Food Research International 43 (2010) 797–804

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Time dependence of the aroma pattern emitted by an encapsulated essence studied by means of electronic noses and chemometric analysis

Silvio D. Rodriguez [a], María Eugenia Monge [a], Alejandro C. Olivieri [b], R. Martin Negri [a], Delia L. Bernik [a,*]

[a] Institute of Physical Chemistry of Materials, Environment and Energy (INQUIMAE), Department of Inorganic, Analytical and Physical Chemistry (DQIAQF), School of Sciences, University of Buenos Aires, Ciudad Universitaria, Pabellón 2, C1428EGA Buenos Aires, Argentina
[b] Department of Analytical Chemistry and Chemistry Institute of Rosario (IQUIR-CONICET), Faculty of Biochemical and Pharmaceutical Sciences, University of Rosario, Rosario, Argentina

ARTICLE INFO

Article history:
Received 1 June 2009
Accepted 29 November 2009

Keywords:
Aroma release
Electronic noses
Chemometric analysis
Essence encapsulation

ABSTRACT

This work studies the changes along days of the aroma released from a flavour encapsulated in a polysaccharide gel matrix using the electronic nose methodology. The purpose is to explore the capacity of the sensor array to assign a pattern of aroma to the corresponding release day within a total period of five days. Different procedures of data treatment and analysis are compared in order to achieve the maximum of information of the system under study in conditions where the number of measurements is limited. Raw and normalized sensor signals are processed using various unsupervised and supervised data analysis algorithms such as Principal Component Analysis, Kohonen-Self Organizing Maps, Cluster Analysis, Multiple Discriminant Analysis and two types of Artificial Neural Networks (BP-ANN and RBF-ANN). Accurate assignation of the number of release days is obtained with a successful classification up to four classes associated to samples at increasing days of aroma release. The relative advantages and drawbacks of the different procedures and data manipulations are discussed.

1. Introduction

It is of high relevance for many flavouring processes to follow the evolution of the aroma released by an encapsulated essence. Volatile compounds throw off the matrix in a differential manner according to their relative concentrations and affinities towards the encapsulation material, providing the aroma of the system. In many cases additional contribution to the aroma variation comes from changes of the matrix itself during the release. Examples are all the processes related to non-stabilized gels, such as those occurring during a sol–gel transition, and those related to the aging of the matrix due to drying or oxidative degradation. Flavour release is a dynamic process which involves not only time dependent aroma changes depending on both the essence and the matrix components (Boland, Buhr, Giannouli, & van Ruth, 2004; Madene, Jacquot, Scher, & Desobry, 2006) but also a dynamic perception, and must be studied using dynamic methods. These facts also imply a particular attention to the data treatment and analysis, as reviewed by Piggott (2000) and references therein.

An electronic nose (e-nose) is an instrument composed by an array of non-specific gas sensors coupled to a pattern recognition system which allows to discriminate odours using chemometric data analysis (Gardner & Bartlett, 1999). The methodology is relatively simple because does not require separation and identification of the odour components, which makes the technique appropriate for real time quality control (Branca, Simonian, Ferrante, Novas, & Negri, 2003; Cosio, Ballabio, Benedetti, & Gigliotti, 2007; Gan, Che Man, Tan, Noraini, & Nazimah, 2005; Hai & Wang, 2006) and as a non-destructive method to control fruit maturity and cultivar processes (Bhattacharyya et al., 2007; Hernández Gómez, Wang, Hu, & García Pereira, 2008; Pathange, Mallikarjunan, Marini, Keefe, & Vaughan, 2006) among other applications in food control (Peris & Escuder-Gilabert, 2009). E-noses have been seldom used for studying the aroma release kinetics of flavours (Baranauskienė, Rimantas Venskutonis, Galdikas, Senulienė, & Šetkus, 2005; Deisingh, Stone, & Thompson, 2004). Therefore, we performed a series of studies exploring the potentiality and limitations of these sensor devices detecting changes in odour patterns along time (Lovino, Cardinal, Zubiri, & Bernik, 2005; Monge, Bulone, Giacomazza, Bernik, & Negri, 2004; Monge, Bulone, Giacomazza, Negri, & Bernik, 2004; Monge, Negri, Giacomazza, & Bullone, 2008).

Due to the high sensitivity of the sensors to register subtle changes, the electronic nose provides indisputable evidence about whether an odour has changed or not, irrespective of which particular component of the odour is the responsible for the change in the aroma fingerprint. When looking for slight differences, all the features of sensor signals will have importance when analyzing

* Corresponding author. Tel.: +54 11 4576 3378; fax: +54 11 4576 3341.
E-mail address: dbernik@qi.fcen.uba.ar (D.L. Bernik).

0963-9969/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodres.2009.11.022

data: signal intensity and signal ratios. In this regard, raw data treatment, total number of measurements and the mathematical algorithms used (supervised or not) will play decisive roles in the resultant discrimination capacity of the sensor device.

Since it is not the scope of this work to identify which components have changed in this way altering the aroma pattern, no validation with conventional analytical methods was done. In addition, the intention is to see which of the afore mentioned chemometric methods yield the best results when using a limited number of measurements: this would mimic a routine situation such as doing a quality control in real time. Thus, herein we examine qualitatively the time dependent pattern of the aroma released from an encapsulated essence 5 days onward since encapsulation and evaluate the performance of different chemometric approaches, stressing that a successful discrimination of the samples will depend on the choice of the appropriate mathematical methodology for data treatment.

