Full Paper

Monitoring the Freshness of Moroccan Sardines with a Neural-Network Based Electronic Nose

Aziz Amari [1], Noureddine El Barbri [1], Eduard Llobet [2], Nezha El Bari [3], Xavier Correig [2] and Benachir Bouchikhi [1],*

1 Sensor Electronic & Instrumentation Group, Faculty of Sciences, Physics Department, University Moulay Ismaïl, B.P. 11201, Zitoune, Meekness, Morocco

2 MINOS, Microsystems and Nanotechnologies for Chemical Analysis, Universitat Rovira i Virgili, Avda. Països Catalans, 26, 43007 Tarragona, Spain

3 Biotechnology Agroalimentary and Biomedical Analysis Group, Faculty of Sciences, Biology Department, University Moulay Ismaïl, B.P. 11201, Zitoune, Meekness, Morocco
E-mail: eduard.llobet@urv.cat

* Author to whom correspondence should be addressed; E-mail: benachir.bouchikhi1@caramail.com

Abstract: An electronic nose was developed and used as a rapid technique to classify the freshness of sardine samples according to the number of days spent under cold storage (4 ± 1°C, in air). The volatile compounds present in the headspace of weighted sardine samples were introduced into a sensor chamber and the response signals of the sensors were recorded as a function of time. Commercially available gas sensors based on metal oxide semiconductors were used and both static and dynamic features from the sensor conductance response were input to the pattern recognition engine. Data analysis was performed by three different pattern recognition methods such as probabilistic neural networks (PNN), fuzzy ARTMAP neural networks (FANN) and support vector machines (SVM). The objective of this study was to find, among these three pattern recognition methods, the most suitable one for accurately identifying the days of cold storage undergone by sardine samples. The results show that the electronic nose can monitor the freshness of sardine samples stored at 4°C, and that the best classification and prediction are obtained with SVM neural network. The SVM approach shows improved classification performances, reducing the amount of misclassified samples down to 3.75 %.

1. Introduction

Fishing is an important industry in Morocco, and the waters off the coast of this country are rich in fish. The fish catch includes sardines, tuna, mackerel, anchovies and shellfish. Sardines represent 80 % of all fish caught each year, and, according to the statistics of the National Office for Fishing (2004), 42.14 % of all fish exported. Much of the catch is processed in Morocco –fresh, frozen or canned– for export. One of the unique characteristics of fish as food is that it is a highly perishable commodity [1]. In fact, the time spent after its catch and its temperature 'history' are very often the key factors determining the final quality characteristics of fish or fish products. Consequently, there is a need for a technique to monitor freshness and quality of fish (Sardine). As an important indicator of fish freshness, analysis of odours has traditionally been performed either by sensory panels or by gas chromatography coupled to mass spectroscopy, which are time consuming and costly [2, 3]. To this end, the electronic nose has proven to be a rapid and non-destructive method for evaluating volatile compounds, which are derived from spoilage odours in fish [4, 5].

The two main components of an electronic nose are the sensing system and the automated pattern recognition system. The sensing system can be an array of different sensing elements (e.g. gas sensors), where either each element measures a different property of the sensed odour or, more often, the sensors respond to a complex odour with overlapped sensitivity, which results in the sensor array producing a characteristic signature or pattern. Thus, presenting the odour of many samples of sardines at different times of cold storage to the sensor array will result in a database of signatures being built up [6]. This database of signatures will be then used to train the pattern recognition system. The goal of this training process is to configure the recognition system to produce unique classifications of each sample of sardines, so that an automated identification can be implemented. The pattern recognition system comprises the feature extraction step, which extracts useful information from the sensor responses and the pattern recognition algorithm. In supervised pattern recognition, input patterns are learned and associated with an odour class. In an unsupervised pattern recognition technique, the multidimensional configuration space from the pre-processor is converted into a feature space. The unknown sardine sample is identified by comparison to a knowledge base, from a previous learning scheme. The pattern recognition algorithms have become a critical component in the successful implementation of gas sensor arrays and electronic noses. Among many pattern recognition techniques, we are interested in those based on artificial neural networks (ANN).

Similar to the brain of mammals, an ANN is composed of neurones, regarded as the processing units, and the massive interconnection among them. ANN have the unique ability to learn from examples and to generalise, i.e., to produce reasonable outputs for new inputs not encountered during a learning process [7]. Electronic noses that incorporate ANNs have been demonstrated in various applications. Many ANN configurations and training algorithms have been used to build up electronic noses [8] including probabilistic neural networks (PNN), backpropagation-trained feed-forward

networks, fuzzy ARTMAP neural networks (FANN), Kohonen's self-organising maps (SOMs), learning vector quantization (LVQs), Boltzmann machines, Hopfield networks and support vector machines (SVM).

In this study, we introduce an electronic nose system capable of discriminating sardine samples according to the number of days spent under cold storage. The system is composed of an array of six commercially available metal oxide semiconductor gas sensors and an artificial neural network as the pattern recognition algorithm. Three different neural networks (PNN; FANN and SVM) have been evaluated in the classification of data obtained with our electronic nose measurement system.

