Copyright © 2018 American Scientific Publishers
All rights reserved
Printed in the United States of America

SENSOR LETTERS
Vol. 16, 1–5, 2018

Detection and Classification of Hydrolyzed Dairy Additives Using Electronic Noses

J. Vorobioff[1,*], M. Adaro[2,3], N. Boggio[1,4], J. Magallanes[1,4], A. Boselli[1], A. Lamagna[1,4], and C. Rinaldi[1,4]

[1] *Comisión Nacional de Energía Atómica, Av. Libertador 8250, 1425 C.A.B.A, Argentina*
[2] *Laboratorio de Bromatología, Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis*
[3] *INFAP-CCT-San Luis CONICET, Chacabuco y Pedernera (5700), San Luis, Argentina*
[4] *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290 (1425) Ciudad Autónoma de Buenos Aires, Argentina*

(Received: xx Xxxx xxxx. Accepted: xx Xxxx xxxx)

In the present work the classification of dairies additives in a rapid and simple way, is proposed. Measurements were made by means of an Electronic Nose developed in our laboratories, which we named 'Patagonia E-Nose.' This E-Nose is composed of SnO_2 sensors located in a thermally stabilized chamber which improves the repeatability of the measurements. Samples of various hydrolyzed dairies were measured using air as a reference gas. Then, the integrals of the signals were analyzed using a combination of different multivariate chemometric methods such as linear discriminant analysis (LDA), principal components analysis (PCA), back-propagation neural network and different classifiers. Also, different algorithms were implemented and compared by calculating the number of correctly classified samples of each method. 99.4% of correct classifications were obtained by using cross-validation and selecting the most appropriate algorithms. The results indicate that the samples were correctly classified through the implementation of a simple and low cost measurement protocol.

Keywords: Dairies Ingredients, Electronic Nose, Metal Oxide Gas Sensors, Protein Hydrolysates Cheese, Multivariate Analysis.

RESEARCH ARTICLE

1. INTRODUCTION

Dairy Ingredients such as hydrolysates are additives used extensively in food industry due to their nutritional or functional properties (solubility, emulsifying power, foaming capacity). In the hydrolysis of proteins to peptides or amino acids, by the action of proteolytic enzymes, the final composition and hence the use of the hydrolysates will depend on the protein source, the type of protease used, the hydrolysis conditions and the hydrolysis grade obtained in the reaction.[1]

In protein hydrolysates, a variety of functional characteristics, such as low viscosity, higher capacity stirring, dispersion and high solubility are enhanced, which give advantages for using them in many food products.[2–6]

One of the most important uses of protein hydrolysates is presently as a nitrogen source in the formulation of diets intended for infant and sick adults. These enteric products are designed to be absorbed in the intestine without prior digestion in the stomach and are essential to the treatment of stomach disorders or problems of the intestinal mucosa, as well as infants with malabsorption or malnutrition syndromes with allergy symptoms in most cases.[7]

The characteristics that must satisfy these proteins hydrolysates in the case of an enteric diet are not to produce osmotic imbalances or allergies, have a high nutritional value not lower than the starting protein and have an acceptable taste.

The degree of hydrolysis is defined as the percentage of broken peptide bonds in relation to the original protein. It is a fundamental property of the additive and will precisely determine other characteristics of it and their possible uses. The degree of final hydrolysis is determined by the process conditions applied such as: substrate concentration, enzyme/substrate ratio, incubation time and physicochemical conditions as pH and temperature. Another factor that will also determine the degree of hydrolysis is the nature

*Corresponding author; E-mail: vorobioff@cnea.gov.ar

of the enzyme, characterized by its specific activity and type of activity. Then the nature of the used enzyme will not only influence the degree of hydrolysis, but also the type of peptides produced.[1]

Hydrolysates produced for foods can be grouped into hydrolysates with low degree of hydrolysis, 1% to 10%, to improve the functional properties; hydrolysates with varying degrees of hydrolysis, for use as flavoring and extensive hydrolysates, which hydrolysis degree exceed 10%, for use in specialized nourishment.[1] The raw material used for producing hydrolysates become from animal, plant or bacterial origins. The protein hydrolysis is usually carried out in a reactor controlling stirring, pH, temperature and time processing. Processes for preparation of hydrolysates and their applications are explained in the literature.[1, 8]