2. Experimental

2.1. Experimental details

The flavouring sample is a tutti-fruti essence provided by International Flavours and Fragrances Argentina, composed by 80% of limonene. The e-nose prototype used in the present work was developed at our laboratory (Lovino et al., 2005). The array is composed by 10 commercial sensors based on polycrystalline tin dioxide having different sensitivities, provided by Figaro Inc.: TGS 831 (Sensor 1); TGS 813 (Sensor 2); TGS 825 (Sensor 3); TGS 832 (Sensor 4); TGS 880 (Sensor 5); TGS 826 (Sensor 6); TGS 816 (Sensor 7); TGS 842 (Sensor 8); TGS 823 (Sensor 9) and TGS 800 (Sensor 10). For each sensor, a voltage proportional to the respective electrical conductance is digitalized (12 bits resolution with voltages within 0–5 V). All the digital outputs are simultaneously recorded and stored on a laptop. The only pre-processing step used was data normalization, when used. The noise/signal relationship was about 5×10^{-3}, while most of the experimental details have been reported in previous articles (Branca et al., 2003; Monge, Bulone, Giacomazza, Bernik, et al., 2004, Monge, Bulone, Giacomazza, Negri, et al., 2004; Monge et al., 2008; Lovino et al., 2005). The reproducibility was tested by comparing the individual signals of the quintuplicate samples at the day 0 of the release.

The values of the sensor's signals after reaching a plateau are used for the analysis after the subtraction of a baseline (the sensor's response to air). All the experimental aspects related to pectin matrix preparation have been extensively described elsewhere (Monge et al., 2008).

Quintuplicate of the gels with the encapsulated essence were initially kept closed for 3 h and then one of the replicates was opened and measured with the e-nose. Along the consecutive days, the replicates were opened and measured. Once a given replicate was opened, it was kept opened and stored at 33 °C to be able to follow the aroma release kinetics in the next days. The released was followed since a "first day" (about 3 h since the gel was prepared) up to the "fifth day" (about 120 h since preparation). Only one replicate is opened (and measured) at the first day. At the second day other replicate is opened and both samples are measured. The number of days that the sample was left opened is referred as D. A measurement associated to $D = 0$ means that the measurement was performed when immediately opened to the air. One associated to $D = 1$ indicates that the sample was measured after being 24 h opened approximately (independently of the number of days elapsed since prepared as far as it remained stored in a closed vial). This procedure generates 20 measurements when using quintuplicates. Experiments were performed following the aroma release up to 5 days since gel preparation in each experiment. Two experiments were performed using quintuplicates, one with quadruplicates and one with triplicates thus generating a set of 73 measurements.

2.2. General methodology and procedure for analysis

As outlined the aim of the work is to discriminate samples according to the number of days that the sample was releasing the aroma to the open air. The number of outputs (target groups) used for grouping the data is indicated as N_g in Tables 1–3.

For this purpose we followed the methodology illustrated in Fig. 1. Each e-nose measurement provides a vector $\mathbf{s}$ (the sensors signals, size $N \times 1$), where N indicates the number of sensors of the array. These vectors can be used as obtained (raw data) or they can be normalized to obtain new vectors $\mathbf{s}'$, whose elements are given by $s'_n = s_n / \left(\sum_{r=1}^{N} s_r^2\right)^{1/2}$. The parameter $NF = \left(\sum_{r=1}^{N} s_r^2\right)^{1/2}$ is the so-called Normalization Factor, an indicator of the intensity of the odour detected by the e-nose. Therefore, all sensor's responses are converted into [0, 1] after normalization. Additionally, when using CA, K-SOM or the supervised methods two other alternatives are possible: (a) the sensor data set are directly used and (b) the sensor data are used as inputs for PCA and then the PCA outputs are used as inputs for the other methods.

In this work the time dimension is discrete and represents the criteria for grouping. In the unsupervised methods it is not possible to incorporate this information. In the supervised methods this information is provided in an implicit manner in the training data set.

2.3. Unsupervised methods

2.3.1. Principal Component Analysis (PCA)

PCA is a very well-known unsupervised method for determining similarity of the input data by comparing the relative location of its associated so-called principal components in a PCA-plot (Johnson & Wichern, 2002). If M measurements are performed with an e-nose of N sensors, then the input data are introduced as a matrix $\mathbf{X}$ of size $M \times N$, where each row consists of the N signals for each of the M measurements. In the present work are $N = 10$ and $M = 73$ and the covariance matrix was used.

Input data are extracted in PCA in a new base of the dimension N: the so-called principal components $\{PC_1, \ldots, PC_N\}$. A new matrix $\mathbf{Y}$ of size $M \times N$ is obtained where each row consists of the N principal component for each of the M measurements. The first components usually explain great part of the total variance. In those cases the data points can be qualitatively discriminated by observing how they group in a 2-D PCA-plot (e.g. PC_1 vs. PC_2) or 3-D plot (PC_1 vs. PC_2 vs. PC_3). In these plots each point is associated to one measurement with the e-nose.

2.3.2. Kohonen-Self Organizing Map (K-SOM)

SOM belongs to the category of competitive learning methods and is based on unsupervised learning. It can be thought of as way of projecting multiple dimensions onto a two-dimensional output allowing the user to visualize the groupings and relationships of the odours or chemical volatile compounds (Bermak, Brahim Belhouari, Shi, & Martinez, 2006; Sinesio et al., 2000).

K-SOM is a neural network based on a single layer of neurons arranged in a box having on its top a two-dimensional response plane (Kohonen, 1988). The usual topology of the Kohonen network adopts a toroidal boundary condition, implying that the network at its right or top edge is continued (in a computational sense) at its left and low edge, respectively, and *vice versa*. During the training of K-SOM, the same vector $\mathbf{u}$ (containing the variables assigned to a given sample) is input to all neurons, and

Table 1
Results of CA-PAM using the sensor's signals as inputs (with and without normalization).

