UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Improving MS-Sensor technologies for food quality assessment

PhD Thesis

Maria Vinaixa Crevillent

Supervisors

Associate Professor Jesus Brezmes Llecha

Professor Xavier Correig Blanchar

UNIVERSITAT ROVIRA I VIRGILI

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Als de casa,

Al Gabriel

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Data does not equal information; information does not equal knowledge; and, most importantly of all, knowledge does not equal wisdom.

We have oceans of data, rivers of information, small puddles of knowledge and the odd drop of wisdom.

Henry Nix, 1990

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI

Departament d'Enginyeria Electrònica, Elèctrica i Automàtica

Improving MS-Sensor technologies for food quality assessment

Presented by Maria Vinaixa Crevillent
as a fulfillment of the requirements for the
Ph.D degree at Universitat Rovira i Virgili,
Tarragona, March 2008

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
Departament d'Enginyeria Electrònica,
Elèctrica i Automàtica

Dr. JESÚS BREZMES LLECHA and **Prof. XAVIER CORREIG BLANCHAR**, Associate professor and Professor, in the Departament d'Enginyeria Electrònica, Elèctrica i Automàtica at Universitat Rovira i Virgili

CERTIFIE:

The Doctoral Thesis entitled: **"IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT"**, presented by **MARIA VINAIXA CREVILLENT** to receive the degree of Doctor of the Rovira i Virgili University, European PhD, has been carried out under our supervision, in the in the Departament d'Enginyeria Electrònica, Elèctrica i Automàtica at Universitat Rovira i Virgili, and all the results presented in this thesis were obtained in experiments conducted by the above mentioned student.

Tarragona, November 2007

Dr. Jesús Brezmes Llecha

Dr. Xavier Correig Blanchar

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Preface

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

ACKNOWLEDGEMENTS

Generalment, en una tesi s'acostuma a fer una pàgina d'agraïments. Cal reconèixer que om ha pogut fer una tesi doctoral gràcies a que tot un seguit de persones, administracions i institucions l'han acompanyat i recolzat durant el procés. Ben sovint aquests procés és de tot menys fàcil. Es sol arribar al final amb ganes d'acabar i de passar pàgina, d'enllestir les coses i deixar-se estar de complicacions. Fer una pàgina d'agraïments, creieu-me, pot arribar a ser complicat. Per raons de temps, es podria caure en la temptació de tendir finalment a la simplificació i no agrair la tesi. Vaig rumiar força si volia que en aquesta tesi hi fos present una pàgina d'aquests caire. A la fi, vaig decidir fer-ho. No podia deixar d'escriure la única pàgina que de ben segur serà llegida. A més a més tampoc em volia deixar escapar l'oportunitat d'incloure la única pàgina que no ha de cenyir-se als encorsetaments i cànons d'una tesi científica. Em prenc la llicència doncs d'escriure aquestes ratlles a la meva manera, tal com raja, parlant des dels sentiments. És per això que en aquesta pàgina voldria mencionar a tot un seguit de gent que, en la majoria dels casos, malgrat no haver intervingut en el més estricte àmbit científic, m'han recolzat d'una altra manera. Per mi això ha estat decisiu, sense vosaltres aquesta tesi no hauria anat a parar a bon port. Sé que segurament seré poc justa, sóc ben conscient em deixaré certes persones amb les que he treballat i mereixen ser-hi a la vegada que seria igualment injust mencionar-ne d'altres.

Als meus directors, al Jesús per haver-me escoltat tots aquests any, per la seva empatia i per tots els "the" que m'ha corretgit; al Xavier per haver-me esperonat; per tantes i tantes discussions, per engrescar-me i abocar-me a tants projectes. A tots dos, moltíssimes gràcies per la paciència, pel vostre esforç i per haver-me supervisat aquesta tesi. Sé que de vegades no ha estat fàcil. Gràcies!

A l'Eduard qui mai ha tingut un no per a les meves preguntes i em va ajudar de manera decisiva al començament d'aquesta tesi. Gràcies per fer de corrector i crític dels meus treballs.

I am so indebted with Rasmus Bro, who without knowing me allows me to joint his group for a three-month stage. In this stage I took my first insight in to multi-way data analysis. They were really a three inspiring months in the Chemometric group at the Royal Veterinary and Agricultural University during winter 2005 where the large group of PhD students and scientists provided a stimulating environment for working with and discussing. A special thank is expressed to Thomas Skov and Frans van den Berg for their kindly help and support during this period.

A los compañeros del 334; al Raúl quien estuvo al principio a mi lado con Matlab y con tantas otras cosas ; a Àlex....por ser Patinet, por los asteriscos chuecos, los chilaquiles, las charlas y utilizarme de bola desestresante...; a Edwin ufff…., Amigo Mío no tengo palabras, gracias por

escucharme, por los consejos, las piscinas, las nadadas hasta la boya, la playa, los cafés; a Edgar...gracias por polemizar, por las risa y por tu buen humor. A Radu por su presencia a mis espaldas durante tanto tiempo y con quien tanto nos reímos y peleamos. Gracias chicos del 334....sabéis que siempre os echaré de menos. Vaya equipo!. Quiero también dar las gracias a Stella por todas las "envidias" que me da cada día, por la compañía y sobretodo por compartir estos últimos tiempos de escritura de la tesis y poder hacer terapia de grupo y también por algunas ideas en el formato final. Al Fer por los buenos ratos, sus alacranes, la maña, la pericia y el hacer las cosas fáciles.

A Mariana, gracias por tu compañía, los cafés lentos, los cigarrillos compartidos al principio, los que no compartimos después, gracias por tu criterio científico, tu honestidad y sinceridad, gracias por trabajar conmigo y gracias por brindarme tu amistad y las llaves del hotel. Un guiño especial a Triffon, mi ídolo en esto de la investigación.

A la gent que s'ha incorporat més recentment com la Roser, gràcies per la teva energia, vitalitat i la teva empenta. A Belén, mi compañera en este último mes, a la que quisiera agradecer todas las conversaciones científicas, su compañía y su ánimo.

Als companys i ex-companys del departament, a la Margarita i el David pels esmorzars i per aguantar-me el rotllo cada dia, al Vilanova per haver-me atès sempre que l'he necessitat, a l'Anna per algunes hores intempestives i totes les converses que compartirem a la secretaria, a la Sofia per la bona acollida al principi, a la Mireia, al Maré, al Badi al Deosdad i al Carmelo. A alguns dels que ja no hi són com el Miquel, el Dani, el Benito, l' Eliana, l' Eli, l' Ivan i Mikhail. Als "chicos checos" Lukas i Zdenek.

Als les manes Imma, Encarna, MªCarme; als de Barcelona, Chewie, Xavi, Dear, Patri, Morrisson, Maribel, Gabriel; a les comanyes de Tarragona, Clara, Esther, Nuria i Patri; a les xiques del poble, Joana, Gemma, Carme, Flari, Cris, Esther; a les d'Ascó Cinta, Marina i Amparo.

Als les meves arrels, Pare, Mare, Iaia, Belén, Moya, Marga, Miquel, Padrina, Tío Marsal, D.J, Ramón, Txema, Jacob, Andreu, Anna, Padrí, Tieta Annita, Ioli, Moni, Marcel i Abraham.

Al Gabriel,

MOLTÍSSIMES GRÀCIES A TOTS!

Mariona, Novembre 2007

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

TABLE OF CONTENTS

SUMMARY

The food industry is demanding fast screening methodologies for analyzing odour and flavour in order to guaranty food quality and safety. These methodologies should allow high throughput with sufficient accuracy and reproducibility. Even though sensory analysis and instrumental techniques such as gas chromatography-mass spectrometry (GC-MS) remain powerful and trusted tools for resolving this type of problems, setting up such methods requires a considerable amount of work, time and investment. In one hand, instrumental techniques are expensive to operate since they require a lot of effort and considerable skill on method development. On the other hand, sensory analysis performed by a panel of experts is a costly process for industries because it requires trained people who can work only for relatively short periods of time. There are also other problems, such as the subjectivity of the human response to odours and the variability between individuals. Consequently, there is an enormous demand for an instrument that can overcome the problems of human sensory panels and the main weaknesses of instrumental techniques. Early in the eighties, a new concept known as the *"electronic nose"* (artificial nose) which seem to fulfil all these requirements was developed. With the term *"electronic nose"* it is understood an array of chemical gas sensors with a broad and partially overlapping selectivity for the measurement of volatile compounds within the headspace of a sample combined with multivariate statistical tools that allows to classify the samples according to their volatile composition. Since the design of the first *"electronic nose"*, which was based on gas sensors, hundreds of articles related to this configuration have been published. They cover fields of applications as diverse as the food industry, the environment and medicine, among others. Despite some promising and impressive results obtained with such configuration, certain limitations have appeared. They include loss of sensitivity in front of polar compounds such as ethanol and relatively short life of some sensors among others. Recently, manufacturers of analytical instruments and a number of research laboratories have become interested in direct injection-mass spectrometry as a potential means of obtaining more rapid and easy solutions able to overcome some of the drawbacks of the classical*"electronic nose"* approach. In the late nineties the so-called MS-Sensor or *"new*

generation electronic nose" was developed. The basic working principle of MS-Sensor systems is based on the introduction of volatile components extracted from the headspace of a sample into the ionization chamber of a mass spectrometer. The mass spectrum resulting from the ionization and fragmentation of this extract constitutes a very complex ionization pattern that can be seen as a "fingerprint" which is characteristic of the matrix being analyzed. These ionisation patterns or signatures are further processed by pattern recognition engines to perform classification, recognition and, to a limited extent, quantification tasks such as in the case of the gas sensor based *"electronic nose"*. To date, direct coupling of a mass spectrometer to techniques such as static headspace (SH) or dynamic headspace (DH) extractions, or solid phase microextraction (SPME), has led to the development of fast and reliable methods of headspace characterization. This new system for rapid analysis is currently subject to a growing interest and is starting to find a commercial outlet in the gas-sensors market. At the same time, the MS-Sensor is a challenging approach for food quality purposes since it allows simultaneous determination of compounds in food matrices and complex mixtures with a high sample throughput.

This thesis is devoted to study the possibilities and capabilities of MS-Sensor devices in several food quality related questions such as the determination of rancidity levels in crisps, the detection of fungal spoilage in bakery products, the monitoring of sardine freshness under cold storage, the classification of virgin olive oils according their organoleptic properties or their origins and the discrimination of two Iberian ham qualities. It has been widely demonstrated that the MS-Sensor profile can be considered as a useful fingerprint signature for the characterization of the targeted quality problem and, as in certain cases, even for quantification of several parameters correlated with this problem.

Despite the demonstrated viability of MS-Sensor devices in such applications, this approach stills suffer from some weakness that may influence their performance. The main drawbacks are the inherent high dimensionality of the data response matrices and the low selectivity of m/z ratios "pseudosensors" used as variables in these matrices. Besides the development of the applications already mentioned, this thesis is also aimed to figure out different strategies for overcoming these two issues. New algorithms for variable selection

or variable reduction have been developed in order to approach high dimensionality problems. Low selectivity of m/z fragments has been overcome by using new pattern recognition methodologies based on the use of multi-way algorithms.

MS-Sensor devices produce the so-called second order data where a series of values such as mass spectra are measured over time (each sample is measured as a function of two sources of variability: time and the mass to charge ratio of its main ions or m/z). The classical MS-Sensor approach uses the averaged mass spectra over time as a matrix response for further pattern recognition. At this way the second order data response is reducet to a first order data or data vector (i.e., a single mass spectrum per sample). Thus, models subsequently used to extract useful information of the data are bilinear. Nevertheless, averaging mass spectra may lead to a loss of potential information in further processing of the data. This thesis is aimed to study whether the pattern recognition models could benefit from preserving the second order data instead of averaging it. Any set of second order data can be arranged in a three-way array. Thus, a set of MS-Sensor measurements (m/z vs. time) obtained by measuring different samples can be arranged in a three-way array. Multi-way based algorithms are specially suited for handling second order data. Besides the applications mentioned above, this thesis attempts to study whether arranging MS-Sensor data in a three-way array layout and processing it using three-way algorithms allows for an improvement from the classical bilinear approach.

The use of three-way algorithms has been conduced in two of the above mentioned applications. It has been demonstrated that, although the use of temporal information through multi-way algorithms is not always translated into an improvement of classifier models when compared with the averaged two-way modelling, at least they lead to simpler and more parsimonious models.

In addition, several issues related to the use of MS-Sensor in food analysis have been studied: the use of different headspace sampling techniques, the comparison of MS-Sensor systems performance against classical metal oxide seminconductors gas sensors (MOX) based *"electronic nose"*; the application of new algorithms for pre-processing MS-

Sensor signals; the correlation of MS-Sensor response and the well-established methods to assess the quality property under study, etc.

MS-Sensor devices are a powerful set-up, capable of producing large amounts of highly selective information. Optimal use of this device implies both, a correct use of analytical techniques (sample handling and instrumental) and a rational use of subsequent data analysis approaches. That can be only attained if the analytical people in charge of the experimental set-up work side by side with data analysis and software developers. This thesis tries to make this close collaboration a little easier with the reports generated on different applications.

ISBN:978-84-691-9752-3/DL:T-126-2009

RESUM

La indústria alimentària necessita desenvolupar mètodes d'anàlisi per tal d'analitzar i caracteritzar la fracció volàtil dels seus productes i els seus aromes. Aquests mètodes han de complir cert prerequisits per tal de poder-se implementar. Han de ser suficientment precisos i reproduïbles a la vegada que han de poder donar resposta a gran quantitat de mostres en un temps limitat. Habitualment, els mètodes basats en tècniques instrumentals com la cromatografia de gasos acoblada a l'espectrometria de masses (GC-MS) o els mètodes basats en l'anàlisi sensorial són els més emprats per a aquests tipus d'anàlisi. Malgrat que aquestes tècniques estan força establertes encara presenten certes desavantatges. Molt sovint el desenvolupament de mètodes d'anàlisi basat en les tècniques instrumentals implica un considerable esforç i es requereix personal format per a tal fi. La resposta que s'obté d'aquests tipus d'anàlisi sol ser un llistat de pics corresponents a totes les substàncies detectades en l'espai de cap de la matriu que es vol analitzar i la seva corresponent concentració. Malgrat que aquesta informació pot ésser molt valuosa, el que molts cops es busca és una resposta més global, més intuïtiva; una resposta que ens doni informació del tipus "aquests producte és diferent al de l'anterior lot o s'assembla més a aquests altre". Per altra banda, l'anàlisi sensorial dut a terme per un panell de cata és un mètode molt costós des del punt de vista industrial ja que requereix disposar d'una sèrie d'experts que només poden treballar en períodes de temps relativament curts. També hi ha d'altres inconvenients com la subjectivitat de la resposta i la variabilitat entre els individus que conformen els panells. Totes aquestes raons fan que hi hagi certa demanda de noves metodologies d'anàlisi que puguin donar resposta a les necessitats de la indústria alimentària i que a la vegada no presentin els inconvenients de les tècniques instrumentals clàssiques i dels panell de cata. En aquests context, a començaments dels anys vuitanta es desenvolupà un concepte nou anomenat *"nas electrònic"*. El terme *"nas electrònic"* fa referència a uns sistema que integra una matriu de sensors de gasos amb selectivitats parcialment solapades i un sistema informàtic que conté algoritmes de reconeixement de patrons i que permeten classificar mostres en funció del seu perfil de volàtils. Des de la seva aparició, s'han publicat centenars d'articles relacionats amb aquest dispositiu en camps tant diversos com l'anàlisi d'aliments, l'anàlisi mediambiental, i l'anàlisi mèdica.

Encara que d'antuvi es preveieren uns auguris molt espectaculars per a aquests tipus de dispositius, amb el temps han anat apareixent una sèrie de problemes i inconvenients en la tècnica (pèrdua de sensibilitat dels sensors en presència de compostos polars, derives, baixa selectivitat, etc.) que han provocat que aquesta tecnologia no s'hagi implantat en el mercat tal i com s'esperava. Més recentment, a finals dels noranta, es desenvolupà el que es coneix com a *"segona generació de nassos electrònics"* o també amb el nom de MS-Sensor. Els productors d'instrumentació analítica van focalitzar la seva atenció en aquest nou concepte basat en la injecció directa dels compostos volàtils d'una matriu en la font d'ionització d'un espectròmetre de masses sense prèvia separació cromatogràfica. Aquests dispositiu sembla poder aportar una solució fàcil i ràpida als inconvenients del *"nas electrònic"* basat en sensors de gasos. L'espectre de masses que resulta de la ionització i fragmentació de tots els compostos volàtils presents en l'extracte es pot considerar com una empremta digital característica de la matriu que s'està analitzant. Aquests espectres, es processen a posteriori mitjançant mètodes d'anàlisi multivariant amb el fi de realitzar tasques associades a aquests tipus de sistemes com son la classificació i el reconeixement de noves mostres i dins de certs límits la quantificació d'aquestes. Avui en dia l'acoblament directe entre un espectròmetre de masses i un sistema automatitzat d'espai de cap estàtic (SH) o d'espai de cap dinàmic (DH) o fins l'acoblament a les tècniques de preconcentració com la microextracció en fase sòlida (SPME) ha portat al desenvolupament de mètodes ràpids i robustos per a la caracterització de l'espai de cap d'una matriu.

Aquesta tesi està dedicada a l'estudi de l'aplicació d'un sistema MS-Sensor en diferents tipus de problemes relacionats amb l'anàlisi de la qualitat de matrius alimentàries. Aquestes aplicacions son la determinació del grau d'enranciment de patates fregides, la detecció del creixement fúngic en productes de brioxería industrial, la monitorització del grau de frescor de sardines guardades en fred, la classificació de diversos olis d'oliva en base a les seves propietats organolèptiques i la classificació del pernil ibèric d'acord amb l'alimentació del porc del qual prové. En aquesta tesi es demostra que l'espectre de masses que s'obté amb un sistema MS-Sensor pot considerar-se com empremta digital molt útil i valuosa per a obtenir informació sobre la qualitat de les matrius analitzades per cada una de les diferents aplicacions. Fins i tot en alguns casos, es demostra que aquesta mateixa

empremta pot ésser correlacionada amb els paràmetres d'anàlisi clàssics que s'empren per resoldre aquestes qüestions o que fins i tot aquests mateix sistema es pot fer servir per predir aquests paràmetres.

Encara que la viabilitat del MS-Sensor per a les aplicacions plantejades ha estat demostrada àmpliament en el decurs d'aquesta tesis, aquests sistema encara té petits inconvenients o punts febles que resten per resoldre. Aquests inconvenients poden afectar de manera directa els resultats que se'n deriven. Els principals punts febles del MS-Sensor son d'una banda l'elevada dimensionalitat de les matrius de resposta que se n'obtenen i que és inherent al propi sistema i d'altra banda la baixa selectivitat del diferents fragments m/z considerats com a "pseudosensors" i que s'usen com a variables en aquestes matrius. A més a més de les aplicacions desenvolupades, aquesta tesi té com a objectiu aportar diferents estratègies que serveixin per millorar aquests punts febles. Una d'aquestes estratègies ha consistit en el desenvolupament i l'ús de nous algoritmes per a selecció i la reducció de variables. La baixa selectivitat dels fragments m/z usats com a variables en l'anàlisi multivariant s'ha abordat fent ús de tècniques *multi-way*.

El sistema MS-Sensor produeix dades que es coneixen com dades de segon ordre on la resposta enlloc de ser un escalar és una matriu (cada mostra es mesura en funció del temps i les intensitats de cada un dels fragments de diferent relació m/z escanejats). Malgrat això, la manera tradicional de treballar amb el sistema MS-Sensor és reduir aquestes dades de segon ordre dades de primer ordre, usant l'espectre de masses mig al llarg de la dimensió temporal. Normalment això es fa per tal d'adaptar la resposta als models bilinears que s'empren pel reconeixement de patrons. El fet de fer una mitja de l'espectre al llarg de la dimensió temporal pot fer obviar informació potencialment útil per a l'anàlisi de dades. Aquesta tesi pretén estudiar si els models que s'empren poden beneficiar-se del fet de preservar l'estructura tridimensional de les dades enlloc de convertir aquests a dades bidimensionals fent una mitja de la dimensió temporal. Si considerem un conjunt de dades d'ordre dos acabem obtenint el que s'anomena unes dades tridimensionals o *three-way array*. Un conjunt de mostres mesurades amb un sistema MS-Sensor es poden ordenar com a un *three-way array.* Els algoritmes *multi-way* estan expressament concebuts per tractar

dades de segon ordre. Aquesta tesi també té com a objectiu estudiar si el fet d'usar la resposta d'un sistema MS-Sensor en forma de *three-way array* (és a dir incorporant la informació temporal) conjuntament amb algoritmes de processat *multi-way* permet millorar la capacitat de predicció dels algoritmes bilinears. S'ha trobat que el fet d'utilitzar la resposta *three-way* no sempre comporta una millora directa en la capacitat de predicció del model en comparació amb el model bilinear que utilitza la resposta mitja al llarg del temps. Malgrat això, s'ha demostrat que l'ús d'aquests nous algoritmes permet obtenir models més simples i robustos i per tant podem aconseguir un millor funcionament del sistema i resultats més reproduïbles.

A més a més s'han estudiat diferents aspectes relacionats amb la utilització d'un sistema MS-Sensor per a l'anàlisi d'aliments com son: l'ús de diferents tècniques de mostreig d'espai de cap, la comparació del sistema MS-Sensor amb sistemes clàssics d'olfacte electrònic, l'aplicació i desenvolupament d'algoritmes de preprocessament dels espectres generats, la correlació de les respostes obtingudes amb un sistema MS-Sensor amb mètodes d'anàlisi d'aliments tradicionals, etc.

El MS-Sensor es un dispositiu que produeix una enorme quantitat de dades. L'ús òptim d'aquests sistema requereix, d'una banda, d'un ús correcte dels aspectes instrumentals com son el propi sistema en si i les tècniques de mostreig i per altra banda d'un ús racional de les tècniques d'anàlisi de dades. Això només s'aconsegueix si els analítics que treballen amb el desenvolupament de l'experiment treballen colze a colze i en estreta col·laboració amb la gent encarregada de fer l'anàlisi de dades. Aquesta tesi pretén fer més estret l'espai entre aquestes dues disciplines i dóna les eines per ajuntar i promoure aquesta col·laboració.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

LIST OF PAPERS

The following papers and contributions are discussed in this thesis. They are referred in the text by their Roman numerals I-VIII.

I. **Vinaixa, M**.; Marin, S.; Brezmes, J.; Llobet, E.; Vilanova, X.; Correig, X.; Ramos, A.; Sanchis, V., Early detection of fungal growth in bakery products by use of an "electronic nose" based on mass spectrometry. *Journal of Agricultural and Food Chemistry* **2004,** 52(20), 6068-6074.

II. **Vinaixa, M.**; Vergara, A.; Duran, C.; Llobet, E.; Badia, C.; Brezmes, J.; Vilanova, X.; Correig, X., Fast detection of rancidity in potato crisps using e-noses based on mass spectrometry or gas sensors. *Sensors and Actuators B,* **2005,** 106(1), 67-75.

III. **Vinaixa, M**.; Llobet, E.; Brezmes, J.; Vilanova, X.; Correig, X., A fuzzy ARTMAP- and PLS-based MS e-nose for the qualitative and quantitative assessment of rancidity in crisps. *Sensors and Actuators B,* **2005,** 106(2), 677-686.

IV. Llobet, E.; Gualdrón, O.; **Vinaixa, M**.; El-Barbri, N.; Brezmes, J.; Vilanova, X.; Bouchikhi, B.; Gómez, R.; Carrasco, J. A.; Correig, X., Efficient feature selection for mass spectrometry based "electronic nose" applications. *Chemometrics and Intelligent Laboratory Systems* **2007,** 85, 253-261.

V. Marin, S.; **Vinaixa, M**.; Brezmes, J.; Llobet, E.; Vilanova, X.; Correig, X.; Ramos, A. J.; Sanchis, V., Use of a MS-"electronic nose" for prediction of early fungal spoilage of bakery products. *International Journal of Food Microbiology* **2007,** 114 (1), 10-16.

VI. **Vinaixa, M**.; Brezmes, J.; Llobet, E.; Vilanova, X.; C.Burian; Correig, X., Saint Petersburg-Russia, 1-3 May **2007**; *Proceedings of the 12th International Symposim on Olfaction and "electronic nose".*

VII. El-Barbri, N.; Amari, A.; **Vinaixa, M**.; Llobet, E.; Bouchikhi, B.; Correig, X., Building of a metal oxide gas sensor-based "electronic nose" to assess the freshness of sardines under cold storage, *Sensors and Actuators B,* **2007**, 128 235–244.

VIII. **Vinaixa, M**.; Llobet, E.; Brezmes, J.; Vilanova, X.; Correig, X. Taking advantage of the time dimension on MS-Sensor approaches: Evaluation of Iberian Ham quality using SH/MS and multi-way data analysis, *In preparation.*

LIST OF ABREVIATURES

ANN	Artificial Neural Networks
ADV	Acid Degree Value
amu	atomic mass units
ART	Adaptive Resonance Theory
CA	Canonical Analysis
CI	Chemical Ionization
CP	Conducting Polymer
DH	Dynamic Headspace
DVB/CAR/PDMS	Divinylbenzene /Carboxen /Polydimethylsiloxane
EI	Electron Impact ionization
ev	electron volts
Fast-GC	Fast Gas Chromatography
GA	Genetic Algorithms
GC-MS	Gas Chromatography-Mass spcectrometry
HCA	Hierarchical Cluster Analysis
INDEX	Inside-needle dynamic extraction sampling
KNN	k-Nearest Neighbors
LDA	Linear Discriminant Analysis
LLE	Liquid Liquid Extraction
LLE	Liquid-Liquid Extraction
LVQ	Learning Vector Quantization
MLR	Multiple Linear Regression
MOSFET	Metal Oxide Semiconducting Field Effect Transistors
MOX	Metal Oxide Seminconductors Gas Sensors
NCI	Negative chemical ionization
N-PLS	Multilinear Partial Least Squares
N-PLS-DA	Multilinear Partial Least Squares Discriminant Analysis
P&T	Purge-and-Trap
PARAFAC	Parallel Factor Analysis
PCA	Principal Component Analysis
PCI	Positive chemical ionization
PCR	Principal Component Regression
PLS	Partial Least Squares

PLS-DA	Partial Least Squares Discriminant Analysis
PTV	Programmed temperature vaporization
Py-MS	Pyrolisis-Mass Spectrometry
QA/QC	Quality Assurance/Quality Control
QC	Quality Control
RAFFT	Recursive Algorithm Fast Fourier Transform
RBF	Radial basis function neural network
RMSECV	Root mean square error of cross-validation
RMSEP	Root mean square error of prediction
SA	Simulated Annealing
SAW	Surface Acoustic Wave
SAW	Surface Acoustic Wave
SBSE	Stirr Bar Sorptive Extraction
SGA	Standard Gas Adition
SH	Static Headspace
SHSE	Headspace Sorptive Extraction
SHSE	Static Headspace Sorptive Extraction
SIM	Single Ion Monitoring
SIMCA	Soft Independent Modeling of Class Analogy
SPDE	Solid-Phase Dynamic Extraction
SPE	Solid-Phase Extraction
SPE	Solid Phase Extraction
SPME	Solid Phase Microextraction
SVD	Singular Value Decomposition
SVM	Support Vector Machines
TCA	2,4,6-tricholoranisole
TOF	Time of flight
unf-PCA	unfolded PCA

CONVENTIONS

In this section the mathematical notation used in this thesis is summarized. It is the one commonly accepted by the scientific community.

Italic lowercase letters (e.g., x) indicate scalars (i.e., zero-order data), boldlowercase letters (e.g., $\mathbf{x}$) indicate vectors (i.e., first-order data), bold uppercase letters (e.g., $\mathbf{X}$) indicate matrices (i.e., second- order data), and underlined bolduppercase letters (e.g., $\underline{\mathbf{X}}$) indicate three-way arrays. Transposition of a vector or matrix is symbolized by a superscripted "T" (e.g., $\mathbf{X}^{\mathbf{T}}$). A superscripted "-1" (e.g., $\mathbf{X}^{-1}$) indicates the inverse matrix.

The characters I, J, and$K$ are reserved for indicating the dimension of an array. Mostly, a two-way array or a matrix will be assumed to be of size $I \times J$, while a three-way array will be assumed to be of size $I \times J \times K$. The lowercase letters corresponding to the dimension will be used to designate specific elementsof an array. For example, x_{ij} is the element in the i^{th} row and j^{th} column oft he matrix $\mathbf{X}$. That $\mathbf{X}$ is a matrix follows explicitly from x_{ij} having two indices.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 1

Introduction

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

1. INTRODUCTION

1.1 MOTIVATION

Monitoring food quality from raw materials to finished products is of great importance to the food industry. Analysis of odour and flavour in food has traditionally been addressed either by a trained sensory panel or using GC-MS. Unfortunately; these techniques are generally time-consuming, costly and labour-intensive. None of them allows for on-line or at-line quality monitoring and both require trained personnel to be used.

The food industry is demanding new high throughput methodologies for objective, automated and non-destructive measurements that can characterize odour and flavour in food. New instruments should allow the analysis of a high number of samples within a short period of time with sufficient accuracy and reproducibility.

During the last 20 years there has been a rapid development of a concept instrument called *"electronic nose"* (artificial nose) based on chemical gas-sensor array technology which seems to fulfil all these requirements.

The term *"electronic nose"* is related to an array of chemical gas sensors with a broad and partly overlapping selectivity for measurements of volatile compounds within the

headspace over a sample combined with computerized multivariate statistical data processing tools. Actually, the most widely accepted definition of an *"electronic nose"* was given by Gardner and Bartelet in 1994[1]: "*An instrument which comprises an array of electrochemical sensors with partial specificity and an appropriate pattern recognition system, capable of recognizing simple or complex odours"*.

The *"electronic nose"* is most often used to provide comparative rather than qualitative information. Because data interpretation is automated, *"electronic noses"* are ideally suited for quality control and quality analysis (QC/QA). These instruments are not useful for the identification and quantification of the individual components of a sample. Instead, they do a more comprehensive and holistic comparison of volatile patterns as the human nose does. In this sense, this technique is quite complementary to the more traditional chromatographic techniques that give rise to a response which contains the composition and the concentration of the volatiles present on the headspace of the sample under analysis.

Despite some initial promising achievements, solid state sensor-based systems have not generally lived up to the expectations. Although these systems offer a significant improvement over existing methods and they have been commercially available for several years, their low selectivity and sensitivity, their drift and the short life-time of the sensors have proven serious limitations. Often, the fluctuation in the performance may lead to difficulties when reproducing data results.

Recently, interest in using a mass spectrometry detector as an *"electronic nose"* has grown and now new generations of artificial noses, based on mass spectrometry, are opening new perspectives in the analysis of odours and volatile compounds. These types of *"electronic noses"* allow, by mass spectrometry, the direct characterization of volatile organic components from liquid and solid samples without separation of the headspace constituents. The use of a state-of-the-art analytical technique such as mass spectrometry yields a very reproducible and precise fingerprint of each sample. Thus, the identification of a headspace sample is straightforward, and the comparison of large data sets becomes easy.

The working principle of a MS-Sensor system is based on the introduction of volatile components extracted from the headspace of a sample into the ionization chamber of a mass spectrometer. The mass spectra resulting from the simultaneous ionization and fragmentation of these volatiles constitute a very complex ionization pattern that can be seen as a fingerprint that is characteristic of the headspace matrix being analyzed. These ionisation patterns are then processed by pattern recognition engines to perform tasks associated to *"electronic nose"* systems such as classification, recognition and, to a limited extent, quantification.

In general terms, this thesis is aimed to the study and evaluation of the MS-Sensor approach as a new tool in food quality control. In this context, several applications related to the evaluation of food quality assessment through use of MS-Sensor are attempted. All these studies were conduced in partnership with different food companies and they respond to a real demand from these industries.

In many food applications the goal is to determine "no directly measurable" parameters which are derived from "direct" measurements. Examples of this approach are the determination of rancidity in crisps, the type of pig feeding in Iberian hams, the origin of olive oils or the fungal spoilage of bakery products. All these attributes cannot be measured directly and they are often determined indirectly by the analysis of other parameters with well-known and established analytical protocols which are somehow correlated with the quality parameter that is sought. For example, the measure of a chemical marker such as ergosterol has commonly been used as a method for the quantification of fungal biomass in food. This thesis is a study of the performance of a MS-Sensor approach as a valuable tool for the determination of such "no directly measurable" parameters in five different food quality applications oriented to overcome real problems currently present in food quality industries.

Although the first application of a MS-Sensor device in quality control emerges in the late nineties (the first paper appeared in 1998) there is still a lot of room for improvement on the use of MS-Sensor and related issues. Almost all applications attempted

through this thesis were new to the date of their publication. Some of the subjects were already covered by the use of classical *"electronic noses"* but still remained uncovered from the MS-Sensor point of view. Below we summarize the main applications developed.

The first quality problem attempted is related to the assessment of freshness and spoilage of food using a MS-Sensor instrument. Spoilage of raw material caused by microbiological or fungal activity represents a great problem in the food industry and leads to a considerable waste in storage and handling of food. Two different applications were sought within the framework of food spoilage. The first study was related to the detection of early fungal spoilage in bakery products through the implementation of a MS-Sensor system. Fungal spoilage is an important issue in bakery products. Although its impact from a safety point of view is not significant, the presence of visible colonies on the products degrades the company image in front of the consumer, resulting in economic losses in the medium–long term. A second study dealt with sardine freshness assessment under cold storage. The composition of the fish headspace is a source of information about the freshness degree of a sample. Microbial spoilage produces a wide variety of volatile compounds. Some of them are often used as indicators of spoilage. Nevertheless, due to the high number of volatile compounds involved in the process and to the fact that they change dynamically, the monitoring of fish freshness spoilage can be achieved with a multicomponent approach. Therefore this is a typical MS-Sensor application where mass spectra fingerprinting can give rise to a qualitative analysis of the samples. By measuring the release of volatiles during storage, a time dependent model can be made for the prediction of the shelf life or storage time.

Another food quality aspect dealt within this thesis is related to the oxidation state of lipids. Lipid oxidation plays an important role in food quality since lipids are major ingredients in many food products and are vulnerable to oxidation during storage. Oxidation gives rise to a high number of volatile molecules such as short chain carbonyl compounds as saturated and unsaturated aldehydes and ketones which generally have unpleasant odour and may lead to some organoleptic defects as rancidity. A third

application developed within the framework of lipid oxidation has been the evaluation of a MS-Sensor as a tool for qualitative and semiquantitative measurement of rancidity in crisps.

Finally, a more general issue often sought by food companies is to measure and predict the main odour properties of food products. This is much more complicated since it should be emphasized that any *"electronic nose"* configuration (neither gas sensor based nor MS-Sensor) does not focus on sensing specifically odour compounds. None of these technologies is devoted to selectively respond to those compounds on the headspace that present aromatic activity; instead, both technologies respond to all the volatiles compounds in the headspace as a whole, the odorants and the non odorants. This is why all experiments oriented to assess odour properties from food matrices should ideally be accomplished and correlated with the help of an organoleptic panel. Only in the case where MS-Sensor data is supported by an organoleptic assessment panel, a MS-Sensor instrument could be used for an odour assessment task. Two different applications have been conduced in which odour compounds play a key role. The first one is dedicated to the classification of olive oils according to their different origin and organoleptic properties assessed by a panel. The ultimate goal of second study is the classification of different iberian hams according to their quality due to the different pig feeding routine. The feeding that the pigs have received during their fattening period contributes in a remarkable way to the organoleptic properties of the hams and thus, a MS-Sensor approach can be a useful tool for assessing ham quality.

Despite the great potential of the MS-Sensor approach in the field of food analysis some aspects of the technique need further improvement. The main weaknesses associated to the MS-Sensor approach are the high dimensionality of the measurements and the low selectivity of the m/z fragments used as variables on the matrices for further pattern recognition. These two drawbacks could be responsible for the lack of reproducibility showed by MS-Sensor systems in certain applications. In order to advance in the development of such applications it was necessary to figure out different strategies for overcoming this high dimensionality and the low selectivity. Furthermore, the development of the applications sought also demanded for further research on other MS-Sensor issues such as handling of volatiles, data pre-processing algorithms, pattern recognition

methodologies, etc. Since MS-sensor performance depends, to a very large extend, on the performance of the pattern recognition methods employed, a lot of effort in this thesis has been conduced in the improvement of classical data analysis techniques used in commercial software packages.

When the number of features exceeds the number of measurements available to train the pattern recognition methods, the models used often suffer from the so-called "course of dimensionality" problem, leading to a dangerous situation due to the severe risk of overfitting the model. The way to overcome this problem is to reduce the dimensionality of the data matrix in an intelligent way. Different strategies have been reported for the reduction of dimensionality. These basically consist of either choosing directly among the variables available (e.g., m/z ratios) or to compute new variables called factors (e.g., by performing some transformation of the original variables such as in PCA and selecting among the factors). Methods based on the selection of m/z ratios are interesting because the variables chosen carry straightforward relevant chemical information. In this thesis, new methods for feature selection have been developed and proved in several datasets corresponding to some of the real applications envisaged. It has to be mentioned that this subject remains unexplored up to now and its study represents a completely new development in the MS-Sensor field.

The selectivity of the m/z variables is as important as the variable selection procedure. MS-Sensor m/z variables intrinsically present a low specificity. This fact arises from the high degree of fragmentation obtained with electron beam ionization. One way to tackle this is to change the ionization mode to soft ionization techniques. That could result in a very expensive and not very practical approach from the MS-Sensor point of view since it increments the price and the miniaturization of the device will be difficult. In this thesis an elegant way for overcoming this low specificity is presented. It is based on considering information from the time dimension, i.e., the variation through time of each one of the m/z fragments scanned. Thus, the m/z sensors considered become much more specific since although the same fragment may be the result of fragmentations from different molecules, their temporal variation should be different if these fragment belongs

to different molecules. Multi-way methods are particularly useful for the analysis of batch process data and analytical datasets that intrinsically present more than two sources of variation (i. e., MS-Sensor data, where a response is being measured as a function of two parameters: m/z and retention time). In fact, data stemming from MS-Sensors should be arranged as a multi-way array where the first mode (dimension) represents samples ("rows"), the second corresponds to mass spectra ("columns") and the third to the elution profiles through time ("tubes"). It has to be emphasized that conceptually, the use of multi-way methods for treating data arising from MS-Sensor is a completely new idea that has never been applied before.

Common two-way pattern recognition algorithms applied on MS-Sensor data make use of data matrices in which columns represent each one of the mass to charge ratio (m/z) values scanned and each row describes the m/z intensities for different measurements. These matrices are obtained by averaging mass spectra signals for a given m/z along the retention time considered as valid. Time averaging of the mass spectrum may lead to a loss of potentially valuable temporal information in further pattern recognition. Actually, averaging mass spectra along the detected peak allows converting real three dimensional data (three-way data) to two-way matrices by eliminating the time dimension. Therefore, the real nature of data is not respected and is changed just to adapt the response of the instrument to current available pattern recognition two-way algorithms. Besides overcoming low specificity of m/z sensors, there is a more general reason for using three-way algorithms in pattern recognition analysis of MS-Sensor responses. This reason derives from a conceptually new idea based on the fact that a fingerprint retaining chromatographic-mass spectra data is more likely to contain additional information than the two-way matrix containing the averaged mass spectra. Then, the inclusion of this information in the modelling of this data should lead to an improvement of the pattern recognition since it is assumed that as much information is retained in the data; much better performance of pattern recognition could be attained. Basically there are two ways of considering the time dimension on the data matrix. The first one is based on matricization or unfolding of a three-way array, which means transforming a third order array as a concatenated two-way array or matrix. In this manner, classical two-way algorithms can be

employed to understanding the structure in data. Nevertheless, unfolding may result in information loss and its interpretation becomes impossible especially if the datasets are noisy. A second way is to treat data using multi-way algorithms. Both approaches are going to be elucidated and compared in the framework of this thesis.

The ultimate aim of including temporal information is to consider chromatographic-mass spectra data as the characteristic fingerprint instead of the averaged mass spectra and thus, to use GC-MS as a MS-Sensor approach. The technique which uses full chromatograms as fingerprints and its chemometric comparison is a well-known technique often called "chromatographic fingerprinting". This technique does not include spectral information on the pattern recognition step, the mass spectrum is just considered for identification purposes as classical GC-MS instruments do. Basically, there are two different approaches to performing chromatographic fingerprinting. Main differences among these two approaches rely on the type of information contained on the response data matrix. First approach is to consider in the matrix response the area values of the peaks corresponding to the key compounds that play an important role on the problem envisaged. In this case, peak identification and peak matching is implicitly needed. Second approach consists in to considering just the chromatographic profiles shape on the response data matrix. Again, success of data modelling will be highly dependent of chromatographic resolution. Very often, when trying to resolve full chromatograms from the headspace extracts of complicated matrices such food products, the analysis time increases and can be at least around half an hour. Analysis time and the need for peak matching are main drawbacks associated to chromatographic fingerprint technique.

The approach attempted throughout this thesis is an intermediate concept in between the classical MS-Sensor approach (making use of mass spectra data) and chromatographic fingerprint (making use of chromatographic data). In MS-Sensor approach chromatographic resolution is avoided giving rise to an unresolved single chromatographic peak which integrates the signals of all the volatile molecules present in the headspace. Usually this unresolved chromatographic single peak takes a form of an asymmetrical gaussian peak which is used to determine the averaging time intervals. Even when

chromatographic resolution is avoided, some kind of diffusion can be observed on this asymmetrical peak which shows some retention potentially able to report additional information that may help for further modelling of data in classification or prediction tasks. Then, treating MS-Sensor response using three-way algorithms enables to consider this information in further chemometric analysis. What is really envisaged is to prove if the minimal diffusion observed in this asymmetric peak gives rise to an improvement of pattern recognition modelling; to explore until which extent is profitable to include chromatographic information on the data response array; and to determine which is the best way to be include this information, either in an unfolded way or considering three-way methods. The question to answer is, would it be possible to improve MS-Sensor pattern recognition performance making use of chromatographic information through the use of multi-way methodologies even when full chromatographic resolution is avoided such as in case of MS-Sensor.

Up to now, time information has been never included as an additional source of information for the MS-Sensor response. For this reason, multi-way data analysis has not yet been used on MS-Sensor pattern recognition. A significant part of this thesis has been dedicated to investigate whether the introduction of the retention time really benefits or not the pattern recognition performance of a MS-Sensor device. Conceptually, that is a completely new idea. The overall concept is to increase the resolution capability of the MS-Sensor system through the use of the temporal dimension. In few words, since MS-Sensor gives rise to high overlapping peaks, the response of low concentrated analytes can be masked by the presence of more concentrated analytes. Considering both, retention and spectral information time may lead to find the presence of these low concentrated compounds in which the properties of the sample under analysis lies in many occasions.

In summary, this thesis is basically dedicated to the evaluation of a MS-Sensor approach as a valuable tool for the five different food quality applications already presented. Finally, it is very important to emphasize that this thesis is not only devoted to solve the quality problems formulated in each one of these applications; it is also devoted to being a useful guide for clarifying the questions that should be addressed when a new

application using a MS-Sensor is sought. Thus, it is also oriented to give a response to other questions related to the use of a MS-Sensor system, such as the study of new pattern recognition techniques to improve MS-Sensor performance (i.e., variable selection or multi-way based methodologies); the comparison of MS-Sensor systems performance against classical MOX *"electronic noses"*; the application of new algorithms for pre-processing MS-Sensor signals; the correlation of MS-Sensor response and the well-established methods to assess the quality property under study, etc. Finally, another big goal of this thesis has been to study whether the inclusion of the retention time reports additionally valuable information for improving pattern recognition and consequently MS-Sensor performance.

1.2 OBJECTIVES OF THE THESIS

The goals of the thesis are going to be described according to their nature, since some are related to a quality application and some to other MS-Sensor related fields that must be developed in parallel to the quality applications envisaged.

On one hand, the main aim of this thesis has been to explore the possibilities and capabilities of the MS-Sensor approach as a valuable tool for the food quality assessment applications already mentioned. The overall concept has been to meet the quality requirements and criteria demanded by food industries. Therefore, the goals related to food quality applications are the following:

1) To assess whether the MS-Sensor approach is able to establish a semi-quantitative measurement of the rancidity level in crisp products without any previous oil extraction step.
2) To study the applicability of a MS-Sensor system for the early detection of fungal spoilage on bakery products.

3) To prove the performance of a MS-Sensor instrument in the discrimination of olive oils from different cultivars and to assess their quality according to their organoleptic characteristics.

4) To study of the feasibility of using a MS-Sensor system to characterize Iberian ham according to the feeding that pigs receive, since the pig"s diet contributes in a remarkable way to the sensorial characteristics and thus to the quality of hams.

5) To test the performance of a MS-Sensor system in the assessment of sardine freshness and as a tool for spoilage monitoring under cold storage conditions.

As stated previously, this thesis is not only aimed to developing food quality applications. An additional aim is to give a general overview on how a new application should be addressed. Therefore, some goals of this thesis are related to MS-Sensor methodologies rather than application specific issues. These goals have been developed in parallel to the applications envisaged since to achieve the performance levels required in the quality assessment studies demanded a careful refinement of the initial approaches. These goals can be summarized in the following points:

6) Comparison between MS-Sensor and MOX *"electronic nose"* performances in food quality assessment.

7) Study of different sampling techniques to be coupled to a MS-Sensor system. Determination of the parameters to be taken in to account when selecting among different headspace techniques.

8) Study and evaluation of the optimal data pre-processing algorithms: (normalization, alignment) and their influence in the overall pattern recognition performance of a MS-Sensor approach.

9) Evaluation of different pattern recognition algorithms in real applications: linear methods (LDA, PCA, PLS, PLS-DA), non-linear methods (Fuzzy ARTMAP) and the concatenation of these methods: PCA-Fuzzy ARTMAP, LDA-Fuzzy ARTMAP.

10) Evaluation and development of variable selection methods in a MS-Sensor approach. To prove whether the pattern recognition analysis may beneficiate from the application of the variable selection algorithms developed. Comparison of variable selection against variable reduction.

11) Evaluation of whether the introduction of the retention time as a second source of variation improves the performance of MS-Sensor devices. Study of the three-way data analysis in data steaming from MS-Sensor approach.

1.3 ORGANISATION OF THE DOCUMENT

This has been the first chapter of the document. The thesis is comprised by six additional chapters that are briefly commented below:

The State of the art in MS-Sensor technology is presented in **Chapter 2**. This chapter starts with an introductory part where some basic definitions are illustrated. An overview of the whole concept of *"electronic noses"* and MS-Sensor is documented and the strengths and weakness of both approaches are discussed. A comparison bettween both approaches is also given. The chapter continues outlining current research issues in MS-Sensor technology reviewing different aspects covered by other research groups working in this field. The chapter also includes a table summarized review. It is composed by several tables including the applications attempted to the date which have been used a MS-Sesnor approach. The year of publication, the type of sample, the type of study, the MS-Sensor configuration system and the pattern recognition algorithms used are detailed in these tables. Finally, the last part of this chapter tries to show a comprehensive overview on MS-Sensor data analysis. The current state of the art in the pre-processing and pattern

recognition processing techniques currently available or used on MS-Sensor devices is documented. Furthermore an overview of the models used in this thesis is also given.

Chapter 3 tries to give a generic overview of the experimental set-up and common protocols or methodologies used in all food quality application subsequently developed in this thesis. Furthermore, this chapter is aimed to show the experimental procedures carried on in order to decide among fundamental issues set-up when attempting a new MS-Sensor application (handling volatiles, data processing, data preprocessing, development of new pattern recognition algorithms, etc). A more detailed and concise explanation on the experimental set-up and protocols used for each one application can be found elsewhere on the papers.

Chapter 4 gives a summary of the experiments conducted throughout this thesis. This chapter is divided in five different parts that cover each one of the five different applications developed. Each part will contain a brief introduction of the problem to be solved, a brief description of the experimental work conduced and the most relevant results obtained.

Chapter 5 presents a general discussion based on the experience gathered on the five applications described in the previous chapter. The discussion can be used as a basic guide where the most important concepts that must be taken into account when new applications are tackled are fully documented. This chapter deals on some issues that are not discussed in depth in published papers. Even though they are omitted, they are considered highly relevant to reach the results presented in the previous chapter. Topics such as sampling, preprocessing, variable selection and multi-way data analysis are fully developed and discussed in this chapter.

The final conclusions arising from the studies performed in this doctoral thesis are presented in **Chapter 6**. Perspectives and future work are presented on **Chapter 7** where a discussion about MS-Sensor industrial applications is also included. Moreover, in this chapter some ideas about what would constitute and ideal MS-Sensor are presented.

Finally, bibliographic **References** are enclosed. The publications derived from the work done are grouped under the **Appendix.** They have been sorted by the year of publication.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 2

State of Art

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

2. STATE OF ART

2.1 FLAVOUR AND ITS ROLE IN FOOD QUALITY

An old saying says that "we are what we eat". Our nutritional status, health, physical and mental faculties depend on the food we eat and how we eat it. Access to high quality food has been main endeavour of humanity from the earliest days of human existence. From a very simplistic point of view the food quality concept can be considered as a complex characteristic of food that determines its value or acceptability to consumers.

Flavour perception is based on two components: taste and aroma. Taste arises from the presence of non volatile compounds, which interact with sensors of the mouth and on the tongue and give rise to the basic tastes of sweet, sour, salt and bitter. The flavour of food cannot be defined by taste alone. What is far more important are the many hundreds of volatile compounds that are responsible for the aroma of a given product. These are the compounds that define the nature of food and its product identity, and they contribute to consumer preferences between brands of products. Volatile compounds are also invariably responsible for the occurrence of off-flavours taints, which arise due to the chemical or biochemical changes of microbial actions or contamination.

Besides safety, quality attributes include: nutritional value, organoleptic properties such as appearance, colour, texture, odour, taste and functional properties. Industry must play its role in assuring food quality and safety through the application of quality assurance and risk-based food safety systems utilising current scientific knowledge. The implementation of such controls throughout production, handling, processing and marketing leads to improved food quality and safety, increased competitiveness and a reduction in the cost of production and wastage. As the commercial environment of the food industry becomes more and more competitive, there is an increasing need for the maintenance and improvement of food and drink quality. As stated before, when consuming foods, colour, texture and flavour are all important in the consumers appreciations and expectations of a product.

However, it is flavour which is accepted as the most important sensory characteristic associated to food quality. Monitoring the flavour quality from raw material to finished products is of prime importance to the food industry. For example, raw materials should be monitored to ensure they have typical flavour quality, are free from taints, and that no deterioration or contamination with off-flavour chemicals has occurred during transport. Also, packaging materials should be checked for residual solvents to ensure that they won"t generate off-flavours in the finished product. Occasionally, packaging materials are not adequately cured prior to their use, and small amounts of solvent associated with the manufacturing of the packaging might remain. Monitoring flavout/off-flavour development during various processing steps should be conduced to ensure processes are operating correctly. For example, high levels of heat treatment during processing can cause thermal degradation reactions to occur in a food matrix resulting in the development of off-flavours[2]. That is why the food industry needs from techniques for the assessment of the flavour quality from raw materials through to final product.

2.2 FLAVOUR MEASURING

Nowadays, volatile analyses in the food industry rely on two classical techniques, conventional GC-MS or quality sensory analysis. Unfortunately, both techniques are too

time consuming and labour intensive for routine quality control applications. When thinking in industrial quality controls applications, the need for speed and high troughput analysis significantly impact the type of analytical test and the instrumentation that can be used.