2. Experimental

2.1. Electronic nose set-up

The electronic nose system was designed to be able to control the quality of sardine samples during cold storage (up to 15 days). Three main parts compose this system: The sampling system, the sensor chamber and the data acquisition system. An overall view of the system is shown in figure 1. The sampling system consists of a dynamic headspace sampling. The sensor array comprised six metal oxide gas sensors (TGS type from Figaro Inc.), a capacitive humidity sensor PHILIPS H1 from Philips and a temperature sensor LM35DZ from National Semiconductor. The identification codes of the Figaro sensors that were used are the following, where the target gases as suggested by the manufacturer are indicated: TGS 823 (Alcohols, Xylene and Toluene), TGS 825 (H_2S), TGS 826 (NH_3), TGS 831 (Chlorofluorocarbons), TGS 832 (Halocarbons) and TGS 882 (Alcohols) [9].
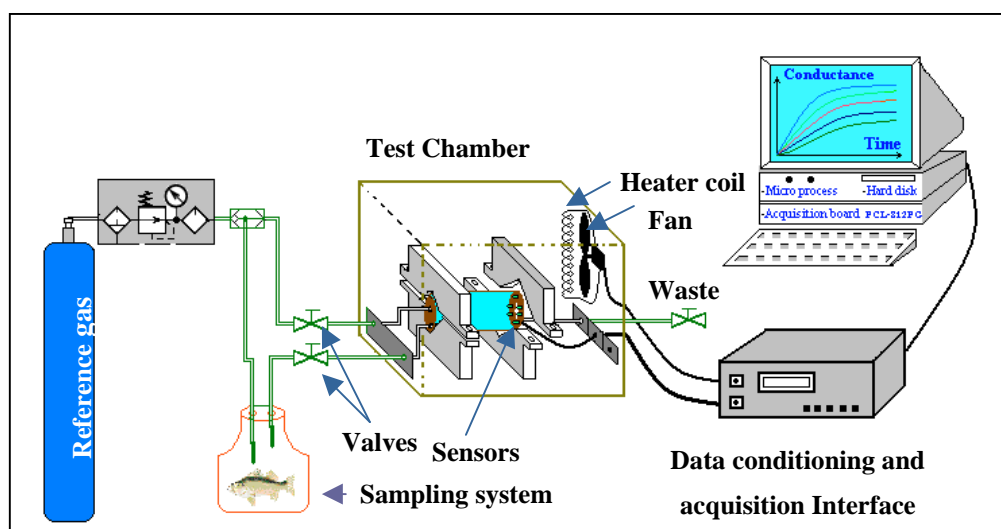

Figure 1. Structure of the experimental system.

Exposure of a metal oxide semiconductor sensor to an odour (e.g. fish odour) produces a large change in its electrical resistance, the circuit conditioning used to measure the conductance [10], is shown in fig.2 (a). The relationship between the sensor resistance and the voltage monitored is expressed by the following equation:

$$R_S = R_L \left(\frac{V_C}{V_{RL}} - 1 \right)$$ (1)

The humidity and temperature sensors are used to monitor the condition of the experiments. The electrical response of the eight sensors was monitored by an acquisition data board (PCL-812) installed in a Compatible Personnel Computer. Pure nitrogen was used as carrier gas. A fan and a heating coil were incorporated into the sensor chamber to keep constant the temperature between 32°C and 33°C.

2.2. Procedure

Sardines used in the experiments have been bought from the local fish market of the city of Meekness (Morocco). The samples were placed in plastic boxes and kept under cold storage at 4 ± 1°C before being analysed. Samples were analysed after 1, 3, 5, 7, 9, 11, 13 and 15 days of cold storage. Ten samples per each storage day were measured, in total 80 sardine samples were analysed. Each sample of sardine was taken from the refrigerator and kept in a sealed sampling vessel 40 min before every experiment. This ensured the homogeneity of sample preparation. For each day of experiment, the sensors were supplied with a 10 V circuit voltage and a 5 V heating voltage 50 minutes before the first measurement in order to stabilise them, and were let powered for all the measurements. The measurement procedure went through the following cycle: first, the carrier gas (nitrogen) was allowed to flow through the sensor chamber for 50 minutes so that the sensors could stabilise their baseline resistance. Afterwards, a valve allowed the nitrogen to flow through the sampling vessel and the sensor chamber for another 50 minutes. The response of the sensors was collected and stored every 1 s. After each measurement, the sensor chamber was opened and cleaned with nitrogen gas to regenerate the sensors' baseline. Right after that, a new measurement with another sardine sample can be started.