Cluster Analysis. Partition Around Medoids (CA-PAM)				
N_g	Non-normalized signal inputs: **s**		Normalized signal inputs: **s**′	
	Number and percentages of errors*	ASW	Number and percentages of errors*	ASW
2	4 (5%)	0.70	30 (41%)	0.26
3	12 (16%)	0.56	34 (46%)	0.28
4	34 (46%)	0.50	39 (53%)	0.25

ASW: Average Silhouette Width.
* In this case an error is a sample that was clusterized in the wrong group. This calculation is possible to assess because the identities of samples are known.

Table 2a
Artificial Neural Networks: conditions used for training and validation.

N_g	Total number of measurements (M)	Total number of measurements used for training	Total number of measurements used for validation	Output classes	Distribution of the number of training measurements among the output classes	Distribution of the number of validation measurements among the output classes
2	73	58	15	$D = 0$	25	7
				$D \geqslant 1$	33	8
3	73	58	15	$D = 0$	25	7
				$D = 1$	12	3
				$D \geqslant 2$	21	5
4	73	58	15	$D = 0$	25	7
				$D = 1$	12	3
				$D = 2$	8	2
				$D \geqslant 3$	13	2

Number of data for each D: $D = 0$, 32; $D = 1$, 15; $D = 2$, 10; $D = 3$, 9; $D = 4$, 5; $D = 5$, 2.
N_g = Number of target outputs.

Table 2b
Artificial Neural Networks: recovered parameters of the optimized networks.

ANN								
N_g	Non-normalized signals				Normalized signals			
	Inputs: **s**		Inputs: principal components		Inputs: **s**′		Inputs: principal components obtained with normalized signals	
	BP	RBF (σ)	BP	RBF (σ)	BP	RBF (σ)	BP	RBF (σ)
2	10:8:8:2 tan:tan:tan:log $\mu = 0.005$	0.5	3:12:2 lin:tan:log $\mu = 0.005$	0.5	10:13:2 lin:tan:log $\mu = 0.005$	0.1	7:14:2 tan:tan:log $\mu = 0.005$	0.1
3	10:12:3 tan:log:log $\mu = 0.005$	0.09	3:10:3 lin:log:log $\mu = 0.05$	0.1	10:12:3 tan:tan:log $\mu = 0.005$	0.06	7:12:3 lin:tan:log $\mu = 0.005$	0.07
4	10:7:7:4 log:lin:log:log $\mu = 0.005$	0.05	4:10:4 log:tan:log $\mu = 0.05$	0.01	10:7:4 lin:tan:log $\mu = 0.05$	0.01	5:6:6:4 tan:lin:tan:log $\mu = 0.05$	0.05

BP (Back Propagation ANN): the sequence of numbers indicates the number of neurons in each layer and the number of layers. For example 10:8:8:2 indicate four layers and the respective number of neurons in each one. Transfer functions: tan = tangential sigmoid; lin = linear; log = logarithmic sigmoid. μ = initial increment of the learning process. RBF (radial basis function ANN): σ = spread value.

Table 3
Results of the different supervised methods. The number of classes is indicated by N_g.

Supervised methods												
N_g	Non-normalized signals						Normalized signals					
	Inputs: **s**			Inputs: principal components			Inputs: **s**′			Inputs: principal components		
	LDA (%)	BP (%)	RBF (%)	LDA (%)	BP (%)	RBF (%)	LDA (%)	BP (%)	RBF (%)	LDA (%)	BP (%)	RBF (%)
2	93	92	97	93	93	97	96	92	95	95	87	95
3	88	87	83	84	80	80	87	79	87	73	84	87
4	88	79	80	75	81	71	83	79	78	73	80	75

%: Average percentage of success.

a comparison is made between its elements and the neuron weights (vector **w**). The vector **u** may be composed of the raw sensor signals **s**, its normalized version **s**′, or a vector of PCs, as explained above.

The parameters used in training the K-SOM for the presently analyzed data were: number of cycles, 1000, dimensions of the neuron map, 20×20, form of function $a(\cdot)$, a triangle with a cut-off distance of six neurons. These parameters were tuned on a trial

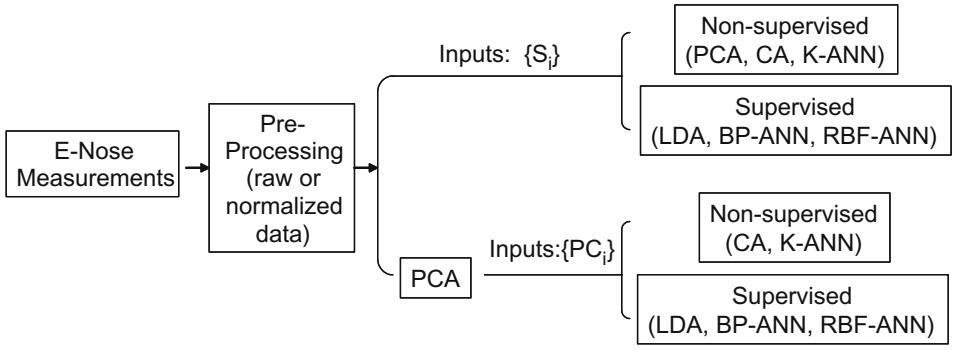

Fig. 1. Scheme indicating the different combination of data treatment used in this work.

and error basis, until the best separation among groups was achieved.

2.3.3. Cluster Analysis (CA)

In Cluster Analysis the input data can be the set of sensor's signals **s,** its normalized values, **s**′ or the set of principal components associated to those vectors as well. The number of desired clusters for grouping the data (*clusterization*) must be fixed previously. CA can be performed using different algorithms being the Partitioning Around Medoids (PAM) the one used in the present work (Kaufman & Rousseeuw, 1987). As outputs, CA-PAM indicates not only by which data input are the clusters composed, but provides the so-called *Average Silhouette Width* (ASW). Following the criteria for the goodness of the *clusterization* indicated by Struyf, Hubert, and Rousseeuw (1997), the ASW must be larger than 0.25 and grouping is perfect when ASW is equal to one; the higher the ASW the better the fit.

