{nj} \tag{5}$$

Once the regression coefficients a_{ip} have been computed on the known data (referred to as supervised learning), then

they can be used to form the classification functions which predict the group membership of unknown response vectors (referred to as cross-validation). Fig. 2 shows the results of applying DFA to the response (fractional change of conductance) of 12 tin oxide gas sensors sampling the headspace of three different coffees [15]. Plots of the first two discriminant functions show reasonable separation of the three groups. The observed classification rate was 81% when half the data set was used for crossvalidation.

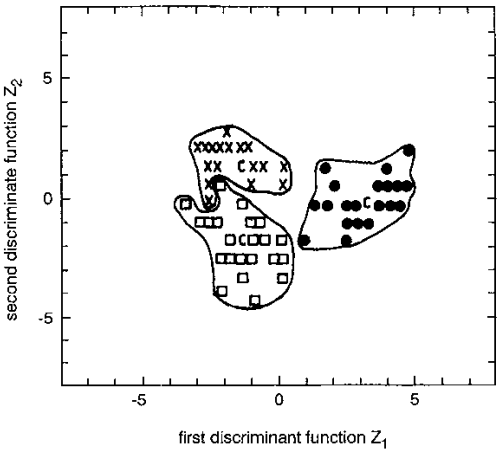

Fig. 2 *Results of linear DFA on analysis of three commercial roasted coffees using a 12-element tin oxide electronic nose*
(Reprinted from Gardner, J.W., *et al., Sens. Actuators B, Chem.*, **6**, pp. 71–75, © 1992 Elsevier Science, with permission)
C centroid
× coffee a
● coffee b
□ coffee c

2.2.3 Nonlinear calibration methods: Most chemical sensors are nonlinear. The use of linear techniques is acceptable when the concentration vector is fairly constant. In this case, the nonlinear part of the response may be linearly approximated. When a wider region of the response space is covered, or the multicomponent gas mixture problem is being considered, a nonlinear pattern analysis method is required. Nonlinear parametric techniques, such as multivariate adaptive regression splines (MARS), have been used to analyse simple gas mixtures with some success [16]. In a real electronic nose application, complex odours can contain hundreds or even thousands of odour molecules that interact. This leads to a highly nonlinear response space, which cannot be described by a straightforward PDF. More interest has thus been shown in the application of nonparametric techniques, which do not make underlying assumptions about the probability distribution. We discuss some of these in the following Section.

2.3 Nonparametric analysis techniques

2.3.1 Principal components analysis: PCA is a linear supervised method that has been used to discriminate the response of an electronic nose to simple and complex odours (e.g. alcohols, beers, coffees). The method consists of expressing the response vectors r_j in terms of a linear combination of orthogonal vectors. Each orthogonal (principal) vector accounts for a certain amount of variance in the data with a decreasing degree of importance. The pth principal component can be expressed as

$$X_p = \alpha_{1p}r_{1j} + \alpha_{2p}r_{2j} + \ldots + \alpha_{ip}r_{ij} + \ldots + \alpha_{np}r_{nj} \quad (6)$$

The variance of each principal component is maximised under the constraint that the sum of the coefficients of the orthogonal vectors $\alpha_p = (\alpha_{ip}, \ldots, \alpha_{ij}, \ldots, \alpha_{np})$ is set to unity and the vectors are uncorrelated. Since there is often a high degree of sensor collinearity in electronic nose data, the majority of the information held in the response can be displayed using a small number of principal components. This means that an n-dimensional problem can be described by a two- or three-dimensional plot. Fig. 3 shows the results of applying PCA to an array of 12 tin oxide sensors. Since metal oxide sensors respond in a similar manner, over 80% of the variance is described by the first two principal components [11, 17]. Three distinct groups are apparent and are associated with lagers (A), beers (B) and spirits (C). The discrimination between beverages within a group seems to be harder, but some success has been reported for the analysis of Japanese beers and whiskies [18].

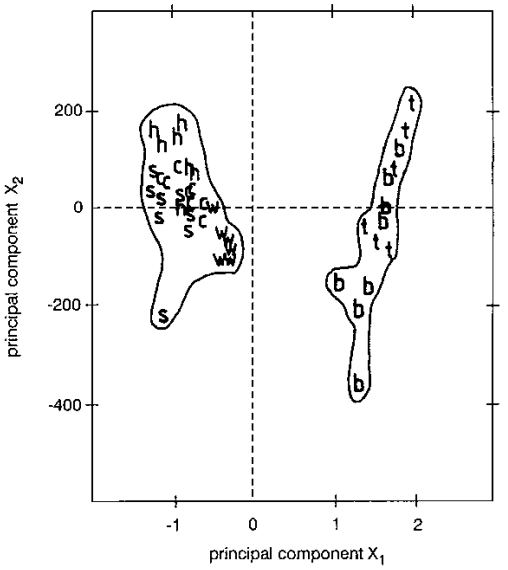

Fig. 3 *Results of PCA analysis of response of 12-element tin oxide electronic nose to two beers (labelled w and s), two lagers (labelled h and c) and two spirits (labelled t and b)*
(Reprinted from Gardner, J.W., *et al., Sens. Actuators B, Chem.*, **4**, pp. 108–116, © 1991 Elsevier Science, with permission)
C centroid
× coffee a
● coffee b
□ coffee c

2.3.2 Cluster analysis: CA is an unsupervised technique for enhancing the differences between the response vectors in n-dimensional space and identifying clusters or groups to which unknown vectors are likely to belong. To do so a multidistance metric d_{ij} is calculated between data points i and j according to the expression

$$d_{ij} = \left(\sum_{k=1}^{N} (r_{ik} - r_{jk})^N \right)^{1/N} \quad (7)$$

If N is set to 2, the Euclidean (linear) metric is used. The proximity of all points relative to each other is then found by computing a so-called scalar similarity value S_{ij}, such as

$$S_{ij} = 1 - \frac{d_{ij}}{\max\{d_{ij}\}} \quad (8)$$

This is also called complete linkage. Other definitions can be considered for the similarity value but the choice of metric and linkage has a marginal effect on the results. The technique has been applied to semiconductor gas sensors (both metal oxides and conducting polymers) [19–21]. Fig. 4 shows the results of a CA (Euclidean metric, complete linkage) on the response of a metal oxide electronic

nose to different alcohols [11]. The dendrogram connects up response vectors with the nearest similarity value.

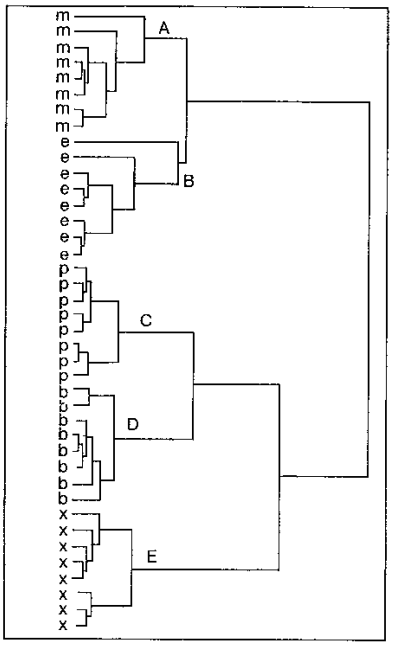

Fig. 4 *Dendrogram showing results of CA on response of 12-element tin oxide electronic nose to samples of methanol (m), ethanol (e), propanol (p), butanol (b) and methyl-butanol (x)*
Using Euclidean metric and complete linkage
(Reprinted from Gardner, J.W., *et al., Sens. Actuators B, Chem.*, 4, pp. 108-116, © 1991 Elsevier Science, with permission)

2.3.4 Artificial neural networks: The nature of electronic nose data is such that it is often desirable to use a more powerful pattern recognition method. More powerful methods should be able to cope with nonparametric nonlinear data, and have further advantages over more conventional methods (such as learning, self-organisation, generalisation and noise tolerance). This has led to the rapid and widespread application of artificial neural networks (ANNs) to the analysis of patterns generated by electronic noses. ANNs consist of parallel interconnected, and usually adaptive, processing elements. Furthermore, ANNs are attractive as they, to a certain extent, mimic the olfactory system [22]. The processing elements represent the biological olfactory cells or neurones, and their interconnections, the synaptic links.

Multilayered feedforward networks: Since three-layered networks have sufficient computational degrees of freedom to solve any classification problem [23], most workers have adopted this topology. In a network, the processing elements are organised in three distinct groups of elements: input, hidden and output layer neurones. The number of input nodes corresponds to the number of sensors in the array. The number of neurones in the hidden layer is to be determined experimentally and the number of odours analysed generally determines the number of output neurones. The interconnection topology and learning rules of the neurones determine the performance of a particular network.