2.3. Data acquisition and pre-processing

Data acquisition was controlled from a Compatible PC and the signal of each sensor was recorded as a function of time. The baseline conductance of the gas sensors, $G_{0i}(t)$ $(i = 1,...,N)$, and $N = 6$, was defined as the steady-state conductance when pure nitrogen flowed through the sensor chamber. The acquisition started before the nitrogen flow was derived via the sampling vessel. The time evolution of every sensor signal, $G_{Si}(t)$, where $G_{Si}(t)$ is the conductance of sensor i while the headspace of a sardine is being sampled, was recorded until all sensors reached the steady state. The response of the sensors was acquired and stored every 1 s for 50 minutes. The associated baselines were then subtracted for each sensor and the difference signal $G_i(t)$, was calculated as:

$$G_i(t) = G_{si}(t) - G_{0i}(t) \qquad (i = 1,...,N)$$ (2)

The data set consisted of 80 measurements, which were assigned to 8 different classes according to the number of days the sample had spent under cold storage (1, 3, 5, 7, 9, 11, 13 and 15). Each class comprised a total of 10 measurements (i.e. 10 samples per category were measured). To analyse the response of the electronic nose, four parameters were extracted from each sensor conductance transient $G_i(t)$ (Fig. 2(b)):

-    The average value of the conductance measured in the interval [0 - 15 min], $G_{0(0 \to 15\,min)}$;

- The slope of the dynamic conductance calculated from the conductance curve within the interval [15 - 35 min], $\left(dG/dt \right)_{(15 \to 35\,\text{min})}$;

- The area below the conductance curve between 15 and 40 minutes calculated by the trapeze method, (Trapeze area conductance)$_{15 \to 40\,\text{min}}$;

- The steady-state conductance calculated from the average value of the conductance during the last 5 minutes $G_{S(45 \to 50\,\text{min})}$.

- Given that 4 features were extracted from the response of each sensor and that there were 6 sensors within the array, each measurement was described by 24 features, the data matrix had 80 rows (i.e. measurements) and 24 columns (i.e. variables).

| (a) | (b) |
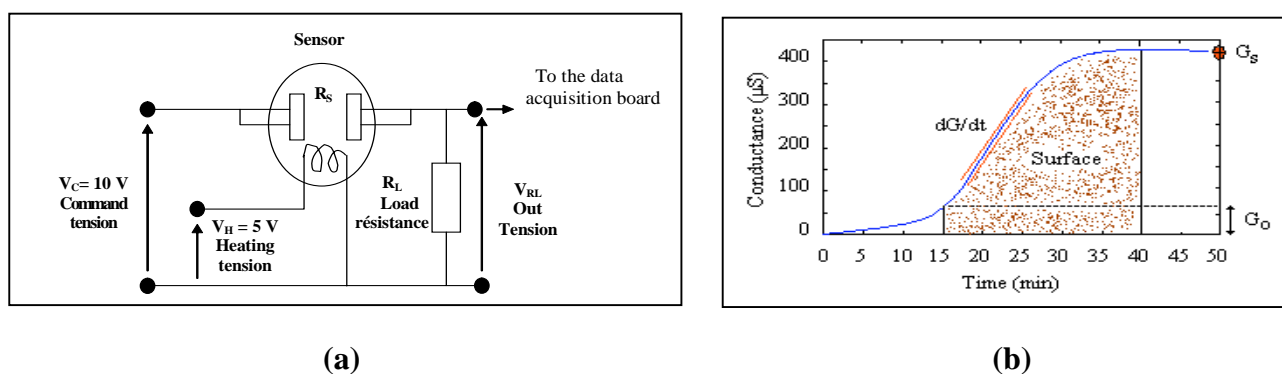
Figure 2. a) Electronic circuit to monitor the sensor response, b) Typical conductance transient response $G_i(t)$ from semiconductor gas sensors and the 4 response features extracted.

3. Artificial Neural Networks

Neural networks offer a chemometric technique of great potential for the treatment of signals generated by electronic noses based on sensors that can provide non-linear responses. Several types of classifiers based on different ANN-algorithms such as multi-layer perceptron (MLP), radial basis function (RBF), self-organising map (SOM), PNN, FANN and SVM have been used in many electronic nose applications [11-15]. The nature of an ANN-algorithm is usually classified as a supervised or unsupervised method.

In a supervised learning ANN method, a set of known odours are systematically introduced to the electronic nose, which then classifies them according to known descriptors or classes held in a knowledge base. Then, in a second stage for identification, an unknown odour is tested against the knowledge base, now containing the learnt relationship, and then the class membership is predicted.

For unsupervised learning, ANN methods learn to separate the different classes from the response vectors without any additional information. In other words, unsupervised methods discriminate between unknown odour vectors without being presented with the corresponding descriptors. These methods are closer to the way that the human olfactory system works. They both use intuitive associations with no, or little, prior knowledge.

In order to evaluate their performance in the classification of sardine samples, three supervised learning approaches (PNN, FANN and SVM) were applied to analyse the data set described above.

3.1. Probabilistic neural network (PNN)

The probabilistic neural network was developed by Donald Specht. This network provides a general solution to pattern classification problems by following an approach developed in statistics called Bayesian classifiers. Bayes' theory, developed in the 1950's, takes into account the relative probability of events and uses a priori information to improve prediction. The probabilistic neural network uses a supervised training set to develop distribution functions within a pattern layer [8, 16]. The Probabilistic Neural Networks (PNNs) approach combines both Bayes theorem of conditional probability and Parzen's method for estimating the probability density functions of random variables. Unlike other neural network training paradigms, PNNs are characterised by high training speed and their ability to produce confidence levels for their classification decisions. For theoretical background and details of the algorithm, the reader is referred to [16-18]. The PNN was first introduced by Specht [18], who showed how the Bayes–Parzen classifier could be broken up into a large number of simple processes implemented in a multilayer neural network which could be run independently in parallel.