2.2.1 Sensory analysis

The human nose is still widely used as an analytical tool. For example, the nose is routinely used to assess the quality of food stuffs, drinks, perfumes and other household products. Flavour and odours released from food matrices are usually complex mixtures of many hundreds of chemical species. Just subtle changes in the relative amounts of these species can be detected as a change in odour. Faced with this complexity, a traditional instrument such as the human nose is perhaps the most appropriate, with panels of experts being frequently used to access odour quality. These panels are expensive to train and maintain, and give subjective assessments which can be adversely affected by external parameters such as illness, mood or fatigue. Moreover, sensory analysis by a panel of experts is a costly process for industries because it requires trained people that can work for only relatively short periods of time. The practical applications of panels are severely limited by the fact that the sense of smell is subjective, tires easily and is therefore both expensive and difficult to standardise. Moreover, subjectivity and variability among individuals often difficult the performances of the panel test.

2.2.2 Instrumental based methodologies

The development of capillary chromatographic columns has provided the flavour chemists with a technique capable of tacking complex mixtures of flavour compounds, quantifying the individual components and, when linked to a mass spectrometer, identifying them as well. Using this technique, the volatile composition of most foods and drinks have been identified and tabulated. However, this technique is still at a research and development

stage. Sample extraction and preparation time and GC analysis runs can be long. The equipment is expensive and requires a trained and skilled technician for both operation and interpretation of the results. Also it is relatively fragile and not ideally suited to the rigors of a food process hall or QC laboratory. Furthermore, additional efforts are necessary in order to relate the GC/MS to the human perception of the aroma.

2.3 THE *"ELECTRONIC NOSE"*

As stated before, the two traditional approaches are time consuming and costly, none of them allows for on-line or at-line quality monitoring and both require trained personnel. The food industry is demanding new methods with high throughput for objective, automated and non-destructive techniques that can characterize odour and flavour in food. These new methods should allow analyzing high number of samples within a short period of time with a sufficient accuracy and reproducibility. One analytical tool that has been proposed to address the need for routine quality testing in food industry and at the same time meets the quality control requirements is the "*electronic nose"*.

2.3.1 Definition

During the last 20 years there has been a rapid development of the "*electronic nose"* (artificial nose) concept. Persaud and Dodd[3] first reported the design of an "*electronic nose"* using chemical sensors and pattern recognition in 1982. With the term *"electronic nose"* is understood an array of chemical gas sensors with a broad and partly overlapping selectivity for measurements of volatile compounds over a sample combined with computerized multivariate statistical data processing tools. Unlike most existing chemical sensors which are targeted to a specific chemical compound, sensors in an *"electronic nose"* are not specific. The *"electronic nose"* is both a chemical sensing and a data analysis system that can, to some extent, discriminate between different simple or complex odours. Actually the most widely accepted definition of "*electronic nose"* was given by Gardner

and Bartelet in 1994[1]:*"An electronic nose is an instrument which comprises an array of electrochemical sensors with partial specificity and an appropriate pattern recognition system, capable of recognizing simple or complex odours"*".

The term *"electronic nose"* comes from a certain parallelism to the measurement concept of the instrument and that of the mammalian olfactory system. In the latter, upon being sniffed through the nose, or through the retro-nasal pathway when a product is tasted, volatile compounds reach the olfactory epithelium which is an area of approximately 5 cm^2 located in the upper nasal cavity. There, the interactions of odorants with the appropriate chemosensory receptors, olfactory neurons (~10^7 belonging to ~10^3 different classes) produce electrical stimuli which are transmitted to the brain. A pattern recognition process assisted by the memory takes place using all the data in order to identify, classify, or perform an hedonic analysis Evidence exists showing that a single olfactory neuron responds to several odorants and that each odorant is sensed by multiple olfactory neurons. In the same way, *"electronic noses"* base the analysis on the cross-reactivity of an array of semi-selective sensors. Hence, products with similar aroma generally result in similar sensor response patterns (similar "fingerprints") whereas products with different aroma show differences in their patterns (different "fingerprints"). Just like the human olfactory system, *"electronic noses"* do not need to be specially designed to detect a particular volatile. In fact, they can learn new patterns and associate them with new odours through training and data storage functions as humans do. However, training of *"electronic noses"* based on sensory panel classifications is required in order to obtain odour-meaningful classifications. The sensitivity of *"electronic noses"* is often similar to that of human noses but humans are specially gifted in sensing specific compounds (e.g., thiols, biogenic compounds, pyrazines, thiazoles, some aldehydes). The biological sensitivity can go down to ppt levels with a response time in the order of milliseconds whereas instruments barely go under ppb levels with a response time in the order of seconds[4].

The most important feature of the sensors inside an *"electronic nose"* is their partial specificity; the response of a sensor is always the sum signal to all the compounds on the headspace it is exposed to at that moment. Thus, it is not possible to associate a part of the

response to a specific compound, especially because sometimes the response to one compound is influenced by the presence of another compound ("cross sensitivities", "matrix effects", and "interaction effects"). Then, the concept of the *"electronic nose"* is therefore to use a relatively limited number of unspecific sensors with overlapping, but nevertheless varying, sensitivities, thus together responding a large range of volatile compounds. This results in a unique response pattern or "fingerprints" when exposed to different odours. Basically that means that it is not a tool to identify the volatile components of the samples as it is done by classical analytical tools such us GC/MS, which allows the determination of the amount and the chemical structure of these volatile and semivolatile analytes. Instead of giving the concentration or identifying which compounds are involved in the headspace of samples, the *"electronic nose"* rather leads to a response such as "the unknown sample is more similar to sample 1 than to sample 2 and differs from the samples belonging to group A".

2.3.2 *"Electronic nose"* block system

Generally, the *"electronic nose"* is developed as a match model for the natural nose comprising the various stages between a bulk of volatile odorants and its recognition, namely interaction, processing and identifications. The parallelism between between biological and artificial nose is illustrated in Figure 2.1[5].

From this figure it can be depicted that, *"electronic nose"* devices are basically composed by three different blocks: the sampling system, the sensory system and the pattern recognition system.

The first block is devoted to sampling volatiles and its main purpose is basically to transfer the headspace of the samples to the sensory system. For the analysis of a given sample with an e-nose, the sample headspace, i.e., the gas volume above the substance to be analysed, or a representative part of it has to be brought into the sensor chamber. The sample headspace contains volatile compounds released by the sample. The measure of this

headspace using chemical sensors reveals information about the nature or the composition of the sample. There are several possible methods for this sampling process, which consists of taking up the sample, conditioning it, and transferring it to the analytical equipment. This should be done with maximum possible efficiency without altering the composition of the headspace. The headspace sampling technique determines to a certain extent the composition and the concentration of the headspace. The design of the sampling system will be strongly dependent on the application. Depending on this application, different sampling methods can be used, which mirror the difficulty of the separation problem to be solved. Traditionally headspace sampling techniques such as DH or SH are commonly used being SH the most frequently used technique where the gas phase of the food sample is in equilibrium with the food matrix before an aliquot of the headspace is transferred to the sensor chamber. Other considerations are sample headspace, equilibration time, sample quantitiy sample surface area, etc. More recently pre-concentration techniques such as SPME or Purge and Trap (P&T) have been attempted in the *"electronic nose"* field in order to improve sensitivity of the sensing process and to take advantage of large sample volumes. There is not room for going into this topic in this document and it has been widely discussed and detailed elsewhere[2, 5-7].

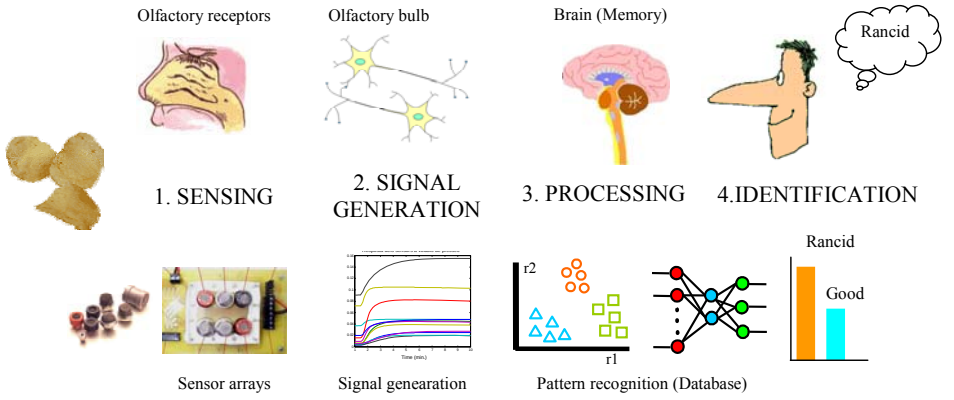

Figure 2.1: Basic diagram showing the analogy between biological and artificial MOX *"electronic noses"*

The second block or sensory system is the electrochemical sensor array. With the term of electrochemical sensor is understood any device able to change its electrical properties when it is exposed to a volatile molecule. There are a wide range of chemosensor

types basically depending on functional principles (chemoresistors such as MOX or CP (conducting polymers); mass sensitive chemosensors such as SAW (Surface Acoustic Wave); chemotransistor such as MOSFET (metal oxide semiconducting field effect transistors), etc.

Details about the types of chemosensors and its functional principle can be found elsewhere[5]. The most frequently used sensors in *"electronic nose"* applications are MOX which operate at elevated temperatures typically ranging between 250-400ºC. MOX consists on a metal-oxide (SnO_2, TiO_2, WO_3, etc) semiconducting film coated on to a ceramic (either alumina or silicon substrate) with an integrated heater. The gas sensing principle is based on the reaction between adsorbed ionized oxygen ant the oxide surface with incoming oxidizing or reducing molecules at elevated temperatures. The response is measured as the change in resistance between two electrodes as a result of the chemical reactions at the surface of the metal-oxide semiconductor. The detection mechanism is represented in Figure 2.2 and it can be described as follows. In an oxidising atmosphere, oxygen atoms (resulting from the decomposition of oxygen molecules in the ambient) or other electron acceptors adsorb in the surface region of the material and trap free electrons from the conduction band of the semiconductor; thus ,the resistance increases. On the other way around, in a reducing atmosphere, the adsorbed oxygen atoms react with the reducing ambient molecules, releasing the trapped electrons to the material and then the resistance decreases. The output signal is derived by a change in conductivity of the oxide caused by the reaction between the adsorbed oxygen and the incoming molecules.

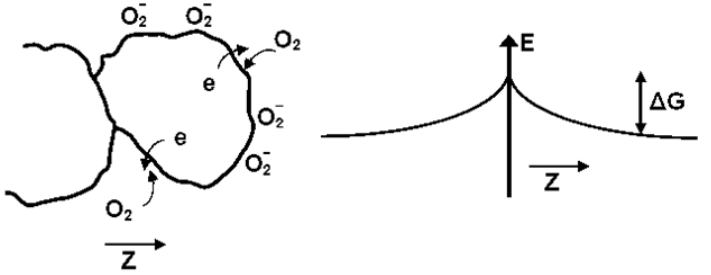

Figure 2.2: Charge Exchange associated with the chemisorption of oxygen at a semiconductor surface and the potencial distribution across a grain junction[8]

The selectivity of the devices can be modified by doping the metal oxide with noble catalytic metals, by changing the working temperature of the sensing element (250-400ºC), or by modifying the grain size. The advantages of MOX are their high sensitivity to most combustible gases including saturated hydrocarbons, NO and CO, but they show poor distinction between different polar compounds, their fast response time and reliability (simple set-up), the good resistance to corrosive gases and humidity, the mechanical strength and the low production cost. The main drawbacks are the relatively poor selectivity, which can be too some extent improved by dopands and temperature adjustment during the measurement, and finally the high power consuming operation temperature[9]. Readers are referred to more specialised literature to a wider review on sensor technologies for *"electronic noses"* [5, 10-13].

Finally, the third block concerns pattern recognition engineering. As stated before, the interaction of volatiles with the array of sensors provokes a series of signals which are then processed by the computer via a pattern recognition program. Pattern recognition is defined as the process of identifying structure in data by comparison to known structure. Patterns are typically described in terms of multidimensional data vectors, where each component is called a feature. The aim of a pattern recognition system is to associate each pattern with one of the possible pattern classes (or simply classes). Obviously, different patterns should be associated with the same class or with different classes depending on whether they are characterised by similar or dissimilar features, respectively. In the case of the *"electronic nose"*, the patterns and the classes are, respectively, the responses of the sensor array to molecules on headspace that give rise to signal in the sensor array.

Basically, there are two main types of pattern recognition techniques: supervised and unsupervised. Both are going to be described in detail later in this chapter. The supervised approach is more common with a knowledge base employed that contains a set of known vector patterns for different classes that have come from an earlier calibration or training procedure. The output of the *"electronic nose"* is then the predicted class to which the unknown input belongs. The different processing stages[14], models and algorithms currently used in *"electronic nose"* instrumentation are excellently reviewed in literature[10, 15-19].

2.3.3 Some considerations about the term *"electronic nose"*

As it has already mentioned, traditionally, a parallelism with the mammal human olfaction has been established in order to explain the basic principle of an *"electronic nose"*. The main similarities rely on the fact that *"electronic nose"* devices incorporate both, an array of chemical sensors as human olfactory receptors in the olfactory region and a pattern recognition engine inspired on brain operation. Humans recognize patterns instead of identifying separately each one of the components conforming an aroma. In other words, the stimuli or signal derived from the interaction of odorants with receptors cells located in bulb an odour causes at the diverse receptors in the human nose to form a kind of a pattern in the brain that can be memorized, compared with others, and eventually recognized.

Nevertheless, the "*electronic nose"* does not exactly mimic the human nose. Not everything that smells intensely must be a good application for an *"electronic nose"* and in contrast, some investigations that have been carried out successfully have no odour at all. This is mainly caused by the different "instrumentation": Human olfactory receptors have developed over thousands of years of evolution and they are optimised to ensure their survival. That is why they are very good in indicating rotten food for example, since they are very sensitive to the gaseous products of decomposition. Other dangers, i.e., carbon monoxide are too new in evolutionary periods of time for human olfactory receptors to adapt to; we do not smell this at all. For an *"electronic nose"*, it is no problem to perceive carbon monoxide, while it can be a problem to monitor the degradation of food in certain cases. Thus, it is mandatory to keep in mind that an absolute odour description is not possible from an *"electronic nose"* point of view. In general, an *"electronic nose"* bases its evaluation upon the analysis of the whole volatile species including odorants and "odourless" molecules. In contrast, the human olfactory sense is based only in the detection of odorant volatile molecules. Furthemore, the molecules producing chemical sensor signals are in general not the same molecules being perceived by the human nose. It even happens that two different samples being well distinguished by the human nose show exactly the same signal pattern on an *"electronic nose"*. Classical instrumental techniques such as GC or GC-MS also encounter this problem. Moreover, some molecules defining the

aroma of a product are present in extremely low concentrations (ppm or ppb) and can even not be detected by GC-MS. Of course, the reverse case also exists. It is possible that strong chemical differences between two products can easily be detected by aid of an *"electronic nose"* whereas human experts hardly recognise any aroma difference. From this point of view the term *"electronic nose"* should not be stressed too much because it may hint at a high comparability concerning possibilities and limitations. Some authors claim for an overstressed use of *"electronic nose"* word[20]. They argue that it must be stressed that the systems that are currently available do not work in the same way, nor detect the same constituents, as the human nose. They suggest to be wisest to use the term *"electronic nose"* only between quotation marks, and have proposed to use instead the term "flavour sensor", or perhaps "odour sensor", "aroma sensor" or "gas sensor". That is from far a most appropriate way to name the concept, even being less grandiloquent and may be also less commercial; to a certain extent it is more difficult to catch the meaning from a non expert audience point of view. My opinion is that the term *"electronic nose"* should be used in a more restricted way and only in studies that support their results on sensory analysis must be allowed to coin it. Nevertheless, scientist should hold a name which designates the overall concept; as stated, some authors have proposed to move the name *"electronic nose"* to other names such as "*gas sensor arrays"*. The term "gas sensor" remains linked to the idea of the sensing part of the overall concept of *"electronic nose"*. I have not been able to find or suggest a name that completely can express the overall *"electronic nose"* concept and finally I decided to write this thesis using the term *"electronic nose"* in inverted commas and italic letters in order to let the reader be aware of the vague and maybe inaccurate term traditionally used for defining the *"electronic nose"* concept.

At present, a technology such as the *"electronic nose"* is unlikely to represent a practical alternative to human sensory screening. However, it may offer a potential means of detecting changes in the headspace of samples. If these changes are related to the quality of the samples under analysis, then the *"electronic nose"* may represent a powerful, fast, feasible and robust approach to monitor and interpret these changes and consequently to assess quality property under study even those changes stems from "odourless" molecules.

In any case, there might be some cases where a strong correlation between the *"electronic nose"* data and the aroma of food exists. In these particular cases, *"electronic noses"* could be very useful in the quality control of foodstuff. Some other cases can also benefit from this technology. As already mentioned before, another challenging situation where the *"electronic nose"* technology is well suited is when detectable changes on the headspace of samples are related to the property under study

Despite some promising and impressive successes, solid-state sensor based systems have not generally lived up to expectations. Problems with drift (short and long term), noise instability due to water vapor, sensor poisoning, the need for time-consuming re-calibration, poor sensor-to-sensor and instrument-to instrument reproducibility have slowed the comercialisation of *"electronic nose"* instruments based on conducting polymers, metal oxide sensors, surface acoustic wave and other type of solid-state sensors. Because of these problems, several manufacturers of *"electronic nose"* instruments have gone out of business, and those that continue to manufacture and market *"electronic nose instruments"* have devoted much of their research to overcoming these deficiencies.

2.4 MS-SENSOR

2.4.1 MS-Sensor definition

Within the context of the *"electronic nose"* definition it is also possible to enclose a emerging technology that uses a mass spectrometer as the sensing part of *"electronic nose"*. Although *"electronic noses"* using gas sensors have been commercially available for several years, their narrow selectivity, their sensitivity towards moisture and the short life-time of the sensors can be serious limitations. Often the fluctuation in the performance may lead to difficulties reproducing data. Then, recently, interest has been focused on a new generation of artificial noses that seems overcome much of the drawbacks described above. They are based on mass spectrometry and they open new perspectives in the analysis of odours and volatile compounds. Due to the physical principles underlying this technique,

this device has great advantages over the solid state *"electronic noses"*, particularly in terms of selectivity and sensitivity. This type of *"electronic noses"* allows the direct characterization by mass spectrometry of volatile organic components from liquid and solid samples without separation of the headspace constituents. Their basic working principle is based on the introduction of volatile components extracted from the headspace of a sample into the ionization chamber of a mass spectrometer. Such devices usually employ quadrupole mass spectrometers in which each m/z ratio acts as a "sensor" that detects any ion fragment with that ratio. Thus, the number of sensors is variable, readily modifiable and, in most cases, high. The mass spectra resulting from the simultaneous ionization and fragmentation of these volatiles constitute a very complex ionization pattern that can be seen as a "fingerprint" and that is characteristic of the matrix being analyzed. These ionisation patterns are then processed by pattern recognition engines to perform tasks associated to *"electronic nose"* systems such as classification, recognition and, to a limited extent, quantification.

In keeping with the traditional *"electronic nose"*, a parallelism between MS-Sensor and human nose would also be possible as showed in Figure 2.3. This holds true in the sense that the MS-Sensor approach has three different blocks: the sampling system, the sensory system and pattern recognition system. Nevertheless the same restriction to the term *"electronic nose"* should be applied to the MS-Sensor concept. Most authors coins the term *"electronic nose"* when referring to MS-Sensor concept[21-28]. As in the case of classical gas sensors based *"electronic nose"* systems, one may found different ways to refer to the mass spectrometry based *"electronic nose"* concept. Terms such as "mass sensors", "fingerprint mass spectra", or sometimes *"new generation electronic noses"* or "MS-based e-nose" have been used without distinction. The same author been fighting against the term *"electronic nose"* finds estrange that once the main *"electronic nose"* manufacturer had himself changed the name for "Sensors Arrays Technology", the initial name was hijacked by the main "fingerprint mass spectra" manufacturers[29]. Following this author's recommendation and in order to provide certain uniformity to the whole manuscript, the term to refer to this approach from now will be MS-Sensor.

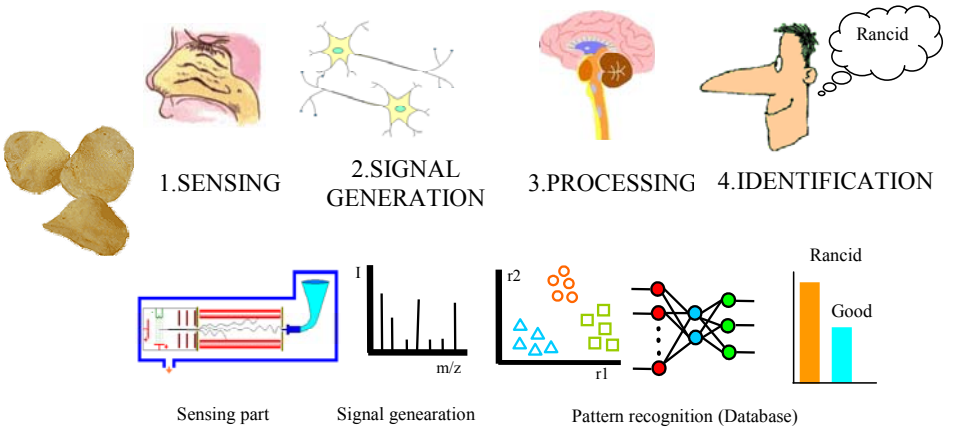

Figure 2.3: Basic diagram showing the analogy between biological olfaction and MS-Sensor

2.4.2 MS-Sensor system blocks

Again, and as in case of classical MOX "*electronic noses*", the MS-Sensor system is also a union of three elements: the sample handling system, the detection system and the data analysis system.

2.4.2.1 Sampling block

The headspace sampling system directs the sampling volatiles coming from the headspace of solids or liquids to the mass detector. Despite the fact that usually benchmarked MS-Sensors are equipped with static headspace autosamplers, a great range of possibilities to introduce more sensitivity and selectivity sampling methods (i.e., P&T, SPME, etc) are open and recently there have been appearing lot of works related to the different options of handling volatiles in the MS-Sensor approach. This section will give more details of the different options in volatile handling techniques.

2.4.2.1.1 Non preconcentrated Static Headspace (SH)

SH is an extraction technique for semi volatile and volatile compounds. When the system is in equilibrium, the composition of the vapour phase is qualitative and quantitaive representative of the composition of the original sample. An aliquot is taken from the gas phase and transferred for analysis. Without further steps like pre-concentration, sample extraction or chemical treatment it is a highly reproducible method of sample introduction to an analytical instrument making a high sample throughput possible. By taking an aliquot of the gas phase, the volatile components in an essentially non-volatile matrix can be investigated without interference. In a closed or static system the gas or vapour phase will be in equilibrium with the condensed phase. Analytes are distributed between the condensed phase matrix and the vapour phase. Conditions are adjusted so that the analyte distribution favours the vapour phase (to increase the concentration of the headspace in volatile compounds). According to Pérès et al[30] five main types of extraction-injection modules have been used for static headspace analysis in MS-Sensor:

1. Systems using gas syringes: an aliquot is taken from the headspace and injected by means of a chromatographic syringes[31, 32].
2. Systems using a stream of inert gas: an aliquot of the headspace is transferred by the carrier gas into the ionization chamber of the mass spectrometer for a short period of time. Instruments equipped with this type of extraction-injection have been employed in the context of studies using Time of Flight and Ion Trap mass spectrometers application[33].
3. Balanced pressure sampling systems: the vial is pressurized by the carrier gas to a pressure equal to the carrier-gas inlet pressure of the transfer line. Next, the carrier-gas supply is interrupted by closing a valve and injection is performed by expansion of the gaseous mixture into the transfer line[34].
4. Pressure/loop systems: the sample vial is pressurized by an inert gas to a pre-set value. Then, the vial is opened temporarily toward a sample loop of a gas-sampling valve and not directly to the transfer line. Finally, purging of the sample loop with the carrier gas injects the volatile components into the mass

spectrometer. This is the typical layout for an automatic headspace sampler system which will be detailed in next chapter (Section 3.1.1). This is the technique most commonly used for handling volatiles in the MS-Sensor approach (see the System configuration columns of tables in section 2.8 for getting an idea on how often SH is used in the reviewed MS-Sensor applications). The biggest advantage is its simplicity. There are some parameters to optimize such as the sample temperature, equilibration time and size of the vial[7]. However, because there is no pre-concentration, the sensitivity of this type of instrumentation can prove insufficient for certain applications[35].

SH is a suitable method for routine analysis of most of the volatile compounds because the preparation of the sample is very simple and the extraction is completely automated. Moreover, the cost is low and the reproducibility is high. However, this method is only sensitive for highly volatile compounds or medium volatiles when present in high concentrations. When it is requiered to sample semivolatile molecules in low concentration, then the use of preconcetration techniques is better suited for MS-Sensor.

2.4.2.1.2 Preconcentration techniques

In order to increase the amounts of material injected, different methods of pre-concentration have been proposed. Pre-concentration improves the sensitivity of the system but introduces a supplementary step, which can prove limiting from a temporal point of view and/or generate analytical artefacts (memory effects, bleeding, irreversible adsorption). In this respect, the pre-concentration medium must be chosen with care.

In order to increase the sensitivity, the use of preconcentration techniques coupled to MS-Sensors systems has grown. The P&T and DH techniques are classical methods of pre-concentration of volatile compounds used in a variety of applications. In both systems, the volatile components are purged by a stream of inert gas and trapped onto an adsorbent. The constant depletion of the headspace leads to a displacement of the equilibrium in favour of the desorption of these molecules from the matrix. The trapped molecules are

subsequently desorbed by heating and injected into the ionization chamber of the mass spectrometer. Apart from the choice of the trap, the main parameters to optimize are the temperature of the sample, the equilibration time, the flow rate of the extractor gas and the purge time of the headspace. In the case of P&T, the gas flow is injected through the sample (preferably in liquid or powder form), whereas, in the case of DH, only the headspace is purged with the gas.

Both, DH and P&T has been used by Schaller et al.[31] in order to preconcentrate volatiles above the headspace of swiss emmental cheeses. Pérès et al[36] also used DH for concentrating volatiles coming from Camembert cheeses on Tenax and then for further direct desorption on MS-Sensor. Another work describing the use of a DH as a sampling technique to couple to MS-Sensors has been reported by Boudaoud et al.[37] where they use this system for enabling the rapid characterization of volatile compounds of cork wine stoppers. Although DH improves the sensitivity of the system, it introduces a supplementary step in the method, which increases the time of analysis. Moreover, analytical artefacts (memory effects, bleeding or irreversible adsorption) are generated in some cases.

Another extraction technique is SPME. There are several published works that make use of SPME coupled to MS-Sensors approaches in order to extract volatiles from cheese (emmental[31] or camembert[38]); milk[39, 40]; tomato[41]; unifloral honeys[42]; bakery products[43] This technique has a considerable concentration capacity and it is very simple because, unlike DH, it does not require special equipment [30]. As in case of SH this technique also needs some parameters optimization [44, 36, 45, 42], being the coating of the fiber, the extraction time the most important ones.

SPME is an extraction technique introduced by Pawliszyn and coworkers[46] that has some advantages over conventional sample-preparation techniques such as solid-phase extraction (SPE) and liquid-liquid extraction (LLE). It is solvent-free, usually only small volumes of sample are required, and is easy to use and automate. The SPME process can be divided into two steps: first the analytes are absorbed by a coated fused silica fiber and then

the retained analytes are desorbed into the analytical instrument. The desorption step takes place usually in the injector port of a GC/MS or a MS-Sensor. A more detailed explanation of the headspace sampling procedure will be further given in Section 3.1.3

Since it has the capability to extract and concentrate headspace volatiles from a food matrix, SPME is significantly more sensitive that SH. That has been demonstrated by Penton et al[7]. They showed that SPME with CAR/PDMS fiber was 30 to 50 times more sensitive than SH for measuring important flavour compounds in beer. The SH/MS-Sensor approach is normally applicable for the analysis of analytes volatiles in the low ppm range while SPME can be extended to analytes in the ppb range and even in the ppt range.

A new technique called stir bar sorptive extraction (SBSE) has been presented recently. This technique seems a promising approach when a very high sensitivity is required[47].

There are several published works dedicated to the comparison among different handling headspace techniques for MS-Sensor applications. Schaler et al[31] compared the performance among three headspace (SH, SPME and P&T) sampling methods coupled to MS-Sensor device in order to discriminate cheese ripening age for "Swiss Emmental" cheeses. When SH, SPME and P&T techniques were compared together, measurements performed with SH showed comparable intensities than those done with the two preconcentration methods for small molecular masses, i.e., approx. until 45 amu. When the small molecular masses, i.e., under 45amu, were taken away, the SH exhibits very small responses in comparison with those obtained with the two other techniques. From the two systems tested, the SPME method shows the highest responses. The repeatability of the measurements using the SH and SPME techniques was comparable. In the other hand, a very poor repeatability was observed for measurements performed with the P&T instrument. This poor repeatability can be partly related to the dynamic injection of this system. This observation is in agreement with that of Roussel et al. [48] who found that a dynamic injection gives less repeatable measurements than the static one. They also considered the P&T system as "complex and difficult to master".

There is also an interesting comparison between SH and SBSE in the detection of whiskey adulteration with a MS-Sensor[49]. Different sampling techniques, with and without pre-concentration of volatiles were tested (SPME, SH) by Ampuero et al[42] in order to study the potential of a MS-Sensor for the classification of several Swiss honey samples by their botanical origin. In general, all these studies conclude that the preconcentrated technique offers a better performance for the applications envisaged. Another interesting work concerning headspace sampler technique effects on the MS-Sensor performance was developed by kinton et al[50]. In this study they tried to clarify whether the MS-Sensor approach coupled to different headspace techniques was able to predict TCA (2,4,6-tricholoranisole) levels in wine. The main conclusion was that SH/MS-Sensor was able to predict TCA in the range of ppm levels while the SPME/MS-Sensor allowed a low ppb detection range. They proved that in order to predict TCA concentration in the range of ppt levels in wines it is necessary make use of enrichment techniques followed by thermal desorption. That results in agreement with Hayasaka et al[51] who used SBSE for enabling quantification of TCA in wines.

Marsili et al[39] found also a better "precision of replicates" with the SPME when compared to DH based method for the study of off-flavours in milk. In the same paper, he reported two additional advantages of the SMPE: no carry-over peaks from sample to sample and no background peaks.

Furthermore Castel et al[52] in a recent study compared several headspace sampling methods SH, SPME and SHSE (headspace sorptive extraction, the headspace version of SBSE) for the analysis of different grades of benzoin gums (Siam and Sumatra). SPME using DVB/CAR/PDMS fiber and SHSE seem to be the most suitable techniques to identify volatile compounds of benzoin gums. SH showed less sensitive but represented a good method for the quality control of these gums. For this reason it was applied directly coupled to a MS-Sensor for a rapid differentiation between several benzoin gum qualities.

Although a priori any sampling headspace technique can be used as the sample-handling part of a MS-Sensor, the choice must be made with care and take into account the

type of sample and the method specifications required[53]. SH is a simple fast and automated technique to implement; no sample preparation is needed. However, this technique presents poor sensitivity compared with the techniques involving an enrichment or partition step such as DH, P&T, SPME or SHSE.

2.4.2.2 Sensor block: mass spectrometer and transferline

The mass spectrometer detector is formed by the transferline and the sensing system. The transfer line ensures coupling of the extraction/injection module to the mass spectrometer. This line must be designed to maintain a high-quality vacuum in the spectrometer source while allowing rapid transfer of the extracted molecules between the two modules. Heating is necessary to prevent recondensation of the compounds. In addition, the internal side, whether fused silica or silica coated steel, must be inactivated in order to limit the risks of pollution in the course of analyses (memory defects or loss of molecules through reversible or irreversible adsorption) and to avoid the catalytic degradation of certain compounds[30].

Concerning the mass detector system, several classes of mass spectrometers do exist depending on the ionization technique or the mass analyzer (technique used to separate the ions).

In one hand, most of the benchmarked MS-Sensors devices have an electron beam ionization system and a quadrupole mass analyzer. Electron beam acts ionizing molecules using the electron beam principle. A certain voltage, usually 70eV, is applied in a tungsten filament and a flow of electrons are generated which impact the molecules releasing from the transferline. This impact leads to fragmentation and ionization of these molecules. Finally these ions are separated in the quadrupole mass analyzer according to their m/z ratio. Quadrupole acts as a “mass filter” that combines direct current and radiofrequency potentials on the quadrupole rods. These currents can be set to pass only a selected mass-to-

charge ratio. All other ions do not have a stable trajectory through the quadrupole mass analyzer and will collide with the quadrupole rods, never reaching the detector.

Nevertheless there exist the so-called second generation instruments which used soft ionization techniques and also other type of mass analyzer. Although marginal at present, an interface permitting ""soft"" ionization of the molecules at atmospheric pressure has been proposed[21, 54]. In association with a time-of-flight (TOF) analyzer, this technology offers the advantage of providing more specific information by allowing one to trace the molecular origin of the ions (mass fingerprinting). Another mass spectrometer class implemented in the MS-Sensor approach is the ion trap mass spectrometer[33, 55].

2.4.2.3 Pattern recognition block

Finally, the third part is related to pattern recognition engineering. When performing measurements with a multi-sensor system such as a MS-Sensor, a large amount of data is generated. The user of a MS-Sensor is usually interested in various properties of the samples in question rather than in the sensor mass spectra signals themselves. Thus, there is a need for data processing methods that allows the correlation of these mass spectra into the sample properties of interest. It is to say that pattern recognition will be used to decipher meaningful trends in the mass spectra output. Several chemometric techniques are used for tretaing this output, most of them are available in statistical software packages that are usually included in the instrument. The chemometric techniques used include unsupervised and supervised pattern recognition techniques. The former reveal natural groupings of the samples in the data set and also detect outlying samples. Hierarchical cluster analysis (HCA) and principal component analysis (PCA) are the most commonly used unsupervised pattern recognition. In the supervised pattern recognition techniques, class membership is known in advance. The aim of these techniques is to build a model to discriminate between several predefined classes or to assign an ""unknown"" sample to a given predefined class. Soft independent modeling of class analogy (SIMCA), K-nearest neighbors (KNN), linear

discriminant analysis (LDA) or discriminant function analysis (DFA) are the supervised techniques that are most commonly used[53]as reviewed in section 2.8.

Below, in Section 2.9, pattern recognition models used in this tesis tasks will be briefly described. It is important to realize, however, that the success of any pattern recognition method depends on reliable MS-Sensor data. Such data can be obtained by using a reliable and accurate measurement system, in addition to taking the necessary precautions to prevent unknown variables from affecting the properties of the samples under investigation.

2.5 MS-SENSOR vs. MOX *"electronic nose"* APPROACH

The MOX *"electronic nose"* and the MS-Sensor play distinct roles in odour applications. However, a MS-Sensor appears to offer many advantages over the more traditional MOX technologies. In the e-nose and headspace MS instruments, there are noteworthy similarities and dissimilarities in functional features. In both cases, the sample is introduced as a thermally equilibrated vapour. Both device types also output a vector of responses; however, the physical processes which give rise to those responses are significantly different. In a MS-Sensor, the various compounds in a narrow plug of sample vapour are ionized and fragmented, with each compound producing a characteristic fragmentation pattern. These ion fragments are filtered by the mass analyzer (for example, a quadrupole) then detected by an electron multiplier. The output vector is a mass spectrum. In a classical MOX "*electronic nose"*, a carrier gas sweeps the sample vapor through the chamber(s) where its components interact with sensor substrates. Depending on the sensor type, a sensor property (e.g., its resistance) is altered by the sample to produce a time dependent response. The output vector is assembled from a user specified number of response points (e.g., the maximum) from every sensor. The performance of common multi-sensor arrays is ultimately determined by the properties of their constituent parts. Key parameters such as number, type and specificity of the sensors determine whether a specific instrument is suitable for a given application. In some cases, it may be necessary to find the

optimal set of solid state sensors to solve an analytical problem. As this requires time and effort, the applicability of solid state sensor technology is often limited. With mass spectrometry used as a sensing technique, this limitation does not exist. Array selection and deselection can be done rapidly by changing the scanning method or simply by changing the fragment ions used for pattern recognition, thus adapting the *"electronic nose"* to a specific application. Having this immense potential, it is very useful to have a precise analytical picture of the given application in order to choose the most suitable parameters for the sensor measurements. When purchasing an *"electronic nose"*, the number and type of sensors to incorporate must be specified. With MS-Sensor preconfiguration is unnecessary as mentioned above. Thus, time spent developing methods is not as large as in case of *"electronic nose"* technology. Between one or two experiments are usually sufficient to determine the appropriate mass range and sample heating parameters; a method may be developed and samples run in a single day. Sensitivity and selectivity, amount of sample volatilized, headspace vial temperature, and the scan range determine the sensitivity of a MS-Sensor. The instrument can also be operated in selected ion monitoring (SIM) mode to increase sensitivity. The sensitivity of solid state sensors is determined by their type, the flow rate over the sensor, the analyte, and temperature. MOX are subject to interference from water and alcohol in the sample, reducing not only the sensitivity to other constituents but also the sample throughput by requiring a lengthy recovery time. Moreover, MOX *"electronic nose"* can be poisoned by strongly-adsorbing materials, such as sulphur-containing compounds. Analogous effects are not observed in mass spectrometry. To the extent that a solvent, say, produces a large, even overwhelming signal at one or several m/z settings, the performance at other m/z settings is unaffected. The MOX *"electronic nose"* sensors are known to be subject to short and long term drift due to changes in relative humidity. In addition, individual sensors have to be replaced periodically. Mass spectrometers have been designed to minimize effects of changes in the external environment; a high degree of stability is a requirement for any analytical technique which relies on standard library searching.

As it has been mentioned the MS-Sensor approach offers important advantages over current solid state sensors, including no problem with water, alcohol, sulfur-containing

chemicals and polar compounds in general or poisoning; a linear response to vapor concentrations up to 10^4; much less drift and significant improved reproducibility. Usually it is faster than gas sensor *"electronic nose"* technology allowing for higher throughput. Basically that is due to the fact that in the case of MS-Sensor there is not a need for response recovering since it is immediately done. Generally speaking MS-Sensor takes 3-5 minutes for each sample and that compares favourably to the at least 30 minutes that may take gas sensor based *"electronic nose"*. Maybe the most important benefit of using MS-Sensor is that can be used to determine not only if a test sample is different from a standard sample but also why it is different. That is just because the mass spectra pattern obtained from MS-Sensor have intrinsically chemical meaning since a mass spectrometer generates directly-interpretable chemical information. The use of a state-of-the-art analytical technique such as mass spectrometry yields a very reproducible and precise fingerprint of each sample. Thus identification of a component is straightforward, and the comparison of large data sets easy. The high performance of this technique offers new perspectives in the analysis of odors and volatile organic compounds. Mass spectrometers specifically configured for rapid headspace analysis were already introduced in 1998. Nowadays different manufacturers are dedicated to commercialize MS-Sensor device. In the next section the differences among all commercial prototypes will be reviewed.

2.6 COMERCIALLY AVAILABLE MS-SENSOR SYSTEMS

Table 2.1 summarizes all the MS-Sensor prototypes currently available from a commercial point of view. Undeniably, one of the pioneering companies in the *"electronic nose"* field was Alpha MOS. This company develops *"electronic nose"* equipment based on technologies, the MS and MOX. One of their equipments, the so-called Prometheus is a combination of both technologies MS model (Kronos) and MOX model named FOX. In summary, it allows a split of the headspace extract and a combination of the responses in the pattern recognition. A different company working in the MS-Sensor field is HKR Sensorysisteme GmbH. Actually; they have two MS-Sensor systems. The first one is called MS-Sensor® that uses as mass analyzer a turbo mass Perkin-Elmer. This analyzer

incorporates CI and EI as ionization systems. The second one is called SensiTOF® and uses REFlecttron Time-Of-Fligth (RETOF) technology as the mass analyzer. That implies a spectacular gain in the sensitivity. The Smart Nose-300 instrument is another commercially available MS-Sensory system which combines in a full automated way a quadrupole mass spectrometer with an autosampler from Balzers Instruments Inc. The system is entirelly software controlled

Table 2.1: Commercially available MS-Sensor systems

Trade Mark	Kronos	Prometheus	MS-Sensor®	Sensi TOF®	SMart Nose-300	HP4440 MS
Technology	Alpha M.O.S	KRONOS (MS)-FOX400 (MOX)	TurboMass (Pelkin-Elmer)	R. M. Jordan Company, Inc	Balzers Instruments Inc.	SH(Agilent 7694)GC/MS 6890/Agilent 5973N)
Ionization Source	EI	EI	EI/CI	EI	EI	EI/CI
Massic analyzer	Quadrupol	Quadrupol	Quadrupol	RETOF	Quadrupol	Quadrupol
Rang m/z (amu)	1-200	1-200	1-1200		1-200	
Mode	Scan/SIM	Scan/SIM	Scan/SIM	Scan/SIM	Scan/SIM	Scan/SIM
Authomatic headspace	SH, SPME, SBSE, SPDE, TDA, Baker	SH, SPME, SBSE, SPDE, TDA, Baker	SH	GC-inlet SPME/inlet.	SH, SPME, SPDE	SH, SPME, SBSE, SPDE, TDA
Pattern recognition software	PCA, DFA, PLS, SIMCA	PCA, DFA, PLS, SIMCA	PCA, GDF, RBF neural network	PCA, GDF, RBF neural network	PCA, DFA	Pirouete (infometrix) PCA, DFA, HCA, PLS, SIMCA, KNN

The Quadstar® from Balzer Instruments controls operations of the mass spec. channels and sample measurmentnts. The Smart Nose software processes the raw mass spectrometric data using statistical algorithms such as PCA or DFA yielding a more userfriendly representation of the results. Finally, Agilent Technologies develops the SH4440 MS-Sensor which results from a combination of its own SH autosampler (HP7694) and recent in-house mass spectrometer 5973 MSD. The Infometrix software named Piroutte allows for data processing using classical pattern recognition models. The Agilent 4440

Chemicla Sensor is nowadays commercialized by Gertel GmbH which offers a wide range of possibilities for improving sampling volatiles step (SBSE and SPME) and also incorporates the possibility for CI.

2.7 MS-SENSOR CURRENT WEAKNESS

Despite the great potential of MS-Sensors in the field of food analysis there are some aspects of the technique that still need further improvement.

2.7.1 Drift and long-term signal stability

Even though the problem is less pronounced than with conventional MOX "*electronic noses"*, the question of signal stability is still important with mass spectrometric instruments and still arises for such systems. Signal drift can occur for many different reasons. Largely these reasons are related to the mass detector itself: a) gradual fouling of the source; b) aging of the ion multiplier and maintenance operations (e.g., opening and cleaning of the source, changing of the electro- or photomultiplier, replacement of the filament and repeller); c) vacuum instability and impaired vacuum quality (quality of the carrier gas, introduction of too much material, etc.). On the other hand, the origin of drift can be also caused by the headspace sampler system. Depending on its origin, the drift can be gradual or sudden, linear or nonlinear, and very often difficult to predict. Thus, although various tuning procedures exist, if the signal goes uncorrected during a campaign of measures, the resulting data could be useless. When no corrections are made, the effects may lead to a high irreproducibility for replicate measurements, gradual drift, and changes in sensitivity.

It is important to have procedures to monitor the state of the mass detectors and help correct drift. Little work has so far been published on these issues. A recent review by

Pérez Pavón et al.[56] describes different strategies for overcoming long-term instability signals in MS-Sensors systems either for classification or for quantification problems.

For classification purposes they suggest that the most usual technique for addressing signal stability is Internal Normalization which involves expressing each mass intensity as a percentage of the total sum of intensities. This type of data treatment corrects the irreproducibility of the headspace sampler and sensitivity changes, provided they are constant along the mass axis. This is the much common approach used to overcome this drift since no additional sample manipulation is needed, but it must take in to account that the quantitative information is lost. It is appropriate for many characterization problems in which the relationship among the different m/z ratios carries the relevant information instead of the intensity of these signals. They suggest a second strategy to eliminate long-term signal instability in classification purposes based on Support Vector Machines (SVM) which allows the construction of reliable models using a reduced number of samples.

In the case of quantification, they reviewed three strategies called internal standardization, standard gas addition and calibration transfer. In internal standardization each mass intensity is divided by the intensity of one fragment of an added internal standard. The effect of this type of correction is similar to that of internal normalization, but the quantitative information remains in the signal. Nevertheless, additional sample manipulation is necessary (the addition of the internal standard). Moreover, suitable internal standards are hard to find because they must afford a fragmentation pattern that will not interfere with the sample profile. Marsili[40] proposed to use an internal standard (chlorobenzene) to overcome serious instrument problems in SPME/MS-Sensor processed milk measurements. Furthermore, internal standardization is very difficult in the case of solid samples. Pérès et al[57] proposed a signal drift of a MS-Sensor by a method called standard gas addition (SGA). It consisted on introducing a gaseous He-Xe mixture continuously and independently of the carrier gas into the mass spectrometer source. In this work, the influence of the SGA signal correction on the discriminating power of the data was evaluated from the analysis of three batches of cheeses by DH/MS-Sensor. SGA afforded a good correction of the main types of drift classically observed in mass

spectrometry. They compared the results obtained from SGA relative to the ones obtained through internal normalization. Unlike internal normalization, SGA normalization was carried out relative to the ^{129}Xe, which is independent of the other mass fragments. Hence there were no statistical coelutions between variables. Consequently, the discriminating power of the data was thereby improved. The method of signal correction by SGA normalization showed that the controlled introduction of the standard gas ^{129}Xe could correct effects arising from modifications made to the instrumental parameters of a mass spectrometer. In internal normalization, the n individual variables have a high covariance due to the mode of expression (relative to the percentage of the sum total), which generates statistical links among all the fragments, some of which drift strongly (in particular ions with high abundance). According this study SGA offered a better solution than internal normalization for the correction of spectral fingerprints obtained by DH/MS-Sensor. Finally, Pérez Pavón et al[58] claim for Calibration transfer as an alternative for rectifying drift and sensitivity changes. It is based on the signal variation observed for a set of reference samples. The intensity for each m/z ratio of the prediction set samples was multiplied by the numbers obtained as a way that best approached the intensities of the transferred sample set to those corresponding to the calibration model. Its main advantage is that it corrects sensitivity changes even when they are not constant along the mass axis. The main disadvantage is that an additional set of samples must be analyzed at regular intervals. However, this is not a big drawback for a rapid technique such as SH/MS-Sensor. Signal stability is one of the problems that has not been solved completely to date.

2.7.2 Price and portability

Another important disadvantage of MS-Sensor devices is the difficult portability when on-line or in-situ applications are envisaged. Nevertheless there exists some promising work where some efforts have been addressed to miniaturize parts of the mass spectrometer[59].

Another issue where MS-Sensor devices do not compare favorably to the other techniques is the price because a MS-Sensor is much more expensive that classical "*electronic nose"* techniques. However, their adaptability and versatility allowing the same instrument to be used for more than a single applications compensate in a somehow this price.

2.7.3 Dimensionality of response data matrices

Another clear difficulty that must be addressed is the high dimensionality inherent to performing full scan measurements. Since in MS-Sensor no prior knowledge of the samples is theorically required, it is often necessary to set a wide range of m/z ratios (e.g., from m/z 35 to m/z 300) in order to cover all the fragmentation range of the whole volatile molecules conforming the headspace extract. Therefore, as every mass to charge ratio (m/z) in the mass spectra can be thought of as a sensor, the response matrix will contain as many variables as m/z sensors considered. Thus, the models used may suffer from the so-called course of dimensionality. Course of dimensionality is produced when the number of variables exceeds the number of measurements available to train the pattern recognition methods. Then, there is a dangerous situation due to the severe risk of overfitting the model. That can be addressed by reducing the dimensionality of the data matrix. Bellow, current strategies reported for the reduction of dimensionality are reviewed.

2.7.4 Low specificity of m/z variables in response data matrices

The specificity of m/z ratios used as "pseudosensors" in MS-Sensor approach will definitely influence the performance of data modelling. This poor specificity resides on the high degree of fragmentation obtained with electron beam ionization. One way to tackle that is just changing the ionization system to soft ionization techniques. That could result very expensive and not practical. Another way could be to consider the time dimension variation of each one of the m/z fragments scanned. In this way the m/z becomes much

more specific because even the same fragment may be the result of fragmentations for different molecules, their temporal variation should be different since they come from different molecules. Common pattern recognition algorithms applied on MS-Sensor data make use of data matrices in which columns represent each one of the mass to charge ratio (m/z) scanned and rows hold their intensities. Even when chromatographic resolution is avoided, some kind of diffusion can be observed on this asymmetrical peak which shows some retention potentially able to report additional information that may help for further modelling of data in classification or prediction tasks. Time averaging of the mass spectrum may lead to a loss of this temporal information so that this potentially useful data is not used in the pattern recognition step. Actually, averaging mass spectra along the detected peak allows converting real three-way data to two-way by eliminating the time dimension. Therefore, the real nature of data is not respected and is changed just to adapt the response of the instrument to current available pattern recognition two-way algorithms. The possibility of computing this extra information using multi-way analysis[60] are going to be explored throughout this thesis. Multi-way methods are particularly useful for the analysis of batch process data and analytical data that intrinsically present more than two sources of variation [61] (i. e. MS-Sensor data, where a response is being measured as a function of two parameters: m/z and time). In fact, data provided from the MS-Sensor should be arranged as multi-way array where the first mode represents samples, the second corresponds to mass spectra and the third to the elution profiles.

2.8 THE STATE-OF-ART ON APPLICATIONS

Nowadays, MS-Sensor is been used widely in a broad range of applications being food quality assessment the most common. Environmental analysis and more recently biomedical analysis have been also attempted within the framework of this technology.

The following state of art is aimed to review current trends in the development of research lines in which MS-Sensor have been involved. Here bellow main applications developed within the context of this technology are going to be reviewed. They are

summarized in different tables according to the type of food analysed. At this way, Table 2.2 shows the applications referred to lactic products; in Table 2.3 one may find the applications reviewed referred to meat, eggs and fish products; in Table 2.4, MS-Sensor applications on analysis of products related to seasoning, dressing and spicing are depicted; Table 2.5 holds the reviewed applications related to fruit, juices and soft drinks; Table 2.6 summarizes the applications related to the alcoholic beverages products. Finally, in Table 2.7 the non food applications are gathered. The publications are sorted in each table in chronological order. Also specified are the types of study conducted, the samples used, the MS-Sensor system configuration and the chemometric treatment applied.

Table 2.2: Applications of MS-Sensor in lactic products (chronological order)

Year	Sample	Type of study	System	Data analysis	Ref.
1999	Milk	Study of abused milk	SPME/MS	PCA	39
2000	Emmental cheese	Discrimination of the cheese ripening age	SH, SPME, P&T-MS	PCA	31
2000	Milk	Development of an analytical method for predicting shelf life	SPME/MS	PLS	40
2001	Camembert Cheese	Differentiation of ripening stages, manufacturing process.	SPME/MS	DFA	38
2002	Milk	Determination of trimethylamine	SH/MS	PCA, Univ.cal.	65
2002	Processed cheese, Evaporated milk	Data transferability between two MS-Sensors	SH/MS	PCA	66
2002	Camembert commercial cheese	Characterization of cheeses according manufacturing type and according to ripening stage	SH/MS	DFA	35
2002	Camembert Cheeses	Study of the influence of the signal correction by Standard Gas Addition on the discriminating power of the DSH/MS	SH/MS	PCA, DFA	57
2002	Camembert Cheeses	Characterization of cheeses according manufacturing type and according to ripening stage	DH-MS	Forward stepwise DFA	36
2003	Emmental cheese	Determination of origin	SH/MS	PCA	67
2003	Milk powders	Oxidation of infant milk powders	SH/MS	PCA	23
2004	Lactic acid bacteria	Differentiate bacterial populations in Gruyere cheese samples and to screen for new aroma-producing strains.	DSH/MS	PCA	68

As it is shown in Table 2.2 the MS-Sensor approach has been widely used in lactic derivates industry. To our knowledge, the first reference to point out the concept of using

MS-Sensor was Marsili[39] in 1999 where he claims for the use of a new technique using solid-phase microextraction, mass spectrometry, and multivariate analysis (SPME/MS-MVA) for the study of off-flavors in milk. The same author[40] used SPME/MS-Sensor to predict shelf life of processed milk by means of a PLS model. Marsili has also published several papers on the discrimination of off-flavours in milk[39, 40, 62-64]. In other work the discrimination power of the SPME/MS-Sensor approach was evaluated using five different cheeses of "Camembert" type[35, 38]. Determination of trimethylamine was also assessed by means of a MS-Sensor device[65], this amine was shown responsible for a disagreeable fishy off-flavour in milk. Among all the related papers it should be highlighted an interesting study about data transferability between different instruments[66].