The most popular ANN in odour classification is the three-layered feedforward backpropagation trained network [22, 24, 25]. To train the network, it is necessary to provide it with a number of sample inputs with their corresponding outputs (supervised learning). Each neurone adds its weighted inputs and performs a nonlinear transformation of this sum (generally a sigmoid function is used). The calculation is carried out for each layer feeding the values through to the output layer. During the learning phase, the weights are adjusted to minimise the difference between the actual output and the ideal output. Once the network is trained, it can be used to predict the membership of novel and untrained samples. Fig. 5 shows the topology of a network used to identify alcohols and Table 1 the results of classification. The results show that the network is able to distinguish between these odours. The success of the training process, in terms of a fast rate of convergence and good generalisation, can be affected by the choice of the architecture and initial parameters (e.g. initial weights and learning rate). Since architecture and parameters are to be determined experimentally, much time may be spent searching for the optimal ANN. An alternative method is to use a genetic algorithm (GA) [26, 27] to determine automatically a suitable network architecture (e.g. growing and pruning the network) and a set of parameters (e.g. learning rate, momentum term) from a restricted region of design space. GAs are heuristic search algorithms based on the mechanics of natural selection. The structure and parameters of the neural network (i.e. learning rate, initial weights, number of layers, number of neurones in each layer and connectivity) are coded using binary strings, which are concatenated to form chromosomes. GAs are then applied to search populations of chromosomes. Typical genetic operators can be defined such as parent selection, crossover and mutation. The performance of the network represented by each chromosome c_i is evaluated using a fitness function,

$$F(c_i) = au(c_i) + b \tag{9}$$

where F is the fitness function, u is the objective function that we want to optimise, and a and b are transformation parameters that are dynamically adjusted to avoid premature convergence. The objective function is generally a weighted sum of the various performance measures. In the sensor data classification problem, the performance measures used in the objective function are based on, for example, the network prediction error, speed of convergence, size of the network and degree of generalisation achieved [28].

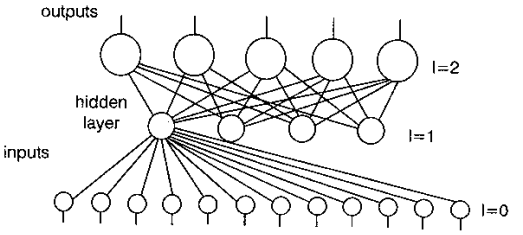

Fig. 5 *Structure of a fully connected three-layer backpropagation network used to process data from a 12-element tin oxide electronic nose for five alcoholic odours*
(Reprinted from Gardner, J.W., *et al., Meas. Sci. Technol.*, 1, pp. 446-451, © 1990 Institute of Physics Publishing, with permission)

Table 1: Results of classification for network shown in Fig. 5

	Actual outputs (g_{ij2})				
Chemical species	Output 1	Output 2	Output 3	Output 4	Output 5
Methanol	0.990	0.023	0.000	0.000	0.020
Butan-1-ol	0.001	0.906	0.002	0.000	0.153
Propan-2-ol	0.000	0.011	0.931	0.000	0.071
2-Methyl-1-butanol	0.000	0.001	0.102	0.949	0.000
Ethanol	0.002	0.009	0.475	0.000	0.907

Reprinted from Gardner, J.W., *et al., Meas. Sci. Technol.*, 1, pp. 446–451, © 1990 Institute of Physics Publishing, with permission

Conventional networks are trained using random weights, and this may be disadvantageous to the overall training process. The main problem with this strategy is that the search for the best set of weights to classify the training patterns and to identify new ones usually starts from a poor point which either slowly, or maybe never, takes us to the desired optimal point. On the other hand, a suitable starting point, dependent on the kind of training data, is desirable. It can speed up training and reduce the likelihood of settling in local minima, resulting in a better performance than common backpropagation networks [29]. A type of fuzzy neural network (FNN) can be used to make use of possibility distributions to determine the initial set of weights [30]. The theory of possibility is more like a membership-class restriction imposed on a variable defining the set of values it can take. Possibility distributions are often triangular and so they are similar in shape to normal distributions with the mean values having the highest possibility of occurrence, which is one. Any value outside the min-max range has a possibility of occurrence of zero. The possibility that a_j is a member of the fuzzy set $X = \{a_1, a_2, ..., a_j, ..., a_n\}$ is denoted by its membership value $M(a_j)$, which depends on the mean, minimum and maximum of the set X. Fig. 6 shows a possibility distribution function and associated membership function. An introduction to the theory of fuzzy logic is given by McNeil and Freiburger [31].

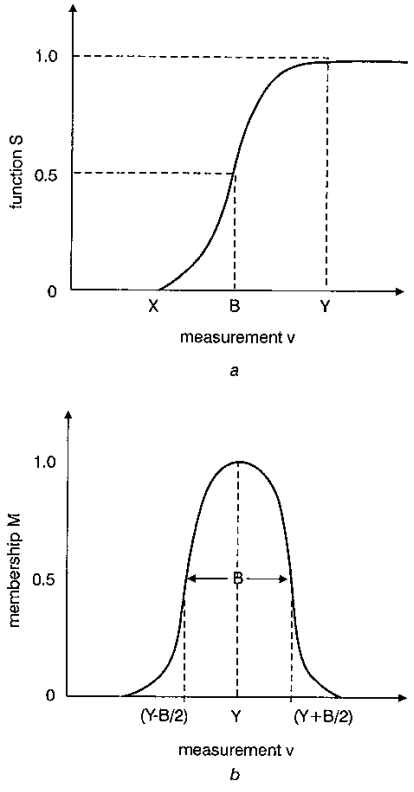

Fig. 6 *Possibility distribution S(v: X, B, Y) used to determine membership value of a measurement v, and membership function M related to S*
S is given by 0 when $v \leq X$, $2(v - X)2/(Y - X)2$ when $X < v \leq B$, $1 - 2(v - Y)2/(Y - X)2$ when $B < v \leq Y$ and 1 when $v > Y$
$M = 1 - S(v: Y, Y + B/2, Y + B)$ when $v > Y$, and $M = S(v: B, Y - B/2, Y)$ when $v \leq Y$
(Reprinted from Singh, S., *et al.*, *Sens. Actuators B, Chem.*, **30**, pp. 185–190, © 1996 Elsevier Science, with permission)

Self-organising maps: Self-organising maps (SOM) applied to electronic nose systems contain a two-dimensional single layer of neurones in addition to an input layer of branched nodes. If the system is left alone in an environment of interest, the learning algorithm of the network processes the sensor outputs step by step, and constructs an internal representation of the environment [32]. SOMs accumulate a lot of statistical information but in an unsupervised fashion. There are m neurones in the neural layer (typically arranged as the knots of a square lattice) and each one has a parameter weight vector $\boldsymbol{V}^{(i)}$ of dimension n, which is the same as the input feature vectors (i.e. the number of sensors). Each neurone is described by the vector s, whose components are the knot co-ordinates in the lattice. The weight vectors are randomly initialised at the beginning. One input vector $\boldsymbol{x}^{(j)}$ is selected from the sample and put into the network, and the squared distances between $\boldsymbol{x}^{(j)}$ and each $\boldsymbol{V}^{(i)}$ are computed using

$$d_{ij}^2 = \sum_{k=1}^{n} \left(x_k^{(j)} - V_k^{(i)}\right)^2 \qquad i = 1, \ldots, m \qquad (10)$$

The minimum distance d_{ji^*} is then determined to obtain the neurone $\boldsymbol{i}^*$ that is the winner over the others. In a winner-takes-all strategy, the winning neurone updates its weights via

$$V_{new}^{(i^*)} = V_{old}^{(i^*)} + \eta \left(x^{(j)} - V_{old}^{(i^*)}\right) \qquad (11)$$

where η is the step gain (or learning rate). All other neurones keep their old weight values. In another strategy, all the neurones that are close to the winner are updated:

$$V_{new}^{(s)} = V_{old}^{(s)} + \eta h_{si^*} \left(x^{(j)} - V_{old}^{(s)}\right) \text{ for all } s \qquad (12a)$$

$$h_{si^*} = \exp\left(-\|s - i^*\|^2 / 2\sigma^2\right) \qquad (12b)$$

h_{si^*} is called the excitatory response and is only appreciable for the neurone that coincides with $\boldsymbol{i}^*$ and its neighbours. σ is the length scale of the proximities to $\boldsymbol{i}^*$ and is generally fixed to a value in the range of 2–5 lattice units. It is desirable that after a number of iterations the weights no longer change. Therefore, if the map is to stabilise asymptotically in an equilibrium state, η must decrease to zero, typically as follows:

$$\eta = \eta_0 \exp(-t/\tau) \qquad (13)$$

Supervised learning, where the SOM is provided with the desired output function, is also possible. This strategy is also known as learning vector quantisation (LVQ) [33]. The number of training patterns to ensure equal accuracy to other approaches could be dramatically decreased because the given calibration data set is not the unique source of information; it is added to the large amount of information collected by the system on its own during unsupervised learning. However, an important limitation of this approach is that lengthy computation is required when it is applied to real problems.

SOMs have been used with some degree of success to classify patterns generated by electronic noses [34–36]. Fig. 7*a* shows the outputs of a sensor array in the presence of two classes of a gas mixture. In Fig. 7*b* the spontaneous clustering of neurones in the SOM network is shown. After a short learning period, there exists a strict correlation between the weights of the network and the assigned classes. The network does not have any direct information about the classes, except for the sensor outputs [37].

Adaptive resonant theory: Adaptive resonant theory (ART) was introduced as a theory of human cognition in information processing [38]. It is based on the fact that a human brain can learn new events without necessarily forgetting events learnt in the past. ART networks are intelligent systems that are capable of autonomously adapting in real time to changes in the environment. They have been

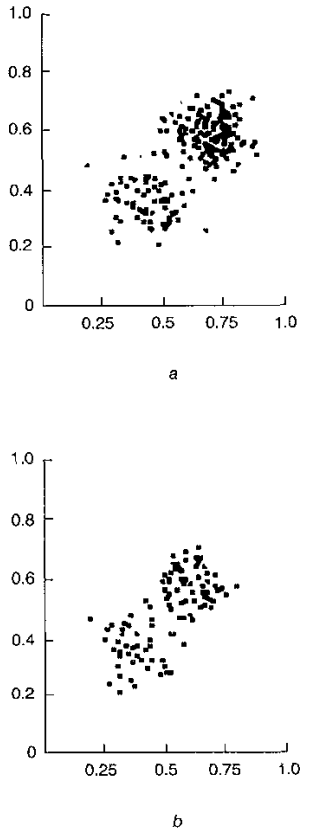

Fig. 7 *Two simulated classes of a gas mixture, projected onto the feature plane. (the two regions correspond to the individual gases), and Kohonen net status after a learning run*
The net provides mapping between the plane and the neurone weights. The inverse mapping is presented here to show the strict similarity between the neurone images and the statistical distribution of classes
(Reprinted from DiNatale, C., *et al., Sens. Actuators B, Chem.*, **23**, pp. 111–118, © 1995 Elsevier Science, with permission)

designed to resolve the stability-plasticity dilemma and so are stable enough to incorporate new information without destroying the memories of previous learning. The interested reader is referred to the work of Carpenter [39] for details. ART networks have been applied to metal oxide sensor based electronic noses. The results are very similar to those obtained with multilayer backpropagation trained networks (BP), but the training time is typically an order of magnitude faster than BP on small data sets [40].

