# FEATURE EXTRACTION METHODS FOR THE ANALYSIS OF AN ELECTRONIC NOSE RESPONSE

Lamagna[1], A; Rodríguez[1], D.F.; Scoccola[1,2], N. N.

[1]Departamento de Física, Comisión Nacional de Energía Atómica, Avda. Libertador 8250, (1425) Buenos Aires, Argentina.

[2]Universidad Favaloro, Solís 453 (1078), Buenos Aires, Argentina.

Reich, S.L.[3]*,

[3]Escuela de Ciencia y Tecnología, Universidad de San Martín, Alem 3901, (1653) Buenos Aires Argentina, Fax: (54) 11 6772 7121, *E-Mail: silreich@cnea.gov.ar

*Received November 05, 2004. In final form May 23, 2005*

## Abstract

*This work deals with different methods for the analysis of the response of an electronic nose. For this purpose we employed data obtained with a prototype of electronic nose, developed in our laboratories, applied to the analysis of a basic ingredient of beer, hop. We aimed to fulfill the two main requirements of the local brewing industry, namely, deterioration of hop's quality, due to aging or poor storage and quick discrimination of the various similar types of pellets usually employed. Different methods for the analysis of the response of each sensor have been studied. A comparison between discrimination obtained with a linear dimensional reduction via Principal Component Analysis and a non linear reduction obtained with Self Organized Maps has also been made.*

## Resumen

*Este trabajo estudia diferentes métodos para la identificación de las características de una muestra medida con una nariz electrónica. Para ello se han utilizado mediciones realizadas con un prototipo de nariz electrónica, desarrollado en nuestros laboratorios, para el análisis del lúpulo que es un ingrediente esencial de la cerveza. El objetivo de este trabajo es la obtención de una identificación de características del lúpulo que sea robusta y orientada a satisfacer dos requerimientos de la industria cervecera local: detección del deterioro de su calidad por envejecimiento o almacenamiento defectuoso y la obtención de una rápida discriminación entre variedades. Se compara la discriminación obtenida mediante una reducción dimensional lineal como es el Análisis de Componentes Principales y la reducción no lineal obtenida mediante mapas autoorganizados.*

## Introduction

Hops, the female inflorescences of the hop plant *(Humulus lupulus* L.), are used in the brewing industry to add bitterness and aroma to beer. Their flowers aroma and bouquet derive from their essential oils that represent among 0.5 - 3% of the dry weight of the cone. Approximately 300 different chemical components [1] characterize these oils. However, due to volatilization during the elaboration process, only a small proportion of them is present in beer.

Hop suffers a continuous process of deterioration after the crop due to oxidative processes when exposed to ambient air. Thus, an early alarm of the onset of oxidative processes is an important parameter for the brewing industry. The essential oils in hops are currently analyzed with gas chromatography coupled to gas spectroscopy [2], although these methods are costly and time consuming. The use of electronic noses is an appealing possibility to overcome these problems

In fact, electronic noses have been widely employed in the food industry as a rather costless and efficient method specially aimed to screening techniques for, among others, quality control purposes [3].

Following requirements of the local beer industries, which demand rigorous techniques and yet not very expensive for quality control in order to successfully compete within the beer market, we studied the feasibility of electronic noses to perform some of the tasks currently undertaken by conventional gas chromatography.

A serious obstacle for the use of an electronic nose for beer control is the difficulty in the detection of some odors masked by the presence of alcohol. Thus as a first step in the feasibility study, we considered the analysis of hops as a raw material, either as flowers or preprocessed pellets.

The two main requirements of the local brewing industry are deterioration of hop's quality, due to aging or poor storage and quick discrimination of the various similar types of pellets usually employed. Besides, it is desirable to allocate future measurements to already defined categories.