2.4. Supervised methods

Supervised approaches assume an *a priori* knowledge of the number of classes as well as the class membership of every sample in the training set. These methods involve training a pattern classifier to relate sensor values to specific odour labels.

We randomly varied five times the selection of the data for training (80% of the data set) and validation (20% of the data set) of the supervised methods.

2.4.1. Multiple Discriminant Analysis (MDA)

MDA is a supervised method based on the determination of a discriminant function which is a combination of the original variables, such that maximize the ratio between-class variance and minimize the ratio within-class variance. The MDA method can be performed with many discriminant functions. The types used in this work are the linear (LDA) and quadratic (QDA) functions. The methods are well described by Berrueta, Alonso-Salces, and Héberger (2007) and Scott, James, and Ali (2007).

2.4.2. Back Propagation Artificial Neural Network (BP-ANN)

In BP-ANN the network parameters are the weights and bias. The number of neurons of the networks satisfy the following condition: (number of network parameters) $\leqslant (1/2) \times$ (number of training data) $\times$ (number of input channels) (Jurs, Bakken, & McClelland, 2000). The Levenberg–Marquardt algorithm, an approximation of the Newton's method, is the training algorithm we used for the BP-ANN (Marquardt, 1963). The μ value, which defines the initial increment of the learning process, was fixed to 0.005 except when indicated.

The architecture of the networks was different depending on the inputs used: the 10 sensors signals or the principal components. The transfer function and their order between network's layers were chosen from every possible combination of logarithmic sigmoid, tangent sigmoid and linear functions, and selected accordingly to their capacity for providing good classification results. The number of neurons and layers are related with the maximum number of parameters that can be used. In the present work, the maximum number of parameters is equal to $M \times N/2$. This condition limits the total number of neurons; therefore different networks were designed containing 3, 4 and 5 layers under this constraint. We performed BP-ANN analysis using 250 and 500 epochs, verifying that the classification results was the same for both cases. All the information about number of data used for training and validating are shown in Table 2a.

2.4.3. Radial Basis Function Artificial Neural Network (RBF-ANN)

The RBF neural network used in the present work was the Probabilistic Neural Network (PNN) (Catelani & Fort, 2002; Specht, 1990). PNN operates by a probability density function for each data class based on an optimized kernel width parameter, called spread (σ). The architecture consists of an input layer, a radial basis function layer, a competitive layer and an output layer. At each neuron in the radial basis function layer, the dot product distance (L) between the inputs and the weight factors in that layer is calculated and the output of the layer is calculated as proportional to $exp(L/\sigma^2)$. The competitive layer sums the outputs from all hidden neuron of each respective data class. The results of the competitive layer are forwarded to the output layer. The parameter that needs to be determined to obtain an optimal probabilistic neural network is σ. This value is crucial because determines the width of the radial basis function. A trial and error method was used to find the best spread value (σ). In addition, the architecture of the networks can be different depending if the 10 sensors signals (non-normalized) or the principal components are used as inputs.

3. Results

According to the experimental setup the greater number of groups that could be differentiated by the e-nose is five, associated to days $D = 0$, $D = 1$, $D = 2$, $D = 3$ and $D \geqslant 4$, respectively. Other possible discrimination groups corresponds to grouping the data into four ($D = 0, 1, 2, \geqslant 3$), three ($D = 0, 1, \geqslant 2$) and two groups ($D = 0$, $D \geqslant 1$). The number of outputs (target groups) used in each case is indicated as N_g in Tables 1–3. The numerical results obtained by the different methods and alternatives are described below. A typical aroma pattern taken at different days of release (D) is shown in Fig. 2.

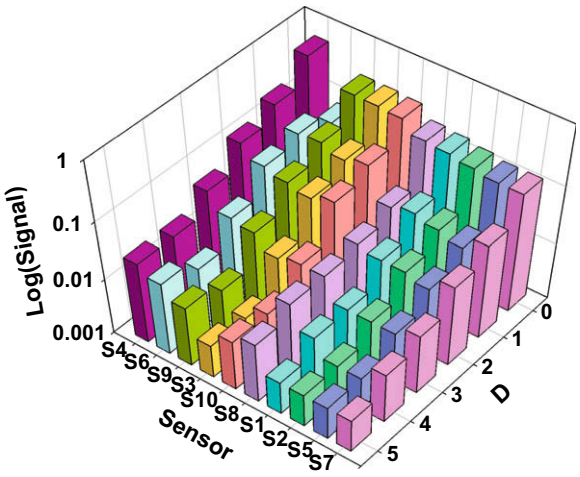

Fig. 2. Typical aroma pattern taken at different days of release (D). The vertical axis represents the common logarithm of the sensor's signal. Each row in the sensor's axis is associated to one sensor of the array.

3.1. Unsupervised methods

Fig. 3 displays the results obtained using the Principal Component Analysis methodology: the PCA-plot shows the two first principal components PC_1–PC_2 both for the non-normalized (3a) and normalized (3b) sensor data. In the first case the samples could be grouped visually into two classes belonging to $D = 0$ and $D \geqslant 1$ (Fig. 3a), while with normalized data the figure shows a continuous path from right to left with the increased number of day of release. Similar results were obtained for 3-D PCA-plots using the three first principal components. Therefore, at first glance it seems that the PCA-plot does not allow to clearly discriminate graphically more than these two main groups of samples.

With Kohonen-Self Organizing Maps (K-SOM) the results obtained are positive in the aspect that a single neuron-sample relationship is obtained, that is: each sample is associated to only one neuron and *vice versa*. However, the K-SOM plot (Fig. 4) shows the samples spread along the plot in which only groups associated with days one, two and three can be individualized only because of the previous knowledge of sample identities. K-SOM maps performed normalizing input data worsens the results, in accordance with the tendency observed for the PCA analysis (data not shown).