3 Dynamic models and system identification: a review

The techniques that are typically used to model the dynamic sensor response are borrowed from the field of system identification. System identification is the process of developing a mathematical representation of a physical/chemical dynamic system using experimental input-output data. The majority of methods that have been developed to study engineering problems assume linearity and stationarity. However, almost all real chemical transducers are characterised by nonlinear dynamics and response drift. This Section starts by reviewing some models for the dynamic response of odour sensors and proceeds to consider the different approaches which can be used to select a model.

3.1 Dynamic models

3.1.1 Linear models: Linear methods have been applied in diverse fields such as econometrics, biological systems and control systems. Their application to the identification of sensor array systems for gas analysis is recent [41]. The objective of the dynamic model is to forecast the output of the sensor from knowledge of the input signals in dynamic conditions (forward modelling). Only the inversion of the model would allow us to identify the input (gases/aromas) given the output signals (inverse modelling). The most common models are ARMA (autoregressive moving average), ARX (autoregressive with extra input; also autoregressive exogenous), ARMAX and Box–Jenkins. These models are of interest in digital signal processing because the time series can be considered to be the output of a linear filter with a rational transfer function. In the following, their mathematical expressions are given ($x[n]$, $y[n]$ and $e[n]$ are discrete input, output and residual term or noise signals, respectively):

$$\mathrm{ARMA}(q,p)\colon y[n] = \sum_{i=1}^{q} \alpha_i y[n-i] + \sum_{j=0}^{p} \beta_j e[n-j] \tag{14}$$

The current value of the output is modelled using q past values of the output and the present and p past values of the noise. Two different submodels of this one can be considered – the autoregressive (AR) and the moving average (MA):

$$\mathrm{AR}(q)\colon y[n] = \sum_{i=1}^{q} \alpha_i y[n-i] + e[n] \tag{15}$$

$$\mathrm{MA}(p)\colon y[n] = \sum_{j=0}^{p} \beta_j e[n-j] \tag{16}$$

Moving average models are also known as all-zero models,

$$\mathrm{ARX}(q,k)\colon y[n] = \sum_{i=1}^{q} \alpha_i y[n-i] + \sum_{k=0}^{r} \gamma_k x[n-k] + e[n] \tag{17}$$

The present value of the output is modelled using a linear combination of the past q values of the output and the present and past r values of the input:

$$\mathrm{ARMAX}(q,k,p)\colon y[n] = \sum_{i=1}^{q} \alpha_i y[n-i] + \sum_{k=0}^{r} \gamma_k x[n-k] + \sum_{j=0}^{p} \beta_j e[n-j] \tag{18}$$

Similar to the previous one but including a moving average term,

$$\text{Box-Jenkins}\,(r,p)\colon y[n] = \sum_{k=0}^{r} \gamma_k x[n-k] + \sum_{j=0}^{p} \beta_j e[n-j] \tag{19}$$

In this model, the prediction of the output is made without the use of past values of the output. It uses present and past values of the input in addition to filtered noise.

State-space models: In the state-space form, the relationship between the input, noise and output signals is written as a system of first-order difference equations using an auxiliary state vector ξ_n. This description of linear dynamical systems became increasingly important after Kalman's work on prediction and linear quadratic control [42]. Insights into the physical mechanisms of the system can usually more

easily be incorporated into space-state models than into the models described previously. The state-space model can be expressed as:

$$\xi_{n+1} = \boldsymbol{A}(\boldsymbol{\theta})\xi_n + \boldsymbol{B}(\boldsymbol{\theta})x[n] + e_p[n] \quad (20)$$

$$y[n] = \boldsymbol{C}(\boldsymbol{\theta})\xi_n + e_m[n] \quad (21)$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are matrices of appropriate dimensions. $\boldsymbol{\theta}$ is a vector of parameters that typically correspond to unknown values of physical coefficients.

3.1.2 Nonlinear models: Chemical sensors are almost always nonlinear for high gas concentrations. Most of them are inherently nonlinear even at low concentrations. Transport, adsorption and reaction processes taking place at the sensor include intrinsic nonlinear dynamics. Thus, an electronic nose instrument can be represented as a nonlinear system.

The analysis of nonlinear systems poses many problems that do not appear in their linear counterparts. For instance, the law of superposition cannot be applied and the addition of two input signals may lead to unknown results. Traditionally, the methods used to identify nonlinear systems are parametric methods that make assumptions about the structure of the system. If the structure is not accurate enough, the model will not work for all inputs. Recently, a few nonlinear time series and other nonlinear models have been proposed. Some of them will be reviewed briefly below.

Nonlinear time series: Some of the nonlinear models are introduced in this Section. The reader is referred to the work of Tong [43] for a more comprehensive survey.

One of the more important classes of nonlinear models is the class of nonlinear autoregression. $y[n]$ is said to follow a nonlinear autoregressive mode of order k if there exists a nonlinear function f such that

$$y[n] = f\left(y[n-1], y[n-2], \ldots, y[n-k], e[n]\right) \quad (22)$$

where $e[n]$ is noise. As a 'dual' to nonlinear autoregressive models, we may have nonlinear moving average models (e.g. of order q):

$$y[n] = g\left(e[n], e[n-1], \ldots, e[n-q], q\right) \quad (23)$$

q being a vector of parameters.

Since the most important linear time series model is the ARMA model, it seems natural to develop a nonlinear generalisation of it. For suitable k and q,

$$\xi_n = \big(1, e[n-q+1], \ldots, e[n], y[n-k+1], \ldots, y[n]\big)^T \quad (24)$$

ξ_n is called a carrier vector. Choosing suitable matrices $\boldsymbol{F}$, $\boldsymbol{G}$ and $\boldsymbol{H}$, we may achieve the nonlinearisation of ARMA models by introducing

$$\xi_n = F(\xi_{n-1})\xi_{n-1} + G(\xi_{n-1})\xi_n \qquad y[n] = H\xi_{n-1} \quad (25)$$

This is formally equivalent, under suitable choices of $\boldsymbol{F}$, $\boldsymbol{G}$ and $\boldsymbol{H}$, to

$$y[n] = \sum_{i=1}^{q} \alpha_i(\xi_{n-1})y[n-i] + \chi[\xi_{n-1}] + \sum_{j=0}^{p} \beta_j(\xi_{n-1})e[n-j] \quad (26)$$

The carrier vector can be regarded as a state vector and the model above as a state-dependent model (SDM) [44].

Functional expansions: Functional expansions were studied by Volterra [45] and Wiener [46]. They are valid representations of nonlinear systems under very weak assumptions (i.e. stationarity). The concept of a functional was introduced to describe the input/output relationship of a system Assuming that $x(t)$ is the input and $y(t)$ the output, then

$$y(t) = F[t; x(t'), t' \leq t] \quad (27)$$

The task of modelling consists of obtaining a mathematical expression for the functional F. This is to identify the input/output map of the system, determining the effect of past values of the input on the output. In the case of a nonlinear time invariant system, F can be expressed as a Volterra functional expansion of the form

$$y[t] = \sum_{n=1}^{\infty} \overbrace{\int_0^{\infty} \cdots \int_0^{\infty}}^{n} k_n(\tau_1, \ldots, \tau_n)x(t-\tau_1)dt \ldots x(t-\tau_n)dt \quad (28)$$

The kernels kn (τ_1, ..., τ_n) constitute the descriptors of the system dynamics. The nth kernel attains the effect of the crossinteraction of n past values of the input on the output. Wiener redefined the basis functionals so that they were orthogonal for white Gaussian inputs.

Block-structured network models: Block-structured network models consist of interconnections of two different classes of blocks – dynamic linear blocks and static nonlinear blocks. This modelling strategy is closely related to the functional expansion method, because a close examination of the relationship between the Wiener kernels is necessary to determine the topology of the network. This method is preferred by some authors to functional expansion because of the difficulty involved in interpretation of the kernels. Furthermore, block-structured models may be related to the inner structures of the system. The reader is referred to the work of Chen *et al.* [47, 48], where a systematic structural classification procedure employing Wiener kernels is reviewed.

Neural networks: In recent years, multilayer perceptrons (series-parallel identification method) and time-delay or recurrent neural networks (parallel identification method) have been proposed for system identification and modelling purposes [49]. It has been proved that the output of an ANN, whose inputs are delayed values of the input signals, can be expressed as an infinite Volterra series [49]. In this case, since the expansion is not limited to the first or second kernels, the network is able to model highly nonlinear relations if there are enough hidden neurones. The output of the network is a nonlinear function of q delayed outputs and p delayed inputs:

$$y[k+1] = f\left(y[k], y[k-1], \ldots, y[k-q], x[k], \ldots, x[k-p]\right) \quad (29)$$

From the point of view of system identification, a multilayer neural network can be assumed to be a nonlinear map. The elements on the weight matrices are parameters, whose optimum values should be found by training the ANN over a training set. Fig. 8 shows the differences between the series-parallel and the parallel identification methods. The stability of the second method, which uses a neural network with feedback, cannot be assured [50–52].