3.2. Fuzzy ARTMAP neural networks (FANN)

The Fuzzy ARTMAP neural network is a supervised pattern recognition method based on fuzzy adaptive resonance theory (ART). It is a promising method since fuzzy ARTMAP is able to carry out on-line learning without forgetting previously learnt patterns (stable learning). It can also recode previously learnt categories (adaptive to changes in the environment) and is self-organising [19]. In its most general form, Fuzzy ARTMAP includes two fuzzy ART modules (ART_a and ART_b) interconnected by an associative memory and some internal control structures that regulate learning and information flow [20]. During supervised learning ART_a receives a stream of input patterns $\{a^M\}$ and ART_b also receives a stream of patterns $\{b^M\}$, where b^M is the correct prediction given a^M. When a prediction by ART_a is not confirmed by ART_b, inhibition of the inter-ART associative memory activates a match tracking process. This increases ART_a vigilance by the minimum amount needed for the system either to activate an ART_a category that matches the ART_b category or to learn a new ART_a category.

3.3. Support vector machines (SVM)

In the last decade, a new classification technique called support vector machines (SVM) has been proposed in the broad learning. SVM, which is based on statistical learning theory (SLT), has recently been introduced as a new technique for solving a variety of learning classification and prediction problems [21, 22]. SVM have been successfully applied to a number of problems ranging from face identification and text categorisation to bioinformatics and data mining.

Support vector machine (SVM) is a statistical classification method proposed by Vapnik in 1995 [22]. The main idea of SVM is to separate the classes with a particular hyperplane, which maximises a quantity called *margin*. The margin is the distance from a hyperplane separating the classes to the nearest point in the dataset (Fig. 3).

In the separable case, a hyperplane, ω separates positive and negative samples. Those data x that satisfy $\langle w, x \rangle + b = 0$ lie on the hypeplane where ω is normal to the hyperplane ($\omega \in R^n$ and $b \in R$), $b/\|\omega\|$ is the perpendicular distance from the hyperplane to the origin. Samples from positive and negative training samples satisfy: $y_i(\langle \omega, x_i \rangle + b) \geq 1$ where $x_i \in IR^n$ and $y_i \in \{-1,1\}$. The decision function is defined as:

$$f(x) = wx + b = \sum_{i}^{N_t} \alpha_i y_i \langle x_i, x \rangle + b \quad (N_t \text{ is the number of training samples}) \quad (3)$$

where α_i are the Lagrange multipliers obtained by solving the optimisation problem [22].
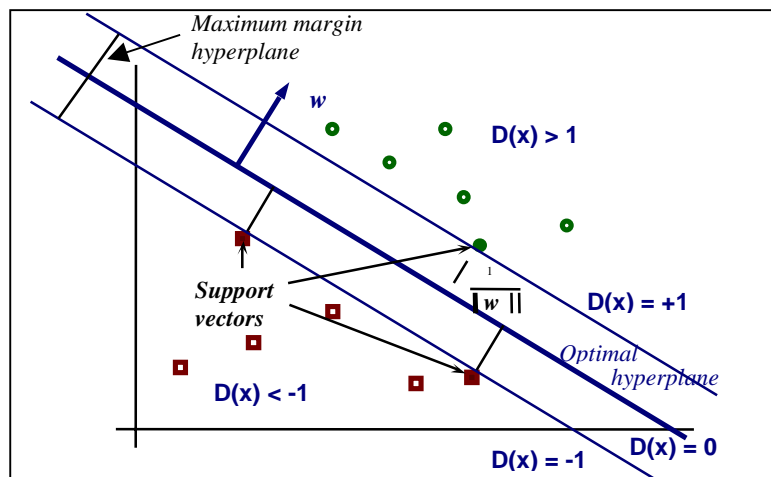
Figure 3. Margin in the case of linearly separable classes.

One class is assigned when the decision function is positive and the other when it is negative. When the classes are non-linearly separable in input space, training patterns are mapped into a higher dimensional space (feature space) $x \in \Re^n \to \Phi(x) \in \Re^f$, where the problem is linearly separable. By introducing kernel functions, *K*, the required scalar products in feature space are calculated directly by computing kernels in input space, which avoids having knowledge about the actual mapping $\Phi(x)$. In this case, the decision function is modified as follows:

$$f(x) = \omega \phi(x) + b = \sum_{i}^{N_t} \alpha_i y_i k \langle x_i, x \rangle + b \quad (4)$$

The positive value of $f(x)$ is associated with 1 and the negative one with -1.