Another important field of application is related to fresh products such as meat, fish and eggs and it is summarized in Table 2.3

Table 2.3: Applications of MS-Sensor in meat, egg and fish products (chronological order)

Year	Sample	Type of study	System	Data analysis	Ref.
2004	Fat samples from egg chicken and pork	Analytical decision maker for PCBs fast screening in food samples	PTV-CI-MS	Univ cal.	54
2002	Fish, egg, chicken, milk,	Determination of spoilage markers in spiked food (Dimethyl sulphide, trimethylamine Diacetyl)	SH/MS	PLS	69
2006	Fish: whiting and mackerel	Volatile compounds to characterize fish spoilage	SH/MS	PCA	70

Products for dressing and seasoning are also good candidates to be measured by a MS-Sensor approach. In Table 2.4 it is reviewed the different applications envisaged with a MS-Sensor on this type of products. In 1998, Elke Anklam[71] used a similar approach to the MS-Sensor in order to distinguish between industrially made vinegar "Aceto Balsamico di Modena" and traditionally produced vinegar "Aceto Balsamico Tradizionale di Modena e di Reggio Emilia" by means of pyrolysis-mass spectrometry (Py-MS). Conceptually, Py-MS was the same device than a MS-Sensor. There are other references where Py-MS has also been used[72, 73]. Nevertheless, these references were not included in the table. Despite

the fact that Py-MS uses approximately the same approach that MS-Sensor do, these last references has not been included in the tables since this device will be not considered as a MS-Sensor. The main reason is that the analyzed fraction is the pyrolisate, a fraction of products derived for a thermal decomposition of the sample under analysis. B. Dittmann et al.[34] used an SH/MS-Sensor device to discriminate between different dosages of garlic flavoring in tomato sauce. This paper relates aroma compounds measured by means of an *"electronic nose"* and headspace fingerprint mass spectrometry. Detection of adulterants[74] and differentiation of sources[74] in olive oil are also demonstrated to be feasible by means of an SH/MS-Sensor.

Table 2.4: Applications of MS-Sensor in dressing, seasoning, and spicing type products (chronological order)

Year	Sample	Type of study	System	Data analysis	Ref.
1998	Aceto Balsamico di Modena	Discrimination between industrially made vinegar from and traditionally produced vinegar	Py-MS	PCA	71
2000	Cloves	Freshness of cloves	SH/MS	PCA	75
2000	Tomato sauce	Discriminate between different dosages of garlic flavouring in tomato sauce	SH/MS	PCA	34
2002	Olive oil	Detection of adulterations	SH/MS	LDA	76
2002	Olive oil	Characterization of olive oils	SH/MS	HCA, PCA	77
2002	Olive oil	Differentiation of monovarieties	SH/MS	LDA	74
2004	Olive oil	Screening for residual benzene hydrocarbons compounds VOCs in olive oil	SH/MS	HCA, PCA, KNN, SIMCA	78
2004	Olive oil	VOCs in olive oil	SH/MS	HCA, PCA, KNN, SIMCA	79
2004	Olive oil	Discrimination between degassed olive oils and pure olive	SH/MS	KNN	80
2004	Unifloral honeys	Classification by botanical origin	SPME/MS, DH-MS	PCA, DFA	42
2005	Olive oil	Discrimination of the origin	SH/MS	PCA, LDA	81
2005	Olive oil	Control of adulteration of olive oil	SH/MS	PLS, PCR, cal. transfer	82
2006	Benzoin Gums	Discrimination from benzoin gums qualities	SH/MS	PCA, HCA	52
2007	Extra virgin olive ois	Study of combinations of MS-Sensor and UV-Vis spectrophotometer for the verification of origin of extra virgin olive oils.	SH/MS	SIMCA	81
2007	Rices	Distinguish scented and nonscented rices	SPME/MS SH/MS	SIMCA	83

Table 2.5 shows the application of a MS-Sensor to the study of quality in fruit, vegetables and derivated products such as soft drinks. The performance of MS-Sensor for tomato aroma profiling was evaluated[41], changes in tomato aroma profiles of two different cultivars were monitored during shelf life. A clear distinction between cultivars based on MS-Sensors measurements was obtained. It should be highlighted the application of an ion-trap MS-Sensor to properly differentiate between grapefruit juices that differ only in the concentration of a single component and to identify this component[33].

Table 2.5: Applications of MS-Sensor in fruit, juices and soft drink products (chronological order)

Year	Sample	Type of study	System	Data analysis	Ref.
2001	Grapefruit juice	Differentiation between juices	SH/MS(ion trap)	PCA, DFA	33
2002	Apple juice	Comparison with sensory results	SH/MS	PCA	86
2002	Fruit flavor samples	Comparison of Different Approaches to rapid screening of headspace samples	SH/GC-MS SH/Fast GC-MS, SH/MS	SIMCA	87
2003	River and tap water	Determination of six VOCs, Calibration Transfer for Solving the Signal Instability in Quantitative Headspace-Mass Spectrometry	SH/MS	PLS, calibration transfer	58
2004	Snake fruit	Study of maturation levels	SH/MS	PCA	88
2004	Tomato	Discrimination of varieties	SPME/MS	HCA, PCA	41
2004	Apple	Determination of firmness and days of shelf life	SPME/MS	PCA, PLS	26
2005	Tomato	Relate the sensory panel scores with the instrumentally measured flavour attributes.	SH/MS	PCA	84

Shelf-life determination of products is another issue in which the MS-Sensor approach has also been used. The MS-Sensor showed to be a feasible device to monitor volatile change profiles from apple fruits during shelf-life[26]. In the same order and in other work[84], eight tomato cultivars harvested at the red-ripe stage of maturity were selected and assessed certain days after harvest by means of descriptive sensory analysis performed by eight trained panellists, QMB e-nose, MS-Sensor and GC-MS. Results demonstrated the potential of the MS-Sensor to complement routine sensory analysis of tomatoes. In a further work[85], the same authors demonstrate that mass spectrometry signatures were useful to relate aroma profile and the preference map to identify the most important aroma characteristics determining consumer acceptance.

The brewery industry may also beneficiate from a MS-Sensor device as it shown in Table 2.6. Kojima et al.[89] investigated the characterization of beer aroma using MS-Sensor for quality control during brewing. Martí et al[53] shows an excellent review of MS-Sensor application to analysis of alcoholic beverages.

Table 2.6: Applications of MS-Sensor in alcoholic beverage products (chronological order)

Year	Sample	Type of study	System	Data analysis	Ref.
2002	Beer	Differentiation of 5 different German beers	SPME/MS	PCA	92
2002	Bourbons	Comparison of SH and SBSE in the detection of whiskey adulteration with a MS-Sensor	SH/MS, SBSE-MS	SIMCA, PCR	49
2003	Wine samples	Prediction of blending rate	SH/MS	PLS	93
2003	White wines	Determination of TCA in wines	SH/MS	PLS	91
2003	Cork wine stoppers	Study of volatile compounds	DH-MS	PCA	37
2003	Wines	Wine Discrimination	SH/MS	KNN, SIMCA	94
2003	Orujo oil	Determination of residual hexane	SH/MS	Univ. cal, PCR, PLS	95
2004	Wine	Differentiation of wines according origin, variety and ageing	SH/MS	PCA, SIMCA	90
2005	Alcoholic beverages	Determination of ageing time of spirits in oak Barrels	SH/MS	PLS	24
2005	Australian White wines	Discrimination according to the varietal origin	SH/MS	PCA, PLS-DA, LDA	28
2006	Riesling wines	Measure sensory attributes in commercial wines grown in Australia. (Combination of MS-Sensor and IR technologies)	SH/MS	PLS	22
2006	White wine	Discrimination between two monovarietal wines	SPME/MS	unf-PCA	96
2007	Red wine	monitor spoilage induced by *Brettanomyces* yeast in two commercial	SH/MS	PCA, SLDA	27

The wine industry has also found a wide variety of MS-Sensor applications. Within this context, a method for wine analysis using a SH/MS-Sensor system and multivariate analysis techniques was successfully applied to differentiate and classify wines according to their origin, variety and ageing[90]. With the same purpose Cozzolino et al[28] evaluated the usefulness of the same system to classify Australian white wine samples according their geographical varietals origin. Martí et al[91] analyzed several white wines using SH/MS to determine 2,4,6-tricholoranisole (TCA). This is an off-flavor that is a serious problem for

the wine industry since it is perceived as a musty, earthy and mouldy aroma that is mainly associated with the use of cork stoppers. Characterization of volatile signatures from cork wine stoppers has also been attempted through use of a MS-Sensor[37].

Finally, Table 2.7 summarizes the non-food miscellanea applications attempted to the date using a MS-Sensor approach.

Table 2.7: Non-food product miscellania

Year	Sample	Type of study	System	Data analysis	Ref.
1999	Diesel fuel	Study of origins (comparison between MS-Sensor and Gas Sensor electronic nose)	SH/MS	PCA, LDA	25
2000	Diesel fuel	Detection of perfumes in fuels	SH/MS	PCA, LDA	99
2001	Watters	Analysis of solvent sample mixtures containing analytes with similar mass spectra	P&T-MS	PLS, N-PLS, PARAFAC, Univ. reg	98
2001	Polymer samples from car interior	Comparative study between modificated MS-Sensor and dedicated one. Quality control of car odor	SH/MS	PCA	97
2003	Urine	Ánalysis of lidocaine in urine	SPME/ion trap-MS-MS	Univ. calib	55
2003	Soil	Pollution by hydrocarbons	SH/MS	CA, LDA, SIMCA	58
2004	Soil	Pollution by hydrocarbons	SH/MS	HCA, PCA, LDA	78
2004	Paper Strip	Analysis of Packaging Materials	SH/MS	SIMCA, KNN	50
2004	Gasoline	Quantification of MTBE in gasoline samples	SH/MS	Univ. cal.	100
2005	Beach sand	Quantification of pollution by hydrocarbons from crude oils	SH/MS	PLS	101
2005	Beach sand	Quantification of pollution by hydrocarbons from crude oils	SH/MS	PLS	101
2005	Bacterial cultures of microbiological pathogens	Discriminate between the different growth phases of Eschericia coli and Staphyloccocus aureus	SH/MS	RBF	102
2005	Urine samples Blood samples	Discrimination among blood with different bacterial infection	SH/MS	Sammon mapping+RBF	103
2006	Pharmaceutical products	Detection and determination of residual solvents	SH/MS	Contour plots + Standard addition quant.	104

From this table it can be derived that besides from food analysis, the MS-Sensor is a useful approach for other fields such as analysis of biofluids, environmental analysis, etc. Within the studies reviewed in this table it should be highlighted a comparative study between a modified MS-Sensor and a dedicated one for quality control of odour cars[97]. From this study it can be deduced that there is not need for a dedicated device since

modified traditional GC-MS provides results as good as the dedicated one. Another interesting study is the analysis of water samples containing similar spectra mixtures of solvents with the use of the MS-Sensor approach for environmental purposes. The most challenging part of this study was the fact that it introduced, for the first time, the idea of taking advantage of the differences in the temporal profile of the analytes in MS-Sensor approach by use of three-way algorithms for quantization purposes[98].

As it has been reviewed, the use of MS-Sensor devices in the framework of food quality assessment has been increasing showing more and more examples of different applications.

2.9 DATA ANALYSIS

An essential step in the analysis with a MS-Sensor is pattern recognition. Although the best performing programs are sophisticated and, therefore, require the operation of skilled personnel, most companies have implemented user-friendly software for basic data treatment in commercially available MS-Sensors. Anyway, a lot of uncovered subjects on data analysis remain to be solved in order to achieve the optimal MS-Sensor performance. Bellow, a review of main subjects on data analysis that are already covered is given.

The multivariate response of a MS-Sensor can be used as a "fingerprint" to characterize a wide range of volatile compound by pattern-recognition means. As illustrated in Figure 2.4 this process can be split into four sequential stages: signal preprocessing, dimensionality reduction, prediction, and validation. The initial block in the figure represents the spectra or raw data response matrix, which typically consists of an averaged mass spectrum that integrates all the volatiles in the headspace.

The first computational stage, called signal preprocessing, serves various purposes; including compensating for sensor drift, appropriate preprocessing and scaling of the data and in general to correct for instrumental and circumstantial variation, interpretation based

on multivariate analysis extracting m/z descriptive parameters and preparing the feature vector for further processing. A dimensionality reduction stage projects this initial feature vector onto a lower dimensional space in order to avoid problems associated with high-dimensional, sparse datasets. The resulting low-dimensional feature vector is then used to solve a given prediction problem, typically classification, regression, or clustering.

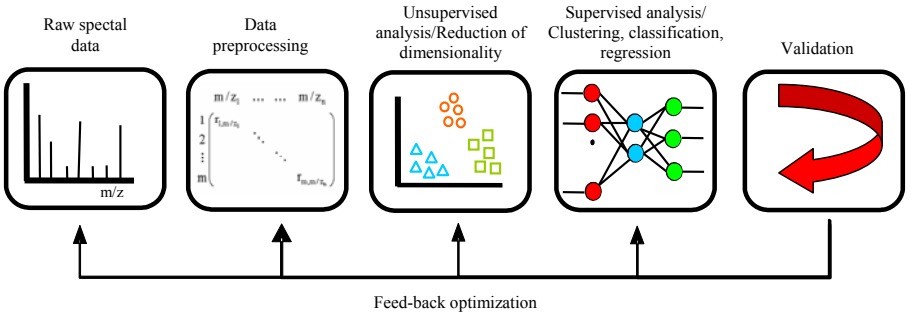

Figure 2.4: Building blocks of the pattern analysis system for a MS-Sensor[19]

Classification tasks address the problem of identifying an unknown sample as one from a set of previously learned categorized samples. In regression tasks, the goal is to predict a set of properties (e.g., concentration, quality) for an analyte, typically a complex mixture. Finally, in clustering tasks the goal is to learn the structural relationships among different odorants. A final step, sometimes overlooked, is the selection of models and parameter settings and the estimation of the true error rates for a trained model by means of validation techniques. Next sections will review the subjects attempted in each one of the different building blocks of pattern analysis. This review is divided in two different blocks according to the three-way or two-way nature of the data under analysis. It has to be emphasized that three-way data analysis has not been used to date in the MS-Sensor approach, thus there are no references un to now

2.9.1 MS-Sensor data

According to Comas[105], Sánchez and Kowalski [106, 107] established a terminology to name and classify the experimental measurements and the analytical instruments that

generate them. When a sample is analyzed, a single value can be obtained (e.g., the area of certain peak); a value over time e.g., a complete m/z monitoring over time, which gives a SIC (single ion chromatogram) or on the other way around an averaged mass spectra over time or a series of values over time (e.g., a mass spectrum over time). Mathematically, these data are arranged as a scalar, a vector or a matrix of values respectively, which we will refer to as "zero", "first" and "second" order data. This data classification is also applied to the analytical instruments that generate them and the calibration methods that use these data. In accordance to this classification, MS-Sensor can be classified as a second-order instrument producing second-order data since the response in ion counts arriving at the detector is measured as a function of retention time or scans and mass to charge ratio (m/z) as showed in Figure 2.5 Then, when a set of samples are available, they can be reshaped as a three-way array.

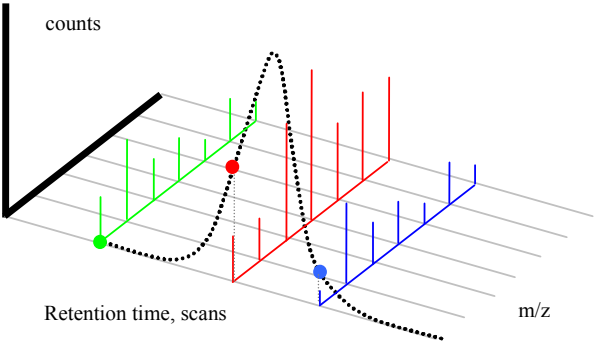

Figure 2.5: Schematic representation of a typical MS-Sensor response

Any set of data for which the elements can be arranged as:

$$x_{ijk}... \; i=1,...,I, \; j=1,...,J, \; k=1,...,K$$

where the number of indices may vary, is a multi-way array. When three dimensions are used, the data can be arranged in a three-way array, where the *ith* index refers to the rows, the *jth* index refers to the columns and the *kth* index refers to the third dimension (or tubes) (Figure 2.6).

Each dimension of this array is named way or mode and the number of levels in the mode is named the dimension of the mode[60].

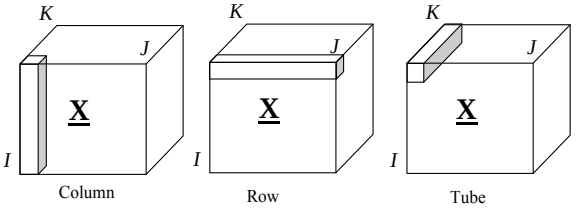

Figure 2.6: Definition of columns, rows and tubes in a three-way array

As mentioned above, in MS-Sensor, the response in ion counts arriving at the detector is measured as a function of the retention time or scans *(K)* and mass to charge ratio (m/z) *(J)*. Then, taking *I* measurements generates a three-way array of size $I \times J \times K$ where samples are in the first model, m/z in the second and time in the third mode. Each horizontal slice through the array is a $(J \times K)$ matrix representing the time history for all m/z variables of a particular batch or sample. Each vertical slice made parallel to the front face of the cube is a $(I \times J)$ matrix representing the intensity values of all the m/z variables or channels scanned in all the samples taken at a common time. A vertical slice made parallel to the side of the cube (the time axis) would represent a $(I \times K)$ matrix of all the time histories of a single m/z variable for all samples (Figure 2.7).

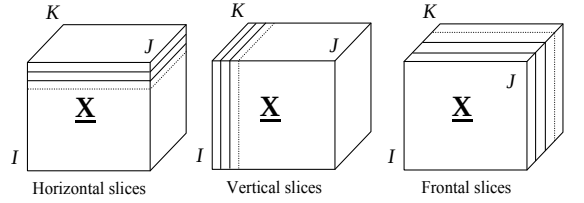

Figure 2.7: Partitioning of a three-way array in slices (two-way arrays)

Nevertheless, MS-Sensor data has never been considered to date as their three-way nature indicates. However, what is always employed is a two-way matrix corresponding to an averaged mass spectrum along the third mode or time dimension. In the MS-Sensor approach chromatographic separation is avoided using either an isothermal elevated temperature program or an uncoated retention gap replacing the traditional chromatographic column.

Figure 2.8 illustrates a typical three-way MS-Sensor data landscape and its corresponding three-way array and how this second order data is reshaped to its two-way matrix counterpart by averaging m/z axis along the time dimension. Information contained in the peak shape, overlapping peaks, or peak shoulders may be useful but is mostly ignored when using the reshaped two-way data matrix instead of the raw three-way array.

Actually, one of the main goals of this thesis is to evaluate weather this information which results invariably lost by averaging time dimension could result in an improvement of the MS-Sensor performance when it is considered in the models for pattern recognition.

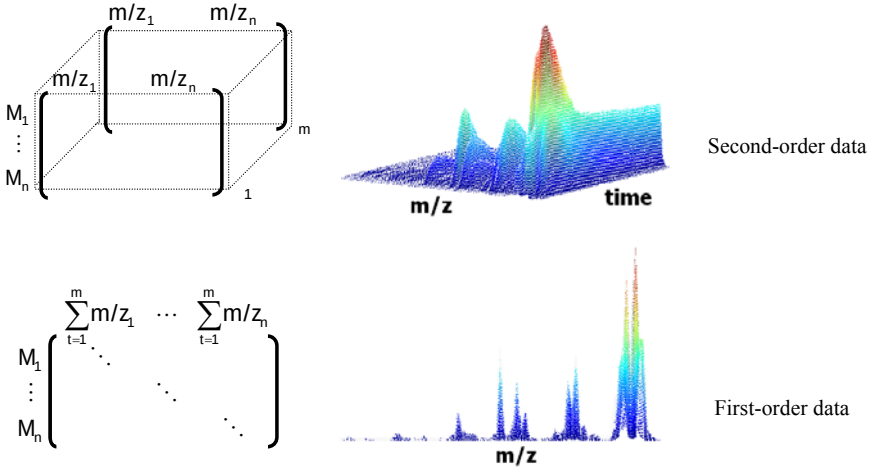

Figure 2.8: Representation of different orders of MS-Sensor data

As it will be further explained there are several ways to consider this time dimension. The first one is by considering the intrinsic three-way nature of the data using the so-called multi-way algorithms. Another technique traditionally used when handling three-way data is by unfolding this three-way array to a two-way matrix. Unfolding (or matricization) involves rearranging a three-way array to form a matrix as it will be shown in next chapter in Section 3.4.1.2.

2.9.2 Two-way data analysis and multi-way data analysis

Data arranged in a two-way matrix structure is used in standard multivariate routines commonly employed to handle MS-Sensor data. Two-way methods, such as PCA can be used for the exploratory data analysis of this type of data. When the relationship between m/z ratio variables and, for instance, concentrations is wanted, techniques such as PLS are commonly employed in MS-Sensor[91]. Furthermore, a PLS based algorithm, the so-called PLS-DA is used when attempting a classification.

In other cases, it might be necessary a description in more directions and a two-way arrangement of the data do not suffices. One example is be the excitation/ emission fluorescence spectra of a set of samples. Each data element can then be indexed by the sample number, emission wavelength, and excitation wavelength, which implies a three-way matrix. Another example includes GC-MS, where the response in ion counts arriving at the detector is measured as a function of GC retention time and m/z. In parallelism with GC-MS data, it could be considered a three-way arrangement of MS-Sensor data and its multi-way analysis. To date, there is only one published paper which consideres the three-way data nature of the MS-Sensor data[98]. Multi-way analysis is the natural extension of multivariate analysis, when data are arranged in three- or higher order arrays. This in itself already provides a justification for multi-way methods. Thus, multivariate methods lead to its extended versions in multi-way analysis; e.g., PCA leads to Parallel Factor Analysis (PARAFAC) or Tucker models, the regression method Partial Least Squares (PLS) leads to multilinear N-PLS, etc.

Multi-way data analysis, originated in psychometrics literature back in the sixties[108]. Nowadays multi-way methods have been applied to a wide variety of problems. Some examples are the decomposition of fluorescence-spectroscopy data of poly-aromatic hydrocarbons[109], the prediction of amino acids concentrations in sugar with fluorescence spectroscopy[110], data exploration of food analysis with GC-MS[111] and sensory data[112] among others. With increasing number of application areas, multi-way data analysis has become more and more popular. A review on multi-way data analysis is not included since

this is not the aim of this thesis. To date there is not any published work about MS-Sensor data analysis through the use of multi-way methods. Readers are referred instead to two reference books where all subjects in multi-way data analysis are fully covered (models, algorithms, constraints, validation, applications, etc) [60, 113].

2.9.3 Data preprocessing

Appropriate preprocessing and scaling of the data to correct for instrumental and circumstantial variation is crucial for reliable data interpretation based on multivariate analysis. The whole mathematical treatment performed on raw GC-MS data prior multivariate analysis is going to be known as *Preprocessing*. The philosophy of the data pre-processing step is to reduce the variation in the chemical data not related to the chemical composition such as analytical variability and concentration effects. Speedy and semi-automazitation with limited human intervention of the mathematical tools involved in preprocessing are highly required[114]. The main preprocessing techniques that have a large influence on analysis results are: Background substraction, alignment and normalization before any multivariate analysis.

2.9.3.1 Alignment

Many of the chemometric techniques available for multi-way modelling rely on trillinearity, a prerequisite seldom met due to the variations in the chromatographic conditions affecting peak position and peak width. One way to tackle this problem is to pre-process the data by some kind of time alignment procedure [115]. Several methods have been proposed for alignment of second order data where the spectral information is used to guide the alignment procedure [115,116,117,118]. Nevertheless, all these approaches require user intervention to set-up the optimal parameters for the alignment algorithms. This is a serious drawback from a MS-Sensor point of view because it does not allow automatic data processing. Recently, a fully automated algorithm called RAFFT (Recursive Alignment Fast Fourier Transform) has been presented by Jason W. H. Wong et al.[119]. This algorithm

makes use of the Fast Fourier transform for rapid computation of the cross-correlation function that enables alignments between a target sample and samples to be optimized. It is based on spectra segmentation models and is developed to offset the need for operating parameters. Minimal segment size is determined automatically by recursive alignment from the full spectrum (global scale) to progressively smaller segments (local scale) until no further alignment is required.

2.9.3.2 Dimensionality reduction

The matrix that results from the preprocessing stage is sometimes not suitable to be processed by a subsequent module due its high-dimensionality and redundancy. Problems with high-dimensional data, known as the "curse of dimensionality" in statistical pattern recognition, imply that the number of training examples must grow exponentially with the number of features in order to build an accurate model. Since only a limited number of examples are typically available, there is an optimal number of feature dimensions beyond which the performance of the pattern analysis model starts to degrade. The notion of the curse of dimensionality was introduced by Bellman[120] as a result of studies in adaptive control processes. The problem stems from the number of data points needed to adequately represent a data set with a high number of features; it is quite possible that within high dimensional data, clusters exist in separate sub-spaces. All classifiers can suffer from the curse[121]. The only practical way to beat the curse is to apply prior knowledge on the data, or to carefully select the minimum number of features to adequately represent the problem.

The problem of redundancy, also referred to as collinearity in chemometrics and statistics, is particularly significant in MS-Sensors instruments due to the cross-selectivity of the m/z variables. When two or more feature dimensions are collinear, the covariance matrix of the entire dataset becomes singular and, therefore, noninvertible, which leads to numerical problems in various statistical techniques (e.g., quadratic classifiers and ordinary least squares).

For these two reasons, a dimensionality reduction stage is required in most cases[19]. Blum and Langley[122] classify feature selection techniques into three basic approaches. In the first, known as the embedded approach, features are added or removed in response to prediction errors of a simple embedded classifier. The second are filter methods and work independently to remove features without knowing the effect on the classification algorithm; typical linear methods used are PCA or LDA. Nevertheless the new computed factors or variables resulting from PCA or LDA variable reduction could be affected by noise. There are different sources of background noise in GC/MS systems. For example, spectral background noise is associated with contaminants present in the ionisation chamber (such as ambient air and contaminants present in the carrier gas). To a higher extent, spectral background noise is due to the counting principle and inherent noise associated with the ion multiplier. Additionally, a baseline drift may appear due to co-eluting compounds, septa and temperature induced column bleed. The selection of an optimal subset of components or factors is not necessarily straightforward because the magnitude of an eigenvalue is not always a measure of its significance for the calibration. Furthermore, unlike m/z ratios, factors have no direct chemical meaning. The third are wrapper methods and evaluate candidate feature sets using a classification algorithm on the training data. The feature subset selection algorithm conducts a search for a good subset using the classifier as part of the evaluation function[15]. In the case of MS-Sensor, selecting from the full spectrum of mass to charge ratios could be challenging because there is considerable overlapping among the spectra and distinctive features can be almost imperceptible. Furthermore, spectra are affected by noise. However, methods based on the selection of m/z ratios are interesting because the variables chosen carry direct relevant chemical information.

To date there has not been published literature dedicated to investigate among the different available methodologies for performing feature selection in an automated way on MS-Sensors data. Current variable selection methodologies rely on choosing directly among m/z factors that holds relevant information for the applications envisaged[123]. That implies to hold a previous knowledge of chemical composition of the headspace being analyzed.

2.9.4 Unsupervised pattern recognition

Only a short description of some of the most frequently used pattern recognition methods for exploratory data analysis is given here. Readers are referred to specialised literature for more information. The models commonly employed to attempt unsupervised pattern recognition are either PCA for two-way data analysis or its three-way counterpart called PARAFAC.

2.9.4.1 PCA

According to Esbensen[124], PCA constitutes the most basic "work horse" of all multivariate data analysis. PCA is a linear combinatorial method which reduces the complexity of the data-set, from the initial n-dimensional space (n m/z pseudosensors) to a few dimensions. The inherent structure of the data-set is preserved while its resulting variance is maximised. In other words, data points will be scaled along new dimensions, linear combinations of the initial dimensions. The magnitudes of the coefficients, in the resulting linear combinations, give an indication of the relative importance of the initial dimensions in the data structure. PCA is performed with no information on the classification of samples. It is based solely on the variance of the data-set. Thus it is an unsupervised method. It should be emphasized that PCA is just a visualization and variable reduction tool. It can not be used as a classifier.

The underlying premise in PCA is that the raw data in X can be decomposed into eigenvectors and associated eigenvalues. Of the several methods available to decompose X, one of the most common is through singular value decomposition (SVD). As shown in Figure 2.9 PCA model makes a bilinear decomposition of the X matrix into a score matrix (**T**) (eigenvectors) and a loading matrix (**P**) (eigenvalues), which describe the original data in a more condensed way. The residuals (i.e., the difference between the original and the reconstructed matrix with the calculated model) are collected in the E matrix. Thus, in PCA each component is the outer product of two vectors (scores pertaining to samples and loadings pertaining to variables). The model is linear in the scores for fixed loadings and

vice versa and therefore bilinear. If a given data set can be modelled by an R-component PCA model, then the $I \times J$ matrix **X** can be approximated as $\mathbf{TP}^{\mathbf{T}}$ where T is the score matrix ($I \times R$) and **P** the loading matrix ($J \times R$).

As already mentioned, the aim of PCA is to retain the main information contained in the original variables in a smaller number of variables, called principal components (PC), which describe the main variations in the data. These PCs are linear combinations of the original variables. Some properties of PC are that they are orthogonal (i.e., uncorrelated to each other), hierarchical (i.e., the first PC retains the maximum information of the data, the second PC retains the maximum information that is not included in the first one, and so on), and are calculated sequentially (i.e., the F-1 component model is a subset of the F component model).

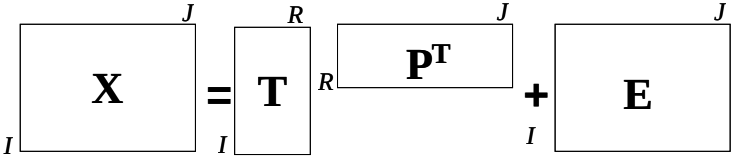

Figure 2.9: Representation of an R component PCA model of the data matrix X

Mathematically, PCA decompositions can be written as Eq 2.1

$$X_{ij} = \sum_{r=1}^{R} a_{ir} b_{jr} + e_{jk} \qquad \text{[Eq. 2.1]}$$

where elements a_{if} , b_{jf} refers to the scores and loading matrices **A** and **B** and where the i, j refer to the dimensions of the matrix and f is the number of factors of the A PCA model.

2.9.4.2 PARAFAC

PARAFAC is a decomposition method for multi-way data that works directly on the multi-way structure. This decomposition is made into triads or trilinear components.

Each component consists of three loading vectors that are treated mathematically in the same way. However, the vector associated with the sample mode is often named the score vector. A PARAFAC model of a three-way array is therefore given by three loading matrices, **A** , **B** and **C**, with elements a_{if} , b_{jf} and c_{kf} , where the *i*, *j* and *k* indices refer to the dimensions of the array and *f* is the number of factors of the model. Figure 2.10 is a graphical representation of the PARAFAC decomposition. The cube E contains the residuals. Mathematically, the PARAFAC model is for *r* components is expressed as Eq. 2.2:

$$X_{ikj} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk} \qquad \text{[Eq. 2.2]}$$

Two of the most important properties of PARAFAC are uniqueness and the fact that PARAFAC is a nested model. Unlike bilinear methods such as PCA, PARAFAC provides unique solutions. This means that, apart from trivial variations of scale and column order, no restrictions are needed to estimate the model. Another difference with respect to PCA is that PARAFAC is a non-nested model (i.e., it cannot be calculated sequentially). This means that the F-1 factor model is not a subset of the F factor model. There are several techniques for determining the number of factors in a PARAFAC model, like residual analysis, visual appearance of loadings, number of iterations of the algorithm, core consistency, etc. For a more detailed view of the PARAFAC model please refer to Bro[110].

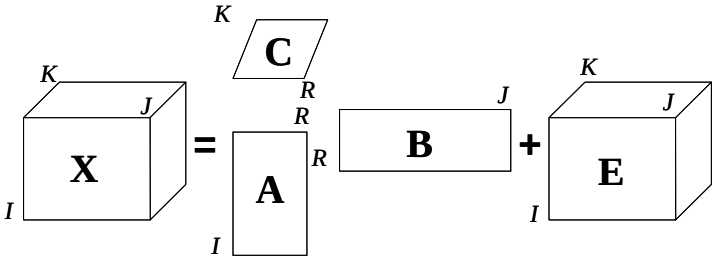

Figure 2.10: Representation of an R component PARAFAC model of the data array X

2.9.5 Regression models

Regression problems constitute a challenging domain for MS-Sensors instruments. The goal of regression is to establish a predictive model from a set of independent variables (e.g., MS-Sensor responses) to another set of continuous dependent variables (properties of the sample under analysis i.e., concentrations. Pattern classification could, therefore, be also treated as a regression[19]. The gold standard model for regression purposes is PLS (Partial Least Squares).

2.9.5.1 Partial Least Squares (PLS)

Partial least squares (PLS) regression is a method for building regression models between independent (X) and dependent (**Y**) variables. It is based on a bilinear decomposition of the calibration matrix and establishing a relationship (regression) between **X** and **Y**[125]. The decomposition is achieved by maximizing the covariance between **X** and **Y**. Thus, an optimal regression is achieved with respect to the prediction error. Factors of **X** and **Y** are computed simultaneously by successively substituting the scores of the **X** matrix (called **t**) by the scores of the **Y** matrix (called **u**), and vice versa, up to convergence. The PLS decomposition can be expressed in matricial form as in Eq. 2.3 and 2.4:

$$\mathbf{X} = \sum_{f=1}^{F} \mathbf{T}\mathbf{W}^{\mathbf{T}} + \mathbf{E} \qquad \text{[Eq. 2.3]}$$

$$\mathbf{Y} = \sum_{f=1}^{F} \mathbf{U}\mathbf{Q}^{\mathbf{T}} + \mathbf{F} \qquad \text{[Eq. 2.4]}$$

T and $\mathbf{W}^{\mathbf{T}}$ are the score and loading matrices of **X**, respectively, while **U** and $\mathbf{Q}^{\mathbf{T}}$ are the score and loading matrices of Y. The columns of **W** are often called the loading weights because they indicate how the t -scores are to be computed from **X** to obtain an orthogonal decomposition. **E** and **F** are the corresponding error matrices.

2.9.5.2 Multilinear Partial Least Squares (N-PLS)

The basis of multi-linear or multi-way partial least squares (N-PLS) regression is the same as that of bilinear PLS but extended to multi-way data. For three-way data, the method is named tri-PLS or N-PLS. This refers to the decomposition of a cube **X** into a set of triads. A triad is the trilinear equivalent of a bilinear factor. It consists of one score vector **t** and two weight vectors - one in the second order, called $\mathbf{w}^J$, and one in the third order, called $\mathbf{w}^K$. The model of **X** can be expressed by Eq 2.5. The term e_{ijk} is the error[125].

$$x_{ikj} = t_i w_j^J w_k^K + e_{ijk} \qquad \text{[Eq. 2.5]}$$

Figure 2.11 represents a two components N-PLS decomposition.

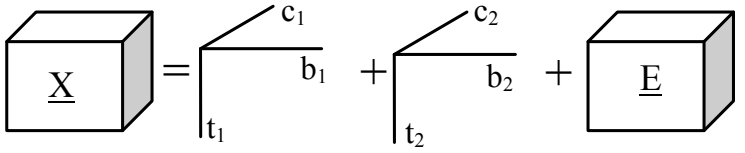

Figure 2.11: Two-component N-PLS decomposition of X array

A more comprehensive explanation of N-PLS can be found elsewhere in Bro[125].

2.9.6 Classifier models

The task of a classifier is to use the feature vector provided by the feature extractor to assign the object it represents to a category. The degree of difficulty of the classification problem depends on the variability in the feature values for objects in the same category relative to the difference between feature values for objects in different categories. The variability in the feature values for objects in the same category may be due to complexity, and may be due to noise. The simplest measure of classifier performance is the classification error rate, the percentage of patterns that are assigned to the wrong category.

The final stage of a pattern classification system is usually making a decision on the class assignment to the input patterns based on measurements taken from the selected features[15].

According to Schaffer et al[126] the ideal classification system should comply all the following requirements:

1) High accuracy: there should be as few misclassifications as possible.
2) Fast: for real-time analysis, the algorithm must be able to produce a classification with the minimum delay.
3) Simple to train: in many applications the database of training patterns will be updated periodically and the classifier retrained. This procedure should be quick and simple to perform.
4) Low memory requirements: when thinking in small portable and systems that may be used as handheld devices the classifier would need to consume only a few resources.
5) Robust to outliers – in uncontrolled environments the algorithm must be able to reduce the potential for misclassifications by being able to differentiate between a pattern on which it was trained to recognise and one that it was not. The assumption is that the system has been trained on all relevant patterns so any ambiguous patterns should be recognised as such.
6) Produce a measure of uncertainty: for many applications the algorithm needs to produce a measure of the match level of the classification, or a statistical measure concerning the certainty of the classification.

In this section only the used classifiers models in this thesis are going to be subsequently described.

2.9.6.1 Linear Discriminant Analysis (LDA):

Fisher"s LDA or canonical variate analysis is a discrimination method that provides classification rules for unknown samples by maximizing the differences between

the classes. This is because it seeks directions in multivariate space that separate the groups as much as possible and uses information along these directions in simple scatter plots. Linear discriminant analysis (LDA) finds a linear discriminant function (LDF) which is a linear combination of the original variables, such that the ratio of the between-class scatter and the within-class scatter is maximized. A more comprehensive explanation about the algorithms used in this classifier model can be found elsewhere[17, 19]

2.9.6.2 Discriminant Partial Least Squares (PLS-DA)

Although PLS was not inherently designed for problems of classification and discrimination, it is routinely employed for those purposes[127] When the property to be predicted (**Y**) is membership to one class, the method is called discriminant PLS and can be used for classification. If there are only two classes, **y** is a vector of ones and zeros depending on the class membership of the sample in **X** (i.e., ones indicate membership to the class and zeros indicate nonmembership). For more than two classes, there are as many **y** column vectors as classes. Each vector has ones for samples belonging to the class and zeros for the other samples. In this case, PLS regression must be done separately between **X** and each **y** column.

2.9.6.3 Neural network: fuzzy ARTMAP

Artificial neural networks (ANN) based classifiers offer a powerful non-linear mapping capability, many types have been employed for *"electronic nose"* data classification including Kohonen networks, learning vector quantization (LVQ) and its variations, MLP with variants of the backpropagation (BP) algorithm and ART(Adaptive Resonance Theory). Adaptive resonance theory-based methods offer significant advantages over the other ANN. They require less training data, because the number of adjustable parameters does not grow as rapidly. ART networks train significantly faster than LVQ networks. While the training time difference is acceptable when one wishes to train using only a few sets of features, if a more extensive feature selection requiring many trainings to compare sets of features is desired, the computational time is significantly higher for LVQ

than for ART networks. ART was developed by Grossberg to address what he termed the stabilityplasticity dilemma. Briefly, other algorithms such as LVQ will adjust their weight vectors when presented with outlier data to accommodate the new data point, which may result in the degradation of the ability of the network to identify existing classes. In analogy to LVQ, ART networks compare the new observation to the winning neuron, and if it is not similar enough, the network architecture is able to adapt. Many variants exist, including ART1, ART2, fuzzy ART, ARTmap, and Fuzzy ARTMAP. Just as the other ANN methods, these ANN also require cross-validation and prediction to prevent overtraining (in supervised methods) and for model validation, respectively[17].

2.9.6.4 PARAFAC-MLR-DA

PARAFAC can be used for quantitative purposes because the score matrix contains information about the differences of composition between the samples. Then it is also possible to use PARAFAC for prediction purposes by simply calculating the scores using suitable trilinear model (PARAFAC-like) of $\underline{\mathbf{X}}$ but such that the scores are predictive of **Y**. Furthermore, when the property to be predicted (**Y**) is membership to one class, it becomes a discriminant model and hence it can be used for classification. The main advantage is that once PARAFAC has been fit to the data, it is ensured that the Y-class distribution is based on the truly underlying phemomena modelled[113]. For quantification purposes Multiple Linear Regression (MLR) is used to build a least squares calibration model from the PARAFAC sample scores. A graphical explanation of PARAFAC-MLR-DA model is given in Figure 2.12

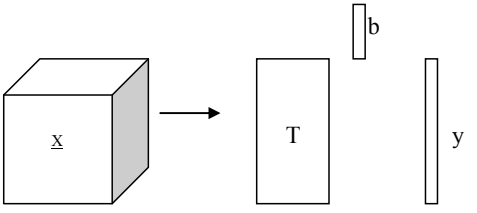

Figure 2.12: Graphical representation of PARAFAC-MLR model

2.9.7 Testing and validation

Cross-validation is a technique to compensate for an optimistic apparent error rate caused by training and testing on the same small data set. The apparent error rate is the percent of misclassified observations. The cross-validation routine omits each observation, one or more at a time (leave-one-out or leave-n-out) and recalculates the classification function using the remaining data, and then classifying the omitted observation. Splitting the data into two sets is used to calculate a more realistic error rate on larger data sets. One set to create the discriminant function and the other set to act as a validation set

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 3 | Experimental and Data analysis

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

3. EXPERIMENTAL AND DATA ANALYSIS

In this chapter, the measurements system layout and the experimental set-up are not going to be described in detail since they differ depending on the application. Furthermore, it is considered that they are fully described in the papers. Instead, this chapter tries to give a general overview to the main methodological steps performed for turning a non-dedicated GC-MS system into a MS-Sensor. The methodologies used for handling volatiles and the steps for optimizing parameters affecting these methodologies are described. This chapter also deals on the methods developed to get the data in a suitable way for further pattern recognition. A general description of the methodological steps performed when attempting data analysis is also given.

3.1 HANDLING VOLATILES

The volatile extraction methodology procedure determines the measurable sample characteristics and is application specific. In the different experiments conduced throughout this thesis it has been used either the SH or SPME as technique s for sampling volatiles.

3.1.1 Static Headspace (SH)

SH can be performed either manually using chromatographic syringes or automatically using an automatic headspace autosampler. Because of the poor repeatability of manual headspace injection, it is recommended that an automatic SH sampler be used. The sample uptake requires the control of different parameters with temperature, pressure, volume and time being the most critical ones. SH sampling can be automated with programmable sample conditioning parameters; in this thesis it has been done coupling a commercial SH HP-7694 (Agilent Technologies) to the inlet of MS-Sensor. It consists of two steps. In the first step, the sample is placed in a vial. The closed vial is then thermostated and, if necessary, shaken for a defined time until equilibrium between phases (solid/gas or liquid/gas) is reached. In the second step, the vial is pressurized and subsequently vented so that an aliquot of the vial"s headspace is introduced into the carrier gas stream and transferred to analysis (to the split-splitless injection port of MS-Sensor). This is performed over a pressure/loop system. In the pressure/loop system the vial is opened toward a sample loop and equilibrated against ambient pressure. Then, the loop is flushed by carrier gas transferring the sample to the MS-Sensor. With a six-port valve controlling the transfer, the carrier gas flows continuously resulting in only a very small pressure peak. The sample volume is defined by the loop geometry. Figure 3.1 shows a schematic set-up of a pressure/loop SH sampler with the centrally heated six-port valve, which introduces an aliquot of the headspace into the carrier gas flow.

With the concentrations of the analyte $\boldsymbol{C_i}$ in the gas phase $[\boldsymbol{C_i}]_g$ and liquid/solid sample phase $[\boldsymbol{C_i}]_s$, the partition coefficient **K** is given by Eq. 3.1

$$K = \frac{[C_i]_g}{[C_i]_s} \qquad \text{[Eq.3.1]}$$

The performance of the SH is limited by this partition coefficient, the vial volume determining the maximum sample amount, and the matrix. This partition coefficient can be modified changing the operational parameters of which temperature, equilibration times and matrix are the most important and must be optimized prior to any application.

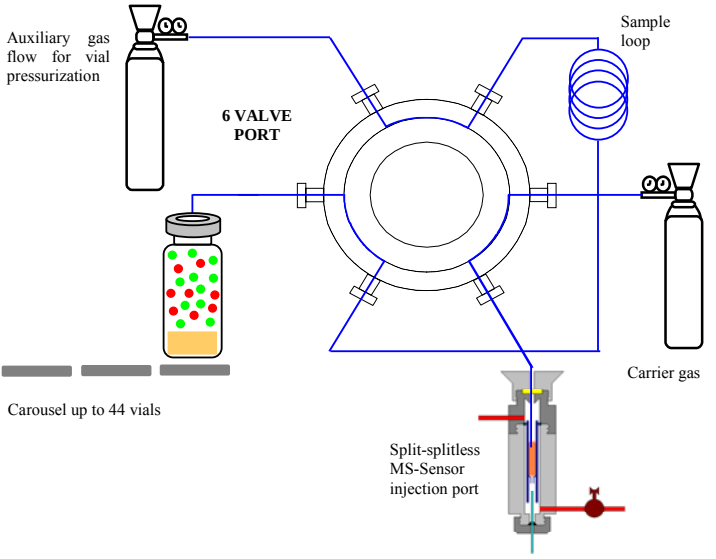

Figure 3.1: Schematic set-up of a pressure/loop SH sampler with the centrally heated six-port valve

In a resolved chromatography, the area of the signal peak A_i for component i is proportional to the concentration in the gas phase $[C_i]_g$ and proportional to the original concentration in the sample (Figure 3.2).

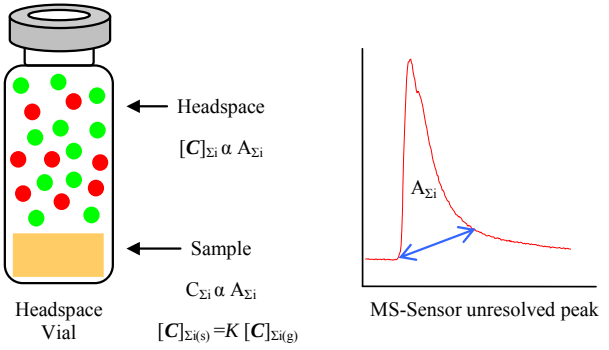

Figure 3.2: Scheme of the signal deriving from sample concentration

In case of MS-Sensor it is going to be considered that the area of the unresolved peak which comprises the contribution of each one of the analytes present in the headspace

is proportional to the concentration of analytes in the original food matrix. The SH parameters has been set-up in all the applications in such way that this area is maximized.

3.1.2 Optimization of SH parameters: oven temperature, vial equilibration time

This section is devoted to describe main experiments carried on in order to set-up the optimal SH parameters influencing its sensitivity namely, oven sample temperature, and vial equilibration time. As an example here bellow are described the experiment performed in order to select the optimal SH parameters in order to distinguish among different species of in-vitro growing fungal cultures on 20-mL headspace vials measured using a SH/MS-Sensor configuration.

To select the optimal temperature, three oven temperatures, 50ºC, 80ºC, and 100°C, were tested. Equilibration time was fixed at 5 min in the three different batches. A preliminary analysis of the results on increasing oven temperature in SH was performed by plotting PCA scores at 50, 80, and 100°C separately. At 50ºC and 80°C those plots did not show any clustering, and samples with fungal contamination and blank vials overlapped. At 100°C, inoculated samples clustered together, clearly separated from blank vials as it shows Figure 3.3. Setting a headspace oven temperature of 100°C permitted the extraction of a larger quantity of volatiles, which enhanced the sensitivity of the system and allows achieving a better discrimination than at 50ºC or 80°C.

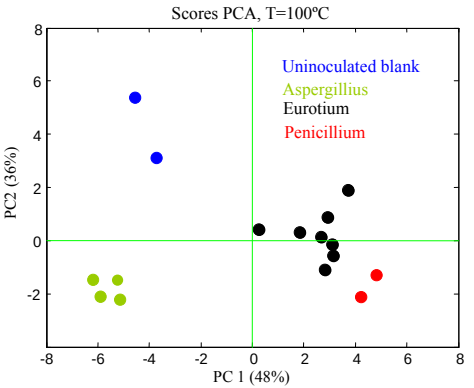

Figure 3.3: PCA scores plot of SH extracts of fungal cultures measured at T=100ºC, t=5min

When oven temperature is set up to 80ºC, most of organic material of the sample (fungal media and fungal biomass) will decompose. Then, volatiles reaching the mass detector will be an extract of all the decomposition products derived from different oxidation processes. That is not at all a representative extract of volatile metabolites present on the headspace. Working at these temperatures may lead to irreproducible injections since the process involving oxidations and thermal decomposition reactions of headspace are extremely difficult to control. Furthermore, it could introduce noise and artefacts in further modelling of data leading to a decrease of performance of MS-Sensor. Therefore it was attempted a second experiment where temperature of vial equilibration was set-up constant at 50°C and only equilibration time was varied.

Equilibration time was increased each 10 min from 10 to 60 min. Injections of different vials from the same fungal culture were used and the total area under the whole MS-Sensor peak was accounted.

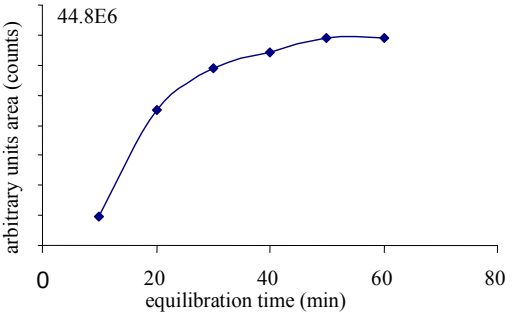

Figure 3.4: Evolution of total area under MS-Sensor curve with equilibration time

Figure 3.4 shows how concentration of volatiles promoting to the headspace (in terms of arbitrary units area) follows a linear dependence with equilibration time until a point where it reaches a wide plateau. At this point it is considered that volatiles reached the equilibrium and achieve their maximum concentration in the headspace, while the composition of volatile patterns remains stable. It is precisely at this point where headspace extracts injections would be reproducible.

According to these results it was decided to increase equilibration time to 50 min in order to ensure maximum concentration of volatile compounds on the headspace and reproducible headspace injections. A second set of measurements were repeated but this time using 50 min of equilibration time and 50ºC of oven temperature. Figure 3.5 shows PCA scores plot for this second set of measurements and again, blank vials are clustering well apart from inoculated vials.