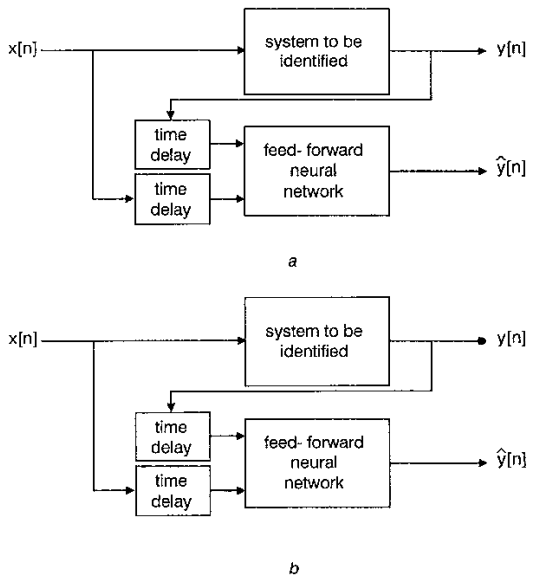

Fig. 8 *Series-parallel system identification method*
a Neural network supplied with lagged inputs and outputs of system to be identified
b Parallel system identification method using neural network with feedback

3.2 System identification: selecting a model

The techniques used to identify a model from measured data typically consist of parametric or nonparametric approaches. With nonparametric techniques, very few assumptions about the system to be modelled are required, and therefore apply more generally. However, parametric techniques can sometimes lead to better results, especially when the amount of data is limited (i.e. short time series). This subsection reviews the different techniques available for model selection.

3.2.1 Nonparametric approach: time and frequency-domain methods: A linear time-invariant system can be described by its transfer function or by the corresponding impulse response. A nonlinear time-invariant system can be described using functional expansions (Wiener kernels). Transfer functions, impulse responses and Wiener kernels may be determined by direct techniques. Such methods are often called nonparametric since they do not explicitly employ a parameter vector in the search for a best description.

Time-domain methods: Time-domain methods include impulse response analysis, step-response analysis and correlation analysis. Impulse response analysis is impractical because many processes do not allow impulse inputs of such amplitude that the error is insignificant compared to the impulse response coefficients. Step-response analysis can furnish some basic characteristics to a sufficient degree of accuracy (i.e. delay time, static gain, dominating time constants). Using correlation analysis, an estimate of the impulse response can be obtained, through the crosscorrelation of input and output signals and the autocorrelation of the input. To identify ARMA models, the estimated autocorrelation and partial autocorrelation functions of the input signal provide valuable information. Autoregressive processes of order 1, 2, ... are fitted successively and the residuals calculated. The partial autocorrelation is the correlation of these residuals and the input signal. If there is a sharp cutoff in the estimated autocorrelation function after lag k, the model can be identified as an MA(k). If the autocorrelation function tails off but the partial autocorrelation function shows a sharp cutoff after lag q, the model can be identified as an AR(q). If both functions tail off, then an ARMA model is to be used. If the autocorrelation function does not tail off or cut off, then the process is nonstationary. If this occurs, the data can be successively differenced until the resulting time series appears to be stationary. In this case, an ARIMA (autoregressive integrated moving average) model is identified. Since to difference the input signal increases the noise level, smoothing of the resulting signal may be necessary. There are many different criteria that can be used to select the order of the model. In general, they do not provide the same model order for the analysed series of data. The reader is referred to the works of Ljung [53] and Diggle [54] for a more detailed discussion.

If the system shows a nonlinear behaviour, it is possible to use either a linear model (this can be a good check for the relative importance of the nonlinear component in the system) or a nonlinear model (i.e. Wiener kernels). The reader is referred to the work of Lee and Schetzen [55], where a nonparametric method based on correlation techniques is introduced for the estimation of Wiener kernels. This method uses Gaussian white noise as the input to the system. The idea of using white noise as a stimulus to identify a system is based on the fact that the system is tested on all the possible inputs regarding values and frequencies (depending on the length of the test).

Frequency-domain methods: The frequency response of a system $H(j\omega)$ may be determined from an estimation of its transfer function $H(s)$ by setting the complex Laplace s parameter to $j\omega$. More commonly it can be determined from the time-domain signals by taking a Fourier transform (continuous or discrete) of the input $x(t)$ and output $y(t)$ signals, namely

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)}$$

where

$$Y(j\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(t)e^{-j\omega t}dt$$

$$X(j\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt \qquad (30)$$

It should be noted that the Fourier transform is a linear integral transform and $x(t)$ and $y(t)$ must be nontrivial (i.e. nonzero) to determine the frequency response using this method. When the input $x(t)$ is a periodic signal, the estimate of the frequency response is only of significance at the frequencies present in the input. When the input is not periodic (i.e. a realisation of a stochastic process), the quality of the estimate falls at those previous frequencies but is a better estimate at the other frequencies. The estimates at different frequencies are asymptotically uncorrelated. This makes the estimate of the frequency response relatively crude in practical situations [53].

Spectral analysis for determining transfer functions of linear systems was developed from statistical methods for spectral estimation. The reader is referred to the work of Brillinger [56] for a detailed account of the method. The only way to improve the poor variance properties of the transfer function estimate is to assume that the values of the true transfer function at different frequencies are related. Since the transfer function estimates at neighbouring frequencies are asymptotically uncorrelated, the variance can be reduced by averaging over these (i.e. using a

window such as Bartlett, Parzen or Hamming [57]). While a broad window leads to biased estimates and low variance, a narrow window leads to unbiased estimates but high variance (i.e. appearance of spurious peaks). Another way of smoothing the transfer function estimate is to split the data set into different subsets. The estimates over different subsets will be uncorrelated, and averages over these can be formed.

3.2.2 Parametric approach: In this approach, a set of candidate models is selected and parametrised as a model structure, using a parameter vector $\boldsymbol{\theta}$. The search for the best model within the set becomes a problem of determining or estimating $\boldsymbol{\theta}$. To do so, two main strategies can be considered: minimising prediction errors and correlating prediction errors with past data.

The first approach employs well known procedures such as the least-squares method and the maximum likelihood method, and is closely related to the Bayesian maximum *a posteriori* estimation. The second approach is based on the correlation between the prediction error and past data. Ideally, the prediction error of a good model should be independent of past data. A pragmatic way of checking this condition is that if the prediction error is correlated with the past data, then there was more information available in the past data about the actual output than was picked up by the model (predicted output). Therefore, the model was not ideal. See Ljung [53] for a detailed review of these methods.

The nonparametric approach introduced by Lee and Schetzen [55] for the estimation of the kernels that characterise a nonlinear system requires long data sequences for optimum performance. Short data sequences lead to significant errors in the estimated kernels. Haber [58] introduced a parametric method to estimate the kernels which reduces their variance, leading to a better estimation when short data series are available.

When using block-structured models, accurate kernel estimation is crucial for the identification of the topology of interconnection. Since the estimation of high order kernels is impractical, especially with short data series, the topology of the system is usually selected from a set of universal representations [59]. This selection can be based on a previous knowledge (or postulation) of the inner characteristics of the system or by performing a structural testing procedure introduced by Chen *et al.* [48]. If the system being studied does not satisfy the test criteria, the structure can be rejected and another selection can be made. On the other hand, if the system satisfies the test it cannot be concluded that it has this specific structure. Once the topology has been selected, the linear time-variant blocks can be identified using crosscorrelation techniques and the static nonlinear blocks are usually identified by fitting a polynomial [60].

4 Intelligent sensor systems

In this Section we briefly review the modelling techniques in the context of what we refer to as 'intelligent sensor systems'. The models and techniques used so far aim to enhance the sensor array selectivity, to reduce the time necessary for calibration (i.e. forecasting the steady-state response using the transient response) and to counteract drift. A summary of the main approaches is shown in Table 2. The main ones are discussed in more detail later. Before applying any technique to dynamically model the sensor system, sensors that are not relevant for the specific application, or that do not work properly, should be

Table 2: Types of modelling approaches in intelligent sensor systems

Modelling technique	Identification	Technology	Application	Ref.
Linear filters, and state-space models	Parametric ARMA, sensor oriented model	Thick-film SnO_2	Calibration time reduction	[7]
	Parametric ARX, sensor oriented	4 QMB polymer coated	Sensor response prediction	[69]
	Parametric Box–Jenkins, sensor oriented	10 MOSFETs 2 thick-film SnO_2	Identification of 2 gases. Drift rejection	[70]
	Parametric AR, sensor oriented	6 QMB polymer coated	Identification of 3 vapours	[41, 71]
	Parametric state-space model, system oriented	4 QMB polymer coated,	Quantitative analysis of ternary mixtures	[72]
	Parametric Box–Jenkins FIR, sensor oriented	2 SnO_2 6 BAW polymer coated	Quantitative analysis of 2 vapours	[73]
Functional expansions (nonlinear)	Nonparametric, correlation techniques, sensor oriented	6 QMB polymer coated	Sensor response prediction	[73, 74, 61]
	Parametric, sensor oriented	4 QMB polymer coated	Sensor response prediction	[69]
Block-structured	Parametric combining correlation and polynomial fitting	6 QMB polymer coated	Structure identification, response prediction	[69, 61]
Neural networks:				
SOM	Non-parametric,	Arrays of SnO_2,	Gas/aroma	[35, 75, 69]
Time-delay	system oriented,	MOSFET and	identification, drift	[8, 73, 76],
ART	adaptive	QMB polymer coated	rejection	[77, 62, 63]
Other techniques:				
Ad-hoc models through gas/aroma or temperature modulation	Parametric, sensor oriented models or FFT techniques	Metal oxides, conducting polymers, QMB polymer coated	Sensor selectivity enhancement, gas/aroma identification and quantification	[78, 79, 5, 6, 80, 4, 81–83, 84, 3]

eliminated. This requires careful exploration or 'pre-analysis' of the system. The use of classical techniques such as PCA may be very helpful in this preliminary stage.