In the case of using the unsupervised Partition Around Medoids-Cluster Analysis (CA-PAM) samples were satisfactorily grouped up to three groups corresponding to release days $D = 0$, 1 and $\geqslant 2$ when using non-normalized signal inputs (see Table 1). Normalization of data induces a much less efficient *clusterization* as can be seen on the columns at the right side of the Table 1, with ASW values near the lower acceptable limit for a good fit (Struyf et al., 1997). The same trend with very similar numeric results is observed when the three first principal components were used as inputs instead of sensor data, normalized or not (data not shown).

Although the ASW values obtained with raw data are within those corresponding to a good *clusterization*, knowing the real identity of each measurement we can calculate the percentage of errors in the clusterization procedure by checking to which group are samples associated in the output of the PAM algorithm. In Table 1 the number of errors (over a total of 73 measurements) and the corresponding error percentage is detailed. It is clear that even with an acceptable ASW of 0.50, the 46% of errors in the *clusterization* with $N_g = 4$ is unacceptable; while *clusterization* with $N_g = 2$ (ASW = 0.70) yields 95% of the samples successfully assigned (5% of error).

The analysis of the influence of normalizing sensor data on sample discrimination is considered in more detail below. In fact we have described that crucially different results are obtained if data are normalized before to be analyzed with the different algorithms. Recalling the PCA analysis, it is worth noting that the first principal component, PC_1, accounts for 94.2% of the total variance in the non-normalized case. PC_1 can be linearly correlated with the Normalization Factor NF as shown in the inset of Fig. 5. This relationship indicates that PC_1 is mainly defined by the intensity of the aroma detected by the e-nose much more than by the "shape" of the fingerprint (the relative relationships between the sensor's signals). Therefore, as the intensities of the signals decrease through the days accompanying the progressive release of the aroma from the encapsulation matrix, a relationship is also expected between PC_1 and the number of releasing days, D. In fact, a step decrease of PC_1 with D was obtained during the first 2 days (Fig. 5), changing thereafter to a less pronounced decrease. The most tender decrease of the intensity after day two results in the poor *clusterization* obtained with CA-PAM (see ASW results for $N_g = 3$ and 4 in Table 1). These relationships obtained between PC_1 and the intensities of the signals and in consequence between PC_1 and D are lost when normalized data are used.

3.2. Supervised methods

While Table 2a shows the conditions used for training and validation of the Artificial Neural Networks (ANN), Table 2b displays the recovered parameters for the optimized networks: number of neurons in the hidden layers, corresponding transfer functions, μ value for BP-ANN and spread value for RBF-ANN. It was possible to classify samples up to four classes ($N_g = 4$; $D = 0$, $D = 1$, $D = 2$ and $D \geqslant 3$) using BP-ANN with average success between 79 and 81% for all cases assayed (Table 3). This percentage is almost the same when classifying into three classes, but significantly increases for classification into two classes ($N_g = 2$; $D = 0$ and $D \geqslant 1$). Similar results were obtained for RBF-ANN.

When using MDA the average percentage of success were much better for LDA than for QDA for all the classes ($N_g = 2, 3, 4$). Moreover, QDA yield worst results than both of the neural networks assayed, which is why QDA method was not included in Table 3. Another aspect is that we obtained better results using raw data, both in the non-normalized or normalized way. LDA brings the best performance of all the supervised methods in the case of three and four classes ($N_g = 3$ and 4, respectively). None training parameter are used by LDA which is very convenient because it is not necessary an optimizing step like the one used with BP-ANN or RBF-ANN.

4. Discussion

The signal's pattern given by a sensor array such as the electronic nose is defined by two main factors: the *shape* (the distribution of signals intensities) and the *intensity* (the numerical value of the signals). Although both factors can change in a multicomponent aroma release experiment, the present results indicate that intensity is a key factor when tutti-fruti essence is released through several days from pectin gels. The intensity factor, which accounts for the aroma concentration in the gas phase through the release, is lost when using normalized data. We call the attention at this point because normalization is regularly used as a tool for smoothing effects due to routine errors in sample manipulation, as already mentioned by Peris & Escuder-Gilabert in a recent review (2009). In the present work normalizing data induces a noticeable decrease in the quality of sample grouping when using unsupervised methods. As example, the ASW values of CA-PAM