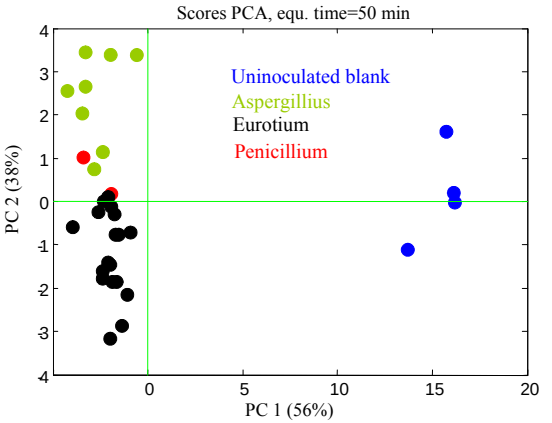

Figure 3.5: PCA scores plot of SH extracts of fungal cultures measured at T=50ºC, t=50min

Despite the results stated were considered quite acceptable, headspace conditions remained little bit harsh. Considering a real MS-Sensor application, an equilibration time of 50 minutes is not feasible because it do not allow fast response in few minutes. In the other hand, it was considered that overheating vials to 100ºC allows to a classification of samples biased towards the oxidized metabolites instead of the real headspace extract. Anyway, the analyst should take a compromise between these two parameters when working with SH. Usually, when working with MS-Sensor, the main aim is the classification of samples into different groups. In such case it is important to maximise differences among this groups, even if samples are slightly denatured during the analysis (by temperature, changes in pH, etc.) as long as the same treatment is used for all samples.

3.1.3 Solid Phase Microextraction (SPME)

One of the most important limitations of MS-Sensor based on SH is the lack of sensitivity to some volatile and rather semivolatile compounds. The use of a

preconcentration techniques such as SPME instead of the non-preconcentrated SH commonly employed with such MS-Sensors systems could cross this drawback. SPME is an innovative, solvent free technology that is fast, economical, and versatile. SPME is a fiber coated with a liquid (polymer), a solid (sorbent), or a combination of both. Sampling SPME procedure is a two step methodology as detailed in Figure 3.6.

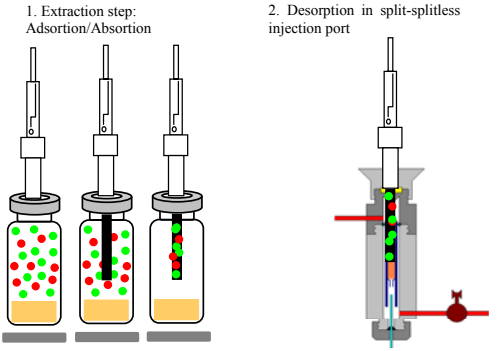

Figure 3.6: Two step headspace SPME sampling procedure

In a first step, the so-called extraction (adsorption or absorption) step, the fiber is exposed to the headspace of the matrix to be analysed. The analytes migrate to the fiber by absorption in the case of liquid coatings or adsorbing in the case of solid coatings. With SPME, the amount of analyte removed by the fiber is proportional to the concentration of the compound in the sample. This is true when the fiber and the sample reach equilibrium. Extraction time is critical for the sample to establish equilibrium with the SPME fiber coating. Extractions typically take 15-20 minutes, it will depend on the size of the compounds, fiber coating, type of extraction used and sample concentration. It can be shortened when analyzing small compounds (<150 MW); using thinner, absorbent type fiber coatings; or working with high concentration samples (high ppb or ppm range). Once the volatile has been trapped on the fiber, the second step is the thermal desorption of these analytes on the split-splitless injection port of the MS-Sensor.

3.1.4 Optimization of SPME parameters: fiber coating and equilibration time

As well as in case of SH, there are also different parameters that can affect SPME performance such as the coating of the fiber, the time and temperature of absorption and

desorption steps, modification of pH, addition of salt to the sample and sample volume. The modification of each one of these parameters should be optimized in order to increase the amount of analyte extracted by the fiber.

The selection of the fiber coating is an extremely important issue and it will definitely influence further analysis and its success. Usually, analyte size, concentration levels and detection limits are the three issues to take in to consideration when selecting the coating of the fiber. Basically there are two different types of fiber, adsorbent and absorbent. Absorbent type fiber extract by the partitioning of analytes into a "liquid-like" phase. It is somewhat like a sponge. The analytes migrate in and out of the coating. The ability of the coating to retain and release the analyte is dependent primarily on the thickness of the coating and the size of the analyte. The polarity of the fiber coating may enhance the attraction of an analyte to that particular coating, but it is the thickness of the fiber that retains the analytes. There is virtually no competition between analytes. It is basically how fast the analytes migrate in and out of the phase. The thicker film coatings have high sample capacity. Adsorbent type fibers extract analytes by physically interacting with them. Adsorbents are generally solids that contain pores or high surface areas. The extraction can be accomplished by trapping the analytes in internal pores. These micro- and meso- pores are ideal for trapping small and midsized analytes and usually retain the analytes until either energy is applied or they are displaced by a solvent. Macropores, primarily on the surface of the material, can also trap larger analytes, but generally retain them through hydrogen bonding or Van der Waals interactions. Because there are a limited number of sites, the analytes can compete. This can result in reduced capacity and/or displacement of analytes with low distribution constants by those with higher distribution constants[128]. In MS-Sensor, there is not a target compound or a targeted family compounds that is wanted to be extracted as in any other analytical method, what it is rather sought is to take the most representative volatile extract of headspace. Instead of a targeted profiling there is a need for a sampling measurement that covers analytes with a wide range of volatilities, polarities and that holds different functional groups. Faced with this complexity, selection of the fiber is not straightforward and often there is certain difficulty in to determine which type of fiber to use. Analyte size, concentration levels, and detection

limits are factors that must all be taken into consideration when selecting fibers. There is not going to be one fiber that will do all the analytes at trace levels to high concentrations. Anyway, there exists some general indications to follow when selecting fibers[128].

Currently a wide variety of coatings with different film thickness for different applications have become commercially available from Supelco. Among these ones, one of the adsorbents type fiber such as the Carboxen-PDMS seems to be the best choice since it is the best for extracting small analytes at trace levels (MW<90). Volatiles are going to be considered to lie within this molecular weight range, nevertheless, the so-called semi-volatiles molecules are out of this rage). In a recent study [129], six types of fibers were tested and the main extraction parameters were studied. Within the fibers tested, the Carboxen-PDMS fiber was the most appropriate in achieving a complete profile of cheese volatile compounds. However, other fibers resulted more appropriate when selective monitoring of substances from certain group was required. The micropores of the Carboxen-PDMS fiber make it an overwhelming choice for the extraction of small analytes at trace levels. Carboxen containing fiber coatings retain more small analytes better than the DVB containing coatings which retain more than the liquid absorbent fiber coatings. Moreover, this fiber exhibit a bipolar behaviour. Anyway it seem that fiber polarity has little effect on the extraction of small analytes but however fiber polarity has a great effect on the extraction of larger, polar analytes. Concerning the linearity range. Carboxen-PDMS has good linearity at trace analyte concentration levels but saturates at high levels with little displacement[66]. According to that, a fiber based on Carboxen seems to be the best choice to couple to a MS-Sensor. Several chromatographic runs were carried on using different fiber coatings on replicates of the same fungal culture as showed in Figure 3.7 which seems to support literature evidences on Carboxen based fibers.

In general, selection of the fiber has done taking in to account the biggest sensitivity. Equilibration time has been determined according a compromise between a reasonable time for manual SPME extraction and the time which maximises the total peak area as in case of SH. Usually a good compromise value is to set 20 minutes extraction time. Desorption time and temperature has been always set-up according the

manufacturer"s recommendation. In general terms, extractions conditions should be selected depending on the goals of the study. Thus, previous study on SPME optimization conditions would be always desirable.

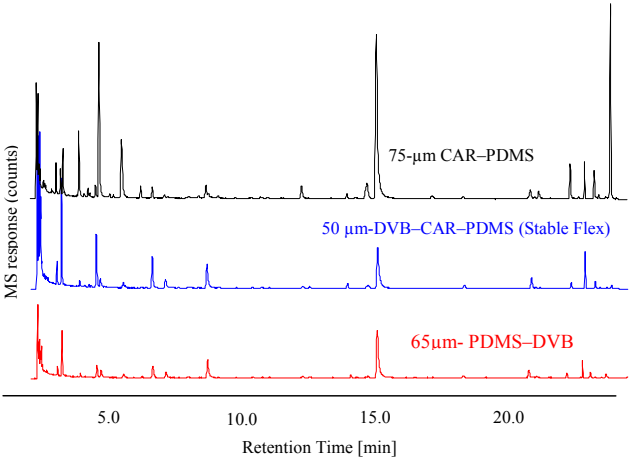

Figure 3.7: Comparison of chromatograms obtained for tested SPME fibres

3.2 TURNING A CLASSICAL NON-DEDICATED GC-MS INTO A MS-SENSOR DEVICE

In order to work with MS-Sensor devices, two different approaches are feasible, using a dedicated MS-Sensor device or upgrading a GC-MS and coupling the appropriate data processing software. It is not mandatory to have a dedicated device and instead, when a GC-MS instrument is available, it can be converted to a MS-Sensor in a straightforward way. It is just necessary to use either an uncoated deactivated retention gap acting as a transferline or an analytical capillary chromatographic column and strong chromatographic conditions (high temperature and high carrier gas flow).

The transfer line ensures coupling of the extraction injection module to the mass spectrometer. As it has been already mentioned in previous chapter, this line must be designed to maintain a high-quality vacuum in the spectrometer source while allowing rapid transfer of the extracted molecules between the two modules.

Holding a chromatographic column it is easy to switch from a MS-Sensor to a conventional GC–MS, simply by changing the temperature program and the column-gas flow. Thus it is easy to obtain a GC-MS screening on headspace previously to any MS-Sensor measurements and most important there is not wasted time when changing from one configuration to another. A second advantage of using a chromatographic column is that a minimum separation can be obtained. This minimum separation may result useful for uncovering minor differences in trace components that would otherwise be masked by other constituents with higher concentrations. In the last papers of this thesis it has been used this configuration because actually this is the effect that is wanted to study. It should be also mentioned that when using a chromatographic column there is an inherent drawback that should be assumed. It is related to the column thermal stress induced by a continuous use of high temperature isothermal conditions. At this way it is promoted the bleeding effect of the column and this may result on an increase of noise in the MS-Sensor signal due basically to signals coming from the ions resulting from this bleeding. In this thesis, a classical non-dedicated GC-MS system (Shimadzu QP 5000, Shimadzu, Kyoto, Japan) has been turn into a MS-Sensor. This system is equipped with an electron beam ionization system acting at 70 eV and simple quadrupol mass filter analyzer which allows analyzing up to 700 a.m.u. Depending on the experiments conduced a 5-m deactivated fused silica column or a normal chromatographic column at isothermal conditions (T=250ºC) has been used, either a Carbowax-10 (30m×0.25mm i.d., 0.25µm film thickness) or a Equity-5 (30m×0.25mm i.d., 0.25µm film thickness) from Supelco. The coatings from these two columns are ideally suited for volatile compounds. The polarities of targeted compounds range between polar and mid-polar compounds which cover the volatiles analysed in the different applications envisaged.

3.3 OBTAINING SUITABLE DATA FOR PATTERN RECOGNITION

3.3.1 Two-way data matrix response

Commercially available software packages from GC-MS manufacturers do not include pattern recognition algorithms as an option. In particuar, GCMSsolutions, which is

the software supplied with GC-MS QP-5000 do not have multivariate analysis tools included. Actually, GCMSsolutions software is conformed by two different parts. First one is the real time part which is dedicated to the hardware control. Second part is the post-run software. It allows chromatograms and spectra processing in a classical way including algorithms for basic preprocessing and further peak area integration or univariate calibration tasks.

In order to obtain a suitable data matrix for further data analysis, the averaged mass spectra for each measurement have been exported in text format using GC-MSsolutions software which using export options tool. Afterwards, it has been used a set of written-in-house routines based on MATLAB® (The MathWorks, Inc., Natick, Massachusetts) in order accommodate these text files in a matricized way. All pattern recognition process was also performed using standard Matlab routines that were programmed and modified according to the different strategies envisaged.

3.3.2 Three-way array response

As it has been explained above, to obtain the two-way matrix response was feasible since GC-MSolutions software allows, for each measurement, exporting the averaged mass spectra in a text file. Then, matricizing this spectra was straightforward done using standard Matlab based routines. However, it was from far much more complicated to obtain the raw three-way data. Generally this is not an option included in commercial software packages. In the particular case of GC-MSsolutions software this is not an included option. Nevertheless there are some manufacturers such as Agilent technologies that have developed macro programs which are able to export a list of each one of the mass spectra recorded at the time points of the whole chromatographic run. Then, after some basic processing it is feasible to load this data in to Matlab and reshape it to a three-way array.

There are basically two reasons why this option is not fully implemented in commercial software packages. First one is that this is not a usual demand in GC-MS user"s community. As it has been previously discussed classical GC-MS users seeks for a reduction of the data to a more understandable format such as lists of peak areas with matching library identifications compounds. Chemometric comparison of chromatographic profiles is still not a technique widely used in routine laboratory. In spite of some promising benefits of this technique, it is still a young technology that remains on the research field. Second reason is a computational memory issue. Retaining all possible m/z intensity values at each scanned point would lead to a huge, sparsely populated matrix which causes a high computational cost and it involves the need for computers and programs able to handle this computational memory effort.

In spite of the GCMSsolutions software did not allow for any option to export the entire three-way data, from the beginning it was suspected that it should be possible somehow. For each measurement run, GCMSsolutions generates a data file which contains the whole raw data information which can be displayed under user"s demand. GC/MSSolutions software includes an OLE (Object Linking and Embedding, Microsoft©) automation interface. This automation interface consists in an objects hierarchy with several properties and methods that allows to access to the information content on the generated files. Hence it is allowed to any other software to communicate with GCMSsolutions and solicit any kind of information included in its objects. Basically, the automation interface is composed by five main objects, one of them contains the three-way data.

A software platform called IDAMAT programmed on Visual Basic 6.0 was created. This platform was conceived as a communication tool between GCMSsolutions and Matlab.

Figure 3.8 represents the basic working principle of created software. IDAMAT is in the central point of the software architecture. It takes the responsibility of creating objects that will further constitute the communication interface between MATLAB and

GCMSsolutions. Once these objects have been created the required information may be obtained. It is necessary to set-up several parameters defining range of mass spectra and time points. Then a new event is generated, this new event will be the command that will start several actions that deal with the building up of raw data matrix. Afterwards, this matrix may be exported to text file or even directly loaded in to a Matlab workspace.

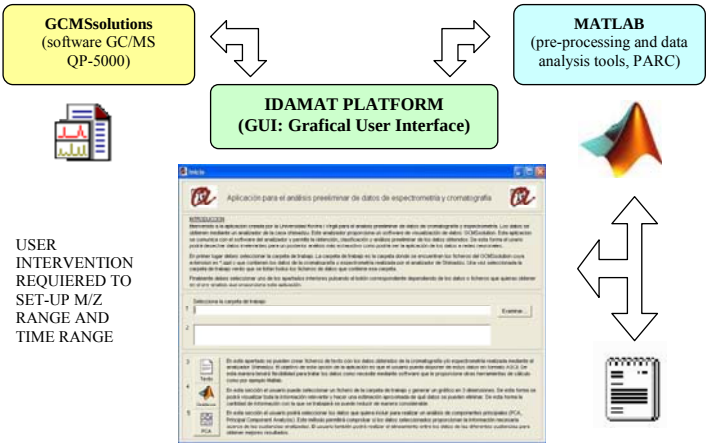

Figure 3.8: Figure representing the working principle of IDAMAT

3.4 PATTERN RECOGNITION

This section is not devoted to a complete overview of the whole pattern recognition steps performed when attempting the different quality problems envisaged. A detailed description on the algorithms used is avoided since it has been already given in previous chapter. Rather than describing each one of the algorithms, this section tries to give a summarized general scheme for all the steps included in the pattern recognition process and an overview of the strategies used in each step.

Otherwise specified, all the algorithms developed within this thesis were built using written-in-house and standard MATLAB® routines (The MathWorks, Inc., Natick,

Massachussets). Most of the algorithms used for pattern recognition were inspired in routines from the PLS-Toolbox® developed by Eigenvector Research Inc[130].

3.4.1 Preprocessing

Once the raw data is available for being loaded in to the Matlab workspace it is converted to a dataset which means that all the information referred to this data (samples labels, m/z labels, scales, title, etc) can be retained and used in a handling way. The first step of the whole pattern recognition process is preprocessing. When comparing whole MS-Sensor data matrices, instead of information on a limited set of peaks, additional variation may obscure the relevant information. This extra variation results from differences in sample preparation and measurement conditions. Before exploratory analysis can be performed, the data should be corrected for this unwanted variation, since it largely influences the outcome of the analyses.

3.4.1.1 Object-wise normalization

Progressive pollution of the source, maintenance operations, or a vacuum quality affected by simultaneous introduction of large amounts of material can lead to signal drifts. Thus, although various tuning procedures exist, if the signal goes uncorrected during a programme of analyses, the resulting data may be weakened or even valueless. Accordingly, it is important to have procedures to monitor the state of the mass detectors and help correct drift. In this way, internal or external normalization is a simple operation, routinely employed in the context of data analysis to eliminate effects linked to variations in the intensity of the signal (sensitivity, quantities injected).

The internal normalization of a given mass spectrum involves expressing the abundance of each mass fragment with respect to the sum of the fragments or with respect to a reference fragment derived either from the product or from an internal standard.

Normalization with respect to the sum of the mass fragments has been the method of choice for internal object-wise normalization in this thesis. It is done by using relative intensity counts instead of absolute counts in the mass spectra response. The mass peak of the ion showing the highest intensity is set to 100 and relative intensity is determined for each of other mass peaks that conforms the mass spectra. This normalization allows comparing among different mass spectra regardless of the amount of a sample introduced to the mass spectrometer. It offers the advantage of being independent from absolute intensities which may extremely vary between measurements. Thus, normalization to the base peak makes comparison of spectra much easier.

There is External normalization of the data has been carried out with respect to an m/z=40 (argon), an ion naturally present in all the spectra and displaying a variance independent of the products analyzed.

3.4.1.2 Unfolding

Unfolding has been considered also as a preprocessing. It is introduced at this point because it is sometimes necessary to conduce variable-wise normalization. As it has been previously reported unfolding (or matricization) involves rearranging a three-way array to form a matrix[113, 131]. This is done by combining two of the original modes while keeping the other fixed. Unfolding a three-way structure of $I \times J \times K$ dimensions can be done in three different ways, depending on which modes are combined: the second and the third *($I \times JK$)*, the third and the first *($J \times KI$)* or the first and the second *($K \times IJ$)*.

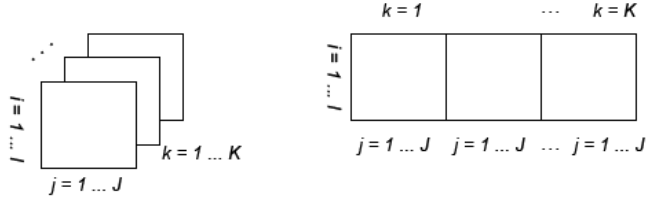

Figure 2.11: Unfolding of a three-way array by combining the second and the third modes[132].

Figure 2.11 shows graphically the process of unfolding a three-way array by concatenating the second and third modes. This type of unfolding has been used in this

thesis in order to keep the sample mode (mode 1) and to combine the mass spectral and chromatographic modes (modes 2 and 3)

3.4.1.3 Variable-wise normalization

In this thesis it has been used basically two different approaches for variable-wise normalization: Centering and Autoscaling

To remove a constant offset, the data can be translated along the coordinates" origin. In two-way data (matrices), the procedure used in this thesis has been column meancentering, where every data on this matrix x_{ij} is centered by subtracting the column mean $\bar{x}_j$ according to Eq. 3.2:

$$x_{ij(centered)} = x_{ij} - \bar{x}_j \qquad \text{[Eq. 3.2]}$$

where *i* is the row index, *j* is the column index, and $\bar{x}_j$ is the column mean calculated from Eq 3.3 :

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \qquad \text{[Eq. 3.3]}$$

Centering a three-way array may be applied across any of the modes. For example, centering the first mode can be done by unfolding the array to a $I \times JK$ matrix (Figure 2.12), and then column mean-centering as described above. If centering is to be performed across more than one mode, one has to do this by first centering one mode and then centering the outcome of this centering[133].

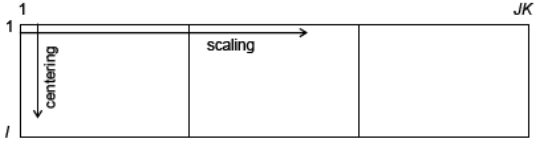

Figure 2.12: Scheme of a three-way unfolded array, showing the directions along which scaling and centering should be performed[132].

When range scale differences among variables are elevated, this range differences will manifest themselves in a subsequent modelling of the data, where the variables with little variation will not be modelled to any significant degree. Centering across the first mode will not remove this scale differences but will move the variation of each variable to the zero level. As the differences in the scales between variables will be arbitrary, it could be useful to scale the data so that each variable ahs the same initial standard deviation (and remove different measurements units). This can be achieved by scaling the data within the second (variable) mode. The entire process of centering across the first mode and scaling within to unit standard deviation within the second mode) is known as autoscaling and it has been used widely throughout this thesis. Autoscaled data have column mean zero and column unit variance according to Eq 3.4 and Eq 3.5:

$$x_{ij\,(autoscaled)} = \frac{x_{ij} - \bar{x}_j}{s_j} \qquad \text{[Eq. 3.4]}$$

where,

$$s_j = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2}{n-1}} \qquad \text{[Eq. 3.5]}$$

In three-way analysis, scaling must be done on the rows of a matrix and not across the columns, as is the case with centering (i.e., one has to scale whole matrices instead of columns) (Figure 2.12). Not all combinations between centering and scaling are possible when working with three-way data. Generally, only centering across both modes is straightforward or also scaling within one mode combined with centering across the other mode which is exactly what e.g., autoscaling amounts to. For example, to remove differences in magnitude between the m/z of a set of m/z-time matrices, one can scale the three-way array within the m/z mode. To do this, the three-way array has to be unfolded in such a way that m/z (variable *j*) is in the rows and then the rows are scaled. This is because all the columns where variable *j* occurs have to be scaled.

Scaling disturbs centering across the same mode but not other modes and cencetering across one mode disturbs scaling within all modes. Autoscaling is a special

case where scaling within the second mode to a standard deviation of one does not affect centering across the first mode. The reason is that the scaling is specifically performed relative to the center of the columns (standard deviation are based on centered data). Hence, any change in the offset is inmaterial for the standard deviations. Scaling by other means that standard deviation will not have this property. For a most widely description on centering and scaling in three-way arrays refer to Bro and Smilde[60, 133].

3.4.1.4 Alignment

Although MS-Sensor data is considered to be reproducible under optimized injection conditions, small differences in retention time between different injections will always occur. These differences may arise from small variations in the carrier flow, changes in the columns during operation, temperature variations, drift in the detectors and other unknown factors that influence retention time reproducibility.

Dealing with two-way data analysis, the signals compared were an averaged mass spectrum, then, these signals did not suffer from misalignments from run to run since time dimension was neglected. It was at the time to include time dimension when it was realised about the importance of alignment of data. Taking advantage of the time dimension in MS-Sensor devices involves the unavoidable evaluation of co-eluted peak shape in the multi-way analysis step. In order to perform direct chemometric analysis of the entire chromatographic data matrices, coeluted chromatographic profiles must be properly aligned to compensate for minor drifts in retention time, either in global or in small sections of the chromatograms. A modification of the RAFFT algorithm introduced by Wong et al.[134] has been considered for the alignment applied in order to overcome retention time shifts prior to modelling both three-way and unfolded data. This algorithm makes use of the Fast Fourier transform for rapid computation of the cross-correlation function that enables alignments between a target sample and samples to be optimized. It is based on spectra segmentation models and is developed to offset the need for operating parameters. Minimal segment size is determined automatically by recursive alignment from the full spectrum

(global scale) to progressively smaller segments (local scale) until no further alignment is required. The modification consisted in the application of this algorithm to each m/z channel instead of doing so to the reconstructed TIC signal. Afterwards, the TIC signal was reconstructed summarizing the intensities of all m/z channels for each scan of the aligned response.

3.4.2 Variable selection

When spectral fingerprints reveal large numbers of mass fragments, part of this signal is often non-informative, it is necessary to select the fragments relevant to the problem in question prior to train any model. Variable selection has been envisaged from different points of views, depending on the available information.

In cases where previous knowledge of headspace composition is available, then the variable selection can be straightforward carried on. Once the key analytes are determined it is easy to set-up the MS-Sensor focused to the highest abundant ions of this key analytes acquiring the mass spectra on the SIM mode. Thus it is required the use of GC-MS or other volatile screening tools for determining which the molecules present on the headspace extract are. Working on the SIM mode is a particular useful approach when using SH/MS systems without a module for pre-concentrating the molecules since the sensitivity of mass spectrometer can be increased.

Nevertheless, the usual situation in a MS-Sensor approach as well as in any *"electronic nose"* approach is to have not a previous vast knowledge on the sample under analysis. Then, variable selection can not be performed as explained above. Thus, throughout this thesis, several strategies have been developed in order to reduce dimensionality in such way that the methods implemented should be able to perform automatically the selection of m/z variables most relevant for the application envisaged without the need of previous knowledge on the exact nature of the analytes.

The first technique developed is based on a PCA or LDA filter variable reduction to reduce the dimensionality of the input matrix to a Fuzzy ARTMAP based classifier. With this reduction in dimensionality the numerous and probably high correlated original variables are transformed into a small number of orthogonal (i.e., uncorrelated) new variables which are a linear combination of the original ones.

Other techniques for variable reduction are based on wrapper methods and evaluate candidate feature sets using a classification algorithm on the training data. The feature subset selection algorithm conducts a search for a good subset using the classifier as part of the evaluation function. In this context, there have been developed two different methodologies. The first one is based on stochastic algorithms such as GA applied to the selection of m/z factors. GA based variable selection procedure was implemented to improve the prediction ability of different PLS models using RMSECV as fitness parameter (For a more detailed description on the algorithm, please refer to ***Paper V***. Second approach used for variable selection consisted on a new method developed for an effective feature selection especially suitable for applications where the dimension of feature space is high, a significant degree of correlation exists between features and some of them are affected by noise, such as in MS-Sensor applications. The feature selection introduced consists of three steps that are run consecutively. The first step helps detecting and removing non-informative, noisy features and is conducted in a supervised way. The second step is aimed at detecting collinearity between features in an unsupervised way. As a result, highly collinear features can be removed. Finally, in the third step, a greedy search method (e.g., a stochastic one, SA) is applied to the reduced feature set, which results from applying the first two steps. The fitness parameter used this time has been the success rate in of prediction of a Fuzzy Artmap classifier. Refer to ***Paper VI*** for a more extent discussion of the algorithms developed. With this approach, the whole variable selection process is time efficient since the first two steps are able to dramatically reduce the number of features at a very low computational cost. The algorithm was first evaluated in a dummy dataset consisting of synthetic mixtures of volatile compounds. This simple database has been used to show that the feature selection process is able to identify a minimal set of fragments that enables the correct discrimination between mixtures using a simple fuzzy ARTMAP

classifier. Furthermore, given the simple nature of the problem envisaged, it was possible to show that the fragments selected "made sense", that is, were characteristic ionisation fragments of the species present in the mixtures to be discriminated. Once the correct performance of the feature selection method was demonstrated, it was applied to the real dataset arranged for evaluating Iberian ham quality.

3.4.3 Exploratory data analysis

Exploratory analysis involves applying chemometric methods to find patterns in the data, i.e., to better understand the structure of the data with the constraint that no previous knowledge of the data is used to find such patterns. For this reason, these methods are often called unsupervised pattern recognition methods. Typical aims of unsupervised pattern recognition methodologies are to reveal groups and trends in the data, to find relationships between samples and variables and to detect outliers. The results of exploratory analysis are often presented in graphical displays, which make it easy to interpret the information. When the purpose of analyzing the data set is to find patterns, relations, differences and agreements between objects and/or variables, then decomposition methods can be used to summarize the data conveniently and explore the data set using plots and figures. Typically, subspace-based methods (projection methods) are used because of their ability in handling collinearity problems. Two-way data exploratory analysis has been always performed by using PCA and PARAFAC has been used in case of three-way arrays.

3.4.3.1 Validation and diagnosis tools for exploratory data analysis models

Data analysis is often based on evaluation of fitted models. These models provide approximations and interpolations of the real system under study. As stated in section above there is often a wide choice of models and the quality of these depends on a number of factors, because mathematical models and statistical models based on real data have

properties depending or not on the system under study but also on the sampling, the properties of the model, the algorithm, etc. Validation can be seen as the part of the analysis where it is investigated if valid conclusions can be drawn for a model. Does the model generalize the data in a parsimonious way, i.e., express the main variation of the data in a way that is simple as possible? Does the model predict efficiently? Has the model been influenced by certain samples or variables in such way that these, rather than data in general are described?

3.4.3.2 Model complexity. Choosing the right number of components in exploratory models

The so-called scree plot[135] has been widely used for detemining the number of components in PCA. This method can be modified for finding the appropriate number of components in PARAFAC. The original scree plot gives the eigenvalue of the cross-product matrix sorted by size as a function of the number of components. There is a an empirical rule which says that a large break in this function-going from a steep slope to a much lower decrease- often represents the "effective" dimensionality. It allows an interpretation of which components are well enough above the noise levels for being useful. The firs eigenvalue equals the number sum of squares explained by the first component, etc. A certain cuttoff in the scree plot determines which components are too small to be used. Usually, the number of components is chosen where the plot levels off to a linear decreasing pattern. Thus, no more than the number of factors to left of this point should be retained. The scree plot with eigenvalue of PCA can also be shown cumulative because the eigenvalue are independent and add up to 100% of the sum of the squares. In other words, this plot is the total residual variance plot. Often, for determining the right number of components is not enough with scree plot. However, interpretation of the components is often the key for choosing the correct number of components in a PCA. Thus, scores, loading and residuals are three golden standard plots that have been employed all together to decide the proper number of components retained in PCA.

A plot similar to the scree plot in PCA modelling can be made for PARAFAC by plotting the sum of squares of the individual components. However, in this case, the cumulative plots cannot be made directly because the variances of the individuals components are not additive due to the fact that PARFAC is a not nested model, it is to say, the estimates of *R* components change when going to a (*R+1*) components model. Furthermore, the sum of squares of the one component model may not equal the size of the largest component in two-components model. Hence, the scree plot is not directly useful for PARAFAC models. On the other hand it can be constructed by plotting the explained or residual sum of squares of one component model, a two component model, etc. This will provide similar information to the two-way ordinary scree plot, with the exception that factors change for every model since PARAFAC is not sequentially fit. Usually the results from cross-validated fit are also showed in the same bar plots since from the fit percentage is difficult to see what the appropriate model dimensionality is.

Determining the proper number of components in PARAFAC modelling is an extremely crucial issue for obtaining a good fit of the data under PARAFAC assumptions. A part from the scree plot, there are several ways to determine the correct number of components in PARAFAC[113]. In this thesis, we used split-half analysis and residual analysis. Split-half analysis involves dividing the data into two halves and then making a PARAFAC model on both halves. Because of the uniqueness of the PARAFAC model, if the correct number of components is used, the same loadings in the nonsplit modes on both data sets should be obtained. Residual analysis can be performed as in bilinear models. Main parameters used for residual analysis are summarized bellow.

3.4.3.3 Residual and influence data analysis

From any PCA or PARAFAC model, two statistics can be defined: Hotelling T^2 and Q. As it has been explained all models whether they are two or three-way are built as follows:

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} \text{ or } \underline{\mathbf{X}} = \underline{\hat{\mathbf{X}}} + \underline{\mathbf{E}} \qquad \text{[Eq. 3.6]}$$

The main goal of the data analysis is usually to find X, but the residual E can give important clues to the quality of this model. Q is simply defined as the sum of squares of each row (samples) of E. The Q statistic indicates how well each sample conforms to the model. It is a measure of the difference, or residual between a sample and its projection into the components retained in the model. Q bar plots for each one of the samples have been used to give an idea on how these residuals are distributed. They should follow a random distribution. A larger residual variance in an object indicates that the model does not fit or explain this object as well as others and thus may indicate that the object is a potential outlier.

Influence analysis has been assessed using the Hotelling T^2 statistics. It is calculated for the systematic part of the variation of the data. The Hotelling T^2 plot therefore represents the projection of each measurement onto the hyperplane defined by the components of the model. As well as in the Q statistics, T^2 can be summarized as bar plots for each sample. Objects or samples having distances much larger than the remaining ones identify samples with larger influence and hence potential outliers. On the other hand, a small distance indicates a variable that is not relevant and may be only contributing noise.

A graphic frequently used for assessing outliers has been the T^2 vs. Q. It is a very informative plot since it can easily identify potential outliers in the upper right side of this figure.

3.4.4 Supervised pattern recognition: classifier models and regression models

In this section the classification and regression methods used in this thesis are not going to be described since it is considered that they are widely described in literature. Refer to the Chapter 2 for a summary of the models. Instead, the general strategy to fit this classifier or regression models is going to be detailed.

Classification methods use a training set (i.e., a set of samples belonging to a known class) to derive a classification rule that enables us to classify new samples with unknown origin in one of the known classes. For this reason, these methods are often called supervised pattern recognition methods. The validity of the classification rule is assessed with a test set. There is a wide range of classifier methods used within this thesis: LDA, PLS-DA, N-PLS-DA, unf-PLS-DA, PARAFAC-mlr-DA and Fuzzy Artmap Classifier models. If prediction/quantification is attempted, then models applied are PLS, unf-PLS or N-PLS.

The general procedure to built up any classifier model is as follows: Dataset is split in two different parts, the training dataset and the test dataset. In a first stage, the so-called training phase the models are fitted using the training dataset. These models are internally validated using a leave-one-out cross-validation method. This internal validation is mainly used to gives an idea on how the model fits the data in the training phase and also it is used to set-up the essential parameters need for fitting the models such as number of components. Finally, to complete the modelling procedure, it is imperative to confront the models developed with testset data, i.e., data derived from samples unseen by the model during the training phase. This will be the evaluation phase. Only from this final evaluation we can obtain realistic and honest indication of the model with respect to how well it explains, generalizes or predicts the real-world problem that is aimed to describe. At this point it should be mentioned that, when choosing data values to leave out, it has it has to borne in mind that different approaches will yield different results. For example, leaving out only replicates, a repeatability error will be obtained whereas if the left-out samples are from new experiments performed by new laboratories, a reproducibility error is obtained.

In fact, in cases where the number of samples are too low, it has been attempted an n–fold training and validation process. It consists in to randomize the response matrix and afterwards, to split sequentially this matrix in to the training and validation dataset. At this way, all the samples takes part on the training and on the validation of the models. The models parameters are calculated according to the n averaged models.

3.4.4.1 Complexity of the models. Tools for choosing the right number of components in classifier and regression models. Residual analysis. Validation and diagnosis tools

The training phase of the classifier models has been internally evaluated using a cross-validation routine. The process is as follows: given *n* measurements in the training matrix, the model is trained *n* times using *n*−1 vectors. The vector left out is then used for testing the model. Performance during training has been estimated using the root mean square error of the cross-validation (RMSECV). RMSECV tells us how well a given model fits the data in the calibration step and it is defined as in Eq 3.7:

$$\mathrm{RMSECV} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_{i,\mathrm{ref}})^2}{n}} \qquad \text{[Eq. 3.7]}$$

where $\hat{y}_i$ are the values of the predicted classes with the samples left-out in each iteration of the cross-validation process. The $y_{i,ref}$ is vector with the property that is wanted to predict in case of regression model. In case of a classifier it is a vector composed by dummy variable defining current classes; n is the number of calibration samples.

The evolution of RMSECV with the number of components has been the tool used to assess the optimal number of factors for model fitting. This value decreases when the number of factors is increased and either increase again or stabilises if more factors are added. In practice, a minimum in RMSECV values may not always be present or sometimes the minimum is marginally better than simpler models. In such cases, it becomes more difficult to decide on the number of components and prior knowledge as well as other model results (residual analysis) and the overall aim of the analysis have to be taken in to account. Choosing too many factors can result in an over-fit the training set, which results in lower prediction ability.

The calibration variance accounted in the training phase either for Y or X-block in classifier modes is also accounted according to Eq 3.8:

$$Var_{(x_{cal},y_{cal})} = 100 \times \left(1 - \frac{SS_{\text{mod}}}{SS_{tot}}\right) \qquad \text{[Eq 3.8]}$$

where SS_{mod} corresponds to cross-validated residuals and SS_{tot} refers to the sum of squares of treated data. Cross-validated residuals were determined as the discrepancy between the fitted value and the real value left out in the full cross-validation routine.

In this thesis, residuals obtained from a test-set or from a cross-validation are used instead of fitted residuals. Above it has been explained how residuals obtained from the training phase are accounted. The residuals may also be obtained from the validation step with the test-set dataset. The residuals associated to the Y-block are summed to a number of regression statistics such as percentage of variance explained, root mean squared root of prediction (RMSEP), etc.

The prediction error (RMSEP) has been estimated as stated in Eq. 3.9

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_{i,\text{ref}})^2}{n}} \qquad \text{[Eq 3.9]}$$

where, this time, $\hat{y}_i$ represents the values of the predicted class of samples in test set, $y_{i,ref}$ is the current class for these samples and n represents number of samples in the test set. RMSEP is an expression of the error that can be expected when using the calibrated model and has been used for assessing the ability of the model fitted to generalize to a new dataset.

The percentage of explained variance in validation either for Y or X-block has been accounted as:

$$Var_{(x_{cal},y_{cal})} = 100 \times \left(1 - \frac{SS_{pred}}{SS_{tot}}\right) \qquad \text{[Eq 3.10]}$$

where SS_{pred} now corresponds to the sum of the squared predicted residuals. The validation variance for the Y-block expresses how well the model will perform with new data. For the X-block it indicates how well the validation data have been projected onto the classifier models already fitted during training.

For the particular case of classifier models the predicted y-value results in a continuous variable that can be interpreted as a class similarity index. Each calculated class prediction value can be compared with a Bayesian distribution curve to determine for a given predicted y-value the probability that this value belongs to that original class. For each sample, either in prediction or in the test set, a table of probabilities is obtained. A threshold of "predicted y" is then determined above which a sample is considered to be a member of the class. Class assignments are done comparing against this value. In order to compute the certainty of each class assignment, values for reduced residuals Q and T^2 (Hottelling"s) statistics are used. The larger the values, the less certain are the class assignment

Once the class assignment is feasible, it is possible to account for the success rate either in the training (using the cross-validated predictions) or in the test set. The success rate is used for both linear and non-linear classifier models and it is computed as success rate in predicting the correct class for the model fitted.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 4

Results

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

4. RESULTS

4.1 ASSESSING ONSET OF RANCIDIY IN CRISPS

4.1.1 Introduction

Two very important issues for potato crisps producers are the detection of rancidity and its associated off-odours/flavours, and the estimation of shelf-life of the product. There are basically two reasons to monitor to what extent oil has undergone oxidation: The first one is an obvious economic advantage when crisp producers can appropriately determine the useful life of frying oils. Premature discarding of the oil results in economic losses and, on the other hand, overuse of frying oil greatly affects the quality of fried products and causes undesirable nutritional effects. The second reason is related to consumer acceptance of the products, highly correlated to the presence of off-flavours compounds in potato crisps.

The methods currently in use to assess rancidity need a previous extraction step. Oil extraction is a very time-consuming, complex and labour-intensive step for routine quality control applications. Furthermore, the solvents or the methods used can induce oxidation and interfere with the final results. Since the crisp industry demands a large number of samples to be analysed and a high sample throughput, there is a need for faster and simpler methods to assess crisp rancidity and off-odours/flavours. In this context, the use of *"electronic nose"* technologies would be of great help.

4.1.2 Goals

To perform a comparative study between gas sensor and MS-Sensor approaches for assessing rancidity in crisps.

- Once the optimal technology has been selected, to determine whether this technology is able to perform a semi-quantitative assessment of the rancidity in crisps for shelf-life determination purposes.
- To correlate *"electronic nose"* data with well-established methods such as Rancimat and ADV test.
- To determine whether the correlation models improve their performance when a variable selection or variable reduction algorithm is applied.

4.1.3 Results

Paper II, reports about the design and use of two different *"electronic noses"* to assess rancidity directly from potato crisps, as an alternative test to currently available methods (either the Rancimat or the ADV test). The two *"electronic noses"* were based on fingerprint mass spectrometry and an array of metal oxide gas sensors, respectively. This technology allows for rancidity measurements without any previous oil extraction step. This simplifies sample preparation, avoids the use of solvents to extract oil and speeds up the measurement process.

A four class categorization of crisps samples according to four artificially achieved rancidity stages was envisaged for semi-quantitative prediction of rancidity. Both MS-Sensor and MOX array approaches were used to classify crisps according to four stages of oxidative rancidity. The MS-Sensor reached a 100% success rate in this classification and the success rate of the MOX array only reached a 68% success rate (these are cross-validation results). However, if a less ambitious task is attempted, i.e.,, the classification between only two groups, fresh and rancid samples, both technologies achieve a 100%

success rate. Thus, it is demonstrated that MS-Sensor technology outperforms the more traditional MOX array approach when discriminating among different rancidity stages.

Furthermore, a cross-validated LDA technique is used to explore the data clustering ability and the influence of headspace conditions. It has been highlighted the importance of proper optimization of headspace parameters and its high influence on the result. Changing some dynamic headspace parameters such as conditioning temperatures of crisps leads to a better performance of the MOX array approach. In the case of MS-Sensor, adjusting SPME parameters (temperature, fiber coating, extraction time, etc) becomes a key step to ensure reliable headspace extracts (and therefore to achieve a good success rate on the analysis). A mild heating of the crisps (up to 70 ºC) is necessary for a headspace representative of their rancidity stage.

Both, MS-Sensor and the MOX array approaches have been found sensitive enough and suitable for semi-quantitatively assessing rancidity in potato crisps. The results obtained from both technologies were found in very good agreement with the more classical methodologies used in crisps production such as Rancimat and ADV tests. While using traditional methods to assess rancidity measurements take several hours to complete, a measurement with the MS-Sensor or the gas sensor array takes 25 and 40 min, respectively. Therefore, the assessment of crisp rancidity using e-nose technology could become a routine test in quality control laboratories of crisp producers.

Paper III reports a new method based on the use of an SPME/MS e-nose to assess oxidative and hydrolytic rancidity in crisps. Unlike more traditional methods such as the Rancimat and ADV tests, this system did not rely on the previous extraction of oil from the samples; instead, it directly evaluated volatile compounds from the headspace of the crisps. This results in easier sample preparation, avoids the use of organic solvents for oil extraction and speeds up the analysis. The time needed to analyse a sample with the MS-Sensor system was around 28 min. This includes 20 min for the SPME, 3 min for acquisition of the mass spectra and 3 min for cleaning the SPME fibre. This compares

favourably with the Rancimat method, which can take several hours to produce a result if the oil under analysis is of good quality.

The effectiveness of the MS-Sensor in the assessment of the quality of crisps was demonstrated in two different applications. First, the MS-Sensor was used to classify the crisp samples according to four pre-established rancidity stages. The system input the data from the MS-Sensor (pre-processed by a PCA) into a fuzzy ARTMAP classifier. The success rate in classification was estimated to be around 93% (test-set validation results). The system could always perfectly discriminate between fresh and rancid crisps. In the second step, the MS-Sensor was used to predict the results of the ADV and Rancimat tests by building quantitative PLS models. The correlation between MS-Sensor and the ADV and Rancimat tests was also investigated. A good correlation existed between MS-Sensor and the Rancimat and ADV tests (the correlation coefficients were 0.98 and 0.97, respectively). For the particular applications developed here, a reduction in the number of input variables results in more parsimonious models with higher predictive ability. The use of a PCA to reduce the dimensionality of input variables resulted in improved performance of the fuzzy ARTMAP classifier. The GA-based variable selection led to a significant improvement of the PLS models that were built to predict the results of the ADV and Rancimat tests.

According to these results, the SPME/MS-Sensor approach might become an alternative tool for the assessment of oxidative and hydrolytic rancidity in crisps.

4.2 DETECTION OF UNWANTED FUNGAL GROWTH IN BAKERY PRODUCTS AT EARLY STAGES

4.2.1 Introduction

The investigation of the detection of fungal volatiles as a method for mould detection in food substrates started in cereals. For a long time, human olfaction has been

used as the primary criterion for the acceptance of grain for human consumption in many countries. It is known that fungi produce volatile compounds during both primary and secondary metabolism that can be used as markers to detect food spoilage, unwanted fungal growth. In that sense, MS-Sensor is a valuable candidate as a technique for the detection of fungal growth spoilage in bakery products.

Fungal spoilage is an important issue in bakery products. Although its impact from the safety point of view is not significant, the presence of visible colonies on a product degrades the company image in front of the consumer, resulting in economic losses in the medium–long term. Furthermore, microbial spoilage can induce nutritional losses, off-flavours, and formation of mycotoxins or potentially allergenic spores. Therefore, it can lead to an organoleptic deterioration of already marketed bakery products, which indeed threatens consumers" confidence level. This is the reason for a growing need to find a method to conveniently assess the degree of fungal growth in bakery products at a very early stage and before it becomes visible. Some companies use the measurement of water activity of the final products as an index for fungal spoilage prediction and rejection of product batches. However, there is a need for more reliable alternative methods.

4.2.2 Goals

- To set-up a MS-Sensor system able to predict fungal growth in bakery products
- To evaluate which sample system (either the SH or SPME) performs better for fungal growth assessment
- Evaluation of the performance of the optimal MS-Sensor configuration for monitoring early stages of fungal growth. Two different situations were studied: The first one was to study whether the MS-Sensor could achieve in-vitro fungal growth monitoring. The second was to prove whether the MS-Sensor could predict fungal growth on bakery product analogues (on in-situ samples). This last experiment was done in order to simulate the performance of the MS-Sensor approach in a more realistic application.

- Use the configuration set-up above in order to predict mould growth (ergosterol content) of fungal contaminated samples of bakery products.

4.2.3 Results

In ***Paper I*** the design, optimization, and evaluation of a mass MS-Sensor for early detection of unwanted fungal growth in bakery products is presented. Seven fungal species (Aspergillus flavus, Aspergillus niger, Eurotium amstelodami, Eurotium herbariorum, Eurotium rubrum, Eurotium repens, and Penicillium corylophillum) were isolated from bakery products and used for the study.

A first study was dedicated to the selection of the optimal sampling system. Two sampling headspace techniques were tested: SH and SPME. The aim was to discriminate inoculated samples from blank vials and to evaluate whether the instrument could classify samples according to their genera and specie. Cross-validated models based on principal component analysis (PCA), coupled to discriminant function analysis (DFA) and fuzzy ARTMAP were used as the preferred data treatment algorithms. The SPME technique gave better results than the SH technique when coupled to our MS-Sensor. While SH shows a good performance in the discrimination between blank and inoculated vials, this success rate dramatically decreased when a classification based on genera or fungal species was envisaged. On the other hand, the MS-Sensor ability for discriminating among samples according to their genera and specie was highly successful when using SPME as the sampling technique. It achieved better repeatability and it was more sensitive due to its ability to concentrate volatile analytes. SPME allowed distinguishing between fungal genera or even between several species while the SH only achieved acceptable discrimination capability in the case of discrimination between blank and inoculated vials.

On the basis of these results, the SPME/MS-Sensor combination was chosen as the optimal configuration for attempting the second study aimed to evaluate whether this approach was able to monitor in-vitro fungal growth. SPME/MS-Sensor monitoring was performed on in-vitro fungal cultures between 48 and 168 h after inoculation. By using

DFA-fuzzy ARTMAP algorithms on the PCA-reduced response matrix, the system achieved a 100% success rate after 48 h from inoculation when attempting to discriminate both between inoculated and blank samples and among fungal genera. This demonstrates that the SPME/MS-Sensor was a suitable tool for on-line in vitro monitoring and early detection of unwanted fungal spoilage.

These results encouraged us to proceed to the third study where an in-situ fungal growth monitoring was attempted. This time a more challenging data set was measured since the samples were bakery product analogues. In situ monitoring was performed 24, 48, 72, 96, and 168 h after inoculation. Distinction between blank and inoculated samples reached an 88% success rate 24 h after inoculation, 98% after 48 h, and 100% after 72 h. Moreover, a general guessing on volatiles resulting from fungal growth and sporulation was also conduced through the interpretation of loadings plots and some non-published chromatographic runs. Volatiles found and suggested by the models such as 3-methyl-1-butanol, 2-methyl-1-propanol, and 1-octen-3-ol, etc resulted in good agreement with the literature and were found to be indicators of fungal presence.

From the results obtained on the monitoring of in vitro fungal growth, it can be concluded that during the first 24 h fungi are mainly producing CO_2 and common metabolites associated with primary fungal growth and structural formation such as 3-methyl-1-butanol, 2-methyl-1-propanol, and 1-octen-3-ol indicative of fungal presence. Species identification may not be possible at this early stage, since the compounds produced in the highest amounts are similar for different species. The volatile pattern profile is, therefore, very similar and does not allow discrimination between species. Characteristic volatiles that might allow species classification are mainly produced during secondary metabolism. After 48 h of incubation, the system is able to predict fungal genera with a 78% success rate, which implies that the secondary metabolism has started. Sporulation happens 72- 96 h after inoculation depending on fungal species, which leads to an increase in several volatile compounds generating different pattern profiles for each fungal genus or species. This allows the best discrimination results in our model.

In ***Paper V*** measurements on bakery inoculated analogue samples were performed 2, 4 and 7 days after inoculation. Ergosterol content was also analysed using HPLS measurements of each vial. PLS models were used to estimate the levels of ergosterol using MS-Sensor as X-block independent variables or descriptors. Ergosterol content clearly increased with time for all fungal species tested. After 48 h ergosterol levels were negligible for some species (A. niger, E. rubrum, P. corylophilum), while after 7 days larger values were found. The highest levels of ergosterol were found for E. herbariorum and E. repens, regardless of the incubation time. Similar values were found for both Aspergillus species, while P. corylophilum, the slowest growing, showed low levels. No ergosterol was found in the control samples. At early stages of fungal growth the correlation between fungal activity and ergosterol levels was high and so good estimates of ergosterol in the spoilt samples could be achieved by the use of the MS-Sensor. Accuracy on prediction of PLS models was between 87 and 96%, except for samples inoculated with Penicillium corylophilum where the best predictions only reached 46%. The predominant m/z fragments observed in this study may correspond to CO_2 (m/z=44) plus other small fragments frequently found in the mass spectra of low molecular mass compounds commonly produced by fungi in their early growth such as 3-memthyl-1-butanol that was revealed as one of the indicators of fungal activity at very early stages.

4.3 MONITORING FRESHNESS OF SARDINES UNDER COLD STORAGE

4.3.1 Introduction

Freshness, defined as the number of storage days at a certain temperature, is the single most important attribute when assessing fish quality. It has been reported that there are a lot of volatile compounds emitted when fish degrades. This fact gives rise to the overall odour of fish when it is degraded. The concentration of some of these volatiles increases with time as the fish spoils; indeed some of these compounds are often used as indicators of spoilage. Due to the high number of volatile compounds involved in the process, and to the fact that they also change dynamically, the measurement of fish

freshness over a long storage period can be achieved with a multicomponent approach. The large number of compounds whose concentrations are only partially correlated make this application particularly appealing for sensor arrays of partially selective chemical sensors. The aptitude of some of these *"electronic noses"* to capture the changes in fish headspace and to correlate these changes with the freshness has been already demonstrated in several applications using different kinds of chemical sensors such as MOX. Nevertheless, when using an *"electronic nose"* to detect dynamical changes in the headspace, great attention must be paid to ensure that these detected changes are not based solely on a comparatively trivial parameter such as humidity or water activity. For many applications *"electronic nose"* measurements should ideally be accomplished by other reference methods able to ensure that changes measured in the headspace are related to the phenomena under study, i.e., volatiles resulting from fish spoilage; otherwise there is a serious danger that the species measured by the *"electronic nose"* could be not correlated with the off-odours pretended to be measured. The characterization of the headspace of fish using the signal pattern obtained by a MS-Sensor device is a recent approach which could allow fish quality analysis by means of chemical composition methods. The patterns produced by MOX array sensors have no chemical meaning; however a MS-Sensor generates directly-interpretable chemical information. Thus, MS-Sensor is a valuable tool to support MOX array results.