Moreover, attention is paid nowadays to the development and production of food products based on scientific requirements to human diets. Special attention should be paid to high-protein cheeses' products dairies.[8]

Recently, the food industry began to commercialize enzymatically modified cheeses, which are products whose flavor intensity is several times the original cheese and consequently can be used to flavor foods and to accelerate the ripe of other cheeses.[9] Depending on alternative procedures such as the final degree of hydrolysis, substrate concentration, incubation time, physicochemical conditions such as pH and temperature, etc., result in multiple situations of compromise from the operational point of view.[10–12] Due to a large number of variables involved in this process and the essential need to perform measurement efficiently with the proper economy in the number of trials, time and costs, the Patagonia E-Nose designed in our laboratories was used. Patagonia is composed of SnO_2 sensors in a thermally stabilized chamber which improves the repeatability of the measurements.[13, 14]

The aim of this work is to analyze and evaluate modified dairy additives to incorporate them into different foods composed of flour and/or milk. According to our knowledge, no bibliographic records on the use of electronic noses applied to the study of the odors properties of protein hydrolysates combined with the use of chemometrics as artificial neural networks are known.

2. EXPERIMENTAL DETAILS

2.1. Samples

The following samples with different degree of hydrolysis were analyzed (see Table I): additives for food composed by flour (Sample No. 1 and 2) and additives for food dairy (Sample No. 3 and 4). As is usual in this type of measurements with electronic nose it is necessary to perform a blank,[14] in this case without hydrolyzing proteins. In this way, a set of control (named Control No. 1 to 4) was tested, consisting of non hydrolyzed samples in order to compare with those hydrolyzed additives.

Table I. Resume of the samples analyzed.

Type of additive	Control No	Sample No
Food flour type 1	1	1
Food flour type 2	2	2
Food dairy type 1	3	3
Food dairy type 2	4	4

2.2. Experimental Design

Measurements were performed with the Patagonia Electronic Nose (E-Nose) developed in our laboratories, shown in Figure 1(a).[14–19] It is assembled for 8 commercial SnO_2 metal oxide thin film semiconductor as gas sensors placed inside a heated chamber that allows to reach thermal stability and to improve the repetitive measurements. The experimental set up is shown in Figure 1(b). During operation, the sensors operate at 380 °C, while the chamber containing them is heated to 40 °C. As described in Table I four samples were measured using air as a carrier gas.

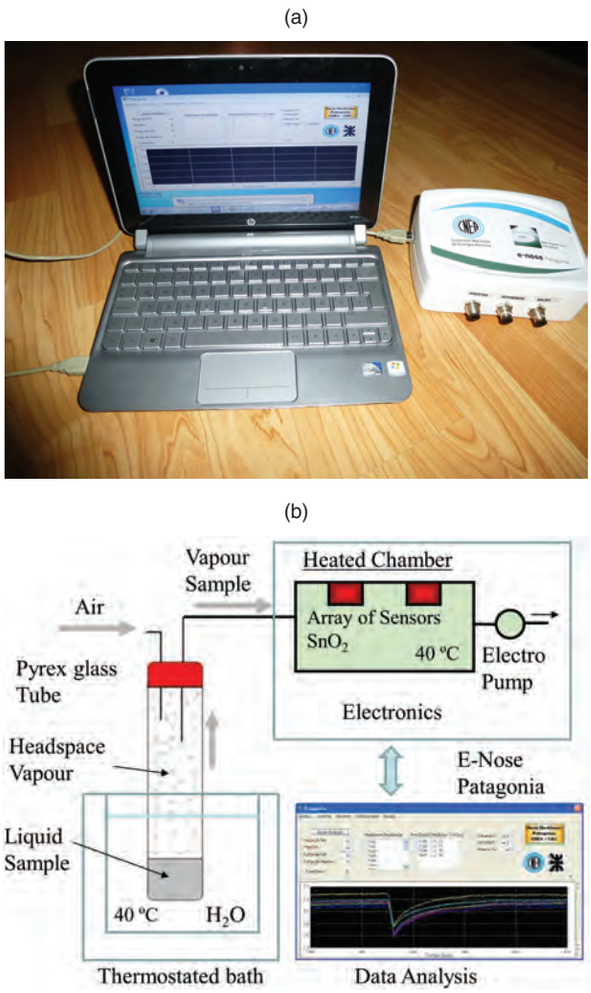

Fig. 1. (a) Patagonia E-nose with notebook associated and software. (b) Experimental set up for the measurements of dairy additives.