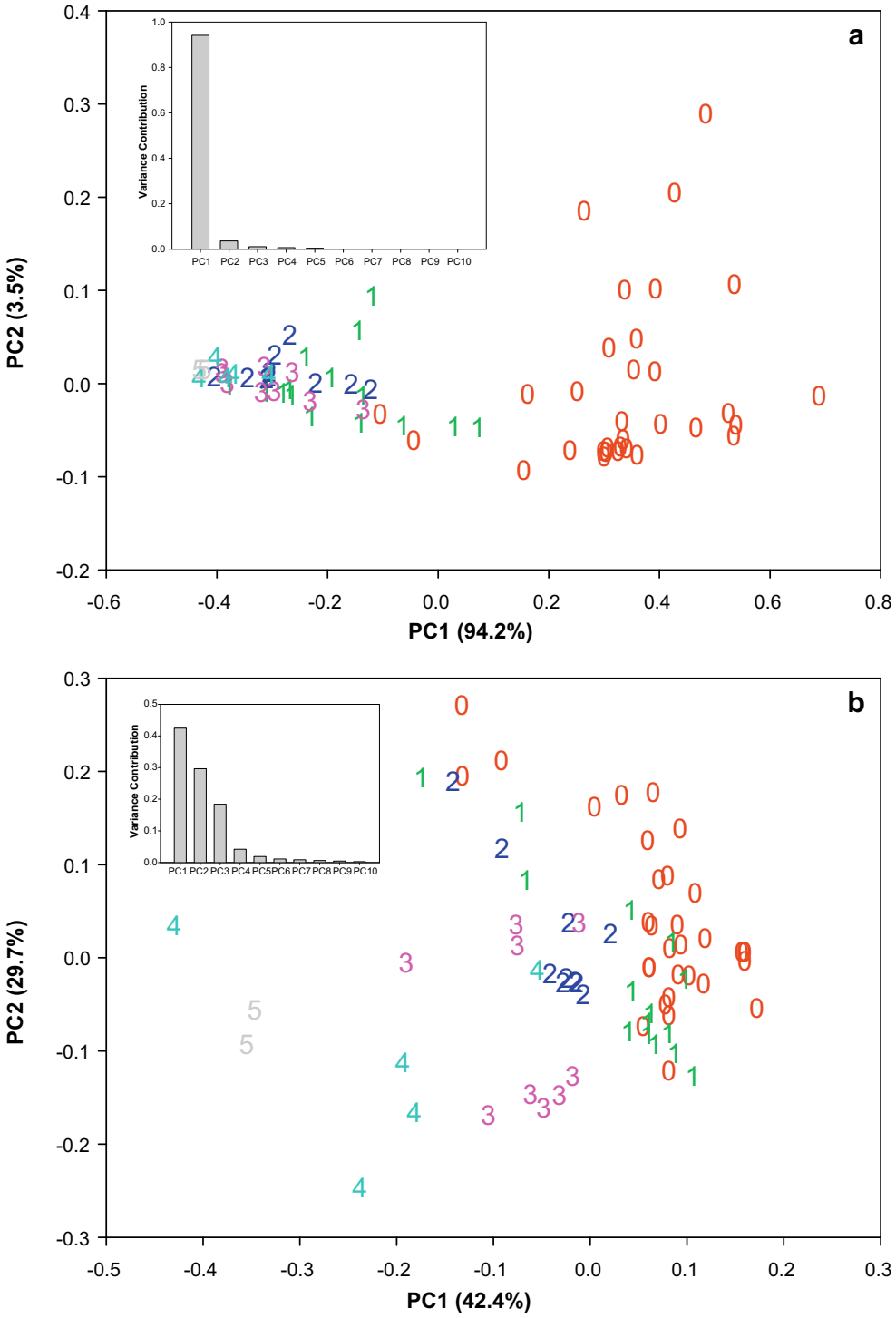

Fig. 3. PCA-plots using non-normalized (a) and normalized (b) sensor's signals as inputs. Each measurement is represented by a number indicating the associated value of D. (0): $D = 0$, (1): $D = 1$, (2): $D = 2$, (3): $D = 3$, (4): $D = 4$, (5): $D = 5$. The insets show the relative contribution of each principal component vector to the total data variance.

obtained in the case of four clusters ($N_g = 4$ in Table 1) are about 0.50 and 0.25 for raw and normalized data, respectively. The results of the K-SOM also show that grouping is much worst for the normalized case with no possibilities of clusterizing data, which becomes completely randomized.

This fact is not observed when using supervised methods: in those cases the use of either normalized or non-normalized data in the analysis does not alter the results. Therefore, it is necessary to distinguish between supervised versus unsupervised, when discussing the discrimination efficiencies of normalized versus non-normalized data because the obtained results were different. In fact, good results were obtained with normalized-supervised methods. In those cases, the lost of information (intensity) due to normalization is compensated by the training process, which recognize changes of the aroma shape, as each class is associated to a given release day (D) or range (class $D \geqslant 3$).

Hence, the results obtained herein are relevant for cases in which the intensity of the odour decreases or increases along time,

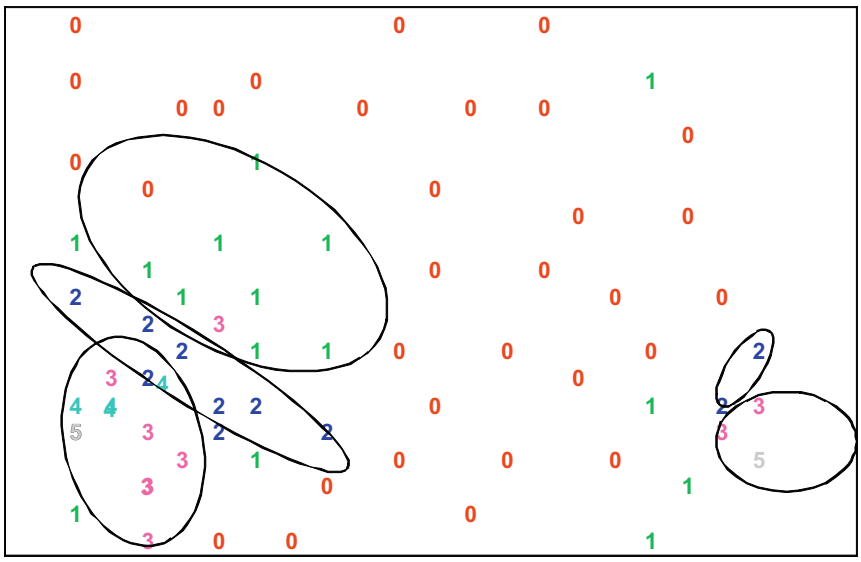

Fig. 4. Kohonen-maps using non-normalized sensor's signals as inputs. Each measurement is represented by a number indicating the associated value of *D*.