4.3.2 Goals

- To use the MS-Sensor approach as a reference method in order to check the results obtained by a MOX based *"electronic nose"* for assessment of sardine freshness.
- To determine chemical markers of sardine's spoilage during 11 days of cold storage at 4 ◦C.

4.3.3 Results

In ***Paper VII,*** an SPME/GC-MS approach was used to asses the chemical changes of the headspace of sardines during eleven days of cold storage. There were ten different compounds identified (trimethylamine, acetone, 2-butanone, ethanol, 2-butanol, 3-methyl-1-butanol, ethylacetate, dimethylsulfide, dimethyldisulfide, dimethyltrisulfide).

A MS-Sensor using the SH method was implemented and a PCA was performed using as inputs the intensities of the three main fragments of the ten previous identified compounds. This multivariate approach was considered to evaluate whether some correlation existed between the spoilage of sardines and the change in headspace composition. Sample clusters are well apart according to the number of days of storage. The fact that most of the volatile molecules identified by SPME/GC-MS have numerous mass fragments in common increases the difficulty to scrutinize which volatiles allow for cluster separation according to different spoilage stages. The poor specificity of m/z fragments can be explained by the low molecular mass of volatiles sampled by means of static headspace and the high degree of fragmentation obtained with electron beam ionization.

Despite this poor specificity, several conclusions can be derived taking into account the scores plot. Temporal separation at early and medium stages of spoilage (Days 1–7) are related to the high presence of m/z ratios 45–47. These fragments stem mainly from sulphur compounds and alcohols such as 2-butanol and ethanol. All these compounds seem to be responsible for clearly separating the most spoilt samples. That makes sense, since these are the compounds that experience the highest increase at Day 11. The early days of cold storage seems to be represented by m/z fragments 41–43, 55, 58, 59, 70 and 72. Then it can be deduced that substances which highly influence at the beginning of fish spoilage are trimethylamine, acetone, 3-methyl-1-butanol and ethylacetate.

Changes in the headspace composition according to the number of conservation days have been demonstrated by identifying the main compounds and their evolution via a SPME/GC-MS study. Compounds found resulted in good agreement with previous reported works. These compounds prove that the changes in conductance of the six MOX sensors conforming the array used in the freshness assessment of sardines can be justified by an increase in the concentration of volatiles generated by sardines as a function of storage time, or the occurrence of new species in the headspace of sardines. The results obtained with the MOX sensor array are supported by those obtained with a MS-Sensor screening

technique (i.e., they are quite similar). This is how it has been demonstrated that the *"electronic nose"* set-up in this paper is feasible for this particular application.

4.4 USING A MS-SENSOR DEVICE FOR OLIVE OIL AUTHENTICATION PURPOSES

4.4.1 Introduction

Olive oil authentication has become a problem of increasing importance in the food industry, from both a commercial and a health perspective. Discrimination of olive oils from different cultivars or different sensory characteristics is an important issue in the food industry since the concession of a Protected Designation of Origin (PDO) requires perfect knowledge of the physical and chemical characteristics of the oil. This concession is also a crucial issue from economical point of view.

The complex flavour of virgin olive oil is mainly produced by volatile and phenol compounds. One of the most important tests to judge the quality of the oil is the organoleptic assessment. Most of the analytical approaches to discriminate olive oils from different cultivars are based on the application of multivariate analysis to composition data of virgin olive oils, currently obtained by chromatographic techniques such as gas chromatography coupled with mass spectrometry. However, the sample preparation in these methods (purification, derivatization, etc...) is too time-consuming for routine use in the food industry or by regulatory inspection. Considering changes in the volatile headspace of olive oil samples as a measure that integrates aroma sensations and could distinguish different oil cultivars, MS-Sensor technology can be used as a discriminatory analysis tool that requires minimal analysis time and gives results that are easily interpreted.

A drawback from EI ionization mode MS-Sensor (the most common types) is the inherent low specificity of m/z "pseudosensors". As it has been mentioned in previous chapters, this low specificity arises from the high degree of fragmentation obtained with

electron beam ionization. A way to approach this specificity could be to consider the time dimension variation of each one of the m/z fragments scanned. Thus, the m/z sensors considered become much more specific because despite the same fragment may come from fragmentations of different molecules, their temporal variation should be different since they come from different molecules.

In the traditional MS-Sensor approach, an averaged mass spectrum along the detected peak is the data to be considered for pattern recognition. Nevertheless, considering this averaged mass spectrum may lead to losing useful temporal information. Even when chromatographic resolution is avoided, a sort of diffusion is observed on the isothermal peak. This fact allows us to consider the possibility of computing this extra-information by considering the three way nature of the data using trilinear algorithms such as N-PLS or PARAFAC. Multi-way methods are particularly useful for treating data with more than two sources of variability such as the data generated by MS-Sensor devices, where the response in ion counts arriving at the detector is measured as a function of time and mass/charge ratio. In fact, data provided from a MS-Sensor should be arranged as multi-way array where the first mode represents samples, the second one corresponds to mass spectra and the third to the elution profiles. Exploiting differences in the time response of the analytes can enhance the subtle variations in the spectra and therefore the classification performance may be improved. To date, the application of second-order methods to classifier models in MS-Sensor devices has not been reported in the literature.

4.4.2 Goals

- To explore whether the MS-Sensor approach succeeds differentiating olive oil according to its geographical origin and according to its organoleptic profile.
- To apply multi-way classifier algorithms to MS-Sensor response data and compare the performance of two-way versus three way classifier models for the olive oil application.

4.4.3 Results

In ***Paper VI*** the discrimination among five olive oils stem from different origins was attempted. These oils were also organolleptically assessed by an expert"s panel. Previous exploratory data analysis was conduced using the two-way PCA approach and the three-way PARAFAC. The same tendencies were found for both the PCA and the PARAFAC model. The five types of oils cluster well apart. Separation according to their geographical origins can be easily guessed. In the same models olive oils appear to cluster with certain gradation of the scores of a panel indicating oil quality.

In the second part, a 5-category classification was envisaged using either two-way PLS-DA or three way related methods such as PARAFAC-mlr-DA or N-PLS-DA PLS-DA models. Both response matrices, either two or three-way, were split up into two parts: a training set and a test set.

In the training phase a 5-category classification was envisaged according to each type of olive oil and four of the six measures in each category were used. The remaining two vials unseen by the model were used as a test set for further validation. Prior to any calculation, data were centered across the first mode for the three-way array and meancentered in the case of the two-way matrix. The training phase of the models was evaluated using a cross-validation method, and RMSECV was used to assess the optimal number of factors for model fitting.

Once the model had been trained, categorisation for test set samples was attempted. The predicted y-value from the calibrated models results in a continuous variable that can be interpreted as a class similarity index. Each calculated class prediction value can be compared with a Bayesian distribution curve to determine for a given predicted y-value the probability that this value belongs to that original class. Furthermore, a threshold of "predicted y" is determined above which a sample is considered to be a member of the class. These thresholds are also calculated for the three models and class assignment is done comparing against this value.

Regardless of which classifier method employed, all of them achieved a 100% success rate in the classification of the test samples. The N-PLS-DA model resulted less prone to overfitting in terms of variance explained in Y-block. Concerning the number of components needed to fit each model, the PARAFAC-MLR-DA built the simplest model and consequently the most parsimonious solution. Furthermore, due to its uniqueness property using PARAFAC simultaneously provide a model that uniquely describes which latent pehnomena are crucial for describing the variations in the dependent variable Y-block. Nevertheless, taking a look to RMSECV and RMSEP, we can conclude that despite the fact that all the models are able to classify new samples in a correct way, PLS-DA seems to outperform three-way methods. The small differences between the averaged two-way model and the trilinear model demonstrate that both can be sensible approaches for olive oil discrimination. In this particular application, the introduction of the third dimension does not improve the performance of the MS-sensor device. However the parsimonity observed in the PARAFAC model and the fact that the performance remains quite acceptable leads us to consider that this model is more robust and may outperform classical approaches in a broader set of applications.

4.5 ANALYSIS OF IBERIAN HAM QUALITY THROUGH USE OF A MS-SENSOR DEVICE

4.5.1 Introduction

The products derived from Iberian pigs constitute an important economic activity in Spain and much attention has been devoted to the study of dry cured ham in recent years. According to the rearing system, Iberian hams can be classified into two basic categories: acorn hams (from pigs fattened outdoors, feeding based on acorn and pasture land) and fodder hams (from pigs fattened indoors, feeding based on concentrated feed). The feeding that pigs receive contributes in a remarkable way to the sensorial characteristics of hams

such as flavour; those most appreciated being that ones coming from pigs fattened with acorn.

It has been previously reported that the volatile compound composition of ham is markedly affected by pig feeding. Currently, the characterisation of Iberian ham as a function of their diet is an issue to which many analytical efforts have been addressed. Nowadays, manufacturer producers need fast and reliable methods to distinguish the feeding that the animal has received. This method will help regulatory authorities and even final consumers to avoid frauds in the Iberian ham commercialization. Since the measure of volatile profile of Iberian ham seems to be informative from its quality state, MS-Sensor could be a valuable tool for general assessment of these properties.

As stated in the previous section, considering the time dimension in MS-Sensor measurements through the use of multi-way methods can enhance the pattern recognition step and therefore MS-Sensor performance. Including the time dimension in the modelling of the data could result in a simple way to tackle the poor specificity of m/z sensors. This poor specificity together with a high dimensionality are some of the most important issues that need to be dealt in the MS-Sensor approach.

4.5.2 Goals

- To study the feasibility of a MS-Sensor device to classify Iberian Hams according to pig feeding.
- To apply multi-way classifier algorithms to MS-Sensor response and compare the performance of two-way versus three way classifier models in Iberian ham discrimination.
- To compare quantitative PLS, unf-PLS and N-PLS models for predicting composition parameters of hams such as NaCl, a_w, and *X*.
- To develop a reliable strategy to reduce dimensionality in a comprehensible and easy way

4.5.3 Results

In ***Paper VIII*** the performance of PLS-DA and models containing temporal information such as unf-PLS-DA and N-PLS-DA are compared, proving the advantage of the three-way approach against classical two-way algorithms in the classification of two different hams qualities in accordance to the type of feeding the pigs receive.

Additionally, composition parameters such as water content (χ), water activity (a_w) and salt content (NaCl) are correlated to MS-Sensor measurements using PLS, N-PLS and unf-PLS regression models. Again, N-PLS compares favourably respect to two-way methods in this quantification problem.
Since multi-way data analysis implies the use of the time dimension, any unwanted variation along retention time has to be addressed. In order to achieve the alignment of MS-Sensor signals along the time dimension a recent published algorithm (Recursive Aligment Fast Fourier Transform, RAFFT) has been adapted and applied.

From the results derived from this study it can be concluded that, in this particular case, pattern recognition on MS-Sensor data clearly beneficiates from including temporal information either for quantification or classification purposes, since improved quantification and classification can be observed by using three-way methods.

Nevertheless, the way in which temporal information is introduced into the model is an important issue. In this context and in the particular case under study it has been demonstrated that the data generated by the MS-Sensor is better arranged in three way arrays rather than in an unfolded way. The main reason is the lack of trilinear constraints in the unf-PLS model, since the information across scans (temporal information) is not used to stabilize the solution. Unfolding methods lead to more unfavourable results for several reasons such as producing complex models, with many parameters which increase the risk of poor predictive capability, and the generation of less robust, interpretable and parsimonious models. Therefore, it is reasonable to use multi-way methods as they have the

potential to simplify the interpretation of the results and provide more adequate and robust models using fewer parameters.

The specific results about ham quality classification or the prediction of composition parameters obtained by the different models built can be a little optimistic, since replicates of the same ham have been always used in the models. So, to some extend, we have been evaluating how well a sample predicts its own replicate. However, since the same strategy was employed to assess the performance of N-PLS-DA, unf-PLS-DA and PLS-DA, the results showing that N-PLS-DA leads to more parsimonious models with higher prediction ability remains fully consistent.

In ***Paper VI***, a new strategy for feature selection was introduced and applied to the Iberian Hams dataset. The performance of this new strategy was previously demonstrated in a synthetic solvent mixture dataset with a well-known composition of its headspace. This simple database has been used to show that the feature selection process thereby developed was able to identify a minimal set of m/z fragments that enables the correct discrimination between mixtures using a simple fuzzy ARTMAP classifier. Furthermore, given the simple nature of the problem envisaged, it was possible to show that the fragments selected "made sense", that is, they were characteristic ionisation fragments of the species present in the mixtures to be discriminated.

Once the correct performance of the feature selection method was demonstrated, it was applied to the Iberian ham dataset. The feature selection algorithm was conducted in three steps. The first two steps were aimed at removing noisy, non-informative and highly collinear features (i.e., redundant), respectively. These two steps were computationally inexpensive and allow a dramatic reduction on the number of variables (near 80% of the initially available features are eliminated after the second step). The third step makes use of a stochastic variable selection method (simulated annealing) to further reduce the number of variables.

For example, applying the method to the Iberian ham dataset resulted in the number of features being reduced from 209 down to 14. Using the surviving m/z fragments, a fuzzy ARTMAP classifier was able to sort ham samples according to producer and quality (11-category classification) with a 97.24% success rate. The whole feature selection process runs in a few minutes in a Pentium IV PC platform.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 5

Discussion

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

5. DISCUSSION

5.1 THE IMPORTANCE ON HAVING PREVIOUS KNOWLEDGE OF THE HEADSPACE COMPOSITION

In accordance with the *"electronic nose"* philosophy, a priori knowledge of the components present in the headspace being analysed should in principle not been required. Nevertheless, MS-Sensor often deals with very complex food matrices and their volatile composition is usually formed by hundreds of substances covering a wide range of polarities, volatilities and solubilities. That is why taking decisions on set-up parameters, i.e., the choice of the sampling technique or the m/z range to be scanned is not an easy task. Therefore, previous knowledge of volatiles could result very helpful in such cases and it greatly simplifies these decisions.

In most of the quality applications studied in this thesis, previous GC-MS analyses of headspace extract of matrices under analysis were used to characterize the substances present in the extract. Thus, in ***Paper II***, for example, some GC-MS measurements of the headspace of fresh and rancid crisps were done for two different purposes. In one hand they were conduced in order to confirm that the rancidity treatment undergone by crisp samples give raise to a different rancidity stage reflected in different chromatographic profiles. On the other hand, characterization of the headspace was sought for determining the volatiles

involved in rancidity appraisal. Figure 5.1 shows the main differences between the chromatographic profile of rancid and a fresh crisp samples. From this figure it can be depicted that different volatile molecules appeared or substantially increased their signal intensity as rancidity develops. Some of these components, such as acetic acid, pentanal, hexanal, heptanal and hexanoic acid, were previously reported in the literature to be present in the chromatographic profiles of rancid crisps. The presence of 2,4-decadienal, which in our chromatograms was present in a similar intensity in fresh and rancid crisps, was also reported as a predominant note in deep-fried potato crisps.

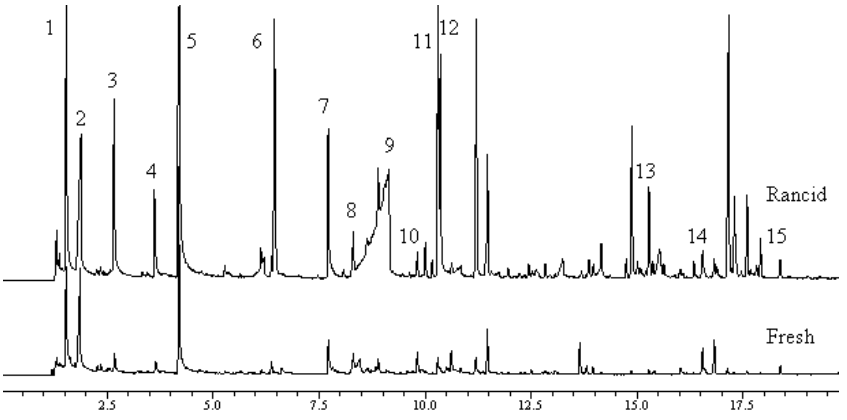

Figure 5.1: Chromatographic profiles identified by GC-MS (1) acetone, (2) acetic acid, (3) pentanal, (4) pentanol, (5) hexanal, (6) heptanal, (7) 2-heptenal,(8) 1-octen-3-ol, (9) hexanoic acid, (10) 3-octen-2-one, (11) 2-octenal, (12) 2,3-octanedione, (13) 2-decenal, (14) 2,4-decadienal, (15) undecane

Another example is ***Paper VII***, where the evolution of the headspace of sardines stored under cold storage from day 4 to day 11 was followed performing SPME/GC-MS analyses. There were ten different compounds identified (trimethylamine, acetone, 2-butanone, ethanol, 2-butanol, 3-methyl-1-butanol, ethylacetate, dimethylsulfide, dimethyldisulfide, dimethyltrisulfide). The main concern in this particular case was to support the results obtained by a MOX *"electronic nose"*. These compounds were found to be present from the fourth day until the day eleven. They proved that changes in the conductance of the MOX *"electronic nose"* developed within the context of this work could be justified by an increasing concentration of volatiles given out by sardines as a function of storage time. Figure 5.2 shows the evolution of these ten components with the

number of storage days. Again, the compounds found to be present in the headspace results in a good agreement with the substances previously reported in the literature.

Once the main volatile analytes present on the headspace were identified, the m/z ratios of their three most intense ions were used as the input variables to perform a PCA analysis of the MS-Sensor response. In this way, the information given by the chromatographic measurements was subsequently used for customising the models used in further pattern recognition. That results in an improvement of model since the inclusion of irrelevant and redundant information is avoided. In this example, it is very clear that pattern recognition may beneficiate from this previous knowledge of samples.

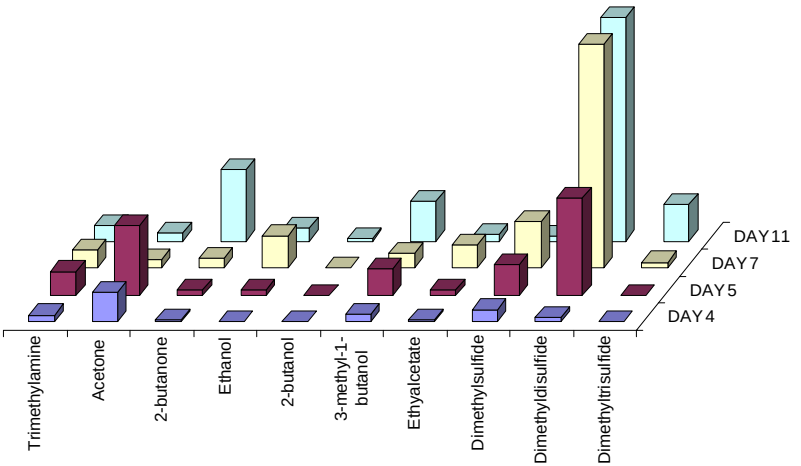

Figure 5.2: Comparative histogram showing the evolution with the number of storage days for the 10 volatiles identified in the headspace of sardine samples by SPME/GC/MS.

Paper IV reveals the importance of the previous knowledge from a pattern recognition point of view. In this paper a new method for an effective feature selection was introduced. The usefulness of the new feature selection method was assessed using a synthetic database. This synthetic database consisted of measurements taken from samples with a well-characterised headspace (the dataset was composed by different dilutions of trichloroethylene, 1-butanol, ethylbenzene and toluene). Since in this particular case the headspace was completely well-known, this database was a good benchmark for the feature selection method introduced. The attribution of chemical sense to the m/z fragments could

be done in a straightforward manner. The results show that the three-step feature selection process introduced was able to find the essential information needed to solve the discrimination problem considered. In the end, only 3 features were selected: m/z=46, which is one of the most relevant mass to charge ratios for toluene and not found for the other compounds in the solvent mixture; m/z=56, which is the most relevant mass to charge ratio for 1-butanol; and finally m/z=106, the second most relevant mass to charge ratio for ethylbenzene.

Once the feasibility of the feature selection algorithm developed was proven, a real classification problem was envisaged. The main concern was to sort eleven types of Iberian ham samples according to their different manufacturers. The dimensionality of the input data was reduced using the algorithm previously developed. This action resulted in the number of features being reduced from 209 down to 14. In principle, some of the selected m/z fragments held a chemical meaning. The m/z fragment 114 selected in the model is present in the mass spectrum of nonanal, and therefore, helps in discriminating between acorn and fodder fed pigs. Other fragments selected such as m/z 77 may arise from aromatic volatiles, m/z 71 from esters, alkanes, propylketones and butanoate, and m/z 45 could be due to the presence of carboxylic acids or alcohols. Finally, the presence of pentylketones and methylketones is revealed by m/z 56 and 58, respectively. All these compounds have been reported to be characteristic of the headspace of dry cured Iberian hams. Nevertheless, in this particular case the selected m/z fragments had not supported chemical evidence such as chromatography and they could be only explained based on previous reported results.

In ***Paper I*** chromatographic runs were also performed (results not shown) on the headspace of in-vitro and in-vivo growing fungal cultures. The main volatiles identified (ethanol, acetone, acetic acid, ethyl acetate, 2-methyl-1-propanol, 2-pentanone, 3-methyl-1-butanol and 1-octen-3-ol among others) resulted in good agreement with the reviewed literature. Again, these compounds were used to support the scores and loading plots of blanks and fungal contaminated vials. Fragment m/z 44 was found to be significant and it was associated to the presence of CO_2. The MS-Sensor differentiates between inoculated

and uninoculated samples mainly on the basis of the production of CO_2 by fungi. Actually, the relationship between the evolution of accumulated CO_2 and fungal growth was reported so far. The presence of other ions such as m/z 41, 42, 43, 45, 46, 55, and 57 can be considered as second-order and less relevant ions. They can be associated to other related fungal metabolites such as ethanol (m/z 45 and 46), 3-methyl-1-butanol (m/z 41, 42, and 55), 2-methyl-1-propanol (m/z 41-43), and 1-octen-3-ol (m/z 57) that were found in previously chromatographic runs. Again, this is in good agreement with the review paper by Magan and Evans[136]. In their study they reviewed the types of volatiles produced by grain spoilage fungi and listed the most common volatiles found and the fungal species involved. The major volatile compounds were found to be 3-methyl-1-butanol, 2-methyl-1-propanol, 1-octen-3-ol, and other 8-carbon ketones and alcohols. The same sort of volatiles was used further in ***Paper V*** which reports about the correlation among MS-Sensor measurements and ergosterol content (as a measure of fungal presence).

From all these examples it can be derived that a previous knowledge of the headspace extract composition clearly beneficiates the method development phase. In one hand it is very helpful as a starting point to set-up the MS-Sensor analysis conditions such as the sampling technique or m/z range to scan. In this way, it is possible to focus the mass spectrometer on the detection of the fragments of the molecule of interest and increase its sensitivity. On the other hand, previous knowledge on the headspace composition is absolutely necessary when assesing the performance of the models implemented in the pattern recognition step, especially when applying variable selection algorithms. It would always be desirable to obtain supporting chemical evidence to validate either the selection of variables using a statistical approach or the loading interpretation of the classifier fitted models.

There are different screening methodologies which can be used to characterize the headspace of a sample. In this thesis the use of GC-MS has proven to be highly valuable since it allows the identification of the headspace products. Even in the worst case where chromatographic measurements are not allowed, the pertinence of the fragments found to be relevant in the modelling of the data should indeed be assessed by the user knowledge of

the product. Hence, an extensive literature investigation about the key compounds for the application under analysis could be a very helpful solution. It is important to note that when working with a MS-Sensor device, even in the optimal case where GC-MS measurements are allowed and the headspace characterization is fully available it is difficult to categorically relate a mass fragment of a spectral "fingerprint" to a specific volatile component since m/z fragments are not selective. That means that the same mass fragments may come from different molecule origins. In spite of that, the knowledge about the ions relevant for a certain application will always delimit the different approaches to the quality problem under analysis leading to a more specific solution.

5.2 HANDLING VOLATILES

Once a previous screening of the headspace has been accomplished, the next step is to select the most suitable sampling technique and/or sampling preparation procedure for the applications under study. As stated before, and according to Marsili[21], sample preparation is complicated by a number of factors: *a)* concentration level: levels of aromatic compounds are generally low, typically in the range of ppm, ppb or ppt and thus it is often necessary to concentrate them by orders of magnitude; *b)* matrix effects: the volatile composition of foods are frequently very complex since they are usually formed by hundreds of substances covering a wide range of polarities, volatilities (volatile and semivolatiles), solubilities and pH. Thus it has to be emphasized that regardless of which sample is on study no single technique will prove optimal for this sample and often it is necessary to make use of more than one single sampling technique in order to perform a full screening of the headspace.

The importance of the sampling step is too often ignored. The success of the analysis will highly depend on the correct choice of this technique. In ***Paper I*** an initial study to select the optimal MS-Sensor sampling technique in the framework of fungal bakery spoilage assessment was conduced. Two different sampling techniques (SH and SPME) were evaluated according to their ability to discriminating among inoculated and

uninoculated vials and according to their ability to distinguish among different fungal genera on in-vitro fungal cultures. Table 5.1 shows a comparison of performances for these two techniques.

Table 5.1: Success Rate Comparison between the two different sampling techniques studied

	Discrimination between	#experimental points	Failures	% Succes rate
SH/MS, 100ºC	fungal growth	16	2	88
	genera	16	7	56
SH/MS, 0 min	fungal growth	32	1	97
	genera	32	13	59
SPME/MS	fungal growth	72	0	100
	genera	72	0	100
	species	72	6	92

The SPME sampling method gave better results than the SH technique, leading to a 100% success in either genera discrimination or in discriminating among inoculated and uninoculated vials. SH performed at 100°C achieved an 88% success rate when trying to determine whether the vial was inoculated or not but the success rate decreased to 56% when attempting to determine the fungal genera. The SPME sampling technique compares favourably against SH when discriminating among genera. Moreover, with this technique the discrimination among different species was also possible. It achieved better repeatability and it was more sensitive due to its ability to concentrate volatiles. The main advantages of SPME against SH have already been reviewed in previous chapters. In general terms, the use of a fiber for the extraction enhances the selectivity and sensitivity of the analysis compared to the most commonly used SH.

Finally it should be mentioned that depending on the headspace sampling techniques, the pattern recognition engine will model either the most volatile part of the headspace extract or the semivolatile fraction. The selection of the sampling technique will lead to an unavoidable bias of the modelling to different types of volatiles present on the headspace. Hence, for example, the study on the quality of hams (***Paper VIII***) is attempted using SH as the sampling technique which means that the further data modelling will be

based on the most volatile analytes. If the SPME technique is the chosen method, then the results are going to be biased toward the semi-volatiles headspace analytes. That can be reflected in Figure 5.3 which shows the differences in raw spectra obtained from the same fungal headspace sample analysed using SPME or SH (T=50ºC & 50 min equilibration time).

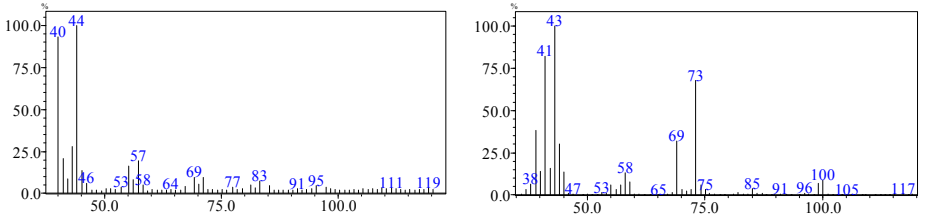

Figure 5.3: Mass spectra differences according sampling methodologie, SH, 50ºC, 50 min on the left and SPME on the right

5.3 MS-SENSOR AGAINST MOX *"electronic nose"*

Despite the fact that this thesis is mainly devoted to the study of the MS-Sensor devices, some studies based on MOX *"electronic nose"* have also been attempted. In ***Paper II***, the performance of both approaches is compared in the framework of crisps rancidity detection. The two systems showed to be able to distinguish between fresh and rancid crisps with a 100% success rate in a cross-validated Fuzzy ARTMAP classifier. When attempting to establish a semiquantitative measurement of the four different levels of rancidity the MS-Sensor was able to maintain the 100% success while the success rate of the MOX *"electronic nose"* decreased to just a 68%. These differences can be basically attributed to a highest readability, robustness and lowest drift attributed to the mass spectrometer detector when compared against the MOX prototype. Part of these differences could also be attributable to the different headspace sampling methodology used in each approach. As it has been mentioned above, data handling is an extremely important part of the whole analytical procedure. The success of the analytical procedure will be in a big part dependent on the success of the sampling methodology. The amount of sample volatilized, the headspace vial temperature, the scan range and the operation mode (either SIM or full scan) determine the sensitivity of a MS-Sensor. The sensitivity of solid state sensors is

determined by their type, the flow rate over the sensor, the analyte, and the operating temperature among other factors.

In ***Paper VII,*** MS-Sensor technology is used as complementary technique to support MOX *"electronic nose"* results, demonstrating that the combination of both techniques should not be ruled out and they can lead to complementary results when approaching a new application.

5.4 PATTERN RECOGNITION

5.4.1 Preprocessing

The first step in data analysis is to pre-process the signals generated by the MS-Sensor. This process transforms the data into the most appropriate form and enhances the features within the data that are useful in the subsequent pattern recognition. The philosophy of the data pre-processing step is to reduce variation in the chemical data unrelated to the chemical composition such as analytical variability and concentration effects: time shifts, baselines, concentration effects, and sensitivity changes related to, for example, fragmentation in the ion source and mass selective detector. Speed and semi-automatization with limited human intervention on the mathematical tools are a high priority.

5.4.1.1 Object-wise normalization

Differences in analyte concentration and ionization efficiency could influence intensities. Hence the normalization of the signal intensity is required. When the problem envisaged does not depend on the concentration of the analytes, the MS-Sensor dataset requires an initial preprocessing for the creation of a data matrix suitable for subsequent analysis. Normalization (object-wise standardization) is an important step in the preprocessing of MS-Sensor data. Depending on which cases, the difference in overall

headspace extract concentrations may be quite large. Multivariate analyses of unprocessed signals of headspace extracts might lead to results that are mainly biased towards the variation in overall sample concentration. Normalization procedures might have a strong impact on data clustering and therefore influence any final interpretation. Figure 5.4 shows PCA scores plot differences for the same fungal data set depending on if absolute intensity or relative intensity was used to get the two-way matrix response.

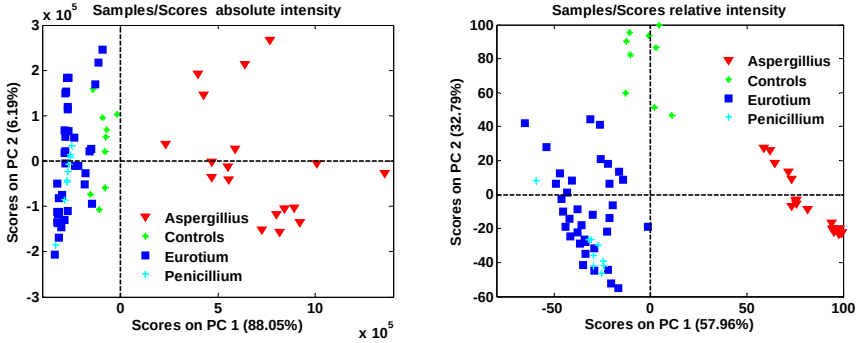

Figure 5.4: Object-wise normalization differences on the same dataset

In this particular case, better clustering results were obtained when relative mass spectra was used. That holds true in the cases where differences among samples are due to the release of new compounds on the headspace that gives rise to different mass spectra fragments. On the other side when the main differences lie on the concentration of the same compounds, then the absolute intensity could be the best choice. The choice of normalization criteria is not arbitrary. The influence of the normalization strategy can be very large, since the range of signal values for different compounds in individual samples is often considerable. Although influential, it is also difficult to determine what type of normalization to use since the optimal approach depends on the context of a specific problem.

5.4.1.2 Variable-wise normalization

Changes in mass spectra peaks originating from compounds at high concentration would result in large peaks, with large absolute fluctuations between samples. This is not

always desired since other compounds giving rise to smaller peaks and peak variation might be equally interesting. Autoscaling can solve this problem by bringing all variables to the same scale. Autoscaling all differences in size between variables are eliminated. In the other way around, autoscalign magnifies the baseline variation since variables (i.e., m/z) representing only noise will also be transformed to the same scale as all other variables. Baseline or noisy m/z fragments will thus become as important as the chemical variation between samples.

Centering converts all the m/z variable intensities to fluctuations around zero instead of around the mean of these variable intensities. Therefore, it adjusts for differences in the offset between high and low abundant m/z. It is used to focus on the fluctuating part of the data, and leaves only the relevant variation (being the variation between the samples) for analysis.

Different pretreatment methods emphasize different aspects of the data and each pretreatment method has its own merits and drawbacks. The choice for a pretreatment method depends on the food quality question to be answered, the properties of the data set and the data analysis method selected. In conclusion, selecting a proper data pretreatment method is an essential step in the analysis of MS-Sensor data and greatly affects the outcome of the analysis. Nevertheless there are no a priori assumptions that can be rule out for applying different data transformation since scaling and normalization are highly problem dependent issues.

5.4.2 Feature selection

Feature selection algorithms have three goals: to improve classification accuracy, to reduce the cost of extracting features and to improve the reliability of the performance. In ***Paper III*** either variable selection or variable reduction techniques have been applied for improving the performance of the models fitted to the crisps dataset. The first technique implemented is based on a PCA variable reduction to reduce the dimensionality of the input

matrix. A Fuzzy ARTMAP classifier using 78 input variables was used, the average classification rates of crisp samples in four rancidity stages was 88%. A PCA filter was used to reduce the number of features to 10; with this input the classification of crisp rancidity improved to 93%. The same criteria was used in ***Paper I*** where besides to performing a PCA variable reduction, the PCA reduced data matrix was used to train a DFA and the range scaled eigenvalues obtained used as input in Fuzzy ARTMAP. These two applications demonstrate that applying a PCA analysis leads to a linear combination of m/z variables that gathers the highest amount of variance and compresses information by eliminating redundancy and collinearity. In this manner, the best combination of m/z variables can be chosen without the need to perform a costly and lengthy initial study to determine the most relevant ion fragments to be monitored. Since a reduction in redundancy and collinearity of data is feasible using this method, an improvement on the performance of further classifier models is expected. Nevertheless, this approach still presents some weakness. A PCA variable reduction approach uses the full spectrum (e.g., including noisy or redundant m/z ratios) to compute the factors. The selection of an optimal subset of PCA-factors is not necessarily straightforward because the magnitude of an eigenvalue is not always a measure of its significance for the calibration. Furthermore, unlike m/z ratios, the PCA computed factors have no direct chemical meaning. That is why a method aimed towards the direct selection among m/z ratios was developed.

In ***Paper II*** different PLS calibration models were built and validated in order to investigate whether the results of a MS-Sensor device correlates or not with the results of the well established analytical methodologies currently used for assessing rancidity in crisps such as Rancimat and ADV tests. The underlying objective was to assess if the MS-Sensor can predict hydrolytic rancidity, oxidative rancidity or both. It is now widely accepted that multivariate calibration techniques such as PLS greatly benefit from an appropriate variable selection. Genetic algorithms have been applied for variable selection in the building of calibration models from spectral data (e.g., PLS) and have shown that they provide better results than full spectrum approaches. After variable selection, the results of the MS-Sensor device improved the correlation coefficients with respect to the original models without applying variable selection.

Finally, in ***Paper IV*** a new method for an effective feature selection is introduced. This is especially suitable for applications where the dimension of feature space is high, a significant degree of correlation exists between features and some of them are affected by noise, such as in case of MS-Sensor. The method is efficient in the sense that after the selection process, only those features that are important for the application considered are retained to build the pattern recognition models and all the process is conducted at a very low computational cost. Using the surviving features, a Fuzzy ARTMAP classifier was able to discriminate ham samples according to producer and quality (11-category classification) with a 97.24% success rate. It was also possible to identify, with a 100% success rate, whether pigs had been fed on acorn or fodder. For the different databases studied, performing variable selection results in a dramatic decrease in dimensionality and an increase in classification performance.

5.5 DOES THE TIME DIMENSION REPORT EXTRA INFORMATION ABLE TO IMPROVE THE MODELLING PERFORMANCE?

In order to assess whether the time dimension of the MS-Sensor approach is able to report extra valuable information, two different examples were studied. In one hand the discrimination of two Iberian hams qualities according to the rearing system that pig had undergone which is studied towards ***Paper VIII.*** On the other hand ***Paper VI*** studies if the addition of a temporal dimension could improve the modelling of MS-Sensor data corresponding to measurements of different virgin olive oils. This section gives an overview of the different ways to include the time dimension on the MS-Sensor response and the models derived in each case.

5.5.1 Exploratory data analysis comparison among models including temporal information and models that do not include it

In ***Paper VIII***, the focus was separating the hams from acorn fed pigs from fodder fed pigs. Exploratory data analysis was performed fitting PCA, unf-PCA and PARAFAC to the MS-Sensor data. The main concern was to study whether the models incorporating the

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-0752-3/DL:T-126-2009

time dimension (either the unfolded or three-way models) give rise to a clearer clustering of data according to the pig feeding than the averaged two way model.

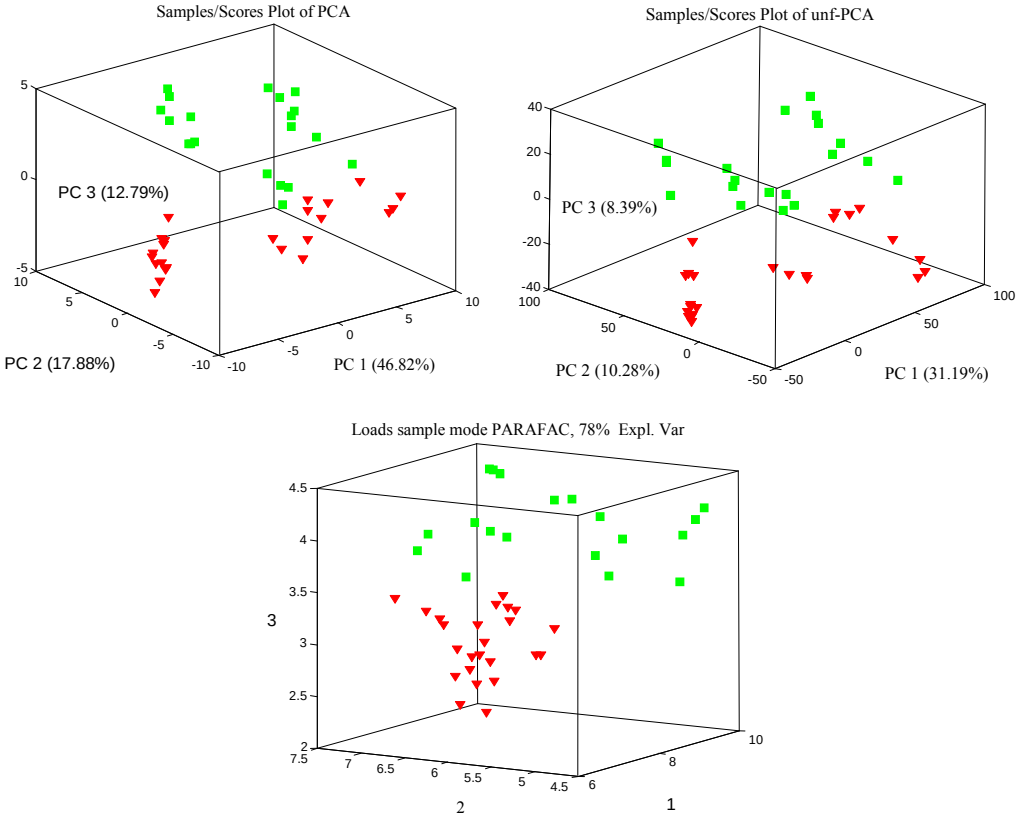

Figure 5.5: PCA, unf-PA and PARAFAC scores plot, acorn (red spots), fodder (green spots)

The number of components selected for each model was determined from a global model inspection using the following parameters: explained variance, size and structure of residuals compared to modelled data and raw data, split-half analysis, and visual interpretation of the scores and loadings. While simple quantitative diagnostics such as crossvalidation seem easier to use, it is important to realize that it only rarely provide correct answers when used for exploratory models as those used in this thesis. It is always necessary to use additional parameters such as the above mentioned for overall validation. Note that percentage fitted variation in itself does not provide a useful measure of appropriateness. For mathematical reasons, a model with more degrees of freedom (e.g.,

unf-PCA compared to PARAFAC) will always fit better. The problem is that the better fit will mostly be fit to noise if the mathematically simpler model (PARAFAC) is already fitting the systematic variation. The degree of overfit can be evaluated by comparing the fit with the cross-validated fit in the scree plot.

The three component model shows a much better clustering of samples according to the type of feeding than two-way models do as it is showed in the scores plot comparison between the three models fitted (Figure 5.5). In all cases secondary effects such as ham origin or production process affect clustering since data is markedly influenced by these effects. This is much more pronounced in the case of the PCA or unf-PCA where the data appears to be clustering according to producers instead of feeding type. Hence it seems that the three-way model is modelling little bit better the underlying phenomena that is wanted to observe.

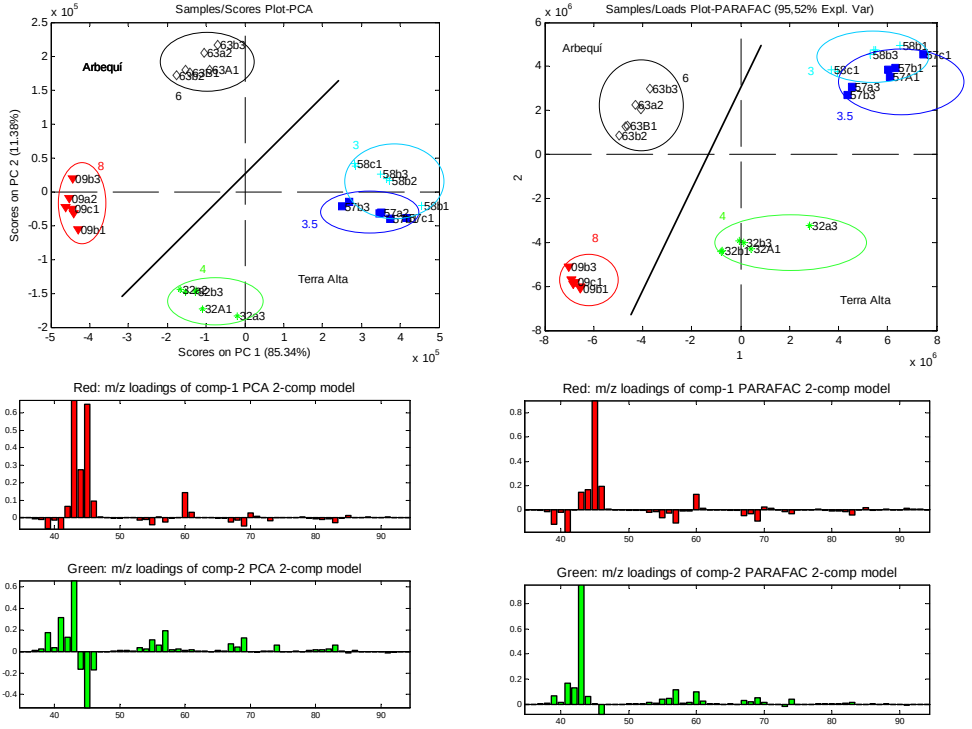

Figure 5.6: PCA scores plot and PCA and PARAFAC loadings (first and second mode)

Again, in ***Paper VI*** exploratory data analysis of MS-Sensor measurements of five different olive oils was attempted using either PARAFAC or PCA modelling. The overall aim was to distinguish among the different olive oil sources. Figure 5.6 shows scores and loadings plot for both. The PARAFAC model includes the time dimension and in the PCA model the time dimension was neglected by averaging the mass spectra response.

From this figure it can be depicted that in both cases a clear clustering of the replicates of the same olive oil are obtained. The two overlapping categories correspond to different samples from the same geographical origin and get similar scores. Component 1 which holds the highest amount of gathered variance is able to sort olive oils according to their organoleptic properties or the scores given by a panel of judges. Moreover, the black line indicates differences on protected regions. Concerning the loadings plot for the first m/z mode, it can be concluded that the same m/z fragments are considered to be relevant for both models in spite of some small differences.

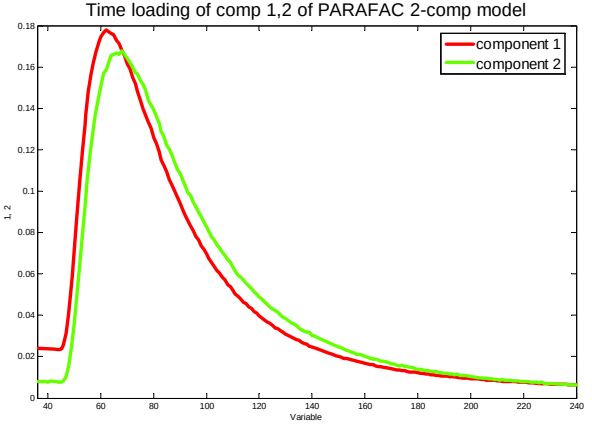

Figure 5.7: Second mode loading matrix for 2-component PARAFAC

It has to be mentioned that PARAFAC modelling has a second loading matrix which handles temporal information. Due to the uniqueness property of PARAFAC this loading matrix is meaningful from a physicochemical point of view. Under the assumption of trilinearity fulfilment and the assumption that a correct number of components are choose, this loading can be interpreted as time dimension profiles as showed in Figure 5.7.

In this particular case, it makes no difference from the unsupervised modelling point of view whether the model includes the temporal dimension or not. In general terms PARAFAC and PCA are modelling exactly the same phenomena and clustering of the data remains the same.

5.5.2 Comparison of classifier and regression models including temporal information and models that do not include it

In ***Paper VIII*** a 2-category classification was envisaged according to the type of feeding that the pigs received using N-PLS-DA, unf-PLS-DA and PLS-DA. Table 5.2 shows the different parameters obtained for each one of the models. Several conclusions can be drawn from this table. The most important conclusion is that N-PLS-DA is the simplest model (lowest number of components) and at the same time it shows the highest predictability in terms of RMSEP (since it has the lowest value). Although PLS-DA is able to fit with the highest percentage of variance in X-block, Y-validation variance decreases dramatically and that is translated into an increase of RMSEP. Thus, the PLS-DA model shows the lowest predictability.

Table 5.2: Characteristics of the N-PLS-DA, PLS-DA and unf-PLS-DA models calibrated for the classification of ham samples according to the type of feeding.

		% Explained variance				RMSEC		Success Rate (%)
	# LV"s	Xcal (Fit)	Xval (Val)	Ycal (Fit)	Yval (val)	RMSECV	RMSEP	
N-PLS-DA	6	72,03	71,61	91,52	77,93	0,220	0,289	100
PLS-DA	13	92,39	86,80	99,23	25,72	0,117	0,429	100
unf-PLS-DA	11	64,39	7,83	99,99	66,41	0,337	0,384	95

Fit in the X-block is better in the case of PLS-DA because of the higher flexibility of this model against to its trilinear generalization. The fit of a trilinear model will be lower per definition that the fit of the corresponding bilinear model because in N-PLS-DA any variation must be consistent over all scans. The PLS-DA model is often overly flexible and

the increased fit to a large extent attributable to fitting the noise of the data and it is not directly translated to an increase on predictability as it can be seen in this particular case. The unf-PLS-DA model even being a bilinear model shows the lowest fit in the X-block. The point is that, by unfolding, no relationship between different scans is imposed and in some way the structure of the data is destroyed. It has to be kept in mind that an unfolded matrix holds all the scans acquired, even the ones that do not carry usable chemical information. An autoscale of the unfolded matrix leads to put on the same level meaningful m/z variables from scans holding signal and from scans just holding noise. Therefore, this matrix contains a lot of irrelevant information not related to the feeding type and that is why it is just fitting 7,83% of the variance in the X-block, which allows to describe 66,41% of the variance in Y. Even with the poor percentage of variance accounted in X-block at prediction, this model it is able to predict the feeding category of the samples better than a PLS-DA does in terms of RMSEP, proving that the inclusion of the time dimension helps to improve the prediction ability of the models.

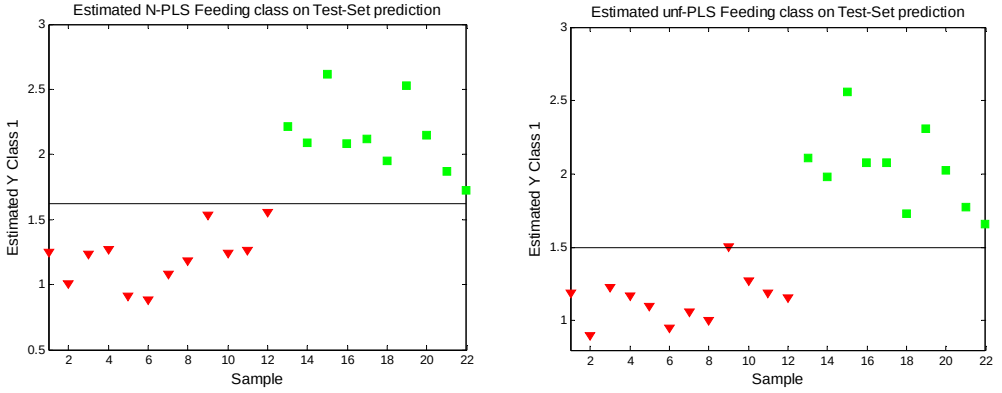

Figure 5.8: Acorn (red spots) or fodder (green spots) class prediction comparison between N-PLS-DA, and unf-PLS-DA approaches for the test set samples.

The predicted Y-value from the calibrated models results in a continuous variable that can be interpreted as a class similarity index. Each calculated class prediction value can be compared with a Bayesian distribution curve to determine for a given predicted Y-value the probability that this value belongs to that original class. For each sample, either in prediction or in the test set, a threshold of "predicted y" can be determined above which a

sample is considered to be a member of the class. These thresholds are calculated for the three models and class assignment is done comparing against this value.

Figure 5.8 shows class predicted values for the entire test set samples for N-PLS-DA and unf-PLS-DA. The black horizontal line is the calculated threshold. As it can be seen, in the case of the unf-PLS-DA model there is a misclassified sample. Even though the predictability of test set samples (i.e., the RMSEP) looks better for unf-PLS-DA than for PLS-DA, the former model presents a misclassified sample while the latter is able to classify all test samples in a correct way.

That can be explained with the certainty of predictions. In order to compute the certainty of each model, values for reduced residuals (Q) and Hottelling's statistics (T^2) are used. The larger the values, the less certain is the class assignment. Q statistics are calculated for the residual part and they simply express the fit of the test data set.

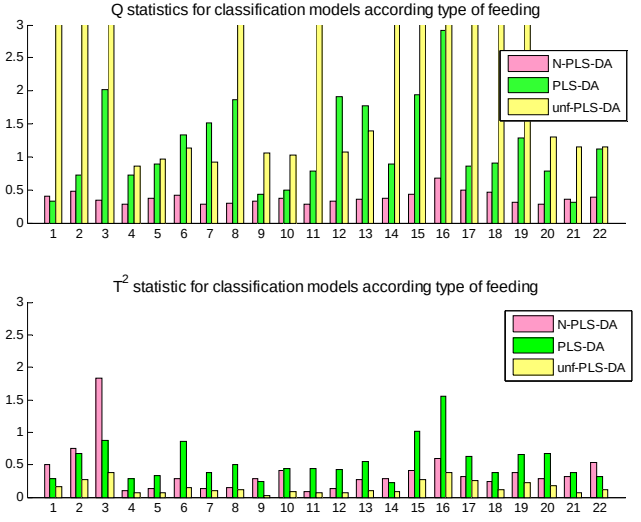

Figure 5.9: Value comparison between N-PLS-DA, PLS-DA and unf-PLS-DA for reduced residuals Q and T^2 (Hottelling"s) statistics.