Discrimination and classification are multivariate techniques concerned with the separations of distinct sets of objects (or measurements) and with the allocation of new objects to previously defined groups. More specifically, the goal of discrimination is to try to describe, either graphically or algebraically, the differential features of the data, finding the best discriminants so as to maximize the separation between classes. On the other hand the goal of classification is to derive a rule that can be used to optimally assign new measurements to the already labeled classes. Although not always possible, the method employed to discriminate may serve as an allocator. It should be mentioned that although the ability of electronic noses to deal with food characterization have been successfully verified elsewhere [4], the implementation of a classification system faces problems arising from false classifications [5] environmental disturbances [6] and correct feature extraction methods [7].

In this work, hops are analyzed through detection of their aroma by an electronic nose prototype (PampaNose 1) [8], developed in our laboratories, based on thin film metal oxide sensors. Several feature extraction methods were applied to the time dependent relative signal of each sensor. In the first method we simply extract the maximum, steady state response that implies just a number per sensor while in the others we implement feature extraction procedures that take into account not only the steady state but also the transient regime of the sensor response. This last step was implemented in two ways, considering the relative signal and its derivative as well as the Fourier transform. The extracted features were then fed into the discrimination procedure. We employed two basic discrimination schemes: linear transformation of the data, such as the popular Principal Components algorithm (PCA) and the non algorithmic transformations of Neural Networks, specifically we used the unsupervised self organized neural networks (SOM) [9]. The PCA is based on a linear transformation which projects the original N-dimensional space into one of lower (two in our case) dimension. On the other hand, the SOM might be understood as a non-linear projection of the probability density of a high dimensionality set of points into a two dimensional array of nodes (neurons). For both approaches, we checked the success of correctly discriminating and classifying by visualizing categories in a two dimensional plane.

We have also studied noisy corruption of the response. This was done, within each designed experiment, blurring the measured sensors signals with noise. In this way we were able to enlarge the data set by the artificially generated new noisy data and, therefore, to study the robustness of the analysis against disturbances.

## Experimental

The electronic nose prototype (Pampa Nose 1.[5] ) employed was developed in our laboratories. It has six metal oxide sensors (MOS) of tin dioxide whose resistivity changes according to the reducing properties of the sensed gaseous sample. Distinct electric behavior of the corresponding sensible film is achieved through doping some of these sensors with Au or with tetra t-Butyl Zinc Phthalocyanine. Two commercial sensors, for temperature and relative humidity control, are included inside the sensor chamber, to keep track of these important parameters during the measurements.

Four different cultivars of hops were selected for this work. They correspond to the usual Argentine varieties employed by the local breweries. For the analysis, 7 grams of each sample were put inside 250 $cm^3$ glass bottles. The bottles were sealed and held at room temperature (20°C±2°C) for 15 minutes in order to reach a stable headspace composition.

The headspace was afterwards fluxed into the e-nose sensor chamber at a velocity of 50 ml/min. The nose response was registered for a time interval of approximately 15 minutes each 0.5 minutes and after that, a 30 minutes purge was done. Ambient air was used as carrier gas. The sampled headspace and the carrier air were alternately switched by a three-way valve according to this time schedule.

In what follows, relative time signals are considered. That is to say, for all points of the sensor time response, the ambient air response after the transient time ($R_0$) is subtracted and the corresponding difference is divided by $R_0$ .

## Results and discussion

We have analyzed the aging and further discrimination for four varieties of hops, preprocessed as pellets, two of them (VA1 and VA2) corresponding to local varieties (Traful and Mapuche) and two types of Cascade pellets (C1 and C2).

The first proposed task is the detection of the aging of the pellets through the analysis of its aroma. Hop's pellets suffer a continuos oxidation process after crop. For industrial purposes it is important that this deterioration is noticed in advance since it affects the final quality of beer. We report the results for one variety of pellets (VA1 or Traful type). For this purpose, samples of this fresh hop's pellets were kept inside a freezer at a temperature of -20 $^{o}$C in closed recipients. They were taken from their recipients and kept in ambient air for a given period of time before the measurement. Four periods of time were considered: very short (up to 8 hours), short (up to 5 days), medium (up to 10 days) and long (more than 15 days). This categories are labeled SS, S, M and L. This process was repeated several times.