4.1 Enhancing the sensor selectivity

To date most of the attempts to use transient information in the sensor signal are based on *ad hoc* models. These models allow for the estimation of parameters that characterise the transient response, conferring some selectivity on the sensors. Generally, an advantage of these models is that they account for physical and chemical properties of the sensing material (e.g. diffusion, reaction). Therefore, some insight into the sensors' dynamic behaviour can be realised. Their main weakness is that transient signals are influenced by previous measurements (memory effect) and by drift (e.g. ageing of the sensor, variations in ambient temperature or humidity). Since these aspects are not considered by the models, the initially learnt pattern recognition ability of a sensor system can deteriorate after a period of time. Fig. 9 shows the PCA results when an array of four thick-film metal oxide gas sensors were used to identify different volatile organic compounds and their binary mixtures [4]. The use of transient signals when the odour concentration varies stepwise helps in the identification task. The identification of single components, using a feedforward back-propagation trained neural network, gave a 76% success rate (using static signals only) and a 100% success rate (using both static and dynamic signals). The success rate in the identification of binary mixtures increased from 75% (using static signals) to 86% (using static and dynamic signals).

4.2 Calibration time reduction

Some applications of sensor response prediction aim to reduce the time necessary to calibrate the sensor array for the gases/odours of interest. Results with ARMA and *ad hoc* multiexponential models applied to the dynamic response of tin oxide sensor arrays have been reported [7]. The dynamic models were used to predict the static response of the sensors to small concentrations of nitrogen dioxide (0–9 ppm). Table 3 shows the relative errors made by the dynamic multiexponential model, which performed better in the extrapolation of the gas concentration. In this application, the prediction of the static response from the initial part of the dynamic response permits a reduction of the calibration time by a factor of four.

Table 3: Relative errors made by a multiexponetial model of a thick-film tin oxide gas sensor in extrapolation of concentration value of NO_2 at different calibration times

Time, s	Error at 1 ppm, %	Error at 6 ppm, %	Error at 9 ppm, %
100	55.2	35.8	13.1
200	17.1	4.7	7.1
400	7.8	2.5	3.7
800	1.3	0.3	0.6

Reprinted from DiNatale, C., *et al., Sens. Actuators B, Chem.*, **24–25**, pp. 578–583, © 1995 Elsevier Science, with permission

4.3 Response models

Dynamic measurements are interesting when the odours or the environmental conditions undergo changes with the same time-scale as the sensor response times, and also in the presence of sensor drift.

Methods of dealing with noise that allow for calculating the impulse response (of linear systems) or the Wiener kernels (of nonlinear systems), using the correlation approach, appear to be useful for constructing models for the sensor

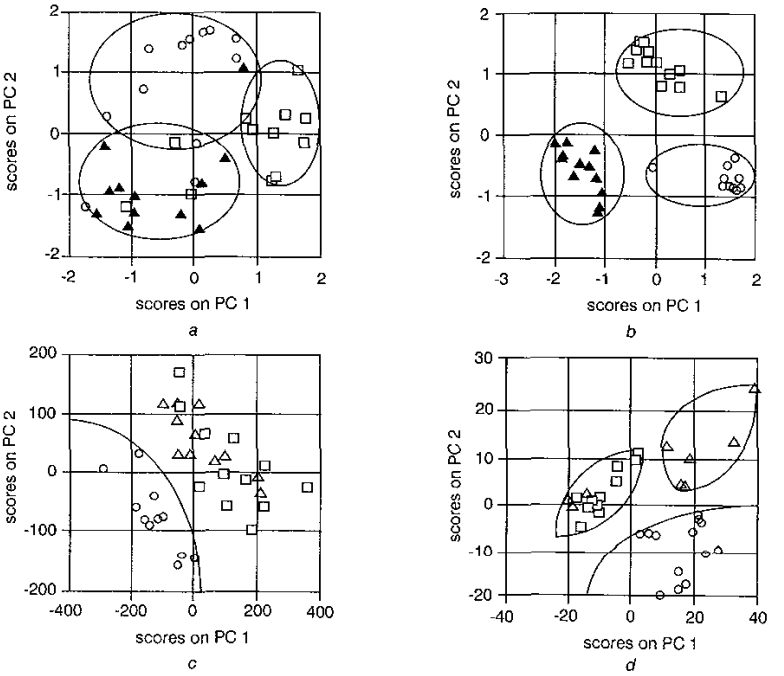

Fig. 9 *Results of PCA analysis of response of 4-element tin oxide electronic nose to three organic volatile compounds using static and dynamic signals and results of PCA of response to binary mixtures using static and dynamic signals*
(From Llobet, E., *et al.*, Proceedings of IEEE Solid-State Sensors and Actuators Conference, Transducers, **2**, pp. 971–974, © 1997 IEEE, with permission)
a, b □ ethanol
○ toluene
▲ o-xylene
c, d □ ethanol, toluene
△ ethanol, o-xylene
○ toluene, o-xylene

response to different odours. Linear filters that use lagged values of the input and the output to characterise the sensor (sensor oriented models) or the sensor array (system oriented models), identified using parametric approaches, such as the least-squares method, are also promising. In [10], Box–Jenkins linear filters were applied to model an array of metal oxide and MOSFET odour sensors in the presence of four alcohols and water vapour. For each sensor five models were created (one for each alcohol and one for water vapour). The classification was done in prediction error space, and the alcohol whose model gave the lowest total squared prediction error for all sensors was identified as the unknown odour (Bayesian approach). Fig. 10 shows the total sum squared prediction error for all sensors and for every model when the measured gas was 1-propanol. The 1-propanol model gives the lowest prediction error in almost all cases, leading to a correct classification, almost always. However, linear and nonlinear models constructed using input-output data (black-box models) do not give any insight into the inner structure of the sensors. In other words, it is not possible to discuss the identified model in terms of physical or chemical properties of the system. On the other hand, block-structured models are more related to the intrinsic characteristics of the sensing mechanisms.

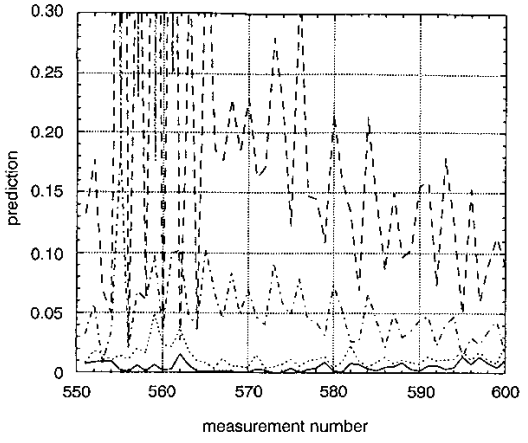

Fig. 10 *Prediction errors for all 5 models when measured gas was 1-propanol*
(Reprinted from Holmberg, M., *et al., Sens. Actuators B, Chem.*, **35–36**, pp. 528–535, © 1996 Elsevier Science, with permission)
——— 1-propanol model
– – – – 2-propanol model
····· 1-butanol model
—·—·— 2-butanol model
—···— water model

Fig. 11 shows the scheme of a two-input block-structured model of a polymer-coated quartz-microbalance sensor in the presence of n-octane and toluene [61]. The impulse response of the two linear blocks, which describe all the memory effects of the system, were estimated using the crosscorrelation approach. The static input-output nonlinearity was estimated by fitting a five-order polynomial. However, this method has not been widely applied because the identification of the model is complicated. In fact, the use of a nonparametric approach, such as the crosscorrelation method, to estimate the impulse response with low errors, requires long data sequences. This can result in time-consuming measurements to identify the sensor array or, even worse, can be impractical in some applications.

4.4 Drift counteraction

Since all of the approaches described above include memory effects, they are generally useful to address the problem of short-term drift (effects in the present response of the system due to measurements in its recent past). Another strategy consists of using SOMs with residual plasticity (i.e. adding a small constant to the expression of η in eqn. 13). This allows the network to deal effectively with small variations in the sensor response [35, 62].

Long-term drift caused by sensor poisoning or ageing implies that the system under identification is nonstationary. All the methods, except the neural network approach, assume that the sensor system is stationary and thus are not suitable to analyse the effects of long-term drift. It has been shown that SOMs with residual plasticity can help to maintain the pattern recognition ability of a sensor system affected by drift [62]. Fig. 12 shows the identification performances of an electronic nose based on six tin oxide gas sensors and static and adaptive SOMs. The gases measured were H_2, CO, CO_2, CH_4, and binary mixtures of both H_2, CO and CH_4, CO. It shows that if an adaptive SOM is used, the identification ability of the electronic nose remains almost unchanged when the drift in the sensor response is up to 20%. However, SOMs with residual plasticity require the frequent measurement of all the patterns. If this requirement is not fulfilled, patterns that seldom occur will be forgotten.