Figure 5.9 shows that 11 of the 22 samples of the test set have very high values for Q in the case of unf-PLS-DA data, hence this model is not able to fit at least half of the

samples on the test set; therefore, it can be concluded that this approach is not able to generalize in a reliable manner. At the same time, and according to the explained variance in prediction, N-PLS-DA shows the best generalisation ability with the test set. T^2 (Hottelling's statistics) expresses unusual variance of the data within the model; it is an indication of how far the projection of the samples are from the multivariate mean. It can be interpreted as a measure of the leverage of samples, i.e., the influence of samples inside the model. Trend on leverage is similar for all samples and for all models and extraordinary variations were not produced in leverage.

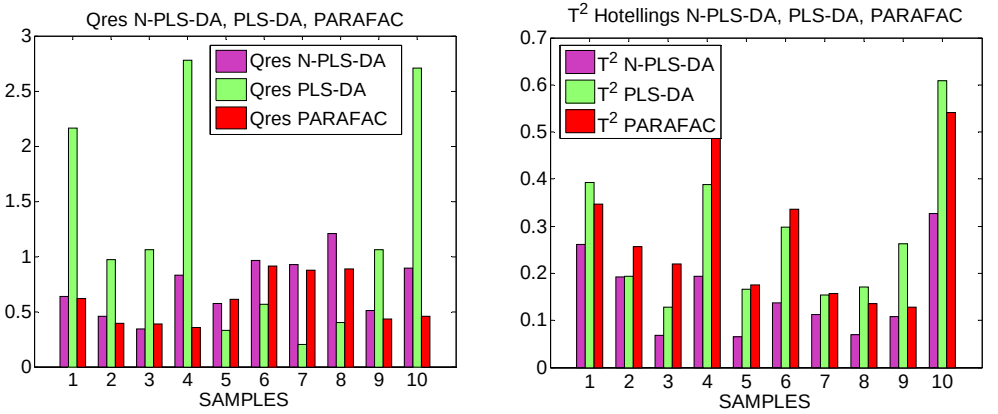

Figure 5.10: Q and T^2 statistics comparison for the classifier models

In ***Paper VI*** three classifier models, namely bilinear PLS-DA, trilinear N-PLS-DA and PARAFAC-MLR-DA (multilinear regression discrimination using loads for first mode as X-block predictors) are applied and compared in the framework of a classical MS-Sensor application such as olive oil discrimination. In this example the use of the unfolded models was not considered. However, a very interesting model was built based on the use of the loadings on the first mode of a PARAFAC in a regression way. It was so due to the uniqueness property from PARAFAC which provides a model that uniquely describes which latent phenomena are crucial for describing the variations in the olive oil headspace. The three attempted models achieved a 100% success rate in the discrimination of the five different olive oils in cross-validated routines. In this case, it could not be clearly stated that the models that included the time dimension outperformed the two-way models in terms of predictability. However, when a residual analysis is attempted (Figure 5.10) one may

consider that N-PLS-DA is less prone to overfitting and the dispersion inside the groups is lower that for the other methodologies. PLS-DA is again pointed as the model having the highest degree of overfitting

In general N-PLS-DA compares favourably to other models for several reasons. It is the simplest one, it is more stable than the rest since it fits data correctly, and it is able to generalize to new samples. Moreover with this model the risk of overfitting is minimized and it shows the highest predictability. Although unfolding may often be helpful as a way to consider the time dimension in the MS-Sensor data response, it has several disadvantages. For example, the three-way structure of the data is lost, the number of variables of the resulting matrix is high because of the combination between the variables of two of the modes, and the resulting models are more difficult to interpret[113].

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 6

Conclusions

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

6. CONCLUSIONS

The suitability of the MS-Sensor approach for food quality control purposes has been widely demonstrated through all the applications developed within this thesis. Below, a summary of the main conclusions derived from the development of each one of these applications is given:

1) A new method to assess oxidative and hydrolytic rancidity based on the use of an SPME/MS-Sensor has been developed. Unlike traditional methods, the new approach does not rely on the previous extraction of oil from the samples. Instead, the method evaluates volatile compounds directly from the headspace of crisps. The effectiveness of the system in the assessment of the quality of crisps has been demonstrated in two applications. First, it was used to classify the crisp samples according to four pre-established rancidity stages achieving a high success rate in classification. The system is always able to perfectly discriminate between fresh and rancid crisps. In a second study, the correlation between MS-Sensor measurements and other traditional methodologies for assessing oil quality such as the acid degree value or rancimat hours were also investigated, leading to a high rate of correlation.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-0752-3/DL-T-126-2009

2) The performance of a MS-Sensor on the detection of in-vivo and in-vitro fungal spoilage on bakery products has also been studied. Within this context, it has also been demonstrated that this approach could be a highly valuable tool for a real quality analysis application in a bakery factory. A SPME/MS-Sensor configuration has been shown to be useful to predict spoilage of bakery products in less than 7 days since inoculation, i.e., before sporulation becomes visible. This period of time would enable sampling of this type of products (shelf life of 5–8 weeks) and rejection of the selected batches before their distribution to markets. In addition, this approach overcomes many of the difficulties encountered on traditional *"electronic noses"*, such as repetitivity, selectivity and drift issues. The instrument offers a fast alternative that can be easily automated and operated. In addition, the MS-Sensor system was further used to build models for prediction of ergosterol as an indication of mould growth.

3) The feasibility of using a MS-Sensor system to assess the freshness of sardines under cold storage has also been reported. An SH based MS-Sensor prototype has been used to evaluate the temporal headspace composition changes with increasing days of storage. In parallel, the results have been compared to the volatiles rising to the headspace of sardines identified by SPME/GC-MS. Results obtained with the MS-Sensor approach have been used as a supporting evidence of the results derived from fish spoilage monitoring using a dedicated gas sensors based *"electronic nose"*. It has been demonstrated that both technologies would be useful for testing the quality of fresh sardines along the logistics chain.

4) A SPME/MS-Sensor configuration has obtained good results in the classification of olive oils samples according to their origins or their organoleptic attributes. Differentiation of virgin olive oils sources using a MS-Sensor approach could be considered a feasible alternative tool to the conventional methodology currently used for this purpose (organoleptic assessment by a panel of experts or chromatographic based methods). Even though these conventional techniques could provide fundamental information about the composition, quality and authenticity of virgin olive oils, the MS-Sensor methodology presents some important advantages. Sample preparation steps are eliminated and this

allows a reduction in the time and cost of the analysis. Hence, the conventional methodology can be considered more convenient when specific characterization studies are performed and the MS-Sensor approach could be the choice for routine control of an elevated number of samples, since the technology allows a higher throughput.

5) A MS-Sensor configuration was successfully used for the classification of Iberian hams depending on the feeding diet of pigs or as a function of the producers. Moreover, correlation of the MS-Sensor measurements with production parameters (a_w, χ and NaCl) was also performed. This methodology reports many advantages in comparison with current analytical methods used for the characterisation of Iberian hams such as GC-MS determination profiles of free fatty acid. Analysis sample time and cost is generally minimized since minimal sample treatment is required.

6) Most of the reported applicability studies of MS-Sensor noses to different aspects of quality assessment in dairy products show satisfactory results. Nevertheless, in most cases the results will have to be confirmed on a larger scale to make sure that the data analysis results obtained, either classifications or quantifications, are still valid with a larger intra-group variability, which is generally found in the case of this type of products.

As stated in previous chapters, the development of these applications intrinsically implies the study of parallel issues related to sample handling, modelling of data, etc. From the entire thesis, some highly important conclusions have also been found. They are summarized as follows:

7) Some published papers state that one of the greatest advantages of the MS-Sensor approach is the fact that it is not necessary to have a previous knowledge of a sample to obtain valid results. However, a conclusion that can be directly retrieved from some of the experiments done in the framework of this thesis is that, even not mandatory, it is very important to have as much information as possible about the samples to be analysed. The possibility of obtaining chromatographic information about the samples under analysis should always be considered. The importance of having as much information as possible

relies on the fact that holding a previous knowledge about the compounds present on the sample is essential to select the optimal sampling technique. It is very useful for guiding the analyst throughout the set-up and optimization of sampling parameters such as the selection of the fiber. It is also very helpful, or even essential, when the modelling of the data is attempted. Validation of models in the variable selection step is also possible if it is hold some previous knowledge of the sample. It would always be desirable to obtain supporting chemical evidence, such as chromatography, to validate the variable selection procedure and the meaning of the fragments selected should indeed be assessed by the user on the basis of the previous knowledge about the sample. This knowledge is also highly valuable when performing loadings interpretation in the exploratory data analysis; in fact, even if it is difficult to relate in a category manner a mass fragment of a spectral "fingerprint" to a specific volatile compound, it is nevertheless possible to deduce the main origin of certain ions when there is previous knowledge of the product.

8) A priori, any sampling headspace technique can be used as the sampling part of a MS-Sensor system. In practice, the choice of the technique should be selected taking into account the nature of the food matrix under analysis and the method specifications required. Different methods of extraction and injection of the volatile components make it possible to account for the specificities of the matrices studied. Therefore, ideally, when a new MS-Sensor method is under development, previous studies on available sampling techniques performance for the particular application study should be carried on. Adequate sampling techniques such as SPME can improve the signal to noise ratio by increasing the volatility or by increasing the concentration of the discriminating compounds relevant to the phenomena under study. The SH is particularly useful for the extraction of high volatile compounds. On the other hand, when dealing with low volatility or semi-volatile compounds preconcentration techniques such as SPME should be rather used. It is recommended the use of a preconcentration technique for MS-Sensors systems when attempting analyses of samples differing mainly in volatile compounds with middle or high molecular masses. For samples differing only in small molecular masses, the non-preconcentrated SH would be preferable. Nevertheless SH lacks from sensitivity. It has been demonstrated that the SPME technique gave better results than the SH technique when

applied as the sampling system coupled to MS-Sensor. It achieved better repeatability and it was more sensitive due to its ability to concentrate volatile analytes. Furthermore, due to its enrichment ability, the SPME technique is less efficient towards small molecular masses but can also extract compounds with higher molecular masses. SPME is ideally suited for coupling to a MS-Sensor system because of its simplicity, and its relatively moderate price. Besides, it can be easily automated for high throughput measurements since it is quite compatible with an autosampler. SPME sampling procedure fits perfectly the conditions required for quality control analysis in the dairy industry because it is solvent free, cheap, easy to use and relatively fast to execute.

9) MS-Sensor based instruments have been found to hold clear advantages over MOX based instruments in some of the applications developed within the context of this thesis, (e.g., in stability and versatility). MS-Sensor devices are much more versatile since the scan range of the detector determines the number of sensors. Thus, when working with a MS-Sensor it is not necessary to change the hardware for each application. In the case of MOX instruments each application requires a different sensor array set-up. Current weakness of MOX *"electronic noses"* include, among others, loss of sensitivity or poisoning in the presence of water and high polar analytes such as ethanol, sensor drift and the difficulty to provide absolute calibration and relatively short life of some sensors. Moreover, MOX *"electronic nose"* need in each new application to do much more considerable method development work than in the case of MS-Sensor. All these drawbacks have limited the potential of MOX *"electronic noses"* especially if they are compared with MS-Sensor devices. Anyway, the most important advantage of the MS-Sensor configuration is that this approach is able to generate directly-interpretable chemical information. Following multivariate analysis of the MS-Sensor data, evaluation of the spectral residuals may provide further insight. For instance, the residual of an out-of spec sample may be searched against a mass spectral library. Classifications performed by the MS-Sensor are based directly on chemical features, and therefore multivariate models become fully interpretable from a chemical point of view. In that sense it has been demonstrated that MS-Sensor can be used as a complementary technique to support MOX *"electronic nose"* results.

10) It has been demonstrated that preprocessing techniques are of paramount importance in the success that can be achieved with pattern recognition algorithms. Differences in data preprocessing methods could severely change the outcome of the data analysis task. Unfortunately, a general rule or recommendation is difficult to give since algorithms and strategies to follow will be highly dependent on the data and the application under study. Thus, in general, preprocessing is difficult to validate, and a correct strategy is by definition not identifiable. Noise correction, mass calibration, intensity normalization, and variable selection methods are crucial and should be carefully evaluated for each application. Different preprocessing techniques are available, and only the expertise of the researcher will determine the choices on the type of preprocessing for every dataset separately. Then, only soft validation is possible, making plots and looking at preprocessed data. Pre-processing in three-way arrays represents a much more challenging task and is not as straightforward as in two-way data. It should be done carefully and taking the considerations already described when centering or scaling of data is envisaged. Furthermore, when working with three-way data, the need for handling trilinearity emerges since three-way algorithms used in this thesis assume data to be trilinear. This property is highly affected by common artefacts present in chromatography such as missalignments from run to run and baselining. In this thesis, some new elegant methodologies for overcoming these artefacts are presented.

11) Among the different preprocessing steps, one of the most studied throughout the development of this thesis has been feature or variable selection. High dimensionality is inherent to MS-Sensor applications where hundreds of variables per measurement (m/z fragments), a significant number of them being highly correlated or noisy, are available. Feature selection is, therefore, an unavoidable pre-processing step if robust and parsimonious pattern classification models are to be developed. It has been demonstrated that a correct choice of the m/z variables that bests describes the application sought is of vital importance for the MS-Sensor to perform well. Course of dimensionality phenomena has been studied through the implementation of new variable selection (stochastic algorithms such as SA or GA) or variable reduction algorithms (PCA, LDA). Some of them are introduced and tested for the first time in the data coming from MS-Sensors. It has been

widely demonstrated that either variable selection or variable reduction can improve the performance of the models commonly used in the pattern recognition step.

12) Most applications with MS-Sensor devices generate vast amounts of data. In order to get a good understanding of the information in this data, a reduction of the information to a reasonable size it is necessary. Multivariate analysis based on projection methods leads to a reduction of the dimensionality and represents a number of efficient and useful methods for the analysis and modelling of complex data arising from MS-Sensors. In general, data description and exploratory data structure modelling has been performed through the use of unsupervised modelling: PCA in two-way data and its counterpart PARAFAC for three-way datasets. The use of these algorithms has been supported on the fundamental assumption that the properties under study, namely the hidden phenomena (rancidity, fungal spoilage…), are related to the directions that maximize variance. This is demonstrated by the fact that relatively a high amount of variance is explained within the number of components retained for each modelling. Both modelling allows extracting and displaying the systematic variation in the data. A PCA or a PARAFAC model provides a summary or overview for all the samples in the data matrix. In addition, natural trends or grouping can also be found. Furthermore, the analysis of residuals is a useful tool for assessing outliers. Hence, such unsupervised projection-based methods represent a solid basis for MS-Sensor data analysis.

13) In this thesis a wide range of classifier methods suitable for MS-Sensor data handling have been presented. In one hand the so-called linear methods such as: LDA, PLS-DA, N-PLS-DA, unf-PLS-DA, PARAFAC-mlr-DA and, on the other hand, the non linear Fuzzy Artmap classifier. When prediction/quantification is attempted, then the models applied are PLS, unf-PLS or N-PLS. Often, it is difficult to rule out a general strategy for selecting which kind of classifier will be best suited for the MS-Sensor data. Mainly it depends on the dataset under study and thus, generalization is not straightforward. Nevertheless, there are general criteria which can be examined when choose among different classifier models. These criteria are: speed of operation, simplicity of use, memory requirements, robust handling of outliers and measures of prediction ability. In general,

using different tools gives corresponding results but with different degrees of accuracy. It is very important to understand the limitations and pitfalls of each one of the methodologies. It is very handy to have a good knowledge of the data set under analysis. The advantage of ANN based methodologies is basically that they are able to handle non-linearity relationships of the data. Nevertheless ANN often lacks from all the diagnostic tools that linear pattern recognition methods have such as score or loadings plots or even residuals analysis tools.

14) Answering the question about whether the inclusion of temporal information on the modelling of data improves the performance of the device is not as straightforward as one may consider a priori. It has been shown that there are different ways to include this temporal information on the response data matrix. In one hand it can be considered the unfolded matrix which contains the temporal data in a matricized way. On the other hand the inclusion of temporal information could also be attained by rearranging data to a three-way array and using further multi-way methodologies. It has been demonstrated that three-way modelling of data leads to better results than using unfolded methods. Hence, it can be concluded that the way in which temporal information is introduced into the model is a key issue to be considered when answering the question formulated above. Generally speaking, it has been demonstrated a most parsimonious performance for the three-way based models than for their unfolded counterparts. Unfolding methods usually lead to more unfavourable results for several reasons such as producing complex models, with many parameters which increase the risk of poor predictive capability, and the generation of less robust, interpretable and parsimonious models. Hence unfolded methods should generally be avoided or, at least, used with great care.

15) On the other hand, when comparing averaged two-way models against three-way models, both give rise to a similar degree of performance in terms of overfitting and prediction capability. No significant differences are found between the performances of both approaches in the application related to olive oils. By performing residual analysis it can be concluded that the two-way model is a little bit more affected by overfitting. The trilinear model tends to preserve explained variance in the test set while its two-way

counterpart does not. Therefore, two-way data modelling leads to models with lower generalisation ability with new samples. That can be attributed to the lack of trilinear constraints in the two-way models coming from the fact that the information across scans (temporal information) is not used to stabilize the solution. Bilinear models are often more flexible and the increased fit is used, to a large extent, to fit the noise of the training dataset. In the case of hams, the three-way modelling of data clearly leads to best results. Model predictability and performance is improved with respect to the averaged two-way model. Then, in general, three-way modelling compares favourably to other models for several reasons since it is the simplest one and it is more stable than the rest because its fits data correctly even though some trilinear constraints are imposed. Because of this, the model becomes much restricted and more robust against overfitting. Then, both three-way and two-way modelling of data have proven to be adequate approaches for the applications envisaged and hence it is not possible to categorically affirm that the temporal information can report extra valuable information into the modelling of data. In general terms, it is reasonable to use multi-way methods as they have the potential to simplify the interpretation of the results and provide more adequate and robust models using fewer parameters. From this point of view, it seems that the introduction of temporal information reports some extra valuable information at least for fitting simpler, more robust and parsimonious models.

16) The use of time dimension information into the modelling of data should be based on what is expected from the data. If there are expected differences in the chromatographic profiles, then three-way modelling is the best choice since it is the only methodology able to handle these differences. Otherwise, if chromatographic profile differences are not expected, then it is not useful to treat this data with a multi-way approach because the third dimension does not hold any information. In this last case, chromatographic information could be omitted by averaging mass spectra along the peak. That could be the reason that explains why in the case of olive oil the inclusion of temporal information did not finally lead to any improvement while in case of Iberian ham it did. The olive oil experiment was performed using the 5-m uncoated fused silica capillary column leading to an unretained, almost symmetrical peak showing no retention of the analytes. Therefore, chromatographic

profiles were the same for all the samples. On the other side, in the case of hams, measurements were performed using a 50-m coated polar column that in spite of being maintained at 250ºC (isothermal) during the whole analysis gave rise to a coeluted asymmetrical peak which shows some diffusion. This asymmetrical shape and the small retention lead to slightly different "chromatographic" profiles. Even minimal, these differences can be used by the three-way modelling approach.

17) The data analysis methods developed within this thesis can be employed in routine investigations and some of them represent major improvements compared to routines on available standard software packages data analysis in MS-Sensor. The limited human intervention required and the extended amounts of chemical information that can be generated, analysed and evaluated using these routines are the major obvious strengths of the methodologies used through this thesis.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Chapter 7

Perspectives and Future Trends

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

7. PERSPECTIVES AND FUTURES TRENDS

7.1 TRANFERING MS-SENSOR TECHNOLOGY TO THE INDUSTRIAL SECTOR

There is no doubt about the great potential of the MS-Sensor approach in the QC/QA of food. The performance and the advantages of this technique have been widely reviewed throughout this thesis. These advances open new perspectives for the rapid characterization of products in a wide variety of applications. Thus, it is expected that the MS-Sensor approach will likely grow in popularity as a routine screening tool for QC monitoring.

Let's consider the role of a MS-Sensor in a QC/QA laboratory. In this type of laboratories instead of looking for a full chromatographic profile that needs from a cumbersome peak matching procedure it is sought comparative rather than qualitative information. Very often, the information required is related to questions such as: is the final product the same as yesterday? Is this ingredient authentic or adulterated? Has this product changed its volatile profile during shelf-life? For answering these questions, it is not

necessary to determine the exact composition of the headspace. In such cases, a MS-Sensor is ideally suited because it offers comparative information. Furthermore, once the device has been trained for a specific application, data interpretation could be done automatically by the software.

The transfer of this type of technology to the industrial sector for the purposes of classification or QA/QC is now entirely feasible, since solutions have been found to solve the problems of instrumental drift. This allows the creation and management of permanent databases. The use of a state-of-the-art analytical technique such as mass spectrometry yields a very reproducible and precise fingerprint of each sample. Thus, identification of a component is straightforward, and the comparison of large data sets easy. Moreover, mass spectrometry is a well-known implemented technology for more that 20 years. That means that much of QA/QC laboratories have trained and skilled staff able to manage it and therefore the integration of a MS-Sensor could be easily done.

An additional advantage for QA/QC laboratories that have already implemented the classical GC-MS technology is that it is not necessary to have a dedicated MS-Sensor device since any GC/MS can be upgraded to a MS-Sensor in a rather straight forward manner by coupling the optimal sampling system and using additional pattern recognition software.

A possible strategy for implementing this technology into the laboratory is to turn-out the classical GC-MS into a MS-Sensor using it as a first-step, fast and reliable screening tool allowing a high productivity screening laboratory. Once the MS-Sensor has been fully calibrated and trained for a certain purpose, new samples can be subsequently analyzed in a high throughput way. In the case of “out of spec” samples, the MS-Sensor approach allows for a fast troubleshooting since the system can be easily switched between MS-Sensor and GC/MS modes. If initial MS-Sensor analysis points out a sample as a potential outlier, then full chromatographic runs might be necessary in order to determine the main causes from this sample behaviour. The main advantages of this configuration is that lot of analysis time could be saved because chromatographic separation is only performed on the “out of spec”

samples and also because minimal sample treatment is required. Thus, the cost of this analysis could be dramatically decreased.

All these advantages are of great value when thinking on an on-line application. For example, in the determination of ham quality according to the type of pig feeding, once the model has been trained, it would be entirely possible to discriminate whether a new sample of ham belongs to a fodder or acorn fed pig in a few minutes. Only in cases where the model indicates a potential problem with a sample (outlier), then a full chromatography of their free fatty acid profiles would be desirable in order to identify the key chemicals responsible for this abnormal behaviour.

7.2 THE IDEAL MS-SENSOR

After having worked with a MS-Sensor in a wide variety of food applications, a "wish list" on what would constitute the ideal MS-Sensor is summarized below:

a) It should incorporate headspace sampling techniques more sensitive than SH.
b) It should use mass detectors with soft ionization modes resulting in a smaller number of m/z fragments and in an increase of their selectivity.
c) It should integrate both chromatographic separation and spectral detection and take profit of them in further pattern recognition. The way for allowing a chromatographic separation without increasing analysis time is through use of Fast-GC. Fast-GC gives more options for data evaluation because it offers the complete chromatographic trace while the MS-Sensor offers only ion ratio information in the classical mode. The system should use pattern recognition algorithms able to handle both time and spectral dimensions given out by a Fast-GC MS-Sensor device. That leads unavoidably to the use of three-way algorithms.
d) Be less expensive that most of the MS-Sensors currently in the market. The miniaturization of mass spectrometers could play an important role in this aspect.

e) Be easy to use; provide easily interpretable results in the order of minutes per sample with minimal user intervention.

These points indicate that there are still some issues where the MS-Sensor approach needs further research. The next section tries to summarize some of these issues.

7.3 FUTURE TRENDS AND FURTHER INSIGHTS IN MS-SENSOR TECHNOLOGY

The MS-Sensor technology applied on food applications must be regarded as being in its early stages. So far, the applications in the scientific literature seem promising for future use in the food industry. Research is rapidly advancing on both instrument hardware and software seeking to enhance the selectivity, sensitivity and reproducibility of this approach. According to the experience gathered through the development of this thesis, the reader can find below the main issues that need further research to improve the usability of such a device.

7.3.1 Sampling

Volatile sampling will definitely be the most crucial part of the analysis when working with the MS-Sensor approach. At least eighty per cent of the success in the whole analytical procedure depends on this part. Most of the works published in the literature do not confer enough importance to this step and there are too few dedicated studies to the sampling of MS-Sensors. Thus, some future work should go in this direction. For example, recently appeared preconcentration sampling techniques such as solid-phase dynamic extraction (SPDE) or SHSE must be studied and their performance when coupled to MS-Sensor devices should be proven. SPDE will allow for the analysis of compounds that present low affinity to the fiber and SHSE will permit detection limits 2-3 orders of magnitude lower than those obtained with SPME.

7.3.2 Detector system

In terms of technical evolution, the use of "soft" ionization methods (ionization at atmospheric pressure, chemical ionization) in association with "high-resolution" mass analyzers (TOF, ion trap) allows to get information concerning the molecular origin of the ions and to deduce the identity of certain compounds present in the headspace. Fingerprints derived from the use of such technologies would enable much more specificity and sensitivity in MS-Sensor technology.

To date, considering the chromatographic resolution into the MS-Sensor response has been avoided mainly because of on the analysis time necessary. Nowadays, fast chromatography allows obtaining fully resolved chromatographic runs within a few minutes. Thus, considering a Fast-GC profile of a headspace extract, temporal information is still preserved. As it has been shown, multi-way analysis gives the chance to handle both, the temporal information and the spectral information as a fingerprint response. Advanced MS-Sensor concepts will bring together these two new issues.

Another important issue to consider for instrumental improvement is the miniaturization of devices. The combination of a Fast-GC followed by a MS instrument is a powerful and sensitive analytical tool but is an extremely large and expensive unit. Miniaturization of mass spectrometers is nowadays entirely feasible, included the miniaturization of TOF mass spectrometers. The volume of the vacuum chamber, the overall size, and the power consumption are drastically reduced with miniaturization. Making a micro GC-MS would be the ultimate challenge for the improvement of MS-Sensor technology.

This thesis has not implemented any of the methodologies already reported to solve the drift problem of the MS-Sensor approach. In this sense, it would be necessary to study to which extent the drift influences the mathematical models used on pattern recognition. It should also be necessary to develop new strategies for overcoming major drift problems.

Another technological aspect that still remains unsolved is evaluating the possibility of the implementation of on-line MS-Sensor monitoring in food quality control lines.

7.3.3 Software improvements

This thesis deals for the first time with the use of pre-processing methods for a MS-Sensor approach. These methods are a key requisite to overcome data uncertainties derived from factors unrelated to the chemical composition and to trilinearity condition fulfilment. Further research is necessary in order to improve some of the developed methods (baselining, background subtraction, alignment tools, etc). The more accurate the data pre-processing step becomes the better performance can be obtained on the pattern recognition phase.

Trilinearity assumptions in PARAFAC are often too strict and not adequate for describing a given dataset. Besides the preprocessing methods already mentioned, the lack of trilinearity can also been dealt by the use of algorithms which are specially designed for handling the situation. PARAFAC2 seems a valuable tool if trilinearity is a problem in one mode. The performance of other methods such as (PARAFAC2, MCR-ALS, TUCKER3) should be proven, since they are able to handle the lack of trilinearity, situation found very often in real datasets where typical non-trilinear chemical situations are found. These are mainly caused by shifts and shape changes in profiles.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

References

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

REFERENCES

[1] Gardner, J. W.; Bartlett, P. N. *Sensors and Actuators B-Chemical* **1994**, *18*, 211-220.
[2] Marsili, R., Ed. *Techniques for analyzing food aroma*; Marcel Dekker, Inc.: New York, **1997**.
[3] Persaud, k.; Dodd, G. *Nature* **1982**, 299: 352.
[4] Ampuero, S.; Bosset, J. O. *Sensors and Actuators B-Chemical* **2003**, *94*, 1-12.
[5] Pearce, T. C.; Schiffman, S. S.; Nagle, H. T.; Gardner, J. W., Eds. *Handbook of Machine Olfaction*; WILEY-VCH: Weinheim, **2003**.
[6] Rouseff, R. L.; Cadwallader, K. R., Eds. *Headspace analysis of foods and flavors: Theory and Practice*; Kluwer Academic/ Plenum Publishers: New York, **2001**.
[7] Penton, Z. In *Varian Application Note, Number 40*.
[8] Madou, M. J.; Morrison, S. R. *Chemical sensing with solid state devices*; Academic Press, Inc. , **1989**.
[9] Haugen, J.-E. In *Headspace analysis of Foods and Flavors. Theory and Practice*; Rouseff, R. L., Cadawallader, K. R., Eds.; Plenum: New York, **2001**, pp 43-55.
[10] Gardner, J. W.; Bartlett, P. N. *Electronic Noses: Principles and Applications*; Oxford University Press, Inc.: New York, **1999**.
[11] Gardner, J. W.; Persaud, K. C., Eds. *Electronic noses and olfaction 2000*; Institute of Physics Publishing: Bristol and Philadelphia, **2000**.
[12] Craven, M. A.; Gardner, J. W.; Bartlett, P. N. *Trac-Trends in Analytical Chemistry* **1996**, *15*, 486-493.
[13] James, D.; Scott, S. M.; Ali, Z.; O"Hare, W. T. *Microchimica Acta* **2005**, *149*, 1-17.
[14] Gutiérrez-Osuna, R.; Nagle, H. T.; Kermani, B.; Schiffman, S. S. In *Handbook of Machine Olfaction*; Pearce, T. C., Schiffman, S. S., Nagle, H. T., Gardner, J. W., Eds.; WILEY-VCH: Weinheim, **2003**.
[15] Scott, S. M.; James, D.; Ali, Z. *Microchimica Acta* **2006**, *156*, 183-207.
[16] L.Hines, E.; Boilot, P.; W.Gardner, J.; Gongora, M. A. In *Handbook of Machine Olfaction*; Pearce, T. C., Schiffman, S. S., Nagle, H. T., Gardner, J. W., Eds.; WILEY-VCH: Weinheim, **2003**.
[17] Jurs, P. C.; Bakken, G. A.; McClelland, H. E. *Chem. Rev.* **2000**, *100*, 2649-2678.
[18] Hines, E. L.; Llobet, E.; Gardner, J. W. *Iee Proceedings-Circuits Devices and Systems* **1999**, *146*, 297-310.
[19] Ricardo Gutierrez-Osuna *IEEE Sensors Journal* **2002**, *2*, 189-201.
[20] Mièlle, P. *Trends Food Sci. Technol.* **1996**, *7*, 432-438.
[21] Marsili, R. In *Flavor, Fragnace and Odor Analysis*; Marsili, R., Ed.; Marcel Dekker: New York, **2002**, pp 349-374.
[22] Cozzolino, D.; Smyth, H. E.; Lattey, K. A.; Cynkar, W.; Janik, L.; Dambergs, R. G.; Francis, I. L.; Gishen, M. *Analytica Chimica Acta* **2006**, *563*, 319-324.
[23] Fenaille, F.; Visani, P.; Fumeaux, R.; Milo, C.; Guy, P. A. *Journal of Agricultural and Food Chemistry* **2003**, *51*, 2790-2796.
[24] Marti, M. P.; Pino, J.; Boque, R.; Busto, O.; Guasch, J. *Analytical and Bioanalytical Chemistry* **2005**, *382*, 440-443.

[25] Feldhoff, R.; Saby, C. A.; Bernadet, P. *Analyst* **1999**, *124*, 1167-1173.
[26] Saevels, S.; Lammertyn, J.; Berna, A. Z.; Veraverbeke, E. A.; Di Natale, C.; Nicolai, B. M. *Postharvest Biology and Technology* **2004**, *31*, 9-19.
[27] Cynkar, W.; Cozzolino, D.; Dambergs, B.; Janik, L.; Gishen, M. *Sensors and Actuators B-Chemical* **2007**, *124*, 167-171.
[28] Cozzolino, D.; Smyth, H. E.; Cynkar, W.; Dambergs, R. G.; Gishen, M. *Talanta* **2005**, *68*, 382-387.
[29] Mièlle, P.; Marquis, F. *Sensors and Actuators B* **2000**, *68*, 9–16.
[30] Peres, C.; Begnaud, F.; Eveleigh, L.; Berdague, J. L. *Trac-Trends in Analytical Chemistry* **2003**, *22*, 858-866.
[31] Schaller, E.; Zenhausern, S.; Zesiger, T.; Bosset, J. O.; Escher, F. *Analusis* **2000**, *28*, 743-749.
[32] Schaller, E.; Zenhausern, S.; Zesiger, T.; Bosset, J. O.; Escher, F. *Analusis* **2000**, *28*, 217.
[33] Goodner, K. L.; Rouseff, R. L. *Journal of Agricultural and Food Chemistry* **2001**, *49*, 250-253.
[34] Dittmann, B.; Zimmermann, B.; Engelen, C.; Jany, G.; Nitz, S. *Journal of Agricultural and Food Chemistry* **2000**, *48*, 2887-2892.
[35] Peres, C.; Begnaud, F.; Berdague, J. L. *Sensors and Actuators B-Chemical* **2002**, *87*, 491-497.
[36] Peres, C.; Denoyer, C.; Tournayre, P.; Berdague, J. L. *Analytical Chemistry* **2002**, *74*, 1386-1392.
[37] Boudaoud, N.; Eveleigh, L. *Journal of Agricultural and Food Chemistry* **2003**, *51*, 1530-1533.
[38] Peres, C.; Viallon, C.; Berdague, J. L. *Analytical Chemistry* **2001**, *73*, 1030-1036.
[39] Marsili, R. T. *Journal of Agricultural and Food Chemistry* **1999**, *47*, 648-654.
[40] Marsili, R. T. *Journal of Agricultural and Food Chemistry* **2000**, *48*, 3470-3475.
[41] Berna, A. Z.; Lammertyn, J.; Saevels, S.; Di Natale, C.; Nicolai, B. M. *Sensors and Actuators B-Chemical* **2004**, *97*, 324-333.
[42] Ampuero, S.; Bogdanov, S.; Bosset, J. O. *European Food Research and Technology* **2004**, *218*, 198-207.
[43] Vinaixa, M.; Mari?n, S.; Brezmes, J.; Llobet, E.; Vilanova, X.; Correig, X.; Ramos, A.; Sanchis, V. *Journal of Agricultural and Food Chemistry* **2004**, *52*, 6068-6074.
[44] Lord, H. L.; Pawliszyn, J. *Anal. Chem.* **1997**, *69*, 3899-3906.
[45] Peñalver, A.; Pocurull, E.; BorrulI, E.; Marcé, R. M. *Chromatographia* **1999**, *50*, 685-688.
[46] Pawliszyn, J. *Solid Phase Microextraction: Theory and Practice*; Wiley-VCH Inc: New York, **1997**.
[47] Hayasaka, Y.; MacNamara, K.; ·, G. A. B.; Taylor, R. L.; Pollnitz, A. P. *Anal Bioanal Chem* **2003**, *375*, 948–955.
[48] Roussel, S.; Forsberg, G.; Grenier, P.; Bellon-Maurel, V. *Journal of Food Engineering* **1999**, *39*, 9-15.
[49] Kinton, V. R.; Pfannkoch, E. A. *Gerstel Application Note 9/2002* **2002**.
[50] Kinton, V. R.; Heiden, A. C.; Gil, C. *Gerstel Application Note 2/2004* **2004**.

[51] Hayasaka, Y.; MacNamara, K.; ·, G. A. B.; Taylor, R. L.; Pollnitz, A. P. *Anal Bioanal Chem* **2003**, *375*, 948–955.
[52] Castel, C.; Fernandez, X.; Lizzani-Cuvelier, L.; Loiseau, A. M.; Perichet, C.; Delbecque, C.; Arnaudo, J. F. *Flavour and Fragrance Journal* **2006**, *21*, 59-67.
[53] Marti, M. P.; Boque, R.; Busto, O.; Guasch, J. *Trac-Trends in Analytical Chemistry* **2005**, *24*, 57-66.
[54] Sandra, P.; David, F.; Tienpont, B. *Chromatographia* **2004**, *60*, S299-S302.
[55] van Hout, M. W. J.; Jas, V.; Niederlander, H. A. G.; de Zeeuw, R. A.; de Jong, G. J. *Journal of Separation Science* **2003**, *26*, 1563-1568.
[56] Pavon, J. L. P.; Sanchez, M. D.; Pinto, C. G.; Laespada, M. E. F.; Cordero, B. M.; Pena, A. G. *Trac-Trends in Analytical Chemistry* **2006**, *25*, 257-266.
[57] Peres, C.; Begnaud, F.; Berdague, J. L. *Analytical Chemistry* **2002**, *74*, 2279-2283.
[58] Pavon, J. L. P.; Sanchez, M. D.; Pinto, C. G.; Laespada, M. E. F.; Cordero, B. M. *Analytical Chemistry* **2003**, *75*, 6361-6367.
[59] Syms, R. R.; T.J.Tate; Ahmad, M.; Taylor, S. *Electronic Letters* **1996**, *32*, 2094-2095.
[60] Smilde, A.; Bro, R.; Geladi, P. *Multi-way Analysis. Applications in the chemical science*; John Wiley & Sons Ltd,: West Sussex, England, **2004**.
[61] Cocchi, M.; Bro, R.; Durante, C.; Manzini, D.; Marchetti, A.; Saccani, F.; Sighinolfi, S.; Ulrici, A. *Food Quality and Preference* **2006**, *17*, 419-428.
[62] Marsili, R. T. In *Headspace Analysis of Foods and Flavours: Theory and Practice*; Rouseff, R. L., Cadwallader, K. R., Eds.; Plenum Press: New York, **2001**.
[63] Marsili, R. T.; Miller, N. In *Food Flavours and Chemistry: Advances of the New Millennium*; Spanier, A. M., Shahidi, F., Parliament, T. H., C. Mussinan, Ho, C.-T., Eds.; RSC: Conti E.T, **2001**, pp 118-129.
[64] Marsili, R. T. *J. Chromatogr. Sci.* **1999**, *37* 17-23.
[65] Ampuero, S.; Zesiger, T.; Gustafsson, V.; Lunden, A.; Bosset, J. O. *European Food Research and Technology* **2002**, *214*, 163-167.
[66] Pillonel, L.; Bosset, J. O.; Tabacchi, R. *European Food Research and Technology* **2002**, *214*, 160-162.
[67] Pillonel, L.; Ampuero, S.; Tabacchi, R.; Bosset, J. O. *European Food Research and Technology* **2003**, *216*, 179–183.
[68] Marilley, L.; Ampuero, S.; Zesiger, T.; Casey, M. G. *International Dairy Journal* **2004**, *14*, 849-856.
[69] Kinton, V. R. *Gerstel Application Note 10/2002* **2002**.
[70] Duflos, G.; Coin, V. M.; Cornu, M.; Antinelli, J. F.; Mallel, P. *Journal of the Science of Food and Agriculture* **2006**, *86*, 600-611.
[71] Anklam, E.; Lipp, M.; Radovic, B.; Chiavarob, E.; Pallab, G. *Food Chemistry* **1998**, *61*, 243-248.
[72] Timmins, E. M.; Howell, S. A.; Alsberg, B. K.; Noble, W. C.; Goodacre, R. *Journal of Clinical Microbiology* **1998**, *36*, 367-374.
[73] Alsberg, B. K.; Goodacre, R.; Rowland, J. J.; Kell, D. B. *Analytica Chimica Acta* **1997**, *348*, 389-407.
[74] Lorenzo, I. M.; Pavon, J. L. P.; Laespada, M. E. F.; Pinto, C. G.; Cordero, B. M. *Journal of Chromatography A* **2002**, *945*, 221-230.

[75] Dittmann, B.; Nitz, S. *Sensors and Actuators B* **2000**, *69*, 253–257.
[76] Marco Lorenzo, I.; Pérez Pavón, J. L.; Fernández Laespada, M. E.; García Pinto, C.; Moreno Cordero, B. *Journal of Chromatography A* **2002**, *945*, 221-230.
[77] Peña, F.; Cárdenas, S.; Gallego, M.; Valcárcel, M. *Journal of the American Oil Chemists" Society* **2002**, *79*, 1103.
[78] Pavon, J. L. P.; Pena, A. G.; Pinto, C. G.; Cordero, B. M. *Journal of Chromatography A* **2004**, *1047*, 101-109.
[79] Peña, F.; Cárdenas, S.; Gallego, M.; Valcárcel, M. *Journal of Chromatography A* **2004**, *1052*, 137-143.
[80] Kinton, V. R.; Whitecavage, J. A. *Gerstel Application Note 1/2004* **2004**.
[81] Oliveros, C. C.; Boggia, R.; Casale, M.; Armanino, C.; Forina, M. *Journal of Chromatography A* **2005**, *1076*, 7-15.
[82] Pena, F.; Cardenas, S.; Gallego, M.; Valcarcel, M. *Journal of Chromatography A* **2005**, *1074*, 215-221.
[83] Laguerre, M.; Mestres, C.; Davrieux, F.; Ringuet, J.; Boulanger, R. *Journal of Agricultural and Food Chemistry* **2007**, *55* 1077-1083.
[84] Berna, A. Z.; Buysens, S.; Di Natale, C.; Grun, I. U.; Lammertyn, J.; Nicolai, B. M. *Postharvest Biology and Technology* **2005**, *36*, 143-155.
[85] Berna, A. Z.; Lammertyn, J.; Buysens, S.; Di Natale, C.; Nicolai, B. M. *Postharvest Biology and Technology* **2005**, *38*, 115-127.
[86] Bleibaum, R. N.; Stone, H.; Tan, T.; Labreche, S.; Saint-Martin, E.; Isz, S. *Food Quality and Preference* **2002**, *13*, 409-422.
[87] Heiden, A. C.; Gil, C.; Ramos, L. S. *Gerstel Application Note 8/2002* **2002**.
[88] Supriyadi; Shimizu, K.; Suzuki, M.; Yoshida, K.; Muto, T.; Fujita, A.; Tomita, N.; Watanabe, N. *Flavour and Fragrance Journal* **2004**, *19*, 44-50.
[89] Kojima, H.; Araki, S.; Kaneda, H.; Takashio, M. *Journal of the American Society of Brewing Chemists* **2005**, *63*, 151-156.
[90] Marti, M. P.; Busto, O.; Guasch, J. *Journal of Chromatography A* **2004**, *1057*, 211-217.
[91] Marti, M. P.; Boque, R.; Riu, M.; Busto, O.; Guasch, J. *Analytical and Bioanalytical Chemistry* **2003**, *376*, 497-501.
[92] Kolahgar, B.; Heiden, A. C. *Gerstel Application Note 11/2002* **2002**.
[93] Kinton, V. R.; Collins, R. J. *Gerstel Application Note 4/2003* **2003**.
[94] Kinton, V. R.; Pfannkoch, E. A. *Gerstel Application Note 2/2003* **2003**.
[95] Pena, F.; Cardenas, S.; Gallego, M.; Valcarcel, M. *Journal of the American Oil Chemists Society* **2003**, *80*, 613-618.
[96] Rocha, S. M.; Coutinho, P.; Barros, A.; Delgadillo, I.; Coimbra, M. A. *Journal of Chromatography A* **2006**, *1114*, 188-197.
[97] Garrigues, S.; Talou, T.; Nesa, D.; Gaset, A. *Sensors and Actuators B-Chemical* **2001**, *78*, 337-344.
[98] Gardner, W. P.; Shaffer, R. E.; Girard, J. E.; Callahan, J. H. *Analytical Chemistry* **2001**, *73*, 596-605.
[99] Feldhoff, R.; Saby, C. A.; Bernadet, P. *Flavour and Fragrance Journal* **2000**, *15*, 215-222.
[100] Pavón, J. L. P.; Sánchez, M. d. N.; Pinto, C. G.; Laespada, M. E. F.; Cordero, B. M. *Journal of Chromatography A* **2004**, *1048*, 133-139.

[101] del Nogal Sanchez, M.; Pavon, J. L. P.; Laespada, M. E. F.; Pinto, C. G.; Moreno Cordero, B. *Analytical and Bioanalytical Chemistry* **2005**, *382*, 372-380.
[102] Yates, J. W. T.; Gardner, J. W.; Chappell, M. J.; Dow, C. S. *IEE Proceedings-Science Measurement and Technology* **2005**, *152*, 97-102.
[103] Yates, J. W. T.; Chappell, M. J.; Gardner, J. W.; Dow, C. S.; Dowson, C.; Hamood, A.; Bolt, F.; Beeby, L. *Computer Methods and Programs in Biomedicine* **2005**, *79*, 259-271.
[104] Pavon, J. L. P.; Sanchez, M. D. N.; Pinto, C. G.; Laespada, M. E. F.; Cordero, B. M. *Analytical Chemistry* **2006**, *78*, 4901-4908.
[105] Comas, E.,*Application of the generalized rank annihilation method (GRAM) to second-order liquid chromatographic data*, Universitat Rovira i Virgili, Tarragona, **2002**.
[106] Sanchez, E.; Kowalski, B. R. *Journal of Chemometrics* **1988**, 254-280.
[107] Sanchez, E.; Kowalski, B. R. *Journal of Chemometrics* **1998**, 247-263.
[108] Tucker, L. R. In *Contributions to Mathematical Psychology*; Holt, Rinehart, Winston, Eds.: New York, **1964**, pp 110-182.
[109] Wentzell, P. D.; Nair, S. S.; Guy, R. D. *Anal. Chem.* **2001**, *73*, 1408.
[110] R. Bro *Chemom. Intell. Lab. Syst.* **1997**, *38*, 149.
[111] Garcia, I.; Sarabia, L.; Ortiz, M. C.; Aldama, J. M. *Analytica Chimica Acta* **2004**, *515*, 55-63.
[112] Pradova, V.; Boucon, C.; Jong, S. d.; Walczak, B.; Massart, D. L. *Anal. Chim. Acta* **2002**, *462* 133.
[113] Bro, R.,*Multi-way analysis in the food industry. Models algorithms and applications*, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark, **1998**.
[114] Christensen, J. H.,*Chemometrics as a tool to analyse complex chemical mixtures–Environmental forensics and fate of oil spills*, Roskilde University, **2005**.
[115] Bylund, D.; Danielsson, R.; Malmquist, G.; Markides, K. E. *J. Chromatography A* **2002**, *961* 237–244.
[116] Nielsen, N.-P. V.; Carstensen, J. M.; Smedsgaard, J. *Journal of Chromatography A* **1998**, *805* 17–35.
[117] Johnson, K. J.; Wright, B. W.; Jarman, K. H.; Synoveca, R. E. *J. Chromatography A* **2003**, *996*, 141–155.
[118] Tomasi, G.; Berg, F. v. d.; Andersson, C. *J. Chemometrics* **2004**, *18*, 231–241.
[119] Jason W. H. Wong; Durante, C.; Cartwright, H. M. *Anal. Chem.* **2005**, *77*, 5655-5661.
[120] Bellman, R. *Adaptive control processes: a guided tour.*; Princeton University Press: Princeton, N.J., **1961**.
[121] Jain, A. K.; Duin, R. P. W.; Mao, J. C. *IEEE T Pattern Anal Machine Intelligence* **2000**, *22*, 4.
[122] Blum, A. L.; Langley, P. *Artificial Intelligence* **1997**, *97*, 245-271.
[123] Dittmann, B.; Nitz, S. *Sensors and Actuators B-Chemical* **2000**, *69*, 253-257.
[124] Esbensen, K. H. *Multivariate data analysis in practice-An Introduction to Multivariate Data Analysis and Experimental Design*, **2002**.
[125] Bro, R. *Journal of Chemometrics* **1996**, *10*, 47-61.

[126] Shaffer, R. E.; Rose-Pehrsson, S. L.; McGill, R. A. *Analytica Chimica Acta* **1999**, *384*, 305-317.
[127] Barker, M.; Rayens, W. *Journal of Chemometrics* **2003**, *17*, 166-173.
[128] Shirey, R. E.; Mindrup, R. F. *Supelco Application Note* **1999**.
[129] Pinho, O.; Peres, C.; Ferreira, I. *Journal of Chromatography A* **2003**, *1011*, 1-9.
[130] Wise, B. M.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W.; Koch, R. S.; Eigenvector Research, Inc., **2005**.
[131] Bro, R. *Chemom. Intell. Lab. Syst.* **1997**, *38* 149-171.
[132] Guimet, F.,*Olive oil characterization using excitation-emission fluorescence spectroscopy and three-way methods of analysis*, Universitat Rovira i Virgili, Tarragona, **2005**.
[133] Bro, R.; Smilde, A. K. *Journal of Chemometrics* **2003**, *17*, 16-33.
[134] Wong, J. W. H.; Durante, C.; Cartwright, H. M. *Analytical Chemistry* **2005**, *77*, 5655-5661.
[135] Cattell, R. B. *Multivariate Behavoiral Research* **1966**, *1*, 245-276.
[136] Pavlou, A. K.; Magan, N.; Sharp, D.; Brown, J.; Barr, H.; Turner, A. P. F. *Biosensors & Bioelectronics* **2000**, *15*, 333-342.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Appendix

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper I

Vinaixa, M.; Marin, S.; Brezmes, J.; Llobet, E.; Vilanova, X.; Correig, X.; Ramos, A.; Sanchis, V., Early detection of fungal growth in bakery products by use of an 'electronic nose' based on mass spectrometry. *Journal of Agricultural and Food Chemistry* **2004,** 52(20), 6068-6074.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Early Detection of Fungal Growth in Bakery Products by Use of an Electronic Nose Based on Mass Spectrometry

MARIA VINAIXA,[†] SONIA MARÍN,[‡] JESÚS BREZMES,*[,†] EDUARD LLOBET,[†] XAVIER VILANOVA,[†] XAVIER CORREIG,[†] ANTONIO RAMOS,[‡] AND VICENT SANCHIS[‡]

Department of Electronic Engineering, Universitat Rovira i Virgili, Avenida Països Catalans 26, 43007 Tarragona, Spain, and Department of Tecnologia dels Aliments, Universitat de Lleida, Avenida Alcalde Rovira Roure 191, 25198 Lleida, Spain

This paper presents the design, optimization, and evaluation of a mass spectrometry-based electronic nose (MS e-nose) for early detection of unwanted fungal growth in bakery products. Seven fungal species (*Aspergillus flavus*, *Aspergillus niger*, *Eurotium amstelodami*, *Eurotium herbariorum*, *Eurotium rubrum*, *Eurotium repens*, and *Penicillium corylophillum*) were isolated from bakery products and used for the study. Two sampling headspace techniques were tested: static headspace (SH) and solid-phase microextraction (SPME). Cross-validated models based on principal component analysis (PCA), coupled to discriminant function analysis (DFA) and fuzzy ARTMAP, were used as data treatment. When attempting to discriminate between inoculated and blank control vials or between genera or species of in vitro growing cultures, sampling based on SPME showed better results than those based on static headspace. The SPME−MS-based e-nose was able to predict fungal growth with 88% success after 24 h of inoculation and 98% success after 48 h when changes were monitored in the headspace of fungal cultures growing on bakery product analogues. Prediction of the right fungal genus reached 78% and 88% after 24 and 96 h, respectively.

KEYWORDS: Electronic nose; mass spectrometry; fungal growth; bakery products; fuzzy ARTMAP; ANN; LDA; PCA

INTRODUCTION

Microbial spoilage is a major problem in bakery products since it can induce nutritional losses, off-flavors, and formation of mycotoxins or potentially allergenic spores. This situation can lead to an organoleptic deterioration of already marketed bakery products, which indeed threatens consumers' confidence and, therefore, results in important economical losses. This is the reason for a growing need to find a method to conveniently assess the degree of fungal growth in bakery products at a very early stage and before it becomes visible (*1*).