As mentioned in the introduction, three alternative feature extraction methods were employed for the sensor signal:

f1) maximum steady peak value of the relative signal response vector,

f2) maximum steady peak value of the relative signal response vector and maximum absolute value of its first derivative,

f3) coefficients of the discrete Fourier transform of the signal obtained by the fast algorithm (FFT).

We discard other methods that would give more than three significant figures for the typical sensor signal that has been sampled at 32 points.

After the feature extraction, the information was reduced using both the PCA algorithm and the SOM method. The PCA algorithm was performed using the covariant matrix and taking into account the first two principal components. For the SOM analysis a 10x13 rectangular network with a Gaussian neighboring function was used. In both cases the input consists of one number per sensor in the f1 case, two per sensor for method f2 and three for method f3 since for the Fourier transform, only three frequencies appear with non negligible amplitude.

Results for the feature extraction possibilities f1, f2 and f3 and for the two above mentioned dimensional reductions are displayed in Figures 1a to 1e. In each case the left panel corresponds to the PCA analysis and the right panel to the SOM method, where only the winner neurons are marked in the plot. Figures 1a and 1b, indicate that the use of the maximum amplitude of the signal, either with PCA or SOM, allows to separate the aging of the hop's pellets in three categories. The first comprises short and very short periods, that is, there is no visible differentiation of hops aroma after 5 days of storage in ambient air. The second region corresponds to medium exposure to ambient air. Finally, a third region corresponding to those samples that were exposed more than 15 days to ambient air can be discriminated. Either for PCA or for SOM projection it is verified that increasing aging of the sample, corresponds to increasing distances in the two-dimensional plane from the fresh hop's locus. The second feature extraction method f2, includes information about the maximum of the first derivative of the signal, a parameter measuring the rise time. Figures 1c and 1d show that, although not very significant, this second method allows a somewhat better discrimination for the PCA projection and almost the same result for the SOM reduction. In any case, the f2 extraction procedure shows a smaller dispersion among the same class while the distance between classes is slightly increased. Figures 1e and 1f show the limitation of using, in this case, the Fourier transform as a feature extraction method from the signal. Probably the approximation of the signal function by its Fourier decomposition introduces uncertainties that blurred the final discrimination.