RESEARCH ARTICLE

Every liquid samples were stored in a Pyrex glass tube immersed inside a thermostatized bath at 40 °C to increase the concentration of volatile compounds in the headspace. A total of eight measurements for each sample were taken.

The following protocol was applied to each measurement:

—An initial air purge is performed during 75 seconds in order to clean the sensors chamber.

—Subsequently, the sample is absorbed in 19 seconds.

—Finally, an air purge is performed during 77 seconds.

The sample and air purge flow are controlled by a micropump at 350 $cm^3 \cdot min^{-1}$.

2.3. Multivariate Analysis of Data

The integrated signals were analyzed using different multivariate chemometric methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).[20]

Geometrically, principal components analysis can be thought of as a rotation of the axes of the original coordinate system to a new set of orthogonal axes that are ordered in terms of the amount of variance of the original data they account for. One of the reasons for performing a principal components analysis is to find a smaller group of underlying variables (named latent variables) that describe the data. In order to do this, we expect that the first few components will account for most of the variance of the original data. The principal component analysis is a variable-directed technique. It makes no assumptions about the existence of groupings within the data and so is described as an unsupervised feature extraction technique.[20]

The term linear discriminant analysis (LDA), although generically referring to techniques that produce discriminant functions that are linear in the input variables, in some sense, maximizes between-class separability and minimizes within-class variability. The axes of the transformed coordinate system can be ordered in terms of 'importance for discrimination.' Those most important can be used to obtain a graphical representation of the data by plotting the data in this coordinate system (usually, two or three dimensions).[21]

The results obtained in two dimensions by method PCA and LDA are classified by the back-propagation neural network and linear classifiers[22–26] (see Figs. 3–5). These methods are provided by the E-Nose-Pat version 1.01 software.

2.4. Classifiers and Cross-Validation

Three matrices of data were analyzed: a matrix of the integral of the signals, another matrix of LDA results and another matrix of PCA results. These data were separated by different classifiers: K-nearest neighbors (K-NN), Linear Classifier and Back-propagation neural networks.[22–26] These classifiers were implemented by libraries in C++ language, Octave (free license) and Matlab.

The K-NN method is used for classification. In this method, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its K nearest neighbors (K is a positive integer, typically small). If $K = 1$, then the object is simply assigned to the class of that single nearest neighbor. The choice of K has a different effect on the K-NN classifier obtained. K-NN is a completely non-parametric approach: no assumptions are made about the shape of the boundary decision. Therefore, we can expect this approach to dominate LDA and logistic regression when the decision boundary is highly non-linear. On the other hand, KNN does not tell us which predictors are important; we do not get coefficients. When the true decision boundaries are linear, then the LDA and logistic regression approaches will tend to perform good. For much more complicated decision boundaries, a non-parametric approach such as K-NN can be superior.[27] In this work, different K values were tested, obtaining better results with $K = 3$.

In general, as we use more flexible classification methods, the training error rate will decrease, but the test error rate will not.[27]

The performance of each algorithm was analyzed by cross-validation.[28] Twenty percent of the total samples were randomly selected and classified. The number of correct classifications (CC) was estimated by calculating the number of samples correctly classified divided by the total number of samples. For statistical purposes, this procedure was repeated 40 times for each algorithm. In this way, the robustness of the algorithms was also checked using various input data.