Classical techniques based on microbiological methods such as CFU (colony-forming units) determination are time-consuming and they cannot give on-line responses. Specific chemical markers such as ergosterol have now become commonly used as a method for the quantification of fungal biomass in food. However, they are nonspecific, they do not provide any information on the species present, and they require a laborious sample preparation (*2*).

On the other hand, it is known that fungi produce volatile compounds during both primary and secondary metabolism that can be used as markers to detect food spoilage, unwanted fungal growth, or even as taxonomic identifiers that can determine the presence of a given species. This idea was initially exploited in the field of cereals (*3*, *4*). Research studies correlated fungal activity with the production of volatile metabolites, CO_2, and CFU. Schnürer et al. (*2*) and Magan and Evans (*5*) reviewed some studies where GC−MS had been used to characterize and analyze volatile profiles of fungal cultures, listing volatiles identified in different growing substrates. More recently, GC−MS intensity peaks from key volatiles have been used by Olsson et al. (*6*, *7*) to evaluate the mycological quality of barley grains and to predict levels of ochratoxin A and deoxynivalenol.

Since the volatile headspace is complex and should be evaluated as a whole, techniques that mimic the human olfactory system (the so-called electronic noses) have already been proposed. Electronic noses based on different types of nonspecific sensors (i.e., metal oxide, conducting polymer, or quartz microbalance sensors) have been evaluated in fungal, bacterial, and yeast monitoring in food such as bakery products (*1*), cereal grains (*5*−*8*), cheese (*9*), water (*10*), bread (*11*), meat (*12*), and milk (*13*). Despite the efforts, e-noses based on nonspecific semiconductor sensors still suffer from serious drawbacks such as poor sensitivity, poor selectivity, and long-term drift. Novel electronic olfactory systems based on mass spectrometry seem to improve drift problems with respect to other classical e-nose technologies. Even so, MS e-nose suffers from low temporal

* Corresponding author (telephone +34 977 559619; fax +34 977 559605; e-mail jbrezmes@etse.urv.es).
[†] Universitat Rovira i Virgili.
[‡] Universitat de Lleida.

10.1021/jf049399r CCC: $27.50
Published on Web 09/09/2004

drift. Nevertheless, this paper does not consider drift as an issue since measurements were done in a short period of time. Preconcentration and extraction techniques such as solid-phase microextraction (SPME) can increase the sensitivity and reproducibility of MS e-noses (*14*, *15*). Since Nilsson et al. (*16*) used SPME as a solvent-free extraction method for analysis of volatile metabolites emitted by *Penicillium* species, many researchers have used SPME as a tool for sampling volatile fungal metabolites (*17*, *18*).

This paper presents a study on the design, optimization, and evaluation of a MS-based electronic nose for early fungal growth detection in bakery products. Two different experiments were carried out. The first one was designed to choose the best sampling technique (SH or SPME) to be coupled to a MS-based e-nose in order to monitor fungal growth. The second study was designed to evaluate the performance of the MS e-nose optimal configuration in early detection of in vitro and in situ growing fungal cultures.

MATERIALS AND METHODS

Samples. Seven fungal species (*Aspergillus flavus*, AF; *Aspergillus niger*, AN; *Eurotium amstelodami*, EA; *Eurotium herbariorum*, EH; *Eurotium rubrum*, EU; *Eurotium repens*, ER; and *Penicillium corylophillum*, PE) were isolated from bakery products. One isolate from each species was used for the present study.

For in vitro studies, 0.90 aw slants of 2% wheat flour agar medium were prepared in 20-mL headspace vials and needle-inoculated with 10^{-6} spores mL^{-1} suspensions (adjusted by use of a Thoma chamber) of the fungal cultures mentioned above. Uninoculated vials with agar medium were used as control blanks (BL).

In situ studies were performed on bakery product analogues prepared as described by Abellana et al. (*19*), adjusted to a water activity of 0.95. Pieces measuring 8 × 8 × 20 mm were introduced in 20-mL headspace vials. Analogues were needle-inoculated at random with the seven cultures mentioned above and, again, uninoculated analogues were used as control blanks. Once sealed, all vials were incubated at 25 °C until their measurement. Incubation periods ranged from 24 h to 7 days depending on the experiments carried out.

MS E-nose Configuration. A Shimadzu QP 5000 GC/MS (Shimadzu Corp., Tokyo, Japan) was used to implement a MS-based e-nose. The instrument was equipped with a deactivated PR-100052 5 m × 0.25 mm ID precolumn (Teknokroma, St Cugat Del Vallès, Barcelona, Spain) that only acted as a transfer line from the injector port to the mass detector. The column was kept isothermal at 250 °C to coelute all volatile components in one single peak. This implies that the components in the headspace of the vials passed directly to the mass detector without any chromatographic separation. In this manner, for any given measurement, the resulting mass spectrum gives a fingerprint that is characteristic of the volatiles present in the headspace of samples. Helium flow was set to 1.4 mL/min. The mass spectrometer operated in the electron impact ionization mode (70 eV) and acquired in a scan range from *m*/*z* 35 to 120 at 0.5 scan/s. Ion source temperature was set at 250 °C.

Two different sampling techniques (SH and SPME) were evaluated as the best candidates to be coupled to the MS-e-nose.

Choosing the Best Sampling Technique. (A) Static Headspace Optimization. Sampling based on a static headspace autosampler was done by coupling an HP-7694 (Agilent Technologies) to the MS-based e-nose. All experiments dealing with MS e-nose optimization were carried out on in vitro growing cultures for 10 days. The main sampling parameters that influence sensitivity when working with the SH technique, namely, temperature of equilibration and vial equilibration time, were modified to improve fungal culture classification and fungal growth detection.

To select the optimal temperature, three oven temperatures, 50, 80, and 100 °C, were tested. For each temperature, a batch of 16 measurements corresponding to two replicates for each of the seven species plus two additional control blanks incubated for 10 days was performed. Equilibration time was fixed at 5 min in the three batches of measurements. The temperature of the loop and the transfer line were always kept at 5 °C above oven temperature to avoid condensation. Each vial was pressurized with helium (i.e., the carrier gas) for 12 s. Then, the 3-mL internal loop of the headspace autosampler was filled with volatiles coming from headspace of the fungal cultures, and finally volatiles were injected into the gas chromatograph. The goal was to determine whether the instrument could distinguish between inoculated and blank vials. A secondary goal was to observe whether the system was able to classify samples according to fungal genera.

In a second experiment, equilibration temperature was kept constant at 50 °C and equilibration time was increased to 50 min. The remaining parameters were kept as described above. A total of 32 samples (four replicates of eight different types of vials) were measured.

(B) SPME−MS E-nose Measurements. Sampling based on SPME was performed with a 75-μm Carboxen/PDMS fiber purchased from Supelco (Supelco Park, Bellefonte, PA). Jelen (*18*) compared four SPME fibers used to perform an extraction of volatile metabolites from fungal cultures. This study showed that the highest amount of isolated volatiles expressed as total peak area was observed for fibers based on Carboxen (CAR/PDMS and CAR/DVB/PDMS). Fibers based on Carboxen are the best choice in terms of sensitivity to extract low molecular weight analytes such as low-chain alcohols, ketones, and aldehydes. The main volatile metabolites involved in early stages of fungal growth that have been cited in the scientific literature belong to this type of molecules. As the goal was to detect the production of these volatiles as early as possible, priority was given to sensitivity and that is why a CAR/PDMS-based fiber was chosen. Prior to any extraction, the fiber was conditioned following the manufacturer's recommendations. In each measurement, the fiber was introduced into the vial and exposed to the headspace of fungal cultures for 20 min at room temperature. Thermal desorption of volatiles trapped on the fiber was conducted for 3 min in the chromatograph injection port at 300 °C. The split valve was closed during desorption. The fiber was always left five additional minutes to ensure its complete cleaning.

Three in vitro replicate vials of each fungal species plus control blanks were prepared and incubated for 10 days. Each replicate was measured three times by the SPME−MS e-nose. Therefore, a total of 72 measurements were performed. The aim was to discriminate inoculated samples from blank vials and to evaluate whether the instrument could classify samples according to their genera and species.

In Vitro Fungal Growth Monitoring. The next goal was to evaluate the performance of the MS-based e-nose to monitor early stages of fungal growth. SPME was used since it was determined that it was the best sampling method. Two replicates of each of the seven fungi plus two control blanks were grown on 2% wheat flour agar. Samples were kept under incubation and extractions were made once a day, obtaining a total of 16 experimental points at 48, 72, 96, and 168 h after inoculation.

In Situ Fungal Growth Monitoring. A final experiment designed to simulate a real application was performed. The aim was to use the final prototype in order to discriminate between spoiled and safe bakery products. Eight blank control vials containing cake analogues, eight replicate vials containing cake analogues inoculated with EA, and four replicates inoculated with ER, EU, EH, AN, AF, and PE were measured. SPME extractions and MS e-nose measurements were performed in every vial 24, 48, 72, 96, and 168 h after inoculation. Overall, 40 experimental points were obtained for each sampling time (1, 2, 3, 4, and 7 days after inoculation). The system was also evaluated as a tool to discriminate among fungal species.

Multivariate and Pattern Recognition Analysis. Data generated by the e-nose device (in any of its different configurations) were collected and processed by use of written-in-house software based on MATLAB 6.5 (The Mathworks, Natick, MA). An unresolved single peak was obtained for each measurement. Averaging mass spectra along the detected peak generated a response spectrum. Since measurements were performed in scan mode from *m*/*z* 35 to 120, the average intensity of each mass could be used as a variable (or sensor). In this manner, an experimental data matrix was built. The number of rows was the number of samples measured in each experimental batch, while the number of columns was 86, corresponding to each *m*/*z* scanned. A

principal component analysis (PCA) was applied to each response matrix, achieving a reduction in dimensionality. By use of the first 10 principal components, 100% of the total data variance was gathered. A reduced response matrix having 10 columns corresponding to scores of the first 10 principal components was obtained. Then, a discriminant function analysis (DFA) was performed on the reduced matrix. Eigenvalues obtained from the DFA were used as input variables to a fuzzy ARTMAP neural network that gave a categorization of fungal cultures according to genera or species depending on the application.

The performance of the model was evaluated by the leave-one-out cross-validation approach. In this method, a different row (measurement) from the original matrix is left out at each iteration. The remaining rows conform the training matrix, which is reduced by a PCA projection, processed by a DFA, and fed to the fuzzy ARTMAP training algorithm after scaling coordinates between 0 and 1. The procedure is then validated with the vector that had been left out. The validation vector (not used for training) is then projected against the PCA model. Then, the PCA scores are projected onto space of the canonical variables of the trained DFA. Finally, the DFA projection coordinates of the validation vector are fed to the neural network model, which produces a classification result. The whole process is repeated N times, N being the number of measurements included in the data matrix, so that each measurement is used in one iteration for evaluation purposes and in $N - 1$ iterations for training. The fact that for each iteration the validation vector is not used in the training process ensures that the vector is completely new to the processing system.

RESULTS AND DISCUSSION

Reduction of *m/z* Variable Dimensionality by Use of PCA. All the results cited bellow were obtained by applcation of multivariate analysis to the response matrix. This matrix was formed by as many rows as experimental measurements made in each study and as many columns as *m/z* variables scanned. According to Dittmann and Nitz (*20*), these *m/z* variables can be used as an array of sensors to emulate a classical electronic nose. In their paper, they claim that, in most cases, it is not useful to work with such a great number of sensors and only a very small number of ion fragments are suitable for setting up a sensor array since meaningless fragments introduce noise into the system. They also consider that there is no way to correctly select *m/z* fragments, unless there is a previous full chromatographic run. Finally, they discuss reliable strategies for selecting the optimal array configuration, based on previous knowledge of the analytes that are important for the application. This previous knowledge is normally based on time-resolved analysis to identify (and quantify) the volatiles present in the headspace of the samples to be studied. This leads, unavoidably, to more traditional analytical techniques such as GC−MS.

On the other hand, applying a PCA analysis leads to a linear combination of *m/z* variables that gathers the highest amount of variance and compresses information by eliminating redundancy and collinearity. In this manner, the best combination of *m/z* variables can be chosen without the need to perform a costly and lengthy initial study to determine the most relevant ion fragments to be monitored. This means that previous fully resolved chromatographic runs can be avoided since no previous knowledge from the samples is required.

The reduced response matrix obtained from the PCA analysis is then used as the input matrix to perform a DFA. DFA is a supervised model that finds a function-based projection that minimizes distances between measurements from the same category and maximizes distance between centroids of each category. Finally the two first factors resulting from the DFA were used as input variables to a fuzzy ARTMAP neural network as described before.

Static Headspace Optimization. A preliminary analysis of the results on increasing oven temperature in SH was performed by plotting PCA scores at 50, 80, and 100 °C separately. At 50 and 80 °C those plots did not show any clustering, and samples with fungal contamination and blank vials overlapped. At 100 °C, inoculated samples clustered together, clearly separated from blank vials. Setting a headspace oven temperature of 100 °C permitted the extraction of a larger quantity of volatiles, which enhanced the sensitivity of the system, allowing it to achieve a better discrimination than at 50 or 80 °C. A cross-validated fuzzy ARTMAP classification of the 16 experimental points at 100 °C achieved an 88% success rate when trying to determine whether the vial was inoculated or not. When attempting to determine the fungal genera, the success rate decreased to 56% (**Table 1**). Measuring samples at 100 °C may accelerate oxidation processes modifying the qualitative volatile pattern profiles. This may be very difficult to control and could introduce noise in our mathematical model.

Therefore we designed a second experiment where temperature of equilibration remained constant at 50 °C and equilibration time was increased. The concentration of an analyte in the headspace usually follows a linear dependence with equilibration time until it comes to a point where the concentration becomes stable. At this point volatiles reach equilibrium and achieve their maximum concentration in the headspace, while the composition of volatile patterns remains stable. That should enhance the repeatability on MS e-nose measurements. Sanz et al. (*21*) studied equilibration time in *Arabica* coffee and they concluded that when the equilibration time increased, the quantity of volatile compounds also increased but in an irregular way, depending on the chemical family considered. Similarly, every analyte from each fungal culture has a different equilibration

Table 1. Success Rate Comparison between the Two Different Sampling Techniques Studied[a]

technique	goal discrimination between	total exptl points	failures	success rate (%)
HS–MS, 100 °C	fungal growth	16	2	88
	genera	16	7	56
HS–MS, 50 min	fungal growth	32	1	97
	genera	32	13	59
SPME–MS	fungal growth	72	0	100
	genera	72	0	100
	species	72	6	92

[a] Data processing was performed with PCA–DFA–fuzzy ARTMAP models. The goal was to classify between inoculated and uninoculated vials (fungal growth) and between genera or species. All tests were performed over in vitro growing cultures.

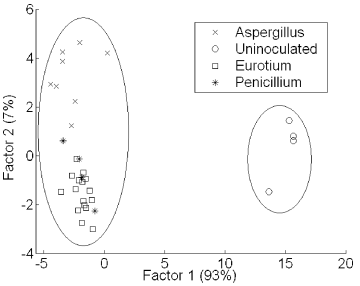

Figure 1. Two-dimensional DFA plot from 10 days in vitro growing cultures measured with the SH–MS e-nose configuration (50 min equilibration time, 50 °C temperature).

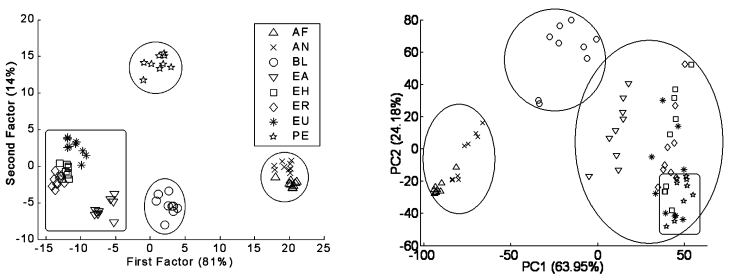

Figure 2. Two-dimensional projections from 10 days in vitro growing cultures measured with the SPME–MS e-nose configuration. (a, left) DFA; (b, right) PCA.

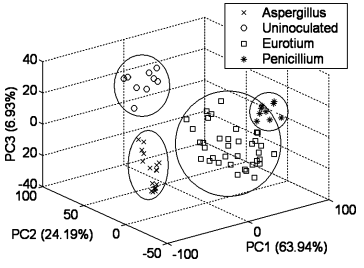

Figure 3. Three-dimensional PCA scores plot from 10 days in vitro growing cultures measured with the SPME–MS e-nose configuration.

time. An equilibration time equal to 50 min was chosen to ensure maximum concentration of volatile compounds on the headspace.

This ensured improved repeatability and sensitivity. Since four in vitro replicates for each type of sample were prepared, 32 measurements were performed. A 2D DFA plot (**Figure 1**) of the reduced response matrix shows two clearly separated clusters corresponding to inoculated and blank vials, respectively. A fuzzy ARTMAP neural network achieved a 97% success rate in the discrimination between inoculated and uninoculated vials. This success rate fell down to 59% when attempting to discriminate among fungal genera (**Table 1**).

Results of SPME Sampling. A total of 72 measurements were performed to test in vitro SPME−MS measurements. Therefore, the resulting data matrix had 72 rows and 86 columns (*m*/*z* ranged from 35 up to 120). **Figure 2**a shows a 2D DFA plot of the restricted response matrix, while **Figure 2b** shows a 2D score plot from the original data matrix. Samples belonging to *Aspergillus*, *Penicillium*, and *Eurotium* clustered together with low dispersion and without overlapping with blank controls, which were clearly separated from the rest. The first two factors from DFA accounted for 95% of the variance in the data. In the case of the PCA, the variance gathered by the two first factors decreased to 88%. In this case, *Penicillium* and *Eurotium* isolates appear to overlap. However, use of the third principal component leads to discrimination between these two genera (**Figure 3**). In both 2D DFA and PCA plots *E. amstelodami* can be distinguished from the other *Eurotium* species. Looking at *Aspergillus* samples, there is a clear separation between species *A. flavus* and *A. niger*.

Applying DFA coupled to a fuzzy ARTMAP neural network model resulted in a 100% success rate when discriminating between fungal growth and blank vials. Also a 100% classification was achieved when classifying fungal genera. The attempt

Table 2. Results Achieved by Use of a PCA–DFA–Fuzzy ARTMAP Model for Classifying Blank/Inoculated Vials (Fungal Growth) and Genera[a]

fungal growth (h)	goal differentiation between	total exptl points	failures	success rate (%)
	In Vitro Measurements			
48	fungal growth	16	0	100
	genera	16	0	100
72	fungal growth	16	0	100
	genera	16	0	100
96	fungal growth	16	0	100
	genera	16	0	100
168	fungal growth	16	0	100
	genera	16	0	100
	In Situ Measurements			
24	fungal growth	40	5	88
	genera	40	17	58
48	fungal growth	40	1	98
	genera	40	9	78
72	fungal growth	40	0	100
	genera	40	9	78
96	fungal growth	40	0	100
	genera	40	5	88
168	fungal growth	40	0	100
	genera	40	5	88

[a] Tests were performed for in vitro and in situ growing fungal cultures, and sampling was always performed for SPME.

to classify samples according to their fungal species reached a 92% success rate (**Table 2**). The system misclassified seven measures out of 72, confusing two EH, three ER, one AF, and one AN. All failures were mistaken between species from the same genus.

Both PCA and DFA plots show a clear distinction between *E. amstelodami* and other *Eurotium* species. These results seem to be in good agreement with previous works where Börjesson et al. (*3*, *4*, *22*) described some differences between the volatile pattern profiles from *E. amstelodami* and other fungal species due to its lower percentage of alcohol release (50% for *E. amstelodami* vs 80% in other fungal species studied).

The SPME technique gave better results than the SH technique when applied as the sampling system coupled to our MS-based e-nose. It achieved better repeatability and it was more sensitive due to its ability to concentrate volatile analytes. SPME allows distinguishing between fungal genera or even between several species. On the basis of these results it was decided to continue the studies with the SPME−MS e-nose configuration since the instrument is meant for the fast detection of fungal growth at early stages.

In Vitro Fungal Growth Monitoring. Table 2 shows the success rates in the classification of in vitro samples between

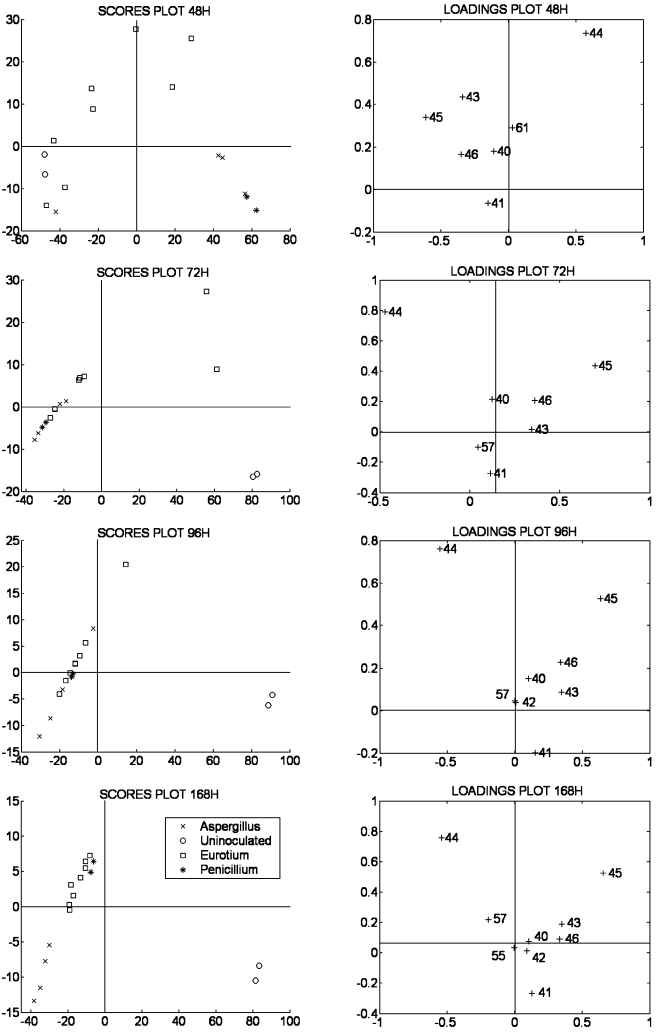

Figure 4. Two-dimensional PCA scores and loadings plots corresponding to 48, 72, 96, and 168 h of headspace monitoring of in vitro growing cultures.

two categories (inoculated and blank vials) and between four categories (corresponding to three fungal genera plus blank vials). These measurements were performed between 48 and 168 h after inoculation. By application of DFA−fuzzy ARTMAP on the reduced response matrix, the system achieved a 100% success rate after 48 h from inoculation when attempting to discriminate both between inoculated and blank samples and among fungal genera. This demonstrates that the SPME−MS-based e-nose is a suitable tool for on-line in vitro monitoring and early detection of unwanted fungal spoilage.

One of the advantages of a MS-based e-nose over e-noses based on gas sensors or other devices is the possibility to obtain structural information from the samples. Intensity on the mass detector is a function of the ion patterns in the fragmentation of each molecule present in the headspace of fungal cultures. Therefore, depending upon the molecules present in the headspace, there will be different mass intensities detected by the instrument.

Making a loadings and scores plot on the response matrix enables us to establish qualitative correlations between samples and variables (*m*/*z* fragments). We constructed PCA models for in vitro measurements at 48, 72, 96, and 168 h. **Figure 4** shows 2D PCA score plots with their corresponding loading plots for each batch performed at the different incubation times. Electron impact ionization mode (EI) causes considerable fragmentation, leading to overlapping fragments and parent ions. Because of the poor selectivity of the *m*/*z* fragments, they cannot be directly correlated with the presence or absence of a volatile. Nevertheless, mapping pattern fragmentation by means of loading plots could give some relevant information. In this case ions corre-

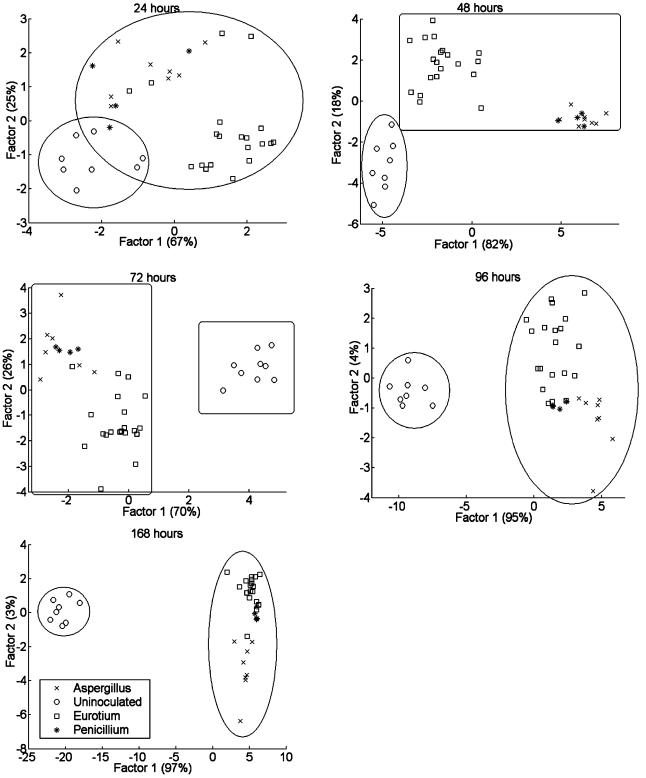

Figure 5. Two-dimensional DFA projections corresponding to 24, 48, 72, 96, and 168 h of headspace monitoring of in situ growing fungal cultures.

sponding to the highest loading values for the first three PC were kept to perform loading plots. **Figure 4** shows an inverse correlation between uninoculated control blanks and *m*/*z* 44. This inverse correlation means that samples with low or no presence of *m*/*z* 44 correspond to blank controls and the rest share a high presence of *m*/*z* 44. Since fragment of *m*/*z* 44 corresponds to the base peak of CO_2, the relationship between the evolution of accumulated CO_2 and fungal growth was found to be significant. The MS-based e-nose differentiates between inoculated and uninoculated samples mainly on the basis of the production of CO_2 by fungi. This is in good agreement with the literature. Börjesson et al. (*3*) measured the concentration of CO_2 produced by *A. flavus*, *A. amstelodami*, *Penicillum cyclopium*, and *Fusarium culmorum* during 14 days of fungal growth and showed a continuous rise in CO_2 concentration. They also studied the volatiles released by six fungal species on grains and found that the relationship between accumulated CO_2 evolution and fungal growth was significant (*22*). The predominating presence of CO_2 in our PCA model is due to the sealing of the vials with silicon septa once they have been inoculated. The presence of other ions such as *m*/*z* 41, 42, 43, 45, 46, 55, and 57 can be considered as second-order and less relevant ions. They can be associated to other related fungal metabolites such as ethanol (*m*/*z* 45 and 46), 3-methyl-1-butanol (*m*/*z* 41, 42, and 55), 2-methyl-1-propanol (*m*/*z* 41−43), and 1-octen-3-ol (*m*/*z* 57). This is in good agreement with the review paper by Magan and Evans (*5*). In their study they reviewed the types of volatiles produced by grain spoilage fungi and listed the most common volatiles found and the fungal species involved. The major volatile compounds were found to be 3-methyl-1-butanol, 2-methyl-1-propanol, 1-octen-3-ol, and other 8-carbon ketones and alcohols.

Anyway, the origin of these fragments cannot be ensured. As **Figure 4** shows, the loading maps remained almost unchanged from 72 to 168 h of incubation. The only remarkable changes in the loading maps appear betweem the plots corresponding to 48 and 72 h of fungal growth. Finally, the presence of ion 55 close to ions 42 and 41 can indicate the raising of 3-methyl-1-butanol after 178 h of fungal growth.

In Situ Fungal Growth Monitoring. Results obtained on the in vitro preliminary experiment encouraged us to study the performance of the electronic nose on in situ growing cultures over bakery product analogues. This is a much more realistic but difficult task since analogues can produce their own volatile pattern profiles; these volatiles produce additional signals in the mass detector that introduce noise into the fungal growth predictive model.

Table 2 summarizes the results obtained when bakery products analogues were measured along the first stages of fungal spoilage. In situ monitoring was performed 24, 48, 72, 96, and 168 h after inoculation. Distinction between blank and inoculated samples reached an 88% success rate 24 h after inoculation, 98% after 48 h, and 100% after 72 h. **Figure 5** shows 2D DFA plots corresponding to fungal cultures sampled

at 24, 48, 72, 96, and 168 h after incubation. In the first plot (24 h after inoculation), blank vials overlap with inoculated samples. As the time of incubation increases, the dispersion between samples belonging to the same class decreases and the difference between classes grows. After 48 h, uninoculated and blank samples can be clearly distinguished. This is in good agreement with our prediction model that achieved a 98% success rate after 48 h. Moreover, samples belonging to the same genus appear to cluster together, a tendency that becomes more pronounced as the time of incubation increases. Once again these results are in good agreement with our predictive model because after 48 and 72 h the instrument achieved a 78% success rate and 88% after 96 h in the prediction of fungal genera

From the results obtained on the monitoring of in vitro fungal growth, it can be concluded that during the first 24 h fungi are mainly producing CO_2 and other common metabolites associated with primary fungal growth and structures formation such as 3-methyl-1-butanol, 2-methyl-1-propanol, and 1-octen-3-ol indicative from fungal presence. According to Börjesson et al. (*3*), it seems that species identification may not be possible at this early stage, since the compounds produced in the highest amounts are similar for different species. The volatile pattern profile is, therefore, very similar and does not allow discrimination between species. Characteristic volatiles that might allow species classification are mainly produced during secondary metabolism. After 48 h of incubation, the system is able to predict fungal genera with a 78% success rate, which implies that secondary metabolism has started. Sporulation happens 72−96 h after inoculation depending on fungal species, which leads to an increase in several volatile compounds generating different pattern profiles for each fungal genus or species. This allows the best discrimination results in our model.

Taking into account the results obtained with the measurements performed on in situ samples, a real quality analysis application in a bakery factory seems feasible. Since SPME sampling time was 20 min and desorption time was 5 min, each measurement took 25 min to be executed. In a real application, many SPME fibers can be used in parallel (e.g., four) and the system can get a throughput of a measurement every 5 min. Moreover, since some quality control departments already have GC−MS equipment, they can convert their units into a MS-based E-nose in a rather straightforward manner, just coupling the optimal sampling system and using additional pattern recognition software.

The next step should be focusing on tuning the prototype to work in a bakery plant, where it would be trained to monitor a certain number of samples from each batch. These measurements made at the quality control laboratory would allow detection of batches likely to be spoiled before their expiration date and thus produce a rejection decision before the batch leaves the production plant.

LITERATURE CITED

(1) Needham, R.; Magan, N. Detection and differentiation of microbial spoilage organisms of bakery products in vitro and in situ. *Proc. ISOEN* **2002**, 240−241.

(2) Schnürer, J.; Olsson, J.; Börjesson, T. Fungal Volatiles as Indicators of Food and Feeds Spoilage. *Fungal Genet. Biol.* **1999**, *27*, 209−217.

(3) Börjesson, T.; Stöllman, U.; Adamek, P.; Kaspersson, A. Analysis of Volatile Compounds for Detection of Molds in Stored Cereals. *Cereal Chem.* **1989**, *66*, 300−304.

(4) Börjesson, T.; Stöllman, U.; Schnürer, J. Volatile Metabolites and Other Indicators of *Penicillium aurantiogriseum* Growth on Different Substrates. *Appl. Environ. Microbiol.* **1990**, *56*, 3705−3710.

(5) Magan, N.; Evans, P. Volatiles as an indicator of fungal activity and differentiation between species, and the potential use of electronic nose technology for early detection of grain spoilage. *J. Stored Prod. Res.* **2000**, *36*, 319−340.

(6) Olsson, J.; Börjesson, T.; Lundstedt, T.; Schnürer, J. Volatiles for mycological quality grading of barley grains: determinations using gas chromatography−mass spectrometry and electronic nose. *Int. J. Food Microbiol.* **2000**, *59*, 167−178.

(7) Olsson, J.; Börjesson, T.; Lundstedt, T.; Schnürer, J. Detection and quantification of ochratoxin A and deoxynivalenol in barley grains by G-MS and electronic nose. *Int. J. Food Microbiol.* **2002**, *72*, 203−214.

(8) Börjesson, T.; Eklöv, T.; Jonsson, A.; Sundgren, H.; Schnürer, J. Electronic Nose for Odor Classification of Grains. *Cereal Chem.* **1996**, *73*, 457−461.

(9) Trihaas, J.; Van den Tempel, T.; Nielsen, V. Electronic nose: Smelling the microbiological quality of cheese, *Proc. ISOEN* **2002**, 380−384.

(10) Shin, H. W.; Llobet, E.; Gardner, E.; Hines, E. L.; Dow, C. S. Classification of the strain and growth phase of cyanobacteria in potable water using an electronic nose system. *IEEE Proc. Sci. Meas. Technol.* **2000**, *147*, 158−164.

(11) Keshri, G.; Voysey, P.; Magan, N. Early detection of spoilage moulds in bread using volatile production patterns and quantitative enzyme assays. *J. Appl. Microbiol.* **2002**, *92*, 165−172.

(12) Vernat-Rossi, V.; Garcia, C.; Talon, R.; Denoyer, C.; Berdagué, J. L. Rapid discrimination of meat products and bacterial strains using semiconductor gas sensors. *Sens. Actuators, B* **1996**, *37*, 43−48.

(13) Magan, N.; Pavlou, A.; Chrysanthakis, I. Milk-sense: a volatile sensing system recognise spoilage bacteria and yeast in milk. *Sens. Actuators, B* **2001**, *72*, 28−34.

(14) Marsili, R. T. SPME−MS−MVA as an Electronic Nose for the Study of Off-Flavors in Milk. *J. Agric. Food Chem.* **1999**, *47*, 648−654.

(15) Marsili, R. *Flavor, Fragance and Odor Análisis*; Marcel Dekker: New York, 2002.

(16) Nilson, T.; Larsen, T. O.; Montanarella, L.; Madsen, J. O. Application of headspace-solid-phase microextraction for the analysis of volatile metabolites emitted by *Penicillium* species. *J. Microbiol. Methods* **1996**, *25*, 245−255.

(17) Demyttenaere J. C. R.; Moriña R. M.; Sandra, P. Monitoring and fast detection of mycotoxin-producing fungi based on headspace solid-phase microextraction and headspace sorptive extraction of the volatile metabolites. *J. Chromatogr. A* **2003**, *985*, 127−135.

(18) Jelen, H. H. Use of solid-phase microextraction (SPME) for profiling fungal volatile metabolites. *Lett. Appl. Microbiol.* **2003**, *36*, 263−267.

(19) Abellana, M.; Benedí, J.; Sanchis, V.; Ramos, A. J. Water activity and temperature effects on germination and growth of *Eurotium amstelodami*, *E. chevalieri* and *E. herbariorum* isolates from bakery products. *J. Appl. Microbiol.* **1999**, *87*, 371−380.

(20) Dittmann, B.; Nitz, S. Strategies for the development of reliable QA/QC methods when working with mass spectrometry-based chemosensory systems. *Sens. Actuators, B* **2000**, *69*, 253−257.

(21) Sanz, C.; Ansorena, D.; Bello, J.; Cid, C. Optimizing Headspace Temperature and Time Sampling for Identification of Volatile Compounds in Ground Roasted *Arabica* Coffee. *J. Agric. Food Chem.* **2001**, *49*, 1364−1368.

(22) Börjesson, T.; Stöllman, U.; Schnürer, J. Volatile metabolites produced by six fungal species compared with other indicators of fungal growth on cereal grains. *Appl. Environ. Microbiol.* **1992**, *58*, 2599−2605.

Received for review April 14, 2004. Revised manuscript received July 21, 2004. Accepted August 1, 2004. This work was funded by INIA (Project CAL01-037-C03-2) and by FIS (Thematic Network C03/08).

JF049399R

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper II

Vinaixa, M.; Vergara, A.; Duran, C.; Llobet, E.; Badia, C.; Brezmes, J.; Vilanova, X.; Correig, X., Fast detection of rancidity in potato crisps using e-noses based on mass spectrometry or gas sensors. *Sensors and Actuators B,* **2005,** 106(1), 67-75.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Sensors and Actuators B 106 (2005) 67–75
SENSORS AND ACTUATORS B CHEMICAL
www.elsevier.com/locate/snb

Fast detection of rancidity in potato crisps using e-noses based on mass spectrometry or gas sensors

M. Vinaixa [a], A. Vergara [a], C. Duran [a], E. Llobet [a,*], C. Badia [b], J. Brezmes [a], X. Vilanova [a], X. Correig [a]

[a] *Department Electronic Engineering, Universitat Rovira i Virgili, Avda. Països Catalans, 26, Campus Sescelades, 43007 Tarragona, Spain*
[b] *Frit Ravich, S.L., Pol. Ind. Puigtió s/n, 17412, Maçanet de la Selva, Girona, Spain*

Available online 2 July 2004

Abstract

Well-established methods to assess rancidity in potato crisps such as the Rancimat or the acid degree value are time-consuming and labour-intensive. Here, we report on alternative methods, based on e-nose technology, to assess rancidity directly from potato crisps without any previous oil extraction step. This simplifies sample preparation, avoids the use of solvents or high temperatures and significantly speeds up the measurement process (from several hours down to 25 min). Two different e-noses were implemented. One was based on SPME coupled to fingerprint MS and the other one was based on dynamic headspace sampling and an array of metal oxide gas sensors. The two e-noses were used to classify crisps according to four stages of oxidative rancidity. While the MS e-nose reached a 100% success rate in this classification, the success rate of the GS e-nose was 68%. These results show that e-nose technology can be a useful tool for the crisp industry.

We show that it is possible to reliably assess rancidity in potato crisps by either a mass spectrometry or a gas sensor-based electronic nose. The two approaches are presented and their performance compared in the framework of this application.

Keywords: Metal oxide gas sensors; Mass spectrometry-based e-nose; Crisp rancidity

1. Introduction

Potato crisps are considered one of the most popular snack products in the world. Usually, they are made by deep-frying fresh potato slices in a vegetable oil bath. The reaction of lipid components with oxygen in the presence of light and heat is a major source of off-odours/flavours in food and, particularly, in potato crisps. During the deep-frying process, vegetable oil is under temperature stress and this can induce onset of rancidity as a consequence of oxidative reactions of lipids present in the oil. From the standpoint of food oxidation, the important lipids are the ones containing unsaturated fatty acids, particularly oleic acid (C18:1), linoleic acid (C18:2) and linoleic acid (C18:3) [1]. Potato crisps are fried in oils that contain a high amount of all of these. Unsaturations are reactive centres liable to be affected by oxidation. So, the greater the number of double bonds, the higher the probability that the fatty acid will react with oxygen to generate undesirable odours and flavours in the product. The oxidation of lipids results in the formation of primary and secondary decomposition products, including hydroperoxides, carbonyls, alcohols, esters, carboxylic acids and hydrocarbons [2], which generally have unpleasant odour and may conduce to rancidity. Various factors can influence the occurrence of rancidity in crisps, such as storage conditions, presence of antioxidants, oil type, time of deep-frying, heat, presence of pro-oxidant metals, oxygen and moisture among other factors.

Two very important aspects for potato crisps producers are the detection of rancidity and its associated off-odours/flavours and the estimation of shelf-life. There are basically two reasons why it is important to monitor to what extent oil has undergone oxidation:

- Previous knowledge, i.e. an estimate, on the useful life of frying oil contributes to reduce the cost of the deep-frying process. There is an obvious economic advantage when crisp producers can appropriately determine the useful life of frying oils. Premature discarding of oils results in economic loss and, on the other hand, overuse of frying oil greatly affects the quality of fried products and causes undesirable nutritional effects [3].

* Corresponding author. Tel.: +34 977 558502; fax: +34 977 559605.
E-mail address: ellobet@etse.urv.es (E. Llobet).

0925-4005/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2004.05.038

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

- The second reason is an ever increasing consumer about quality and safety of food products. According to Marsili [4], the food industry needs the development of equipment and techniques to trace the quality of raw materials and finished products, not only in the production plant, but also during storage and vending. Monitoring of off-odours/flavours during the different processing steps should be conducted to ensure that the processes are being operated correctly. Finished products should be monitored too, ensuring that no off-flavours have developed. All these improvements would greatly contribute to food quality and consumer satisfaction.

Nowadays, there are some 360 procedures to verify the quality of oil either during the process of frying or in finished products [1,5]. However, there is not a reliable, easy-to-use and fast method to determine rancidity in potato crisps. The most well-established methods for the evaluation of rancidity are based on sensory evaluation or chemical analysis. Some of these methods are revised below:

- Sensory analysis: Samples are evaluated by a panel of experts. This is a slow and expensive method. It requires the panel to be integrated by highly trained personnel and, however, results can be somewhat subjective.
- Peroxide value (PV): This method determines all the substances, in terms of milliequivalents of peroxide per gram of sample, which oxidise potassium iodine under the conditions of the test. These substances are generally assumed to be peroxides or other similar products of fat oxidation. The higher the PV, the more oxidised the fat is and the higher the likelihood of off-odours/flavours.
- Acid degree value (ADV): This is a titration method. It obtains the amount of potassium hydroxide required to neutralise the free fatty acids hydrolysed with 95% ethanol. The higher the ADV, the higher the level of free fatty acids present in the oil. Free fatty acids indicate undesirable hydrolysis, which results in flavour deterioration. Che Man et al. [3] showed that ADV was an important indicator of frying oil quality, and highly correlated with the shelf-life of potato chips.
- Iodine value (IV): Indicates the number of double bonds or degree of unsaturation in lipids. It can be used as an estimate of the oxidation stability of a lipid.
- HPLC analysis: Determination of the fatty acid composition of oil. This method provides fatty acid profiles and is more informative than IV.
- IR an UV band absorption of some oxidation by-products like hexanal, pentanal and pentane.
- Methods based on the measurement of some physical properties of oil, such as melting point, solid fat index and refractive index.
- Rancimat test: Measures the susceptibility of oil to oxidation. An oil sample is kept at 120 °C in a vessel where air flows to extract volatiles from the headspace. These volatiles are then collected in water. The conductivity of water is monitored and results expressed as Rancimat hours indicate the time at which oxidative rancidity occurs. Rancidity triggers a sharp increase in water conductivity. Since this test is very informative about the resilience to rancidity of oils, it has become a reference in the crisp industry.

All the methods cited above can be used to assess rancidity in potato crisps, provided that a process to extract oil from the crisps is performed. Oil extraction is a very time-consuming, complex and labour-intensive step for routine quality control applications. Furthermore, the solvents or the methods used can induce oxidation and distort final results. Since the crisp industry demands a large number of samples to be analysed and high sample throughput, there is a need for faster and simpler methods to assess crisp rancidity and off-odours/flavours. In this context, the use of e-nose technologies would be of help.

In the last decade, the use of e-nose technology in many food-related applications has been studied. Electronic noses are multisensor instruments that use a suitable pattern recognition engine to classify complex odour patterns. According to previous works, electronic noses based on metal oxide gas sensors are suitable for the discrimination of different stages of lipid oxidation in oils [6–8]. In the last few years, mass spectrometry-based e-noses (MS e-noses) are becoming an increasingly used alternative (or complement) to gas sensor-based e-noses in food quality applications [4,9]. The use of pre-concentration and extraction techniques such as solid-phase micro-extraction (SPME) have improved the sensitivity and reproducibility of MS e-noses [4,10].

In this work, we report, for the first time, on the design and use of two e-noses to assess rancidity directly from potato crisps, without any previous oil extraction step. This greatly simplifies sample preparation, avoids unwanted artefacts derived from oil extraction and speeds up the measurement process. The two e-noses are based on SPME–MS and an array of semiconductor gas sensors (GS e-nose), respectively. In the next section, details on the e-nose architectures sample preparation and measurements run are given. In Section 3, the results are shown and the usefulness of the methods implemented for the application considered is discussed.

2. Experimental

2.1. Experiment 1

2.1.1. Crisp samples

Four boxes (labelled *A–D*) with 200 g packs (12 packs per category) of potato crisps were prepared by Frit Ravich, S.L. These crisps belonged to the same frying batch of 50% palm and 50% sunflower oil, but they underwent different rancidity accelerating treatments:

- Crisps in box *A* were stored during 28 days in a dry and dark conservation chamber, where their temperature was kept around 20 °C.

- Crisps *B–D* were kept during 14, 21 and 28 days, respectively, in a rancidity accelerating chamber. The chamber was kept at high temperature (around 40 °C) and UV light was used to promote oxidation. As soon as the samples within a given category finished their ageing treatment, they were removed from the rancidity chamber and stored in the conservation chamber to maintain unchanged the rancidity stage reached.

2.1.2. Measurement procedures

The content of each potato pack was split to perform consistent measurements with an MS e-nose and a GS e-nose.

2.1.2.1. Metal oxide sensors-based electronic nose. The electronic nose system was designed to measure volatiles directly from the packs of the crisps. The system consisted of a sensor chamber where seven TGS-type sensors and five FIS sensors were housed, several electrovalves, tubing and a pump (see Fig. 1a). A similar set-up is described elsewhere [11]. The measurement procedure consisted of two steps. In the first step, (measurement phase) the electrovalves were set to form a closed loop between the sensor chamber and the pack containing the crisps under analysis. The air flow (150 ml/min) was re-circulated, which caused a dynamic sampling of the crisps' headspace. During this phase, which lasted 10 min, the resistance of the sensors was acquired and stored for later processing. Finally, in the second step (cleaning phase), the crisp pack was removed and the system was cleaned with dry air during 20 min before a new measurement could start.

After a pack of crisps had been measured by the GS e-nose, 4 ± 0.2 g of the crisps were crushed and put into a 20 ml vial that was immediately capped and sealed with a Teflon septum. A subsequent analysis with the MS e-nose system was run.

2.1.2.2. Mass spectrometry-based electronic nose. A Shimadzu QP 5000 gas chromatograph–mass spectrometer was used to implement an MS e-nose. The separation column was replaced by a 5 m deactivated fused silica column to co-elute all volatile components achieving one single peak for all the components in the headspace of crisps. The column

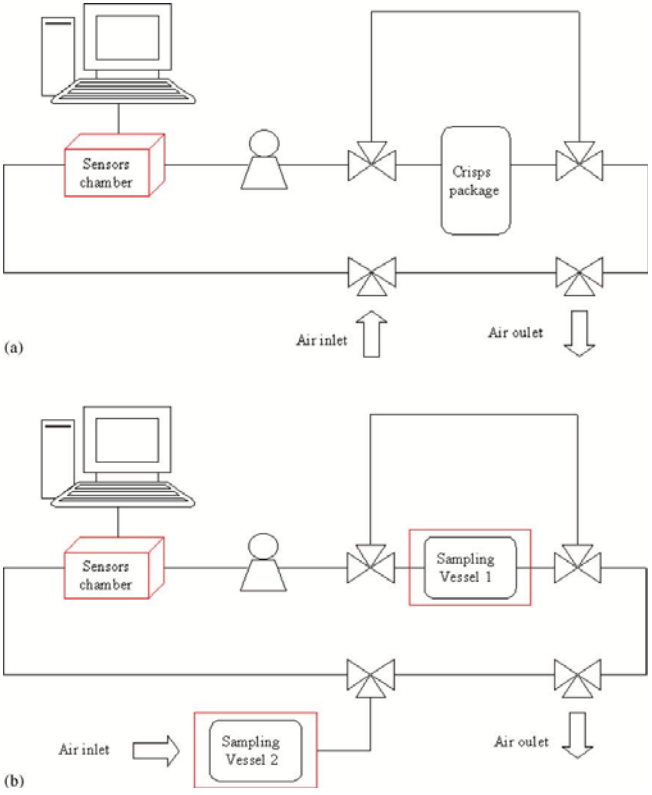

Fig. 1. Block diagram of the gas sensor-based e-nose used in experiment 1 (a) and experiment 2 (b).

was kept isothermal at 250 °C and the helium flow was set to 1.4 ml/min. This implies that the components in the headspace of crisps were directly analysed without chromatographic separation. For a given measurement, the resulting mass spectrum gives a fingerprint that is characteristic of the volatiles present in the headspace of the sample.

The vials that contained the samples to be measured were placed inside a thermostatic bath (50 °C) to promote the presence of volatiles in the headspace. SPME was performed by introducing a 75-μm Carboxen/PDMS fibre into the vial and exposing it to the headspace of crisps for 20 min. Thermal desorption of the volatiles trapped on the fibre was conducted for 3 min in the chromatograph injection port at 300 °C. It was equipped with a 0.75-mm i.d. liner to optimise SPME desorption and sample delivery onto the column The split valve was closed during desorption. The quadrupole mass spectrometer acquired in scan mode, and the mass range used was *m/z* 35 to *m/z* 390 at 0.5 scan/s. To ensure the complete cleaning of the fibre, it was left five additional minutes in the injector port.

2.1.2.3. Rancimat, ADV and chromatographic profiles. Rancimat and ADV tests were performed at the quality laboratory of Frit Ravich, S.L. The chromatographic profiles were obtained at the Gas Sensor Lab of the University Rovira i Virgili, using a Shimadzu QP 5000 GC/MS. After sample preparation (as described above), the SPME fibre was introduced into the GC injection port and thermally desorbed for 5 min at 250 °C onto an Equity-5 poly (5%diphenyl/95% dimethylsiloxane) (30 m × 0.25 mm × 0.25 μm) capillary column, purchased from Supelco Inc. The injector port was also equipped with a 0.75-mm i.d. liner. The GC oven was held at 45 °C during 1.5 min. Then, its temperature was raised up to 250 °C at 6 °C/min rate. Helium at 1.2 ml/min was used as carrier gas. Mass detector was operating in the electron impact ionisation mode (70 eV) with a scan range of 35 to 290 amu. The ion source temperature was kept at 250 °C.

2.2. Experiment 2

A new experiment was performed with an improved version of the gas sensor-based electronic nose. The main differences with the previous system were the use of a 12-element TGS-type sensor array (the seven TGS sensors already used in the first experiment + five TGS sensors added) and a new sample delivery method.

2.2.1. Crisp samples

Four boxes (labelled *A–D*) with 200 g packs of crisps (12 packs per category) were prepared by Frit Ravich S.L. in a similar way to the crisps used in experiment 1.

- Crisps in box *A* were stored during 18 days in the same conservation chamber used in experiment 1.
- Crisps *B–D* were kept during 6, 12 and 18 days, respectively, in the rancidity accelerating chamber used in experiment 1.

2.2.2. Improved gas sensor-based e-nose

The sample delivery system consisted of two temperature-controlled stainless-steel vessels (see Fig. 1b): a sampling vessel and a reference vessel. These chambers were identical and kept heated at 70 °C. To run a measurement, crisp samples (60 ± 1 g) were placed into an aluminium tray and inserted into the sampling vessel. An identical (but empty) aluminium tray was also placed inside the reference vessel. New aluminium trays were used at each new measurement to avoid cross-contamination between samples. The measurement procedure was as follows:

In the first step (concentration phase), the crisps were heated at 70 °C for 30 min inside the sampling vessel, which was kept closed by the electrovalves. This allowed the volatiles from the crisps to concentrate in the headspace. During this phase, clean air flowed at 150 ml/min through the sensor chamber via the reference vessel.

In the second step (measurement phase), the electrovalves were set to form a closed loop between the sensor chamber and the sampling vessel. The air flow (150 ml/min) was re-circulated, which caused a dynamic sampling of the crisps' headspace. During this phase, which lasted 10 min, the resistance of the sensors was acquired and stored for later processing. The use of identical sampling and reference vessels is essential to ensure that sensor responses are solely due to the volatiles in the headspace of the crisps.

Finally, in the third step (cleaning phase) the crisps were removed from the sampling vessel and the system was cleaned with dry air during 20 min, before the concentration phase of a new measurement could start.