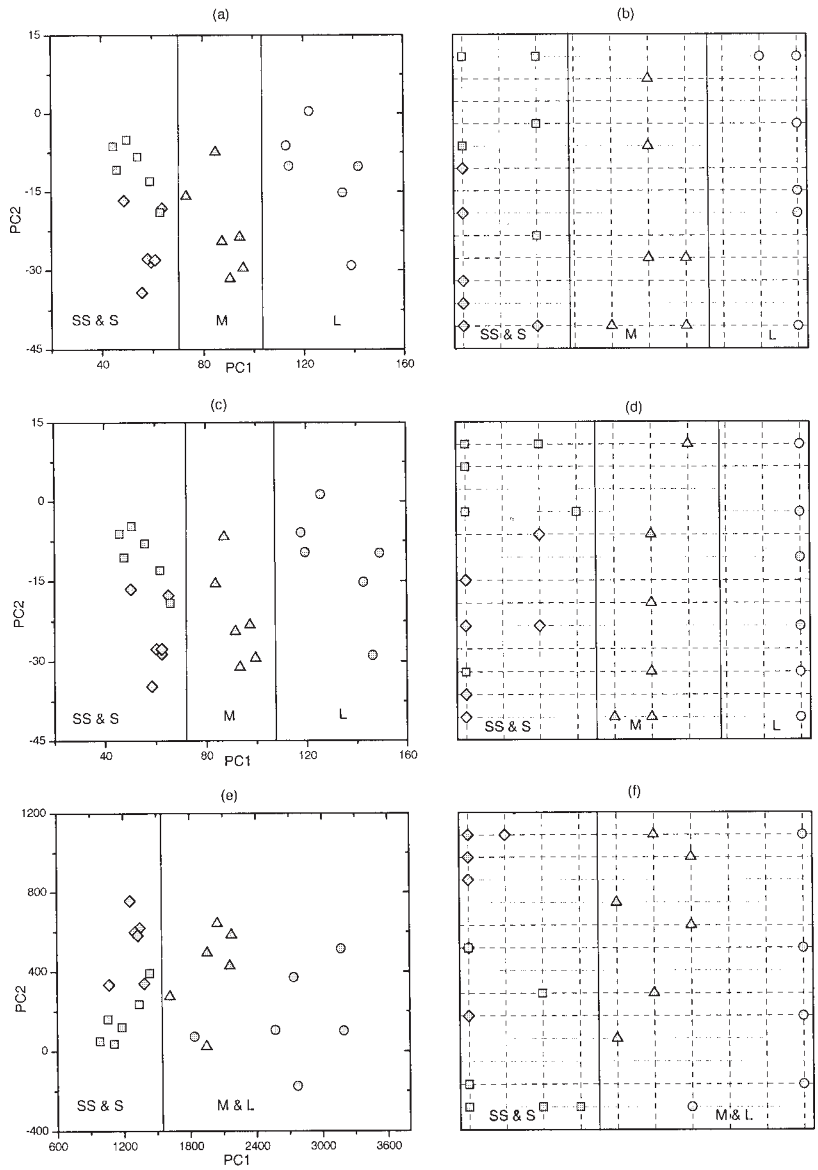


***Figure 1:*** *Different days of storage: square symbols correspond to very short (SS), rhomboids to short (S), triangles to medium (M) and circle to large (L) storage time. Score plot for PCA analysis of the nose response for feature extraction methods f1 (a), f2 (c), and f3 (e). SOM network with the winner neurons for feature extraction methods f1 (b), f2 (d), and f3 (f).*

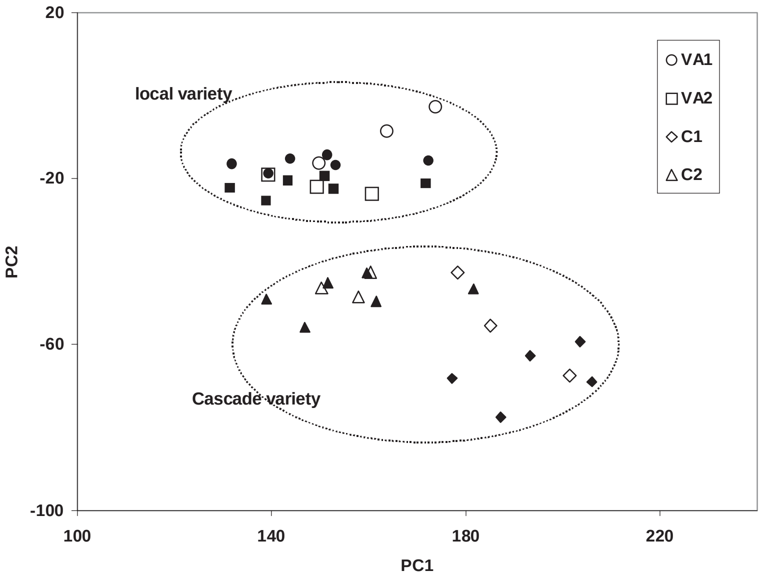


***Figure 2:*** *Score plot for PCA analysis of the nose response for different hop's varieties. Open symbols correspond to measured data while solid represent their noise corruption. Almost 94% of the total variance in the first two Principal Components.*

The second study carried out concerns the discrimination among the four pellets' varieties considered: two of them (VA1 and VA2) corresponding to local varieties (Traful and Mapuche) and two types of Cascade pellets (C1 and C2). For each of them, three different samples were measured in different days. We do not report here consecutive measurements of the same sample, since they show no significant variations among themselves.