Very recently, in some preliminary work, ART neural networks have been proposed to deal with sensor drift [63]. The short-time memory of the network gives it some plasticity to adapt to sensor drift, while the long-time memory may give the necessary rigidity to avoid forgetting previously learnt patterns. ARTMAP (adaptive resonance

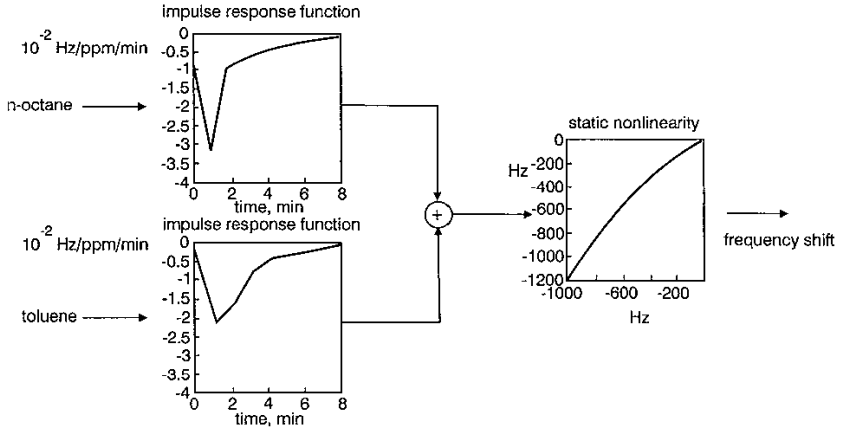

Fig. 11 *Complete scheme of estimated two-input Wiener model of a polymer-coated QMB*
(Reprinted from Davide, F., *et al., Sens. Actuators B, Chem.*, **24–25**, pp. 830–842, © 1995 Elsevier Science, with permission)

theory supervised predictive mapping) and fuzzy ARTMAP are nonparametric, adaptive networks that are well suited to solve pattern classification problems [64, 65]. With other adaptive algorithms, the learning of new events tends to wash away the memory traces of previous, but still useful, knowledge. ARTMAP and fuzzy ARTMAP contain a self-stabilising memory that permits accumulating knowledge to new events in a nonstationary environment [66]. Therefore, the use of these seems to be a promising strategy to address drift in intelligent sensor systems.

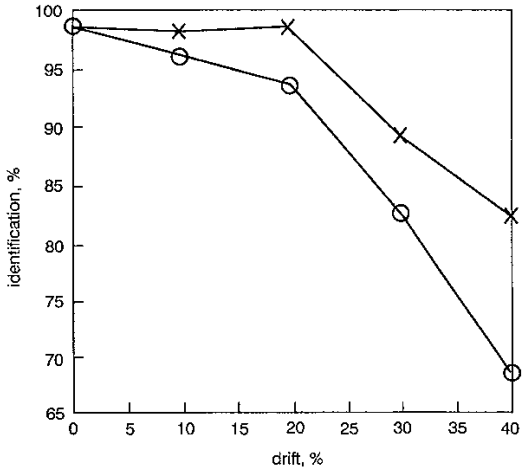

Fig. 12 *Comparison of identification performances of nonadaptive and adaptive SOMs in presence of simulated drift in response of a tin oxide gas sensor array*
(From Marco, S., *et al.*, Proceedings of IEEE Instrumentation and Measurement Technology Conference, pp. 904–907, © 1997 IEEE, with permission)
× adaptive SCM
○ static SCM

5 Outlook and conclusions

There is no universal sensor system that can solve all odour or gas mixture analysis problems. Instead there is a need to employ intelligent application-specific sensor systems that are appropriate to the application. This means building in intelligence through the development of suitable sensor structures, sensor materials and pattern recognition methods [1]. New pattern recognition methods should make use of the transient information in the sensor signal to enhance the identification ability of the system. This requires the use of dynamic models, for the sensor system, which can account for the drift in sensor parameters and thus extend the calibration period.

The importance of many problems associated with current chemical sensor technology is application specific. If the system has to analyse low levels of low reactive species, sensors tend to perform well. If the system has to analyse high levels of reactive species, poisoning of the sensors is likely and drift effects become very significant. The baseline of sensing devices (e.g. metal oxides, polymeric chemoresistors and polymer coated QMB) is sensitive to the operating temperature, the humidity and type of carrier gas [67]. Very often, the sensors require a long recovery time between measurements to reach their baseline. In continuous monitoring or repeated measurement applications, the response of the sensors is influenced by their previous history (short-term memory effect). Under these constraints, the choice of a suitable modelling strategy should be considered carefully:

- Nonadaptive models can be useful when the application implies the analysis of weakly reacting species with systems where temperature and humidity are strictly controlled by the sample delivery system (i.e. drift is likely to be small in such a system).
- Adaptive models are required when analysis of strongly reacting species is to be performed (i.e. the sensors are likely to drift due to poisoning). These models can also handle drift caused by slight variations in the temperature and humidity of the carrier gas.
- Of the nonadaptive models, *ad hoc* parametric models are interesting because they may give some insight into sensor behaviour. The measured parameters can be fed directly into well established pattern recognition systems. Linear filters and nonlinear models can be used to compensate for the short-term drift caused by the memory effect of the array when successive measurements are performed.
- The development of nonlinear, adaptive models in which competition between component gases occurs may best be solved using neural paradigms.
- SOMs with residual plasticity can be a good choice when frequent measurements of all the patterns are performed. When this condition is not fulfilled, the ART approach is a promising one.

These basic ideas are contained in Figs. 13 and 14, where the suitability of a specific dynamic model to a particular type of problem is shown. In Fig. 13 the sensor responses are considered to be linear or quasi-linear in concentration. This is generally the case when the species concentration is low, e.g. for conducting polymer resistive sensors, or when

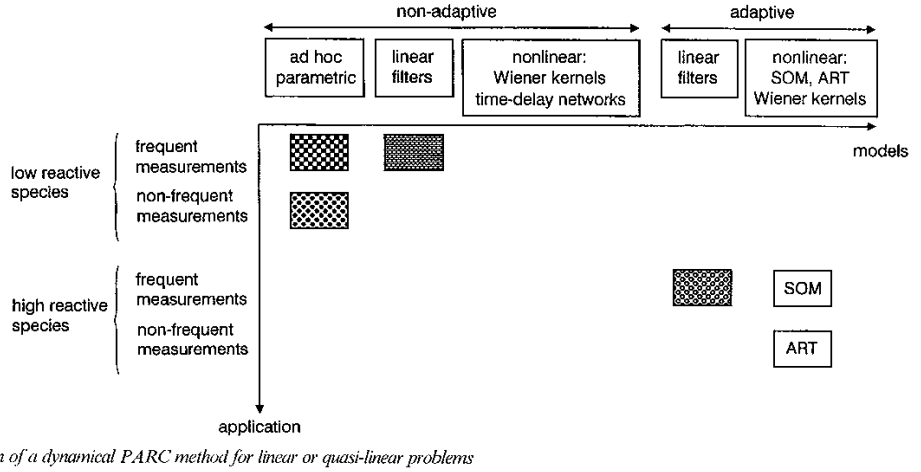

Fig. 13 *Selection of a dynamical PARC method for linear or quasi-linear problems*

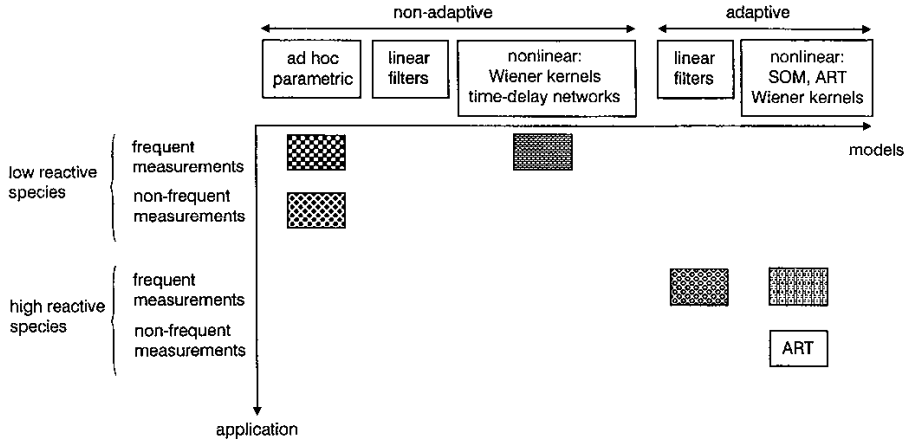

Fig. 14 *Selection of a dynamical PARC method for nonlinear problems*

the concentration range is small and so is step-wise approximately linear. If the sensor response is nonlinear in concentration in a well defined manner, a pre-processing linearisation algorithm can be used [68]. On the other hand, in Fig. 14 the selection assumes that the nonlinear part of the sensor response is important and must be accounted for in the models. The first attempts to use the dynamic sensor signals in electronic noses have essentially consisted of the development of *ad hoc* sensor-oriented parametric models. To develop a new generation of electronic noses, there is a need to extend these models taking into account the effects of environmental variables such as temperature and humidity, and to implement improved adaptive models to counteract sensor drift and poisoning.

6 Acknowledgments

The authors are grateful to Elsevier Science, The Institute of Electrical and Electronic Engineers and the Institute of Physics Publishing and associated authors for the permission to reproduce a table and a number of figures. Appropriate acknowledgment to the specific source is given in the captions where relevant. Dr. Llobet gratefully acknowledges financial support (grant ref. PR1997-0138) from the Spanish Ministry of Education and Culture during his stay at the University of Warwick.