Support vector machines (SVM) were originally designed for binary classification. Currently there are two types of approaches for multi-class SVM. One is by constructing and combining several binary classifiers "one against one or one against all methods", while the other is by directly considering all data in one optimisation formulation [23]. The direct approach constructs a classifier recognising the set of classes: the determination of the hyperplane between these different classes permits to choose a class among the *k* when a new input is presented. In order to carry out a comparative study, we have applied the direct approach that is considered a suitable method for practical use. This algorithm is clearly described by J. Weston in [24]. The kernel function used in this study is polynomial (degree = 3) with an optimal value of the penalty coefficient C = 2^{17}.
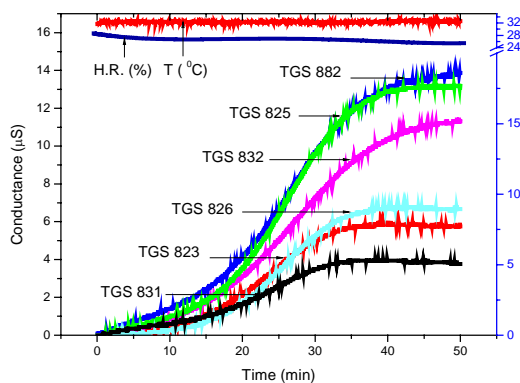
(a)
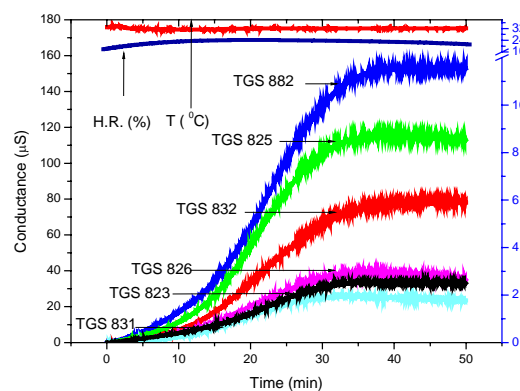
(b)
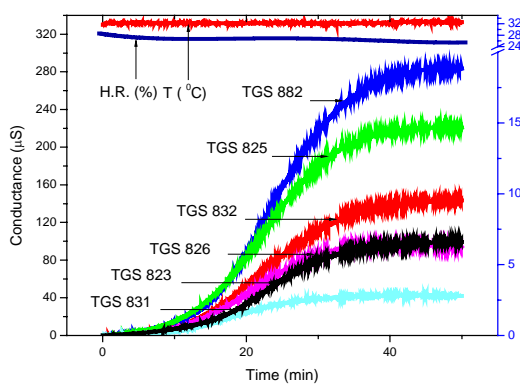
(c)

Figure 4. Time variation of the sensor signals, $G_i(t)$,

(a) day 3; (b) day 5; (c) day 9.

4. Results and discussion

The effect of the number of storage days in a refrigerator, *D*, undergone by the sardine samples on the sensor array response was analysed at first. Hence, every two days, a sample of sardine was taken out of the refrigerator and the array response was recorded as function of time, *t*, until steady state was reached for every sensor. A plot of the time dependence of $G_i(t)$ when the sensors are exposed to a sample of sardine after 3, 5 and 9 days of cold storage is shown in Fig. 4. It can be shown that the sensor signals reach a plateau and the conductance of the sensors dramatically increases with storage days. This indicates that the concentration of volatile compounds, which characterise the spoilage odour of fish [6], increases with the number of storage days. We note that all the TGS gas sensors showed reducing gas behaviour: i.e., sensor conductance increased when fish volatiles were sampled.

However, the major problem with sensors used in electronic noses is their response drift [25]. Sensors can be problematic due to a change of response pattern with time, these drifts must be minimised by measuring dummy samples before the actual samples. The relative humidity changes slightly with time (Fig. 4) and mainly depends on the number of storage days of the fish, but does not affect the results and the conclusions obtained [9]. The temperature inside the chamber remains almost constant in every experiment (about 32 °C).

Radar-like plots with unitary radius were performed in order to see whether pattern differences developed between fish samples with different storage days. Fig. 5 shows a representative case. To construct these plots, the values of $G_i(t)$ were divided by the value corresponding to sensor TGS 882, which showed the maximum signal. This helps to easily visualise the differences among typical patterns. The radar plots show at a glance a clear pattern variation between 3 and 5 or 9 days of storage. The shape of the patterns remains very similar for storage days that are previous to a particular day D_{th} (day3), referred as the threshold day. It is then possible to differentiate between fish of good (day 1 to day 3) (Fig. 5(a)) and moderate freshness (day 5 to day 9) (Fig. 5(b) and (c)) [26]. It can be seen that each storage day has different pattern. The data associated to the threshold day was always coincident with the expiration day indicated by the sardine seller.
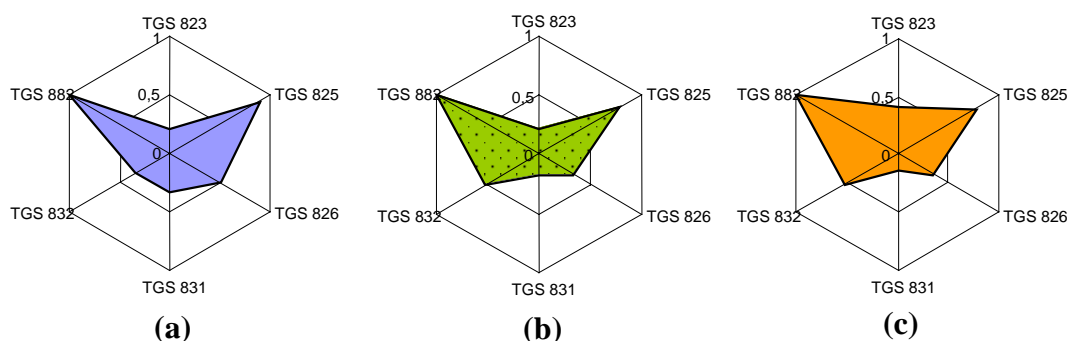
Figure 5. Radar plot for three different storage days of sardines (a) day 3; (b) day 5; (c) day 9. the three plots have radius equal to 1.