Based on the results of the aging analysis discussed above, in this part we have only applied method f1 for the feature extraction from the signal and the linear dimensional reduction PCA as discrimination method. The use of other methods (i.e. SOM as discrimation method) is not expected to lead to significantly different results.

Results corresponding to the measured data are indicated with open symbols in Figure 2 with 94% of the total covariance taken into account with the first two principal components. Once the projection transformation was defined, new data was mathematically generated and then projected onto this space. They were generated adding a random noise (of about 15%) to the already registered signals. The results are plotted with the same convention but with solid symbols. This was done to analyze the possibility to define regions in the (PC1-PC2) plane that can be assigned to a given variety of pellets. It can be seen from Figure 2 that there can be establish two different regions. One for VA1 and VA2 varieties, the local cultivars, while the two Cascade type pellets can be grouped into another region. This discrimination is also supported by gas chromatography studies performed by the laboratories of one local brewery.

## Conclusions

The results presented in this work indicate that the use of the Pampa Nose provides an adequate tool to carry out several tasks concerning the analysis of hop properties by measuring their aromas in ambient air conditions. The aging process was studied via three different alternatives for feature extraction of the sensors signals and two discriminating methods (PCA and SOM). We were able to verify that with both methods it is possible to associate increasing aging of the hops with increasing distance from the fresh aroma locus in a two dimensional plane and thus it is possible to introduce a «cut-off» for undesired hops rather quickly and costless.

Another point to be mentioned is that taking into account the rise time properties together with the maximum relative signal, the discrimination is slightly improved. The use of the Fourier spectrum of the signal implies that at least components for three frequencies are required but in spite of the apparently better information about the relative signal, no achievement above the simpler method is obtained and the same behaviour is observed either with SOM or PCA. We were also able to discriminate between different varieties of hops with a PCA analysis of the nose response and, introducing mathematically noisy corruptions of the signals, we checked the ability of a previously defined (PC1-PC2) plane to allocate different varieties to the corresponding class.

We expect that the present study will allow to implement an adequate and rather simple e-nose analysis to be operated by the brewery industry for hop's quality control.

## Acknowledgments

The authors wish to acknowledge financial support of J.P.A. S.A. (Javier Picco y Asoc.) through PID131-2001 of the National Agency for the Promotion of Science and Technology (ANPCYT) of Argentina.

## References

[1] Peacock, V.E.; Deinze, M.L., *J. M. Soc. Brew. Chem.*, **1981**, *39*, 136-141.
[2] Lermusieau, G.; Bulens, M.; Colli, S., *J. Agric. Food Chem.*, **2001**, *49*, 3867-3874.
[3 Gardner, J.W.; Bartlett, P.N., *Electronic Noses and Applications*, Oxford University Press, Oxford, **1999**.
[4] Bartlett, P.N.; Elliot, J.M.; Darner, J.W., *Food Technol*, **1997**, *51*, 44-48.
[5] Goodner, K.; Glen Dreher, J.; Rouseft, R., *Sens. Actuators B* , **2001**, *80*, 261-266.
[6] Di Natale, C.; Martinelli, E.; D'Amico, A., *Sens. Actuators B*, **2002**, *82*, 155-165.
[7] Distante, C.; Leo, M.; Siciliano, P.; Persaud, K., *Sens. Actuators B*, **2003**, *87*, 274-288.
[8] Lamagna, A.; Reich, S.; Negri, M.; Boselli, A.; Cocco, M.; C. Di Natale, C., *Thin Solid Films*, **2002**, *418*, 42-44.
[9] Kohonen, T., *Self-organizing Maps*. Springer-Verlag, Heidelberg, **1995**.