3. RESULTS AND DISCUSSION

A typical set of signals obtained with the E-nose as normalized resistance versus time is shown in Figure 2. The integrals of these signals were analyzed through the

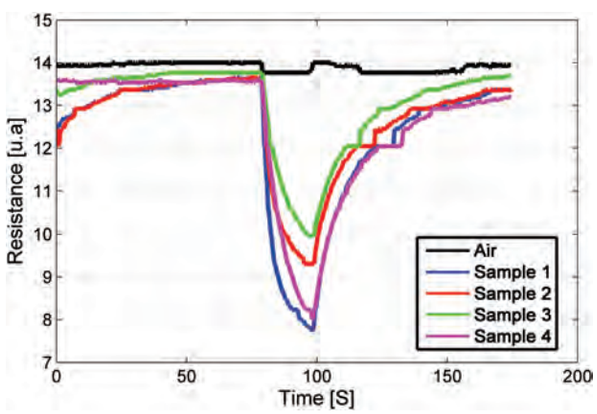

Fig. 2. A typical signal of the Patagonia E-nose. Differences between the four samples and air signals are observed.

RESEARCH ARTICLE

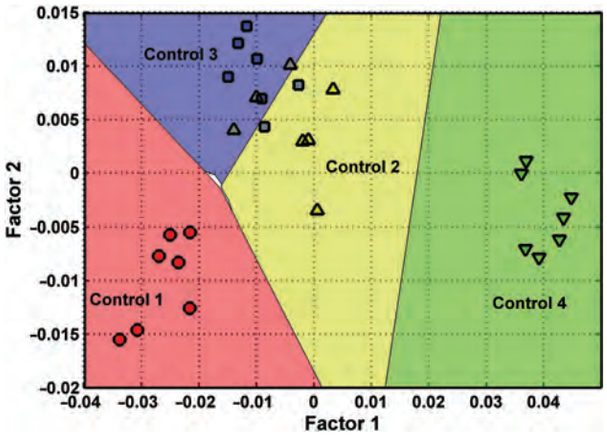

Fig. 3. LDA for control (without hydrolysis). (●) Control 1, (▲) control 2, (■) control 3 and (▼) control 4.

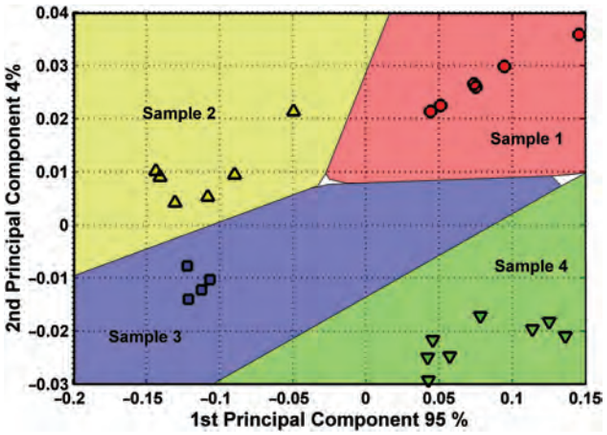

Fig. 5. PCA for samples. (●) Sample 1, (▲) sample 2, (■) sample 3 and (▼) sample 4.

application of different processing algorithms: Linear Discriminant Analysis (LDA), Principal Components Analysis (PCA) and Back-propagation Neural Network.

Figures 3 and 4 show the results obtained by LDA corresponding to non-hydrolyzed (Control) and hydrolyzed samples respectively. As can be seen, an appropriate discrimination of the samples and control system (without hydrolysis) was obtained; this is due to different compositions of the control samples and furthermore, different degree of hydrolysis in each hydrolyzed sample.

Figure 5 shows PCA of the hydrolyzed samples. It can be observed that although the 4 samples are correctly separated and classified, the figure shows that, according to the values of the axes in the PCA (PC1 95% and PC2 4%), samples 2 and 3 present greater similarity and the same behavior is observed for samples 1 and 4.

The different types of samples (with different degree of hydrolysis) and control system (without hydrolysis) allowed the electronic nose to distinguish the differences between them. The hydrolysis process helps the release of volatile compounds of low molecular weight, intensifying the aroma of the sample.