4.2. Data analysis

The data collected with our electronic nose was treated and analysed with three supervised classification methods. These supervised methods were applied with normalised data, for the different classification tasks. To compare the classification accuracy of the three pattern recognition algorithms (PNN, FANN and SVM) used, leave-one-out and *N*-fold cross validation techniques were implemented. For the second validation approach, the data set was divided into *5* subsets (*N*=5), each subset was made of 16 data vectors. The response patterns constituting the training and testing sets were rotated so that we would have unique training and testing sets in every fold. Response patterns were grouped in 8 categories, which corresponded to 1, 3, 5, 7, 9, 11, 13 and 15 days of storage. While most of the samples belonging to the first four classes could be correctly classified, the success rate in classification decreased noticeably for samples belonging to the other classes. This is true for the 5 different training and validation folds available and for the three different classifiers implemented (Table 1).

The performance in identification is shown in Table 2 as a confusion matrix. Rows and columns indicate true and predicted values, respectively. The validation was performed on the whole data set, since the leave-one-out validation technique was used. In the confusion matrix that corresponds to the PNN classifier (Table 2(a)), errors are quantitatively and qualitatively higher than those of the FANN (Table 2(b)), and the SVM methods (Table 2(c)). In particular, when the PNN classifier was used, more samples were wrongly classified and many samples belonging to storage days from 9 to 15 were classified as belonging to class 1 (one day of storage). Besides, the number of classification errors within samples stored from 9 to 15 days is significant, even for fuzzy ARTMAP and SVM based classifiers.

On the other hand, the results reported by El Marrakchi et al. [26] indicate that sardines have a shelf life of 9 days (if kept under cold storage). According to this, class 5 (day 9), class 6 (day 11), class 7 (day 13) and class 8 (day 15) were grouped together in a new class.

In the second study, the three classifiers were used to sort sardine samples in five classes only (i.e. 1, 3, 5, 7 and 9 or more days of cold storage). Comparing the prediction error of the three different methods, which was estimated using the leave-one-out procedure, it was revealed that the PNN method gave poor classification results with a 35 % error. On the other hand, the fuzzy ARTMAP neural network and SVM performed well with classification errors estimated at 5 % and 3.75 %, respectively. These results are summarised in Table 3. It is important to note that the proposed method based on SVM gives the best generalisation with a success-rate in classification of 96.25 %. The confusion matrix for the PNN approach is shown in Table 3(a). Errors are quantitatively higher in this case since 23 samples that had been under storage from 7 to 15 days were classified as fresh sardine samples. Table 3(b) and Table 3(c) show respectively the confusion matrix for the FANN and SVM approaches. Both methods result in a low number of misclassified samples (4 and 3 respectively). Moreover, we note for the two approaches, that the errors are qualitatively negligible in the sense that errors occur between consecutive categories Furthermore, these two methods are able to correctly separate fresh sardine samples from those that had undergone cold storage.

Table 1. Confusion matrix (5-fold cross-validation results) for PNN, Fuzzy ARTMAP and SVM in the classification of sardine samples in eight categories. Rows and columns are for true and predicted values, respectively.

Fold 1

PNN

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			1	1				
D_7				1	1			
D_9	2							
D_{11}	1					1		
D_{13}	1					1		
D_{15}	1					1		

Fuzzy ARTMAP

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5					2			
D_7					1	1		
D_9							1	1
D_{11}							1	1
D_{13}							1	1
D_{15}							1	1

SVM

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5					1	1		
D_7						1	1	
D_9						1	1	
D_{11}							1	1
D_{13}								2
D_{15}								2

Fold 2

PNN

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9	1			1				
D_{11}					2			
D_{13}	2							
D_{15}	1						1	

Fuzzy ARTMAP

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9				1	1			
D_{11}					1	1		
D_{13}								2
D_{15}								2

SVM

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9					1	1		
D_{11}						1	1	
D_{13}							2	
D_{15}								2

Fold 3

PNN

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7			1	1				
D_9	2							
D_{11}	1				1			
D_{13}						1	1	
D_{15}	1						1	

Fuzzy ARTMAP

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7			1	1				
D_9						1	1	
D_{11}						1		1
D_{13}							1	1
D_{15}						2		

SVM

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7			1	1				
D_9				1		1		
D_{11}						1		1
D_{13}							1	1
D_{15}							1	1

Fold 4

PNN

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9	1			1				
D_{11}	2							
D_{13}	1						1	
D_{15}	1						1	

Fuzzy ARTMAP

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9				2				
D_{11}				1		1		
D_{13}						1	1	
D_{15}							1	1