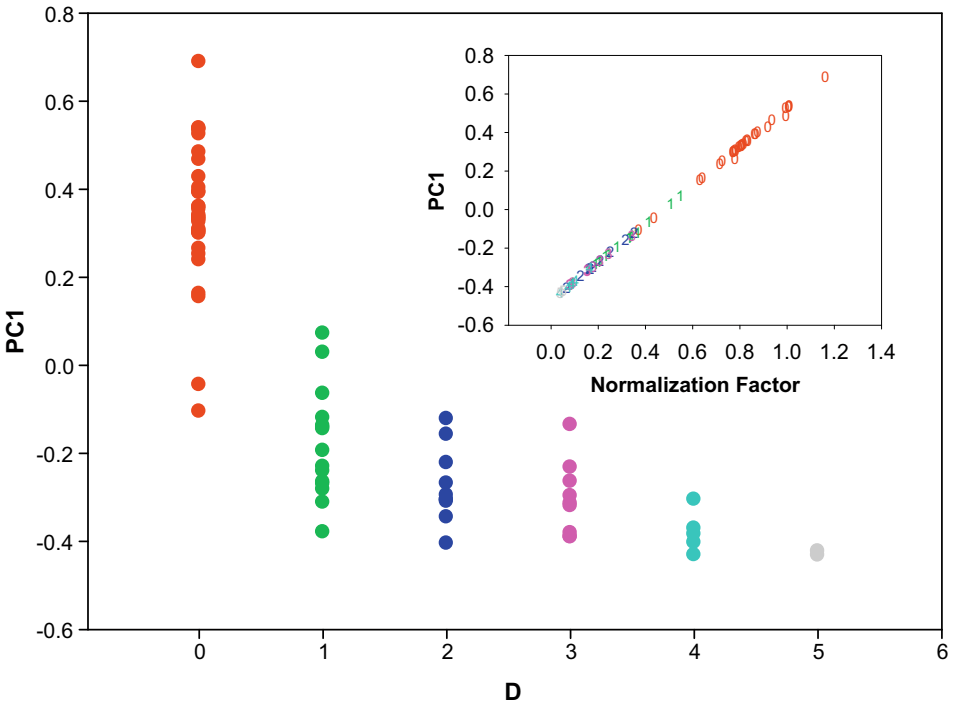

Fig. 5. Analysis of the first principal components (PC_1) vs. *D* (number of days) obtained using non-normalized data as inputs for PCA. The numbers indicates the days that a sample was left open before measurement. Inset: PC_1 vs. the Normalization Factor: $NF = \left(\sum_{r=1}^{N} s_r^2\right)^{1/2}$.

pointing the needs of keeping signal intensity and/or the time dimension variable to help sample discrimination.

The evolution of the aroma pattern can be explained considering the sample composition. On one hand, the essence has an unique majority component (limonene) so that release pattern in terms of sensor's signal ratio changes slightly its shape due to a differential release of the different components along days. On the other hand, after day 0 there is a sharp decrease in signal intensity followed by a soft decrease since day 2 (see Fig. 4), suggesting that after $D = 2$ diffusion becomes the rate limiting step inside the gel until the essence reaches the gel–air interface releasing the aroma to the vapour phase. This type of behaviour in the release pattern of encapsulated molecules in swollen gels has been described before (Bernik, Zubiri, Monge, & Negri, 2006).

No differences in the efficiency of discrimination were observed when using raw data or the relevant principal components as input for CA and/or ANN classification, and raw data were better inputs in the case of MDA. However these results (concerning the convenience of using raw data) applies for the present study but can not be generalized.

LDA and the neural networks are able to correctly classify samples up to four classes ($N_g = 4$). Moreover, LDA yield the best results in classification and is the faster of the supervised methods assayed herein because no training parameter are needed. When

comparing BP-ANN and RBF-ANN by inspecting Table 3 we see that the average percentage of success was similar for both networks. The main difference between BP and RBF is that RBF training is faster which significantly reduces the running time, as already mentioned by Catelani and Fort (2002). In the present case, RBF-ANN requires an analysis time 10 times shorter than BP-ANN.

Last, it was noticeable the bad performance of PCA, one of the most popular unsupervised methods used for visual grouping of samples. This is not a minor point because PCA is still being regularly used; Peris and Escuder-Gilabert (2009) show in their work that more than 70% of e-nose users employ PCA and 22% of them use exclusively PCA as discrimination method in food technology areas such as process monitoring, shelf life investigation, freshness and authenticity assessment. According with this study may run the risk of finding false negative results. The possibility of using supervised methods must be always explored if the amount of data is large enough for training the networks or the algorithm (MDA), a situation which, however, is not always experimentally possible in terms of cost/benefit ratio. Therefore, it is crucial to find a balance between the total number of measurements, the time necessary to analyze data and the necessity of using supervised methods. Fulfilling these equilibrium will make the e-nose technique a more practical and confident tool at the industry and research levels.

5. Conclusions

The electronic nose device allowed a highly satisfactory classification of the samples according to the day of aroma release of an encapsulated essence. Discrimination is clearly improved using raw data and supervised methods of analysis. In the present case Linear Discriminant Analysis (LDA) outperformed the two neural networks assayed, Back Propagation (BP-ANN) and Radial Basis Function (RBF-ANN); therefore, it is suggested the LDA method as the appropriate choice for use in data analysis.

Acknowledgements

RMN, ACO and DLB are members of the *Consejo Nacional de Investigaciones Científicas y Tecnológicas* (CONICET, Argentina). SDR is recipient of CONICET fellowship. This work was supported by the University of Buenos Aires (*UBACyT 2004–2007, Proyecto X267*) and CONICET (PIP 6382 and PIP 5303).

References

Baranauskienė, R., Rimantas Venskutonis, P., Galdikas, A., Senulienė, D., & Šetkus, A. (2005). Testing of microencapsulated flavours by electronic nose and SPME–GC. *Food Chemistry, 92*, 45–54.