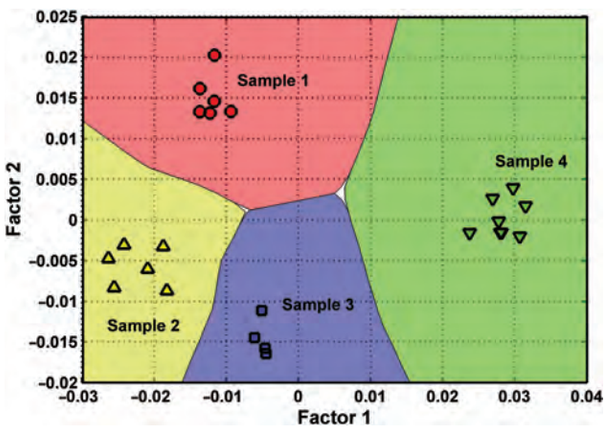

Fig. 4. LDA and neural network for samples. (●) Sample 1, (▲) sample 2, (■) sample 3 and (▼) sample 4.

Table II. Result for the analysis of different classifiers.

	Correct classifications CC (%)	
Algorithms	For control	For samples
K-NN euclidean metric	70.5	98.75
K-NN classifier citiblock metric	75	95.6
PCA + linear classifier	72	99.4
LDA + linear classifier	79	98.75
LDA + *K*-NN euclidean metric	77.5	98.75
LDA + neural network back-propagation	65	80.6

In order to evaluate the capability of the combination of different algorithms for data analysis, correct classifications (CC) were estimated. Table II present the results of the analysis of different classifiers, showing the CC percentage.

According to the results of Table II, high values of %CC for samples were obtained; however, for control system, this number is lower. The most significant %CC value is the combination of PCA with linear classifiers: 99.4% of correct classifications were obtained. This is because the measurements of the control system are more similar between themselves than hydrolyzed samples, affected by different degrees of hydrolysis.

4. CONCLUSIONS

A simple and low-cost method of classification of hydrolyzed dairy additives was implemented using an electronic nose. Different algorithms were implemented and compared by calculating the number of correctly classified samples of each method. By using cross-validation and selecting the most appropriate algorithms, 99.4% of correct classifications were obtained.

The application of the PCA and LDA methods allows a quick visualization of the results in the two-dimensional projection data. Moreover, to strengthen the ability of

RESEARCH ARTICLE

discrimination between samples, different methods were implemented with these classifiers: K-nearest neighbors (K-NN), Linear Classifier and Back-propagation neural networks.

It is also important to admit that depending on the conditions of the hydrolysis, the ability to separate by the E-nose may vary, being necessary to adapt the algorithms to achieve a good separation of samples.

This methodology represents a potential application in the food industry in order to measure and consequently to control the production of hydrolyzed food additives. This work aims to contribute to the study of the functionality of a single ingredient or additive, providing useful tools to predict, control and induce desirable functional characteristics to real food systems.

Acknowledgments: The authors gratefully acknowledge financial support from the National Council of Scientific and Technical Research (CONICET), Comisión Nacional de Energía Atómica and the Facultad de Química, Bioquímica y Farmacia of Universidad Nacional de San Luis from Argentina.