SVM

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9				2				
D_{11}						2		
D_{13}							2	
D_{15}						1	1	

Fold 5

PNN

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7	1				1			
D_9				1	1			
D_{11}	2							
D_{13}	1					1		
D_{15}	2							

Fuzzy ARTMAP

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9			2					
D_{11}					1			
D_{13}							1	1
D_{15}							2	

SVM

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	2							
D_3		2						
D_5			2					
D_7				2				
D_9					2			
D_{11}					2			
D_{13}							2	
D_{15}							1	1

Table 2. Confusion matrix (leave-one-out cross-validation results) for a) PNN, b) Fuzzy ARTMAP and c) SVM in the classification of sardine samples according to the number of days spent under cold storage. Rows and columns are for true and predicted values, respectively.

(a)

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	10							
D_3		10						
D_5			9	1				
D_7	1		1	6	2			
D_9	6			1	3			
D_{11}	6				4			
D_{13}	5						4	1
D_{15}	5						2	3

(b)

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	10							
D_3		10						
D_5			10					
D_7			1	8	1			
D_9				2	3	3	2	
D_{11}					1	7		2
D_{13}					1		4	5
D_{15}							5	5

(c)

	D_1	D_3	D_5	D_7	D_9	D_{11}	D_{13}	D_{15}
D_1	10							
D_3		10						
D_5			10					
D_7			1	8	1			
D_9				2	5	1	2	
D_{11}					3	6		1
D_{13}							9	1
D_{15}					1		4	5

Table 3. Confusion matrix (leave-one-out cross-validation results) for a) PNN, b) Fuzzy ARTMAP and c) SVM in the classification of sardine samples in five categories. Rows and columns are for true and predicted values, respectively.

(a)

	D_1	D_3	D_5	D_7	$\geq D_9$
D_1	10				
D_3		10			
D_5			9	1	
D_7	1		1	6	2
$\geq D_9$	22			1	17

(b)

	D_1	D_3	D_5	D_7	$\geq D_9$
D_1	10				
D_3		10			
D_5			10		
D_7			1	8	1
$\geq D_9$				2	38

(c)

	D_1	D_3	D_5	D_7	$\geq D_9$
D_1	10				
D_3		10			
D_5			10		
D_7			1	8	1
$\geq D_9$				1	39

Table 4 summarizes the results of sardine classification when a 5-fold cross validation procedure was implemented to estimate the success rate of the different classifiers studied. The results show the superior performance of the support vector machine model with a classification success rate of 91.25 % (averaged over the 5 folds), when compared to probabilistic neural network and fuzzy ARTMAP neural network techniques (51.25 % and 88.75 % respectively). These results, although slightly different from those obtained when a leave-one-out cross-validation was used, do not differ in the essential. Namely, SVM, and then FANN are significantly better techniques than the conventional PNN. The SVM method shows improved classification performance, and gives a good estimate of the number of storage days undergone by sardine samples.

Table 4. Results of analysing the fish data using a 5-fold cross-validation. In each fold, 64 patterns were used for training and the remaining 16 patterns were used for testing.

Fold	% Correct PNN	% Correct FANN	% Correct SVM
1	50	81.25	87.50
2	56.25	100	100
3	56.25	93.75	93.75
4	56.25	100	100
5	37.5	68.75	75
Average	**51.25**	**88.75**	**91.25**

5. Conclusions

In this paper, we have demonstrated the feasibility of an electronic nose based on an array of semiconductor gas sensors to detect off-odours from sardines under cold storage. Preliminary results seem very promising for the development of an electronic nose devoted to the monitoring of fish freshness, mainly when a rapid control is needed.

The sensor array coupled with feature extraction and pattern recognition methods can be trained to classify the fish according to the number of days spent under cold storage. In this study, the freshness of sardines has been evaluated with an electronic nose using three pattern recognition methods (PNN, FANN and SVM); the objective being to find the best method able to identify the number of days spent under cold storage with minimal errors. A leave-one-out and an N-fold cross validation techniques have been used in order to obtain a good estimate of the success rate in sardine classification for the different pattern recognition methods studied.

Our results indicate that the performance of SVM was superior to that of PNN and FANN. The superior performance of SVM over PNN or FANN may be due to the following reason: SVM implement the structural risk minimisation principle that minimises an upper bound for the generalisation error rather than minimising the training error.

Acknowledgements

References and Notes

1. Hammond, J.; Marquis, B.; Michaels, R.; Oickle, B.; Segee, B.; Vetelino, J.; Bushway, A.; Camire, M.E.; Dentici, K.D. A semiconducting metal-oxide array for monitoring fish freshness, *Sensors Actuators B* **2002**, 84, 113–122.
2. Garrigues, S.; Talou, T.; Nesa, D. Comparative study between gas sensors arrays device, sensory evaluation and GC/MS analysis for QC in automotive industry, *Sensors Actuators B* **2004**, 103 55–68.

3. Kocsis, N.; Amtmann, M.; Mednyanszky, Z.; Korany, K. GC-MS investigation of the aroma compounds of hungarian red paprika (Capsicum annuum) cultivars, *J. Food Compos. and Anal.* **2002**, 15, 95–203.