7 References

1 GARDNER, J.W., and BARTLETT, P.N.: 'A brief history of electronic noses', *Sens. Actuators B, Chem.*, 1994, **18–19**, pp. 211–220
2 GARDNER, J.W., and HINES, E.L.: 'Pattern analysis techniques' *in* KRESS-ROGERS, E. (Ed.): 'Handbook of biosensors and electronic noses: Medicine, food and the environment' (CRC Press Inc., Boca Raton, 1997), pp. 633–652
3 CAVICCHI, R.E., SUEHLE, J.S., KREIDER, K.G., GAITAN, M., and CHAPARALA, P.: 'Optimized temperature-pulse sequences for the enhancement of chemically specific response patterns from microhotplate gas sensors', *Sens. Actuators B, Chem.*, 1996, **33**, pp. 142–146
4 LLOBET, E., BREZMES, J., VILANOVA, X., FONDEVILA, L., and CORREIG, X.: 'Quantitative vapor analysis using the transient response of non-selective thick-film tin oxide gas sensors'. Proceedings of IEEE Transducers'97, Chicago, USA, 1997
5 AUERBACH, F.J.: 'Pattern recognition using gas modulation'. Proceedings of Transducers'95-Eurosensors IX, Stockholm, Sweden, 1995
6 WILSON, D.M., and DEWEERTH, S.P.: 'Odor discrimination using steady-state and transient characteristics of tin-oxide sensors', *Sens. Actuators B, Chem.*, 1995, **28**, pp. 123–128
7 DINATALE, C., MARCO, S., DAVIDE, F., and D'AMICO, A.: 'Sensor-array calibration time reduction by dynamic modelling', *Sens. Actuators B, Chem.*, 1995, **24–25**, pp. 578–583
8 SCHWEIZER, M., GOPPERT, J., HIERLEMANN, A., MITROVICS, J., WEIMAR, U., ROSENTIEL, W., and GÖPEL, W.: 'Application of neural-network systems to the dynamic response of polymer-based sensor arrays', *Sens. Actuators B, Chem.*, 1995, **26–27**, pp. 232–236
9 ROTH, M., HARTLINGER, R., FAUL, R., and ENDRES, H.E.: 'Drift reduction of organic coated gas-sensors by temperature modulation', *Sens. Actuators B, Chem.*, 1996, **35–36**, pp. 358–362
10 HOLMBERG, M., WINQUIST, F., LUNDSTRÖM, I., DAVIDE, F., DINATALE, C., and D'AMICO, A.: 'Drift counteraction for an electronic nose', *Sens. Actuators B, Chem.*, 1996, **35–36**, pp. 528–535
11 GARDNER, J.W.: 'Detection of vapours and odours from a multisensor array using pattern recognition. Part 1: Principal components and cluster analyses', *Sens. Actuators B, Chem.*, 1991, **4**, pp. 108–116
12 PATRICK, W., BEEBE, K.R., SANCHEZ, E., GELADI, P., and KOWALSKI, B.R.: 'Analysis of multisensor arrays', *Sens. Actuators B, Chem.*, 1986, **9**, pp. 223–234
13 LORBER, A., WANGEN, L.E., and KOWALSKI, A.: 'A theoretical foundation for the PLS algorithm', *J. Chemometr.*, 1987, **1**, pp. 19
14 DRAPER, N., and SMITH, H.: 'Applied regression analysis' (Wiley, New York, 1981)
15 GARDNER, J.W., SHURMER, H.V., and TAN, T.T.: 'Application of an electronic nose to the discrimination of coffees', *Sens. Actuators B, Chem.*, 1992, **6**, pp. 71–75
16 CAREY, W.P., and YEE, S.S.: 'Calibration of nonlinear solid-state sensor arrays using multivariate regression techniques', *Sens. Actuators B, Chem.*, 1992, **9**, pp. 113–122
17 GARDNER, J.W., and BARTLETT, P.N.: 'Pattern recognition in odour sensing' *in* BARTLETT, P.N. and GARDNER J.W. (Eds.): 'Sensors and sensory systems for an electronic nose' (Kluwer, Dordrecht, 1991)
18 ABE, H., YOSHIMURA, T.Y., KANAYA, S., TAKAHASHI, Y., MIYASHITA, Y., and SASAKI, S.: 'Automated odor-sensing system based on plural semiconductor gas sensors and computerised pattern recognition techniques', *Anal. Chim. Acta*, 1987, **194**, pp. 1–9
19 SHURMER, H.V., and GARDNER, J.W.: 'Intelligent vapour discrimination using a composite 12-element sensor array', *Sens. Actuators B, Chem.*, 1990, **1**, pp. 256–260
20 AISHIMA, T.: 'Aroma discrimination by pattern recognition analysis of responses from semiconductor gas sensor arrays', *J. Agri. Food Chem.*, 1991, **39**, pp. 752–756
21 BARTLETT, P.N., BLAIR, N., and GARDNER, J.W.: 'Electronic noses, principles, applications and outlook'. Proceedings of 15th Colloque Assoc. Sci. Int. Café, Montpellier, France, 1993
22 GARDNER, J.W., HINES, E.L., and WILKINSON, M.: 'The application of artificial neural networks in an electronic nose', *Meas. Sci. Technol.*, 1990, **1**, pp. 446–451
23 LIPPMANN, R.P.: 'Introduction to computing with neural nets', *IEEE ASSP Mag.*, 1987, **4**, (2), pp. 4–22
24 SUNDGREN, H., WINQUIST, F., LUKKARI, I., and LUNDSTRÖM, I.: 'Artificial neural networks and gas sensor arrays: quantification of individual components in a gas mixture', *Meas. Sci. Technol.*, 1991, **2**, pp. 464–469
25 NAKAMOTO, T., FUKUDA, A., and MORIIZUMI, T.: 'Perfume and flavor identification by odor sensing system using quartz resonator sensor array and neural network pattern recognition', *Sens. Actuators B, Chem.*, 1993, **10**, pp. 85–90
26 GOLDBERG, D.E.: 'Genetic algorithms in search, optimization and machine learning' (Addison Wesley, Reading, MA, 1989)
27 SRIVASTAVA, A.K., SHUKLA, K.K., and SRIVASTAVA, S.K.: 'Exploring neurogenetic processing of electronic nose data', *Microelectron. J.*, 1998, **29**, pp. 921–931
28 FEKADU, A.A., HINES, E.L., and GARDNER, J.W.: 'Genetic algorithm design of neural network based electronic nose' *in* ALBRECHT, R.F., and STEELE, N.C. (Eds.): 'Artificial neural nets and genetic algorithms' (Springer-Verlag, New York, 1993) pp. 691–698
29 SINGH, S., HINES, E.L., and GARDNER, J.W.: 'Fuzzy neural computing of coffee and tainted-water data from an electronic nose', *Sens. Actuators B, Chem.*, 1996, **30**, pp. 185–190