References and Notes

1. R. Benítez, A. Ibarz, and J. Pagan, Hidrolizados de proteína: Procesos y aplicaciones. *Acta bioquím. clín. latinoam* 42, 227 (**2008**).
2. S. Yin, C. Tang, J. Cao, E. Hu, Q. Wen, and X. Yang, Effects of limited enzymatic hydrolysis with trypsin on the functional properties of hemp (*Cannabis sativa* L.) protein isolate. *Food Chem.* 106, 1004 (**2008**).
3. X. Kong, H. Zhou, and H. Qian, Enzymatic preparation and functional properties of wheat gluten hydrolysates. *Food Chem.* 101, 615 (**2007**).
4. X. Kong, H. Zhou, and H. Qian, Enzymatic hydrolysis of wheat gluten by proteases and properties of the resulting hydrolysates. *Food Chem.* 102, 759 (**2007**).
5. I. Paraman, N. S. Hettiarachchy, C. Schaefer, and M. I. Beck, Hydrophobicity, solubility, and emulsifying properties of enzyme-modified rice endosperm protein. *Cereal Chem.* 84, 343 (**2007**).
6. V. P. Ruíz-Henestrosa, C. Carrera-Sanchez, M. Yust, J. Pedroche, F. Millan, and J. M. Rodriguez-Patino, Limited enzymatic hydrolysis can improve the interfacial and foaming characteristics of beta-conglycinin. *J. Agric Food Chem.* 55, 1536 (**2007**).
7. E. Lebenthal, P. C. Lee, and L. A. Heittinger, Impact of development of the gastrointestinal tract on infant feeding. *J. Pediatr.* 102, 1 (**1983**).
8. L. Nadtochii, L. Zabodalova, and M. Domoroshchenkova, Development of cheese product with hydrolyzed soybean emulsion. *Agronomy Research* 13, 1010 (**2015**).
9. J. E. Kinsella, Functional properties of proteins in foods: A survey. *CRC Crit. Rev. Food Sci. Nutr.* 7, 219 (**2009**).
10. B. Hernández-Ledesma and A. Quiros, Angiotensin-converting enzyme inhibitory activity of peptides derived from caprine kefir. *J. Dairy Sci.* 88, 3480 (**2005**).
11. S. V. Cavalli, S. V. Silva, C. Cimino, F. X. Malcata, and N. Priolo, Hydrolysis of caprine and ovine milk proteins brought about by aspartic peptidases from *Silybum marianum* flowers. *Food Chem.* 106, 997 (**2008**).
12. J. Adler-Nissen, Determination of the degree of hydrolysis of food protein hydrolysates by trinitrobenzenesulfonic acid. *Food Chem.* 27, 1256 (**1979**).
13. J. Vorobioff, D. Rodriguez, N. G. Boggio, and C. Rinaldi, Development of an electronic nose for determining the freshness of fish by the desorption constants of sensors. *Sensor Lett.* 11, 2215 (**2013**).
14. K. Pierpaulli, J. Vorobioff, N. Boggio, C. Rinaldi, C. A., S. Reich, and A. Lamagna, Improvement of the sensitivity of gas sensor by low laser power irradiation, *13th International Meeting on Chemical Sensors-IMCS-13*, Australia (**2010**).
15. S. M. Scott, D. James, and Z. Ali, Data analysis for electronic nose systems. *Microchem. Acta* 156, 183 (**2007**).
16. J. W. Gardner and P. A. Bartlett, Brief history of electronic noses. *Sensors and Actuators B* 18, 211 (**1993**).
17. T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner, Handbook of Machine Olfaction-Electronic Nose Technology, Wiley-VCH, Weinheim (**2003**), pp. 79–104.
18. T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner, Handbook of Machine Olfaction-Electronic Nose Technology, Wiley-VCH, Weinheim (**2003**), pp. 81–84.
19. E. A. Baldwin, J. Bai, A. Plotto, and S. Dea, Electronic noses, and tongues: Applications for the food and pharmaceutical industries. *Sensors* 11, 4744 (**2011**).
20. A. R. Webb and K. D. Copsey, Statistical Pattern Recognition, 3rd edn., Wiley (**2011**), Chap. 5.
21. A. R. Webb and K. D. Copsey, Statistical Pattern Recognition, 3rd edn., Wiley (**2011**), Chap. 10.
22. H. B. Demuth, M. H. Beale, and M. T. Hagan, Neural network toolbox for use with MATLAB: User's guide 9th for version 6.0, https://filer.case.edu/pjt9/b378s10/nnet.pdf (**2008**).
23. B. M. Del Brio and S. Molina, Redes Neuronales y Sistemas Borrosos, 3rd edn., Alfaomega (**2007**), Chap. 2.
24. C. M. Bishop (ed.), Pattern Recognition and Machine Learning, Springer (**2006**), pp. 559–577.
25. F. Vojtech and H. Václav, Statistical pattern recognition toolbox for matlab, Research Reports of CMP, Czech Technical University in Prague No. 8 (**2004**).
26. S. Theodoridis and K. Koutroumbas (eds.), Pattern Recognition, Academic Press (**2009**), pp. 323–341.
27. G. James, D. Witten, T. Hastie, and R. Tibshirani (eds.), An Introduction to Statistical Learning: With Applications in R. Corrected Edition, Springer (**2013**).
28. A. R. Webb and K. D. Copsey, Statistical Pattern Recognition, 3rd edn., Wiley (**2011**), pp. 581–587.