Bermak, A., Brahim Belhouari, S., Shi, M., & Martinez, D. (2006). Pattern recognition techniques for odour discrimination in gas sensor array. In C. A. Grimes, E. C. Dickey, & M. V. Pishko (Eds.). *Encyclopaedia of sensors* (Vol. X, pp. 1–17). American Scientific Publishers.

Bernik, D. L., Zubiri, D. B. V., Monge, M. E., & Negri, R. M. (2006). A new kinetic model of drug release from swollen gels under non-sink conditions. *Colloids & Surfaces A: Physicochemical and Engineering Aspects, 273*, 165–173.

Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A, 1158*, 196–214.

Bhattacharyya, N., Seth, S., Tudu, B., Tamuly, P., Jana, A., Ghosh, D., et al. (2007). Monitoring of black tea fermentation process using electronic nose. *Journal of Food Engineering, 80*, 1146–1156.

Boland, A. B., Buhr, K., Giannouli, P., & van Ruth, S. M. (2004). Influence of gelatin, starch, pectin and artificial saliva on the release of 11 flavour compounds from model gel systems. *Food Chemistry, 86*, 401–411.

Branca, A., Simonian, P., Ferrante, M., Novas, E., & Negri, R. M. (2003). Electronic nose based discrimination of a perfumery compound in a fragrance. *Sensors and Actuators B: Chemical, 92*, 222–227.

Catelani, M., & Fort, A. (2002). Soft fault detection and isolation in analog circuits: Some results and a comparison between a fuzzy approach and radial basis function networks. *Instrumentation and Measurement, IEEE Transactions on Instrumentation and Measurement, 51*, 196–202.

Cosio, M. S., Ballabio, D., Benedetti, S., & Gigliotti, C. (2007). Evaluation of different storage conditions of extra virgin olive oils with an innovative recognition tool built by means of electronic nose and electronic tongue. *Food Chemistry, 101*, 485–491.

Deisingh, A. K., Stone, D. C., & Thompson, M. (2004). Application of electronic noses and tongues in food analysis. *International Journal of Food Science and Technology, 39*, 587–604.

Gan, H. L., Che Man, Y. B., Tan, C. P., Noraini, I., & Nazimah, S. A. H. (2005). Characterisation of vegetable oils by surface acoustic wave sensing electronic nose. *Food Chemistry, 89*, 507–518.

Gardner, J. W., & Bartlett, P. (1999). *Electronic noses*. Oxford: Oxford University Press.

Hai, Z., & Wang, J. (2006). Electronic nose and data analysis for detection of maize oil adulteration in sesame oil. *Sensors and Actuators B: Chemical, 119*, 449–455.

Hernández Gómez, A., Wang, J., Hu, G., & García Pereira, A. (2008). Monitoring storage shelf life of tomato using electronic nose technique. *Journal of Food Engineering, 85*, 625–631.

Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis*. New Jersey: Prentice Hall.

Jurs, P. C., Bakken, G. A., & McClelland, H. E. (2000). Computational methods for the analysis of chemical sensor array data from volatile analytes. *Chemical Reviews, 100*, 2649–2678.

Kaufman, L. & Rousseeuw, P. J. (1987) Statistical data analysis based on the L_1 norm. In Y. Dodge, (Ed.), *Clustering by means of medoids* (pp. 405–416). North-Holland, Amsterdam.

Kohonen, T. (1988). An introduction to neural computing. *Neural Networks, 1*, 3–16.

Lovino, M., Cardinal, M. F., Zubiri, D. B. V., & Bernik, D. L. (2005). Electronic nose screening of ethanol release during sol–gel encapsulation. A novel non-invasive method to test silica polymerisation. *Biosensors & Bioelectronics, 21*, 857–862.

Madene, A., Jacquot, M., Scher, J., & Desobry, S. (2006). Flavour encapsulation and controlled release – A review. *International Journal of Food Science and Technology, 41*, 1–21.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics, 11*, 431–441.

Monge, M. E., Bulone, D., Giacomazza, D., Bernik, D. L., & Negri, R. M. (2004). Detection of flavour release from pectin gels using electronic noses. *Sensors and Actuators B: Chemical, 101*, 28–38.

Monge, M. E., Bulone, D., Giacomazza, D., Negri, R. M., & Bernik, D. L. (2004). Electronic nose screening of limonene release from multicomponent essential oils encapsulated in pectin gels. *Combinatorial Chemistry & High Throughput Screening, 7*, 337–344.

Monge, M. E., Negri, R. M., Giacomazza, D., & Bullone, D. (2008). Correlation between rheological properties and limonene release in pectin gels using an electronic nose. *Food Hydrocolloids, 22*, 916–924.

Pathange, L. P., Mallikarjunan, P., Marini, R. P., Keefe, S. O., & Vaughan, D. (2006). Non-destructive evaluation of apple maturity using an electronic nose system. *Journal of Food Engineering, 77*, 1018–1023.

Peris, M., & Escuder-Gilabert, L. (2009). A 21st century technique for food control: Electronic noses. *Analytica Chimica Acta, 638*, 1–15.

Piggott, J. R. (2000). Dynamism in flavour science and sensory methodology. *Food Research International, 33*, 191–197.

Scott, S. M., James, D., & Ali, Z. (2007). Data analysis for electronic nose systems. *Microchimica Acta, 156*, 183–207.

Sinesio, F., Di Natale, C., Quaglia, G. B., Bucarelli, F. M., Moneta, E., Macagnano, A., et al. (2000). Use of electronic nose and trained sensory panel in the evaluation of tomato quality. *Journal of the Science of Food and Agriculture, 80*, 63–71.

Specht, D. F. (1990). Probabilistic neural networks. *Neural Networks, 3*, 109–118.

Struyf, A., Hubert, M., & Rousseeuw, P. J. (1997). Integrating robust clustering techniques in S-plus. *Computational Statistics and Data Analysis, 26*, 3–17.