4. Di Natale, C.; Olafsdottir, G.; Einarsson, S.; Martinelli, E.; Paolesse, R.; D'Amico, A. Comparison and integration of different electronic noses for freshness evaluation of cod-fish fillets, *Sensors Actuators B* **2005**, 77, 572-578.

5. Olafsdottir, G.; Nesvadba, ; Di Natale, C.; Careche, M.; Oehlenschlager, J.; Tryggvadottir, S.V.; Schubring, R.; Kroeger, M.; Heia, K.; Esaiassen, M.; Macagnano, A.; Jorgensen, B.M. Multisensor for fish quality determination, *Food Sci. Technol.* **2004**, 15, 86-93.

6. Olafsson, R.; Martinsdottir, E.; Olafsdottir, G.; Sigfusson, T.I.; Gardner, J.W. Monitoring of fish freshness using tin dioxide sensors, In *Sensors and Sensory Systems for an Electronic Nose* J.Gardner, P.N. Bartlett, Eds.; Kluwer: Dordrecht, 1992; p. 257–272.

7. Martın, Y.G.; Oliveros, M.C.C.; Pavon, J.L.P.; Pinto, C.G.; Cordero, B.M. Electronic nose based on metal oxide semiconductor sensors and pattern recognition techniques: characterisation of vegetable oils, *Anal. Chim. Acta* **2001**, 449, 69–80.

8. Shaffera, R.E.; Rose-Pehrssona, S.L.; McGill, R.A. A comparison study of chemical sensor array pattern recognition algorithms, *Anal. Chim. Acta* **1999**, 384, 305–317.

9. O'Connell, M.; Valdora, G.; Peltzer, G.; Negri, R.M. A practical approach for fish freshness determinations using a portable electronic nose, *Sensors Actuators B* **2001**, 80, 149–154.

10. Brezmes, J.; Ferreras, B.; Llobet, E.; Vilanova, X.; Correig, X. Neural network based electronic nose for the classification of aromatic species, *Anal. Chim. Acta* **1997**, 348, 503–509.

11. Ciosek, P.; Wroblewski, W. The analysis of sensor array data with various pattern recognition techniques, *Sensors Actuators B* **2006**, 114, 85-93.

12. Chen, W.H.; Hsu, S.H.; Shen, H.P. Application of SVM and ANN for intrusion detection, *Comp. Oper. Research* **2005**, 32, 2617–2634.

13. Santos, J.P.; Garcia, M.; Aleixandre, M.; Horrillo, M.C.; Gutiérrez, J.; Sayago, I.; Fernàndez, M.J.; Arés, L. Electronic nose for the identification of pig feeding and ripening time in Iberian hams, *Meat Sci.* **2004**, 66, 727–732.

14. Llobet, E.; Hines, E.L.; Gardner, J.W.; Franco, S. Non-destructive banana ripeness determination using a neural network-based electronic nose, *Meas. Sci. Technol.* **1999**, 10, 538–548.

15. Brudzewski, K.; Osowski, S.; Markiewicz, T. Classification of milk of an electronic nose and SVM neural network, *Sensors Actuators B* **2004**, 98, 291–298.

16. Masters, T.A. In *Practical Neural Network Recipes in C++*; Academic Press: San Diego; 1993.

17. Specht, D. Probabilistic neural networks for classification, mapping, or associative memory, In *Proceedings of the IEEE Int. Conf. on Neural Networks*, 1988; Vol. 1, p. 525-532.

18. Specht, D. Probabilistic neural networks, *Neural Networks* **1990**, 3, 109–118.

19. Llobet, E.; Hines, E.L.; Gardner, J.W.; Bartlett, P.N.; Mottram, T.T. Fuzzy ARTMAP based electronic nose data analysis, *Sensors Actuators B* **1999**, 61, 183–190.

20. Carpenter, G.; Grossberg, S.; Markuzon, N.; Reynolds, J.; Rosen, D. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of multidimensional maps, *IEEE Trans. Neural Nets.* **1992**, 3, 698–713.

21. Cristianini, N.; Shawe-Taylor, J. *An Introduction to support vector machines*; Cambridge University Press: Cambridge; 2000.

22. Burges, C.J.C. A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discov.* **1998**, 2, 121–167.

23. Hsu, C.W.; Lin, C.J. A comparison of methods for multi-class support vector machines, *IEEE Trans. Neural Nets.* **2002**, 13, 415–425.

24. Weston, J.; Watkins, C. Multi-class support vector machines, In Technical Report CSD-TR- 98-04, Department of Computer Science, Royal Holloway, University of London; 1998.

25. Deisingh, A. K.; Stone, D. C. Applications of electronic noses and tongues in food analysis, *Int. J. Food Sci. Technol.*, **2004**, 39, 587–604.

26. El Marrakchi, A.; Bennour, M.; Bouchriti, N.; Hamama, A.; Tagafait, H. Sensory, chemical, and microbiological assessments of Moroccan Sardines (Sardina pilchardus) stored in ice, *J. Food Protect.* **1990**, 53, 600–605.