30 GUPTA, M.M., and QI, J.: 'On fuzzy neuron models' *in* ZADEH, L.A. (Ed.): 'Fuzzy logic for the management of uncertainity' (John Wiley, New York, 1992), pp. 479–490
31 McNEIL, D.F.: 'Fuzzy Logic' (Touchstone Books, New York, 1993)
32 KOHONEN, T.: 'Self-organized formation of topologically correct feature maps', *Biol. Cybern.,* 1982, **43**, pp. 59–69
33 KOHONEN, T.: 'Learning vector quantization for pattern recognition'. Helsinki University of Technology, Helsinki, 1986
34 LEMOS, R.A., HAKAMURA, M., SUGIMOTO, I., and KUWANO, H.: 'A self-organizing map for chemical vapor classification'. Proceedings of 7th international conference *Solid-state sensors and actuators* (Transducers'93), IEEE, Yokohama, Japan, 1993
35 DAVIDE, F., DINATALE, C., and D'AMICO, A.: 'Self-organizing multisensor system for odour classification: internal categorization, adaptation and drift rejection', *Sens. Actuators B, Chem.,* 1994, **18–19**, pp. 244–258
36 HINES, E.L., GARDNER, J.W., and POTTER, C.E.R.: 'Olfactory feature maps for an electronic nose', *Meas. Control,* 1997, **30**, pp. 262–268
37 DINATALE, C., DAVIDE, F., and D'AMICO, A.: 'Pattern-recognition in gas sensing: well-stated techniques and advances', *Sens. Actuators B, Chem.,* 1995, **23**, pp. 111–118
38 GROSSBERG, S.: 'Adaptive pattern classification and universal recoding, II: Feedback, expectation olfaction and illusions', *Biol. Cybern.,* 1976, **23**, pp. 187–202
39 CARPENTER, G.A., and GROSSBERG, S.: 'A massively parallel architecture for a self-organizing neural pattern-recognition machine', *Comput. Vis. Graph. Image Process.,* 1987, **37**, pp. 116–165
40 GARDNER, J.W., HINES, E.L., and PANG, C.: 'Detection of vapours and odours from a multisensor array using pattern recognition: self-organising adaptive resonant techniques', *Meas. Control,* 1996, **29**, pp. 172–178
41 NAKAMURA, M., SUGIMOTO, I., KUWANO, H., and LEMOS, R.: 'Chemical sensing by analysing dynamics of plasma polymer film-coated sensors', *Sens. Actuators B, Chem.,* 1994, **20**, pp. 231–237
42 KALMAN, R.E.: 'On the general theory of control systems' *in* 'First IFAC Congress' (Butterworths, Moscow, London, 1960)
43 TONG, H.: 'Non-linear times series. A dynamical system approach' (Clarendon Press, Oxford, 1990), Chap. 3, pp. 96–120
44 PRIETSLEY, M.B.: 'State dependent models: a general approach to non-linear time series analysis', *J. Time Ser. Anal.,* 1980, **1**, pp. 57–71
45 VOLTERRA, V.: 'Theory of functionals and of integro-differential equations' (Dover, New York, 1930)
46 WIENER, N.: 'Non-linear problems in random theory' (Wiley, New York, 1958)
47 CHEN, H.W., ISHII, N., and SUZUMURA, N.: 'Structural classification of non-linear systems by input and output measurements', *Int. J. Syst. Sci.,* 1986, **17**, pp. 371–377
48 CHEN, H.W., JACOBSON, L.D., and GASKA, J.P.: 'Structural classification of multi-input non-linear systems', *Biol. Cybern.,* 1990, **63**, pp. 341–357
49 BILLINGS, S.A., JAMALUDDIN, H.B., and CHEN, S.: 'Properties of neural networks with applications to modelling non-linear dynamical systems', *Int. J. Control,* 1992, **55**, pp. 193–224
50 KUSCHEWSKI, J.G., HUI, S., and ZAK, S.H.: 'Applications of feedforward neural networks to dynamical system identification and control', *IEEE Trans. Control Syst. Technol.,* 1993, **1**, pp. 37–49
51 GILES, C.L., KUHN, G.M., and WILLIAMS, R.J.: 'Dynamic recurrent neural networks', *IEEE Trans. Neural Netw.,* 1994, **5**, pp. 153–55
52 QIN, S.Z., SU, H.T., and MCAVOY, T.J.: 'Comparison of four neural network learning methods for dynamic system identification', *IEEE Trans. Neural Netw.,* 1992, **1**, pp. 123–130
53 LJUNG, L.: 'System identification: Theory for the user' (Prentice Hall, Englewood Cliffs, NJ, 1987), Chaps 6 and 8–11
54 DIGGLE, P.J.: 'Times series: A biostatistical introduction' (Clarendon Press, Oxford, 1990), Chaps 2 and 6
55 LEE, Y.W., and SCHETZEN, M.: 'Measurement of Wiener kernels of a non-linear system by using cross-correlation', *Int. J. Control,* 1965, **2**, pp. 237–254
56 BRILLINGER, D.R.: 'Time series: Data analysis and theory' (Holden Day, San Francisco, CA, 1981), Chap. 6
57 PROAKIS, J.G., and MANOLAKIS, D.G.: 'Introduction to digital signal processing' (Macmillan Inc., New York, 1988), Chap. 8, p. 549
58 HABER, R.: 'Parametric identification of non-linear dynamic systems based on non-linear cross-correlation functions', *IEE Proc. D,* 1988, **135**, pp. 405–420
59 KOREMBERG, M.J.: 'Identifying noisy cascades of linear and static nonlinear systems'. Proceedings of 7th IFAC symposium on *Identification and system parameter estimation*, York, 1985
60 WESTWICK, D.T., and KEARNEY, R.E.: 'A new algorithm for the identification multiple input Wiener systems', *Biol. Cybern.,* 1983, **68**, pp. 75–85
61 DAVIDE, F., DINATALE, C., D'AMICO, A., HIERLEMANN, A., MITROVICS, J., SCHWEIZER, M., WEIMAR, U., and GÖPEL, W.: 'Structure identification of non-linear models for QMB polymer-coated sensors', *Sens. Actuators B, Chem.,* 1995, **24–25**, pp. 830–842
62 MARCO, S., PARDO, A., ORTEGA, A., and SAMITIER, J.: 'Gas identification with tin oxide sensor array and self organizing maps: adaptive correction of sensor drifts'. Proceedings of IEEE *Instrumentation and measurement technology* conference, Ottawa, Canada, 1997
63 VLACHOS, D.S., FRAGOULIS, D.K., and AVARITSIOTIS, J.N.: 'An adaptive neural network topology for degradation compensation of thin-film tin oxide gas sensors', *Sens. Actuators B, Chem.,* 1997, **45**, pp. 223–228
64 CARPENTER, G.A., GROSSBERG, S., and REYNOLDS, J.H.: 'ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organizing neural network', *Neural Netw.,* 1991, **4**, pp. 565–588
65 CARPENTER, G.A., GROSSBERG, S., and REYNOLDS, J.H.: 'A fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems', *IEEE Trans. Neural Netw.,* 1995, **6**, (6), pp. 1330–1336
66 CARPENTER, G.A., and GROSSBERG, S.: 'Fuzzy ARTMAP: a synthesis of neural networks and fuzzy logic for supervised categorization and nonstationary prediction' *in* YAGER, R.R., and ZADEH, L.A. (Eds.): 'Fuzzy sets, neural networks, and soft computing' (Van Nostrand Reinhold, New York, 1994), pp. 125–165
67 SRIYUDTHSAK, M., PROMSONG, L., and PANYAKEOW, S.: 'Effect of carrier gas on response of oxide semiconductor gas sensor', *Sens. Actuators B, Chem.,* 1993, **13–14**, pp. 139–142
68 MOORE, S.W., GARDNER, J.W., HINES, E.L., GÖPEL, W., and WEIMAR, U.: 'A modified multilayer perceptron model for gas mixture analysis', *Sens. Actuators B, Chem.,* 1993, **15–16**, pp. 344–348
69 MARCO, S., PARDO, A., DAVIDE, F., DINATALE, C., D'AMICO, A., HIERLEMANN, J., SCHWEIZER, M., WEIMAR, U., and GÖPEL, W.: 'Different strategies for the identifiaction of gas sensing systems', *Sens. Actuators B, Chem.,* 1996, **34**, pp. 213–223
70 HOLMBERG, M., DAVIDE, F., DINATALE, C., D'AMICO, A., WINQUIST, F., and LUNDSTRÖM, I.: 'Drift counteraction in odour recognition applications: Lifelong calibration method', *Sens. Actuators B, Chem.,* 1996,
71 NAKAMURA, M., SUGIMOTO, I., and KUWANO, H.: 'Application of plasma-polymer-film-coated sensors to gas identification using linear filters', *Sens. Actuators B, Chem.,* 1996, **33**, pp. 122–127
72 NAKAMOTO, T., OKAZAKI, N., and MORNZUMI, T.: 'High-speed active gas-odor sensing system using adaptive control theory', *Sens. Actuators B, Chem.,* 1997, **41**, pp. 183–188
73 PARDO, A., MARCO, S., and SAMITIER, J.: 'Dynamic measurements with chemical sensor arrays based on inverse modelling',Proceedings of IEEE *Instrumentation and measurement technology* conference, Brussels, Belgium, 1996
74 DAVIDE, F., DINATALE, C., D'AMICO, A., HIERLEMANN, A., MITROVICS, J., SCHWEIZER, M., WEIMAR, U., GÖPEL, W., MARCO, S., and PARDO, A.: 'Dynamic calibration of QMB polymer-coated sensors by Wiener kernel estimation', *Sens. Actuators B, Chem.,* 1995, **26–27**, pp. 275–285
75 DINATALE, C., DAVIDE, F., and D'AMICO, A.: 'A self-organizing system for pattern classification: time varying statistics and sensor drift effects', *Sens. Actuators B, Chem.,* 1995, **26–27**, pp. 237–241
76 DINATALE, C., MACAGNANO, A., D'AMICO, A., and DAVIDE, F.: 'Electronic-nose modelling and data analysis using a self-organizing map', *Meas. Sci. Technol.,* 1997, **8**, pp. 1236–1243
77 SAUNDERS, B.W., THIEL, D.V., and MACKAY-SIM, A.: 'An artificial olfactory system using tiered artificial neural networks'. Proceedings of IEEE Australian and New Zealand conference on *Intelligent information systems*, 1994, pp. 76–80
78 IPPOMMATSU, M., and SASAKI, H.: 'Dynamic properties of SnO_2 semiconductor gas sensors sensing of hydrogen contained in air', *J. Electrochem. Soc.,* 1989, **136**, (7), pp. 2123–2128
79 ENDRES, H.E., GOTTLER, W., JANDER, H.D., DROST, S., SBERVEGLIERI, G., FAGLIA, G., and PEREGO, C.: 'A systematic investigation on the use of time dependent sensor signals in signal processing techniques', *Sens. Actuators B, Chem.,* 1995, **24–25**, pp. 785–789
80 VILANOVA, X., LLOBET, E., ALCUBILLA, R., SUEIRAS, J.E., and CORREIG, X.: 'Analysis of the conductance transient in thick-film tin oxide gas sensors', *Sens. Actuators B, Chem.,* 1996, **31**, pp. 175–180
81 GARDNER, J.W., BARTLETT, P.N., HINES, E.L., MOLINIER, F., and MOTTRAM, T.T.: 'Prediction of the health of dairy cattle from breath samples using a neural network with parametric model of the dynamic response of an array of semiconducting gas sensors', *IEE Proc., Sci. Meas. Technol.,* 1999, **146**, pp. 102–106
82 NANTO, H., KONDO, K., HABARA, M., DOUGUCHI, Y., WAITE, R.I., and NAKAZUMI, H.: 'Identification of aromas from alcohols using a Japanese-film-coated quartz resonator gas sensor in conjunction with pattern recognition analysis', *Sens. Actuators B, Chem.,* 1996, **35–36**, pp. 183–186
83 SEARS, W.M., COLBOW, K., SLAMKA, R., and CONSADORI, F.: 'Selective thermally cycled gas sensing using fast Fourier transform techniques', *Sens. Actuators B, Chem.,* 1990, **2**, pp. 283–289
84 YEA, B., KONISHI, R., SUGAHARA, K., and OSAKI, T.: 'An advanced gas discrimination method utilizing the periodic operation of a semiconductor gas sensor'. Proceedings of IEEE conference on *Industrial automation and control*, 1995