3. Results and discussion

3.1. Experiment 1

3.1.1. Rancimat, ADV and chromatographic profiles

Fig. 2(a) shows the Rancimat and ADV results for crisp samples *A–D* in experiment 1. The monotonous decrease in Rancimat time combined with an increase in the ADV for samples *A–D* shows that these categories correspond to crisps with increasing oxidative rancidity. Furthermore, the clear differences in Rancimat time between categories suggest that crisps in different categories are in significantly different rancidity stages. While there is an important difference in Rancimat time between samples *A* and *B*, they share an almost identical ADV. This suggests that ADV may not be suitable to assess the early stages of rancidity in crisps.

Chromatographic profiles of the headspace of crisps belonging to class *A* (fresh) and class *D* (rancid) were

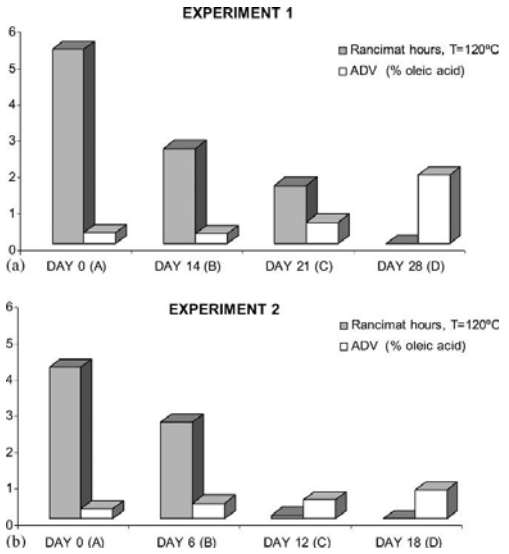

Fig. 2. Results of the Rancimat and ADV tests for potato crisps in eperiment 1 (a) and experiment 2 (b).

measured. These profiles are shown in Fig. 3. Different volatile molecules appear or substantially increase their signal intensity as rancidity develops. These include acetone (peak no. 1), acetic acid (2), pentanal (3), hexanal (5), heptanal (6), hexanoic acid (9), 3-octen-2-one (10), 2-octenal (11), 2,3-octanedione (12), 2,4-decadienal (14) and undecane (15). Some of these components, such as acetic acid, pentanal, hexanal, heptanal and hexanoic acid, have been reported to be present in the chromatographic profiles of rancid chips [12]. 2,4-decadienal [14], which is present in a similar intensity in fresh and rancid crisps, has been identified by GC-olfactometry [13,14] as a predominant note in deep-fried potato crisps.

3.1.2. Mass spectrometry-based e-nose

Nine replicate measurements per sample category were performed. For each measurement, a response spectrum was obtained by averaging mass spectra along the detected peak. The variables selected were from *m*/*z* 35 to *m*/*z* 120. The components identified as indicators of rancidity in the cromatographic profiles have base peaks that lie in the range selected. Therefore, the data matrix, ***R***, consisted of 86 columns (variables) and 36 rows (measurements). A linear discriminant analysis (LDA) was performed on ***R***. This is a supervised method (e.g. the classes to be discriminated are known before this analysis is performed). Geometrically, the rows of the response matrix, ***R***, can be considered as points in a multidimensional space. Discriminating axes are determined in this space in such a way that optimal separation of the predefined classes is attained. Like PCA, LDA finds new orthogonal axes (factors) as a linear combination of the input variables. Unlike PCA, however, LDA computes the factors as to minimise the variance within each class and maximise the variance between classes. The first factor will be the most powerful differentiating dimension, but later factors may also represent additional significant dimensions of differentiation.

The data matrix was mean-centred before the LDA was performed. If this scaling of the data is not performed, there is a risk of LDA ignoring mass intensities with low mean (but important for discriminating the four rancidity classes) in front of mass intensities with high mean. The two first factors accounted for more than 99% of the variance in the data. LDA results are shown in Fig. 4. Replicate samples of a given category cluster together with low dispersion, which shows the good repeatability of the MS e-nose. Fig. 4 shows that crisp samples with increasing rancidity appear ordered from left to right along the first factor. While samples from categories *A* (fresher) and *D* (more rancid) appear in clusters well apart, the clusters of categories *B* and *C* are very near. These results are in very good agreement with the Rancimat tests (see Fig. 1a). For example, while there is a moderate change in the Rancimat time between samples in categories

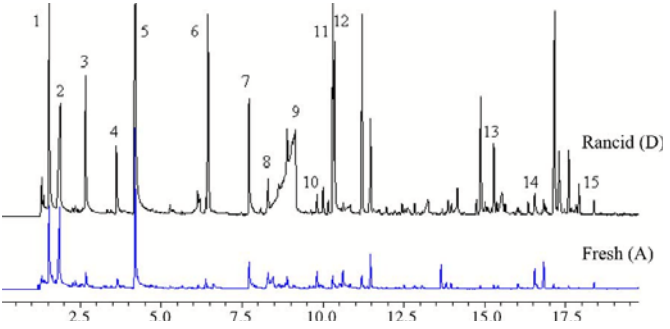

Fig. 3. Chromatographic profiles identified by GC/MS (1) acetone, (2) acetic acid, (3) pentanal, (4) pentanol, (5) hexanal, (6) heptanal, (7) 2-heptenal, (8) 1-octen-3-ol, (9) hexanoic acid, (10) 3-octen-2-one, (11) 2-octenal, (12) 2,3-octanedione, (13) 2-decenal, (14) 2,4-decadienal, (15) undecane.

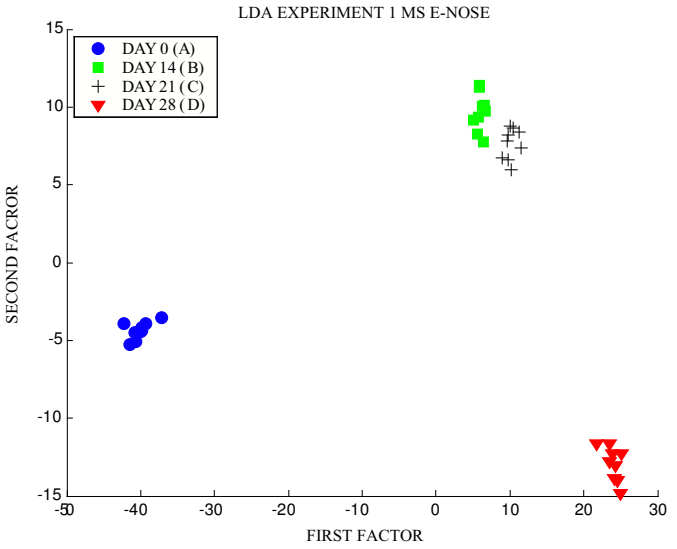

Fig. 4. Results of a linear discriminant analysis for the measurements gathered with the mass spectrometry-based e-nose in experiment 1.

B and C, there is a dramatic change in this parameter between samples in categories *A* and *B* and also between samples in *C* and *D*. Therefore, MS e-nose results are in excellent agreement with Rancimat results.

A fuzzy ARTMAP neural network was used to classify the samples within the four categories of rancidity (*A*–*D*). Because of the limited number of measurements available (36), the network was tested using the leave-one-out cross-validation method. Given n measurements, the network was trained n times using $n - 1$ training vectors. The vector left out during the training phase (i.e. unseen by the network) was then used for testing. Performance was estimated as the averaged performance over the n tests. For each iteration of the cross-validation process, a different row from the data was left out. The remaining 35 rows conformed the restricted data matrix. A pre-processing step was performed on the restricted data matrix, which consisted of computing a 4-class LDA and retaining the two first factors. The scores of the 35 measurements conformed the new data matrix. Therefore, this new matrix had 35 rows and 2 columns. The matrix was then normalised because the fuzzy ARTMAP network needs that its input data lie in the range [0, 1]. Once the data matrix had been pre-processed, it was used to train the neural network model. After the training phase, the network was validated using the vector that had been left out (i.e. validation vector). The procedure was as follows:

Because a LDA had been used as pre-processing, then the scores of the validation vector were calculated by projecting its components onto the space of factors. In the second step, the validation vector was normalised. Finally, the validation vector was input into the neural network model, which produced a classification result. The fact that the validation vector had been left out before any pre-processing of the data ensured that this vector was completely 'new' for the neural network.

The number of inputs to the network was set to 2 (the scores on the two first factors). The number of outputs was set to 4 because a 1-of-4 code was used for the different classes (*A*: 0001, *B*: 0010, *C*: 0100 and *D*: 1000). For example, the activation of the first output neurone (i.e. output pattern 0001), implies that an input vector is recognised as belonging to class *A* (fresh crisps). This approach aimed at identifying rancidity in a semi-quantitative way. The baseline vigilance parameter was set to 0. This is the recommended value for the vigilance since it allows for very coarse categories and the match tracking system will only refine these categories if necessary. The re-code rate was set to 0.5. This value allows the established categories to be modified if there is a persistent attempt to do so (slow recode). The value of the choice parameter was set to 0.1. The Fuzzy ARTMAP network could learn the training set in just one iteration. The number of committed nodes, which play a similar role as hidden neurones in multilayer perceptron networks, ranged between 4 and 6 after the network had been trained. Under these conditions, the success rate reached in rancidity classification was 100%. This shows that the SPME–MS e-nose was able to assess crisp rancidity from the volatiles present in the headspace of the crisps.

3.1.3. Gas sensor-based e-nose

The responses of the 12 metal oxide gas sensors to the different crisp samples were obtained. The feature extracted from each sensor response was the conductance change, defined as $\Delta G = G_{max} - G_o$, where G_{max} is the maximum value of the sensor electrical conductance in the presence of the volatiles from the headspace of the crisps, and G_o is the sensor conductance in the presence of air (i.e. the baseline conductance). The responses of the FIS sensors were very weak compared with the responses of the TGS sensors. Therefore, only the responses of the seven TGS sensors were used for further analysis. A LDA was performed on the mean-centred response matrix. The two first discriminant factors accounted for more than 99% of variance in the data. LDA results are shown in Fig. 5. While measurements that correspond to fresh crisps (class *A*) cluster together, the clusters of measurements corresponding to the remaining three classes appear clearly overlapped along the first and second discriminant factors. A fuzzy ARTMAP was used to classify the samples according to their rancidity stage. The same training and validation techniques employed with the MS e-nose were implemented. The neural network had seven inputs (seven TGS sensors) and four outputs. The number of committed nodes during the repeated training and validation processes varied between 8 and 12. Under these conditions, the success rate reached in rancidity classification was 56%. The samples misclassified belonged to categories *B–D*.

According to these results, the GS e-nose showed lower repeatability and discriminating power than the MS e-nose. However, an important difference between the two e-nose methods lies in sample preparation. While for the GS e-nose volatiles were sampled from the headspace of the crisps at room temperature, the MS e-nose made use of a SPME from the headspace of crisps heated at 50 °C. Therefore, the differences in classification success rate between the two e-noses could be due to significant differences in the headspaces sampled. This is why a new experiment was designed.

3.2. Experiment 2

3.2.1. Rancimat and ADV results

Fig. 2b shows the Rancimat and ADV results for crisp samples *A–D* in experiment 2. The oil used to deep-fry the crisps in experiment 2 had the same composition than the one used in the previous experiment. However its initial stage (class *A*) was, according to the Rancimat test, more evolved towards rancidity. The results of the Rancimat test showed that the classification of samples in four rancidity categories was going to be more challenging here, because samples in classes *C* and *D* had very similar Rancimat and ADV results.

3.2.2. Gas sensor-based e-nose

The GS e-nose with a re-designed sample delivery system was used. The responses of the 12 TGS-type metal oxide gas sensors to the different crisp samples in experiment 2 were obtained. The feature extracted from each sensor response was, once again, the conductance change. Since 12 replicate measurements per category were gathered, the response matrix had 48 rows and 12 columns. A LDA was performed on the mean-centred data matrix. The two first

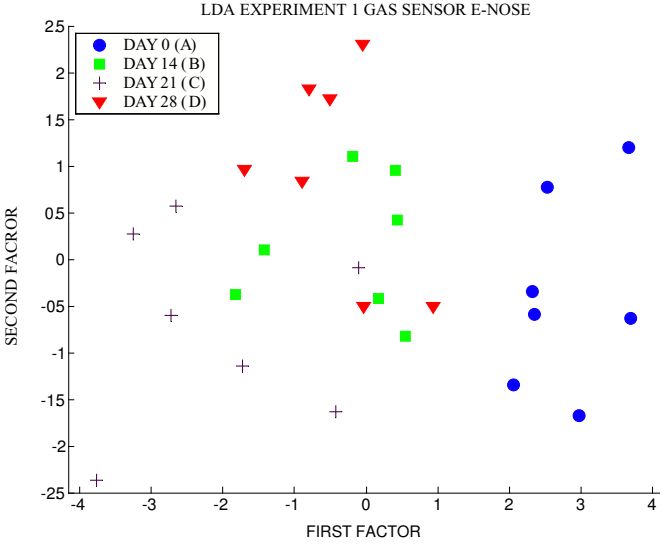

Fig. 5. Results of a linear discriminant analysis for the measurements gathered with the gas sensor-based e-nose in experiment 1.

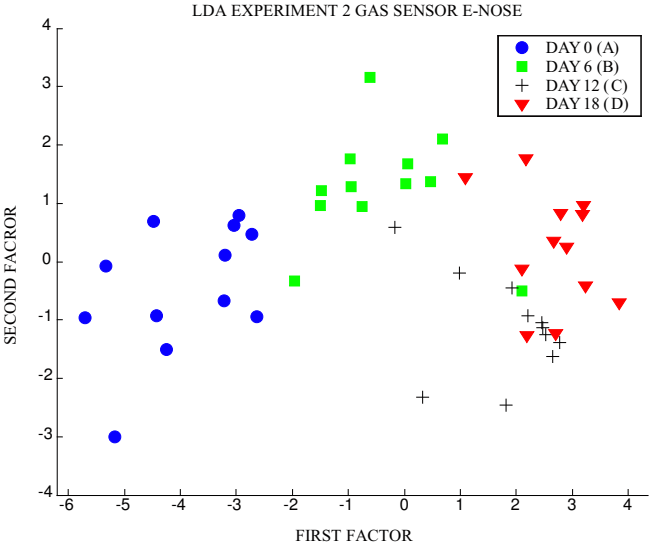

Fig. 6. Results of a linear discriminant analysis for the measurements gathered with the gas sensor-based e-nose in experiment 2.

factors accounted for more than 99% of variance in the data. LDA results are shown in Fig. 6. This figure shows that crisp samples with increasing rancidity appear ordered (with some overlapping) from left to right along the first factor. Overlapping occurs between samples in categories *C* and *D*, which is in good agreement with the very similar Rancimat times found for these categories. These results suggest that it is necessary to heat the crisps for a headspace that is representative of their rancidity stage to develop.

A fuzzy ARTMAP was, once again, used to classify the samples according to their rancidity stage. The same training and validation techniques employed in experiment 1 were implemented. The neural network had 12 inputs (12 TGS sensors) and 4 outputs. The number of committed nodes during the repeated training and validation processes varied between 7 and 10. Under these conditions, the success rate reached in rancidity classification was 68%. Considering that the classification problem envisaged in experiment 2 was more challenging, a 68% rate of successful classifications compares very favourably with the 56% success rate reached in experiment 1. Table 1 shows the confusion matrix for experiment 2. It can be seen that most confusions occur between consecutive rancidity categories (only two samples belonging to class *D* were misclassified as belonging to class *B*).

Table 1
Confusion matrix for the classification of crisps samples in four categories of rancidity (*A*–*D*) in experiment 2, using a fuzzy ARTMAP neural network

	Actual			
	A	*B*	*C*	*D*
Predicted as				
A	9	2	0	0
B	3	8	2	2
C	0	2	8	2
D	0	0	2	8

These promising results show that the GS e-nose with improved sample delivery system is able to perform a semi-quantitative classification of crisp rancidity.

4. Conclusions

In this work, we have reported on the design and use of two e-noses to assess rancidity directly from potato crisps, without any previous oil extraction step. This simplifies sample preparation, avoids the use of solvents to extract oil and speeds up the measurement process. The two e-noses were based on fingerprint mass spectrometry and an array of metal oxide gas sensors, respectively. While a single measurement using either the Rancimat or the ADV test takes typically some hours to complete, a measurement with the MS e-nose or the GS e-nose takes 25 and 40 min, respectively.

Sample conditioning plays a very important role. A mild heating of the crisps (up to 70 °C) is necessary for a headspace that is representative of their rancidity stage to develop. Under these conditions, the MS e-nose and the GS e-nose have been found sensitive enough and suitable for semi-quantitatively assessing rancidity in potato crisps. The results obtained by both e-nose instruments are in very good

agreement with the Rancimat test. Therefore, the assessment of crisp rancidity using e-nose technology could become a routine test in the quality laboratories of crisp producers.

Further work is in progress to analyse the shelf-life of potato crisps.

Acknowledgements

The authors gratefully acknowledge financial support from the Spanish Institute of Agro-Food Technology, INIA, under Project no. CAL01-052-C2.

References

[1] R. Marsili, Controlling the Quality of Fats and Oils, 1993. http://www.foodproductdesign.com.
[2] N. Shen, S. Duvick, P. White, L. Pollack, Oxidative stability and AromaScan analyses of corn oils with altered fatty acid content, J. Am. Oil Chem. Soc. 76 (12) (1999) 1425–1429.
[3] Y.B. Che Man, J.L. Liu, R.A. Rahman, B. Jamilah, Shelf-life of fried potato chips using palm olein, soybean oil and their blends, J. Food Lipids 6 (1999) 287–298.
[4] R. Marsili, Flavor, Fragance and Odor Analysis, Marcel Dekker, Inc., New York, 2002.
[5] AOCS, Official Methods and Recommended Practices of the American Oil Chemists' Society, fifth ed., 1998.
[6] M. Adechy, V.P. Shiers, J.B. Rossell, Study of rancidity and resistance to oxidation in edible oils and fats using electronic nose technology in comparison with conventional analytical techniques, Leatherhead Food R.A. Research Reports, no. 751, 1998.
[7] Y.M. Yang, K.Y. Han, B.S. Noh, Analysis of lipid oxidation of soybean oil using the portable electronic nose, Food Sci. Biotechnol. 9 (3) (2000) 146–150.
[8] M. Muhl, H.U. Demisch, F. Becker, C.D. Kohl, Electronic nose for detecting the deterioration of frying fat – comparative studies for a new quick test, Eur. J. Lipid Sci. Technol. 102 (2000) 581–585.
[9] R.T. Marsili, SPME–MS–MVA as an electronic nose for the study of off-flavors in milk, J. Agree. Food Chem. 47 (1999) 648–654.
[10] E. Schaller, S. Zenhäusern, T. Zesiger, J.O. Bosset, F. Escher, Use of preconcentration techniques applied to a MS-based "Electronic Nose", Analusis 28 (2000) 743–749.
[11] J. Brezmes, E. Llobet, X. Vilanova, J. Orts, G. Saiz, X. Correig, Correlation between electronic nose signals and fruit quality indicators on shelf-life measurements with pinklady apples, Sens. Actuators B 80 (2001) 41–50.
[12] R.E. Shirey, L.M.Sidisky, Analysis of Flavors and Off-Flavors in Foods and Beverages Using SPME, Supelco Application Note, http://www.infonew.sigma-aldrich.com/Graphics/Supelco/objects/8600/8564.pdf.
[13] R.K. Wagner, W. Grosch, Key odorants of french fries, J. Am. Oil Chem. Soc. 75 (10) (1998) 1385–1392.
[14] R.K. Wagner, W. Groshc, Evaluation of potent odorants of french fries, Food Sci. Technol. 30 (2) (1997) 164–169.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper III

Vinaixa, M.; Llobet, E.; Brezmes, J.; Vilanova, X.; Correig, X., A fuzzy ARTMAP- and PLS-based MS e-nose for the qualitative and quantitative assessment of rancidity in crisps. *Sensors and Actuators B,* **2005,** 106(2), 677-686.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Sensors and Actuators B 106 (2005) 677–686
SENSORS AND ACTUATORS B CHEMICAL
www.elsevier.com/locate/snb

A fuzzy ARTMAP- and PLS-based MS e-nose for the qualitative and quantitative assessment of rancidity in crisps

M. Vinaixa, E. Llobet*, J. Brezmes, X. Vilanova, X. Correig

MINOS, Dept. Enginyeria Electrònica, Avda. Països Catalans 26, 43007 Tarragona, Spain

Received 12 July 2004; received in revised form 16 September 2004; accepted 17 September 2004
Available online 19 October 2004

Abstract

A new method based on the use of an SPME MS e-nose, to assess oxidative and hydrolytic rancidity in crisps is introduced. The method can become an alternative tool to the traditional methods such as the Rancimat and ADV tests. Because it does not rely on a previous step of oil extraction from the samples to be analysed, sample preparation is easier and the analysis is speeded up (e.g., the Rancimat method can take several hours to produce a result if the oil under analysis is of good quality). The effectiveness of the e-nose in the assessment of the quality of crisps was demonstrated by developing two different applications. First, the e-nose was used to classify crisp samples according to their rancidity stage. When the e-nose data was processed by a fuzzy ARTMAP classifier, the success rate in classification was estimated to be around 93% (validation results). The system could discriminate 100% of the fresh crisps (i.e., class A) from the rancid ones (B, C and D). The e-nose was trained to predict the results of the ADV and Rancimat tests by building quantitative PLS models. A good correlation existed between our instrument and the results of the Rancimat and ADV tests (the correlation coefficients were 0.98 and 0.97, respectively). The best results were obtained when the dimensionality of input data was reduced by applying variable selection procedures based on principal component analysis and genetic algorithms.

Keywords: SPME-MS e-nose; Fuzzy ARTMAP; Genetic algorithm; PLS models; Crisp rancidity; Rancimat; Acid degree value

1. Introduction

Potato crisps are usually made by deep-frying fresh potato slices in a vegetable oil bath. During the deep-frying process, potatoes and oil undergo a variety of chemical and physical changes. Once fired, crisps contain 2% and 35% of moisture and edible oil, respectively. If they are stored under reasonable conditions, their maximum shelf-life is estimated at 12 weeks, after which, moisture uptake and rancidity development become unacceptable [1]. Oxidation of lipids is one common and frequently undesirable chemical change that may generate products related to the occurrence of rancidity. Rancidity is based on the subjective organoleptic appraisal of off-flavours from food [1]. It conduces to a flavour, aroma and taste deterioration leading to consumers' unacceptability of final products, which results in money loss for the crisp industry. This is why the industry needs to run routine quality tests both on raw materials and finished products.

Two types of lipid oxidation cause the most concern. These are oxidative rancidity and hydrolytic rancidity. Hydrolytic rancidity results in the formation free fatty acids and glycerol from the hydrolysis of triacylglycerols. It is caused by either the reaction of lipids and water in the presence of catalysts or by the action of lipase enzymes [2]. These free fatty acids may be precursors of volatile compounds such as aldehydes, ketones or alcohols that generally cause off-flavours in crisps. Oxidative rancidity results from a more complex lipid oxidation where polyunsaturated fatty acids are converted to hydroperoxides and further to secondary oxidation products [3]. These secondary oxidation products, especially aldehydes, produce the off-flavours associated with rancid oil [1].

Rancidity is a qualitative term or state that is not chemically defined and is not directly quantifiable [4]. As a result, a

* Corresponding author. Tel.: +34 977 558502; fax: +34 977 559605.
E-mail address: ellobet@etse.urv.es (E. Llobet).

0925-4005/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2004.09.015

number of different methods trying to measure intermediate products of lipid oxidation known to be correlated with the occurrence of rancid taste have been developed. The AOCS in its Official Methods and Recommended Practices lists some 365 test procedures to verify the quality of oil either during the process of frying or in finished products [5,6]. Many factors must be considered when selecting the most appropriate tests or series of tests and it is often difficult to predict which method will yield the most meaningful results [2].

The most used test by the crisp industry to assess hydrolytic rancidity is the acid degree value (ADV). The fat in a sample is dissolved into an organic solvent and titrated with sodium hydroxide. The rancidity is calculated as percent of oleic acid. The higher the ADV value the higher the level of free fatty acids (FFA) present in the product. Also FFA level has been used as an indicator of the quality of frying oil and of oil degradation. Oxidative rancidity is generally tested with two types of tests. The first type monitors the extent of oxidation that oil has undergone as the peroxide value (PV). The second type of test uses accelerated ageing conditions to measure the stability of a fat or finished product and also can estimate product shelf-life [2]. An example of these predictive tests is the oxidative stability index (OSI) that has become widely used in quality control by the crisp industry. Rancimat (Brinkman Instruments, Des Plaines, IL) is an automated, commercially available instrument to perform OSI tests. It measures the oxidation susceptibility of oil under oxidative conditions and heat. The test is as follows: air is passed through a sample held at constant temperature and then bubbled through a reservoir of deionised water. Volatiles produced by lipid oxidation are dissolved in water, which increase water conductivity. Conductivity is continuously monitored and the Rancimat value (Rh) is defined as the elapsed time required for water to show a sharp conductivity change.

The problem of well-established methods such as the ADV or the Rancimat test is that they are very time-consuming and labour intensive. Especially the Rancimat can take hours to produce a result and, no matter if the Rancimat and ADV test are used, a previous extraction of fat from crisps must be performed. This implies the use of organic solvents, which could distort final results.

According to Allen and Hamilton [1] a measure of the development of rancidity in crisps or snacks during storage can be gained by carrying out similar tests to those employed in the screening of oils. One of the most important tests to judge the quality of oil in finished products is organoleptic assessment by a panel. To achieve this, people that are involved in the testing should be capable of accurately discriminating the nature and degree of rancidity development. Problems associated to taste panels are difficulty of training, subjectivity, fatigue and low sample throughput. This results in a high cost per sample analysed.

Actually, a compendium of ADV, Rancimat and organoleptic assessment is what is considered to be routine tests for assessing rancidity and oil stability by the crisp industry. However, since the crisp industry demands analysis techniques with high sample throughput, there is a need for faster and simpler methods to assess rancidity and the development of off-flavours. In this context, the use of e-nose technology would be of help.

Electronic noses are multi-sensor array instruments inspired in mammal olfaction and are designed to characterise and discriminate samples by their volatile pattern profiles using a suitable pattern recognition engine [7]. Assuming that changes in the headspace of crisp samples could be an indicator of the degree of fat oxidation, e-nose technology could be used to determine rancidity levels. Furthermore, e-noses have the potential of performing the discriminatory analysis, yielding results that are easily interpreted [8].

A wide range of sensors have been tested along e-nose development but those based on metal oxides appeared most suitable for the discrimination of different stages of lipid oxidation and hence for shelf-life prediction [8,9]. Anyway, e-noses based on non-specific semiconductors sensors suffer from serious drawbacks, being response drift one of the most important. That is why mass spectrometry-based e-noses (MS e-noses) are becoming an increasingly used alternative to e-noses based on solid-state sensors. The use of pre-concentration and extraction techniques such as solid phase microextraction (SPME) have improved sensitivity and reproducibility of MS e-noses [10–12].

This paper reports on the design of an SPME-MS e-nose with the ability of predicting both the hydrolytic and oxidative rancidity in crisps. Our SPME-MS e-nose analyses the volatile headspace of crisps as a whole, without chromatographic separation. Each fragment ion represents a potential sensing element and the intensity of the fragment is equivalent to the sensor signal [13]. For a given measurement, the resulting mass spectrum gives a fingerprint that is characteristic of the volatile compounds present in the headspace of the crisps.

In the first step, the ability of the SPME-MS e-nose to discriminate between crisps in four different stages of rancidity was evaluated. The fingerprint spectra obtained with the e-nose were gathered in a data set and used to build a fuzzy ARTMAP model to predict the stage of rancidity in crisps.

In the second step, partial least squares (PLS) models were developed to predict the results of the Rancimat and ADV tests using fingerprint mass spectra (i.e. *m*/*z* fragments) as descriptor variables. The predictive ability of these models was improved by a genetic algorithm (GA)-based variable selection, which determined the best *m*/*z* fragments to be used as input data. The correlation between the SPME-MS e-nose and the more traditional Rancimat and ADV tests is discussed.

2. Experimental and computational techniques

2.1. Crisp samples

Four boxes (labelled A, B, C and D) with 200 g packs of potato crisps (10 packs per category) were prepared by Frit

Ravich, S.L. These crisps belonged to the same frying batch of 50% palm and 50% sunflower oil, but they underwent four different rancidity accelerating treatments:

Crisps in box A were stored during 28 days in a dry and dark conservation chamber, where their temperature was kept around 20 °C.

Crisps B, C and D were kept during 14, 21 and 28 days, respectively, in a rancidity accelerating chamber. The chamber was kept at high temperature (around 40 °C) and UV light was used to promote oxidation. As soon as the samples within a given category finished their ageing treatment, they were removed from the rancidity chamber and stored in the conservation chamber to maintain unchanged the rancidity stage they had reached.

One of the 10 packs from each rancidity stage (A, B, C or D) was left out to perform Rancimat and ADV tests. The nine remaining packs were used to obtain SPME-MS e-nose measurements.

2.2. SPME-MS e-nose measurements

The MS e-nose was implemented using a Shimadzu QP 5000 gas chromatograph–mass spectrometer. The chromatographic column was replaced by a 5 m deactivated fused silica column to co-elute all volatile components, achieving one single peak for all the components in the headspace of crisps. This column was kept isothermal at 250 °C and the helium flow was set to 1.4 ml/min.

The content of each pack was homogenised and 4 ± 0.2 g of crisps were crushed and put into a 20 ml headspace vial that was immediately capped and sealed with a Teflon septum. Then, the vials containing crisp samples were placed inside a thermostatic bath (50 °C) during 30 min to promote the release of volatile compounds into the headspace. SPME sampling was performed by introducing a 75 μm Carboxen/PDMS fibre into the vial and exposing it to the headspace of crisps for 20 min. Thermal desorption of the volatiles trapped on the fibre was conducted for 3 min in the chromatograph injection port at 300 °C. The port was equipped with a 0.75 mm i.d. liner. The split valve was closed during desorption. The quadrupole mass spectrometer acquired in scan mode, and the mass range used was *m/z* 35 to *m/z* 200 at 0.5 scan/s. To ensure its complete cleaning, the fibre was left five additional minutes in the injection port.

2.3. Rancimat test

Rancimat tests were performed at the quality control laboratory (QC) of Frit Ravich, S.L. They measure the susceptibility of oil to oxidation. A previous step of oil extraction from each sample of crisps is required. Then, the oil sample is kept at 120 °C in a vessel where air flows to extract volatiles from the headspace. These volatiles are then collected in water. The conductivity of water is monitored because rancidity triggers a sharp increase in water conductivity. The time at which this increase occurs is a value (Rh) that indicates the resilience of oil to rancidity. The higher the Rh value is, the more stable oil is and less prone to develop rancidity.

2.4. Acid degree value (ADV)

ADV tests were also performed at the QC of Frit Ravich, S.L. The content of FFA was measured by a standard titration method, which determines acids formed by hydrolysis and by degradation of oxidised fatty acids and glycerides. A previous isolation of fat (centrifugation or extraction) must be performed. ADV obtains the amount of potassium hydroxide required to neutralise the free fatty acids hydrolysed with 95% ethanol.

2.5. Data processing techniques

For each measurement, a response spectrum was obtained by averaging mass spectra along the detected unretained peak. Data from this averaged mass spectrum were processed with a written-in-house software based on MATLAB. Nine replicate measurements in each rancidity stage category (A, B, C and D) were performed with a scan range from *m/z* 35 to *m/z* 200. This range was selected because *m/z* fragments below 35 have less specificity and we did not find in the literature any compound involved in development of rancidity with significant fragments above 200 amu. The data matrix consisted of 166 columns (i.e. the scan range of *m/z* variables) and 36 rows (9 replicate measurements in 4 rancidity stages). Fragments whose intensity was remained constant for the 36 measurements performed were removed from the data matrix. Therefore, the reduced data matrix was conformed by 78 columns (out of the 166 initially scanned) and 36 rows. Prior to perform any analysis, the data matrix $\mathbf{R}$ (36×78) was mean-centred. If this scaling of the data is not performed, there is a risk of ignoring mass intensities with low mean in front of mass intensities with high mean. This data were used to attempt a fuzzy ARTMAP-based classification of the samples in four rancidity stages and also to build quantitative PLS models to predict the ADV and Rancimat results. Interested readers are addressed to [14,15] for a good introduction to the fuzzy ARTMAP neural network and to [16] for a tutorial on the PLS algorithm.

2.5.1. Feature selection

Variable selection on a given set of variables consists of selecting a subset that performs well in a classification/prediction problem. It is a very relevant step in multivariate analysis because the removal of non-informative variables will produce more parsimonious models [17]. MS-based e-noses have as many sensors as the *m/z* variables included in the range of a full scan. However, some authors have reported that is not useful to work with such a high number of sensors as, in most cases, a very small number of fragment ions is suitable for setting up a sensor array [13].

In fact, using all the scan range of *m/z* variables can be a way to introduce background noise. This can lead to a decrease in the reproducibility of MS e-nose systems. A correct choice of the *m/z* variables that best describe the application sought is of paramount importance for the MS e-nose to perform well. In the specific case of crisps, *m/z* fragments coming from the breakdown of molecules involved in the rancidity process should be selected. Dittmann and Nitz [13] consider that there is no way to correctly select *m/z* fragments, unless there is a previous full chromatographic run. In [13] they discuss reliable strategies for selecting the optimal array configuration, based on previous knowledge of the analytes that are important for the application. This previous knowledge is normally based on time-resolved analysis to identify (and quantify) the volatiles present in the headspace of the samples to be studied. This leads, unavoidably, to more traditional analytical techniques such as gas chromatography–mass spectrometry. GC/MS is one of the most powerful techniques that could be used to allow the identification of a high range of volatile compounds involved on rancidity processes. Taking a look to fully resolved chromatograms, volatile pattern profiles from fresh and rancid crisps can be compared. However, obtaining fully resolved chromatograms implies to develop a chromatographic method to analyse the samples. The method is involved, time consuming, with low sample throughput and needs qualified personnel to perform the analysis and interpret results. This is why two different approaches for variable selection were implemented. The methods implemented perform automatically the selection of *m/z* variables that are relevant for our application without the need of previous knowledge on the exact nature of the analytes. This is the usual approach in e-nose applications.

The first technique is based on a PCA variable reduction to reduce the dimensionality of the input matrix. Eklöv and Mårtensson [18] have reported the advantages on performing variable selection in gas-sensor data. Their variable selection consisted in using PCA scores instead of raw data to train a neural network. With this reduction in dimensionality the numerous and probably high correlated original variables are transformed into a small number of orthogonal (i.e. uncorrelated) new variables which are a linear combination of the original ones. Therefore, a PCA was performed to the data matrix and the scores on the first 10 PCs were retained for further processing. Higher PCs were discarded, because it was assumed that they explained a residual variance (e.g. noise). It is now widely accepted that multivariate calibration techniques such as PLS greatly benefit from appropriate sensor selection [19]. Genetic algorithms have been applied for variable selection in the building of calibration models from spectral data (e.g. PLS) and have been shown to provide better results than full spectrum approaches [20]. Therefore, the second technique consisted on a GA applied to the selection of *m/z* factors. The selected *m/z* variables were used to build quantitative PLS models to predict the results of Rancimat and ADV tests.

2.5.2. Calibration and validation procedures

The training phase and the prediction ability of the different models built were evaluated. The response matrix, **R** (36×78), was split into two different matrices: training matrix and validation matrix.

In fact, four different training and their corresponding validation matrices were used (i.e., four-fold training and validation process). In the first fold, replicate measurements 1 and 2 of each rancidity stage were gathered in the validation matrix $\mathbf{V}_1$ (8×78) and the remaining measurements were gathered in the training matrix $\mathbf{T}_1$ (28×78). In the second, third and fourth folds, the validation matrices used replicate measurements 3 and 4 ($\mathbf{V}_2$), 5 and 6 ($\mathbf{V}_3$) and 7 and 8 ($\mathbf{V}_4$), respectively. The corresponding training matrices $\mathbf{T}_2$, $\mathbf{T}_3$ and $\mathbf{T}_4$, were formed by the remaining replicate measurements.

Each training matrix, $\mathbf{T}_i$ (28×78), was used to train a fuzzy ARTMAP classifier to predict rancidity stage or to build a PLS calibration model to predict either Rancimat or ADV results. Each validation matrix, $\mathbf{V}_i$ (8×78), was used to estimate the success rate in classification of the classifier and the accuracy of the calibration models in the prediction of Rancimat and ADV results. It should be noted that measurements in $\mathbf{V}_i$ had not taken part in the training phase, which ensured that the validation measurements were completely new to the models built.

The training phase of the fuzzy ARTMAP and PLS models was evaluated using a cross-validation method. The process was as follows: given n measurements ($n = 28$ measurements within each training matrix), the model was trained n times using $n - 1$ vectors. The vector left out was then used for testing the model. Performance in training was estimated as the averaged performance over the n tests.

3. Results and discussion

3.1. Results of the well established tests

Rancimat and ADV tests were performed on the crisps that had undergone rancidity accelerating treatments. Table 1 summarises these results.

ADV is usually considered to be one of the main parameters for evaluating the quality of oil, specially the state of frying oils. Che Man et al. [21] choose ADV as a measure to correlate with shelf-life of potato chips, because acidity of the medium affects the storage stability of the fried products. They found a high degree of correlation between the shelf-life of potato chips and ADV.

Table 1
Results of the Rancimat and ADV tests run on the four types of crisp samples

Rancidity class	Days under rancidity accelerating treatment	Rancimat, 120 °C (h)	ADV (% oleic acid)
A	0	5.36	0.30
B	14	2.63	0.28
C	21	1.61	0.60
D	28	0.01	1.91

It was expected an increase in the ADV as the permanence inside the rancidity accelerated chamber was increased. In other words, the more a sample is kept in the rancidity chamber, the higher the extent of hydrolysis reaction, which results in increased percentage of FFA. According to [1], the FFA level in the oil of a fryer should not exceed 0.3% and should remain below 0.25% to have an acceptable turnover period. Turnover period is the time taken for the crisps to absorb the oil capacity of the fryer. The lower this time is, the lesser the oil will be degraded. Table 1 shows that hydrolytic rancidity measured as ADV remained approximately near the threshold of 0.3% of FFA in class A and B samples. A sharp change in the ADV resulting from hydrolytic rancidity is observed for samples C and D only. Keeping samples during 2 weeks in the rancidity accelerating chamber did not trigger hydrolytic rancidity. This could be due to the unsuitability of the ADV test to reveal the occurrence of hydrolytic rancidity at this early stage. At early stages of hydrolytic rancidity the level of FFA produced is too low to be accurately measured by this titration test. Finally, the ADV is a simple acid base titration that measures not only FFA but also any acid present in the oil extracted from crisps. Therefore, the presence of other acids can cause interference in the measurement of hydrolytic rancidity.

Unlike for the ADV test, when oxidative rancidity is estimated as Rancimat hours, a quantitative difference arises between samples A and B. This result is not surprising because, in most cases, oxidative rancidity occurs before hydrolytic rancidity. This is because the mechanism of oxidative rancidity is a chemical reaction with lower activation energy than the mechanism of hydrolytic rancidity. In fact, the hydrolysis reaction requires high temperature, presence of moisture or activity of lipase enzyme. That is the reason why hydrolytic rancidity can be minimised by cold storage, controlled transport and careful packaging. On the other hand, oxidative rancidity is not stopped by lowering the temperature during food storage [1].

As a result of this, the Rancimat test is considered by the crisp industry to be more informative than the ADV test in the assessment of rancidity development.

3.2. MS e-nose for identifying the stage of rancidity in crisps

In the first step, the classification of crisps according to their rancidity stage (A, B, C or D) was envisaged. A fuzzy ARTMAP classifier was built using the training matrices $\mathbf{T}_i$ (see Section 2.5.2 for details on these matrices). Before training, $\mathbf{T}_i$ must be normalised because the fuzzy ARTMAP network needs that input data lie in the range [0,1]. The number of inputs to the network was set to 78 (the number of *m*/*z* columns in $\mathbf{T}_i$). The number of outputs was set to 4 because a 1-of-4 code was used for the different classes (A: 0001, B: 0010, C: 0100 and D: 1000). For example, the activation of the first output neurone (i.e. output pattern 0001) implies that an input vector is recognised as belonging to class A (fresh crisps). The baseline vigilance was set to 0. This is the recommended value for the vigilance since it allows for very coarse categories and the match tracking system will only refine these categories if necessary. The recode rate was set to 0.5. This value allows the established categories to be modified if there is persistent attempt to do so (slow recode). The value of choice parameter was set to 0.1. The network could learn the training set in just one iteration. The number of committed nodes, which play a similar role as hidden neurones in multilayer perceptron networks, ranged between 4 and 6 after the network had been trained.

The training and prediction ability of the fuzzy ARTMAP classifier was assessed using the procedures described in Section 2.5.2. Success rate in training (SRT) was defined for each training matrix as the success rate in classification averaged over the 28 leave-one-out tests. Success rate in prediction (SRP) was defined as the success rate in classification of the measurements in each validation matrix. Validation matrices must be normalised before being input to the fuzzy ARTMAP classifier. Each validation matrix, $\mathbf{V}_i$, is normalised for its values to lie in [0,1] using the normalisation coefficients computed from its corresponding training matrix $\mathbf{T}_i$.

The second column in Table 2 summarises the results of the fuzzy ARTMAP classifier. When the 78 *m*/*z* variables were used, the averaged success rate achieved in the classification of crisp samples in four rancidity stages was 88%. It is important to notice that misclassified samples belonged to classes B, C and D. Therefore, the MS e-nose never confused fresh (A) with rancid (B, C, D) crisps. To better understand this result, a PCA was performed on the data. Fig. 1 shows the scores plot of this analysis. It can be seen that samples belonging to fresh crisps (A) cluster well apart from the other samples. The discrimination between fresh and rancid samples is possible along the second PC that explains near 42% of the variance in the data. This explains why the MS e-nose is able to clearly discriminate between fresh and rancid crisps, even at early stages of rancidity (e.g. discrimination between samples A and B). The first PC somewhat explains the difference between rancidity stages B, C and D. From higher to lower values of PC1 the samples tend to cluster according to increasing rancidity stages. However, a strong overlapping exists.

Table 2
Success rate in training (SRT) and validation (SRV) of a fuzzy ARTMAP classifier applied to the discrimination of the rancidity stage in crisps

	Without PCA pre-processing		With PCA pre-processing	
	SRT	SRP	SRT	SRP
Fold 1	89	63	100	100
Fold 2	89	100	93	86
Fold 3	89	88	96	86
Fold 4	82	100	93	100
Average	87	88	96	93

The data matrix was split in four folds to evaluate the training and validation processes.

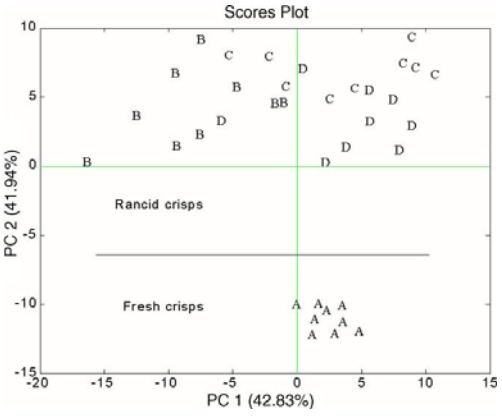

Fig. 1. Results of PCA performed on the MS e-nose data. Samples A, B, C and D correspond to crisps with 0, 14, 21 and 28 days of rancidity accelerating treatment, respectively.

A pre-processing technique that consisted in applying a PCA on the data prior to train the fuzzy ARTMAP classifier was performed to see whether the success rate in the classification of crisp rancidity could be improved. Therefore, instead of using the 78 *m/z* variables, the scores onto the first ten principal components were used. Table 2 summarises the results in training and validation when the PCA-based variable selection procedure was implemented. The average values of SRT and SRP were 96% and 93%, respectively, which compare favourably with the values reached when no dimensionality reduction was applied (88% and 87%). Keeping the number of selected PCs below 10 resulted in lower SRT and SRP. Selecting more than 10 PCs did not improve the results. The discrimination between classes B, C and D improved when the variance explained by the higher order PCs (e.g. 3rd–10th) was used.

The fact that the fuzzy ARTMAP classifier performed better when a PCA-based variable reduction was used is due to the elimination of collinearity among the input variables (while *m/z* variables are highly correlated, PCA scores are orthogonal, and thus uncorrelated). Furthermore, the first 10 PCs (i.e., selected ones) contain important information that allows discriminating the stage of rancidity. On the other hand, higher PCs (i.e., discarded ones) capture a variance that corresponds to background noise.

Dimensionality reduction is advantageous since, according to Eklöv and Mårtensson [18], applying inputs with high dimensionality to a neural network (i.e. the 78 *m/z* variables) can cause at least two problems. First the process of training becomes very time consuming. Secondly the number of measurements has to be very large to be representative of the complete input space in order to get a good model.

3.3. Correlation between SPME-MS e-nose and the Rancimat and ADV tests

To investigate whether the results of SPME-MS e-nose correlate well with the results of the Rancimat and ADV tests, different PLS calibration models were built and validated. The underlying objective was to assess if the e-nose can predict hydrolytic rancidity, oxidative rancidity or both.

PLS is a linear and supervised multivariate calibration method that attempts to find factors (i.e. latent variables), which capture as much variance as possible in the predictor block X-matrix, under the constraint of being correlated with the predicted block Y-matrix.

The X-block consisted in the response matrix acquired with SPME-MS e-nose. For the first calibration model, the Y-block was a column matrix with the results of ADV tests. In this way, the ability of the e-nose to predict hydrolytic rancidity was investigated. For the second calibration model, the Y-block was a column matrix with the results of the Rancimat tests. Therefore, the ability to predict oxidative rancidity was checked. Finally, a third calibration model was built to simultaneously predict hydrolytic and oxidative rancidity. Therefore, the Y-block was a two-column matrix with the results of the Rancimat and ADV tests. The X and Y-blocks were mean-centred before the PLS were performed.

The PLS calibration models were trained and validated using the same procedures described above (see Section 2.5.2). The effectiveness of the training process was evaluated calculating the root mean square error of cross-validation (RMSECV) according to the following equation:

$$\mathrm{RMSECV} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - y_i)^2}{N}} \quad (1)$$

where Y_i is the actual value of the response and y_i the prediction of the PLS-model. Because training is evaluated using a leave-one-out cross-validation, the error is averaged over the N test ($N = 28$).

The selection of the number of latent variables to build the models was done by representing the evolution of RMSECV versus the number of latent variables. An example of this process is shown in Fig. 2. Generally the value of RMSECV decreases when the number of factors is increased and either increases again or stabilises if more factors are added. Choosing too many factors can over-fit the training set, which results in lower prediction ability. For the example shown in Fig. 2, the number of factors selected was 15, which corresponds to the minimum RMSECV.

The prediction ability of the PLS models was estimated using the root mean square error of prediction (RMSEP). This error was computed formally as in Eq. (1). But here, the measurements in the validation matrices were used. During the validation phase, the models under test used the number of factors that had been set in the training phase.

To better assess the accuracy of the different PLS models, regression between the actual ADV and Rancimat results and

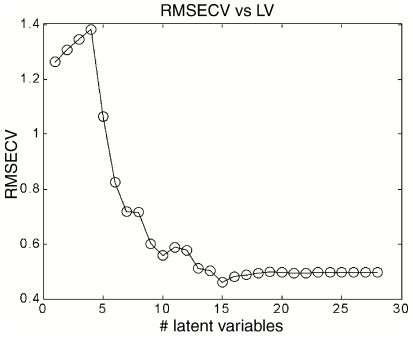

Fig. 2. A typical plot showing the root mean square error of cross-validation (RMSECV) in a PLS model as a function of the number of latent variables. In this example, 15 LVs seems to be a good choice.

those predicted by the models was performed. The correlation coefficient, slope and intercept were calculated. A perfect prediction would yield 1 as correlation coefficient and slope and 0 as intercept.

Table 3 summarises the averaged results of this study (for training and validation) for all the PLS models used to predict ADV, Rancimat or both. Specific or simultaneous predictive models performed similarly. However, the specific models made use of a lower number of LVs (11 instead of 13). The simultaneous prediction of ADV and Rancimat results implies using a model with a more complex structure (with higher number of LVs). MS e-nose results correlate well with the results of the Rancimat tests. (correlation coefficients of 0.97 both in training and validation). On the other hand, the correlation with ADV results is lower (0.94 and 0.93 in training and validation, respectively). The MS e-nose correlates well with Rancimat because these methods evaluate volatile metabolites. ADV is a titration test that evaluates volatile and non-volatile acids. This can explain the lower correlation found between the e-nose and ADV tests.

A genetic algorithm (GA) variable selection procedure was implemented to improve the prediction ability of the PLS models. In this way a restricted set of *m*/*z* variables was identified to build more parsimonious models. When the number of variables is high (i.e. 78 *m*/*z* variables), it is usually out of question to make an exhaustive search because it is a very time-consuming process, given the number of variables to be considered for selection. GAs have been shown to solve this optimisation problem by exploring all regions of the potential solutions and exponentially searching promising areas through mutation, crossover and selection operations applied to individuals (i.e. chromosomes) in a population. The population (i.e. set of solutions) is maintained and manipulated by implementing a 'survival of the fittest' strategy in the search for the optimal solution. Because the next explored point in a solution space is chosen by stochastic rather than deterministic rules, GAs do not need to make assumptions about the characteristics of the problem to be solved and, therefore, apply generally [22].

Possible solutions to the problem envisaged here were encoded as binary strings called chromosomes, where the code 1 means that the gene (*m*/*z* variable) has been chosen to build a PLS model and 0 means that the variable has not been chosen. The initial population (initial set of possible solutions) consisted of 128 chromosomes encoded at random. The size of the population is kept constant from generation to generation. The average number of genes set to 1 within the initial population of chromosomes was 50% (initial terms). This means that initial models made use of half the variables. At each generation, the model structures represented by the population are fully cross-validated. Each member of the population is ranked according to its fitness. Fitness is evaluated as the RMSECV of the PLS model built using only those variables set to 1 in the chromosome. The lower the RMSECV is, the better the fitness. At each generation, half of the chromosomes with the lower RMSECV are allowed to live and breed. Pairs of these chromosomes (i.e. the ones from the half with lower RMSECV) are randomly selected for breeding using a double crossover technique. Double crossover implies that two crossover sites per chromosome are used. Also, some of the genes in the chromosomes are randomly flipped (mutation rate was set to 0.005) after each generation. The chromosomes with better fitness are kept unchanged in the next generation (elitism). An iteration is performed until either the population converges or the maximum number of iterations (set to 300) is reached. The population converges when the percentage of duplicate chromosomes in the population is high (e.g. 75%).

The process of variable selection was conducted for each one of the different PLS models (prediction of Rancimat

Table 3
Training and validation results of the PLS models built to predict ADV, Rancimat and both tests simultaneously

	#LVs	Results on training phase				Results on validation phase			
		r	*m*	*a*	RMSECV	*r*	*m*	*a*	RMSEP
Specific models									
ADV	11	0.94	0.93	0.06	0.22	0.92	0.94	0.04	0.28
Rancimat hours	11	0.97	0.97	0.07	0.50	0.96	0.83	−0.02	0.57
Simultaneous model									
ADV	13	0.94	0.94	0.06	0.22	0.93	0.95	0.01	0.28
Rancimat hours	13	0.97	0.97	0.04	0.51	0.97	1.01	−0.02	0.53

The results shown are averaged over the four training/validation folds.

Table 4
Training and validation results of the PLS models built to predict ADV, Rancimat and both tests simultaneously

	#LVs	Results on training phase				Results on validation phase			
		r	*m*	*a*	RMSECV	*r*	*m*	*a*	RMSEP
Specific models									
ADV	8	0.98	0.97	0.03	0.13	0.97	1.00	−0.02	0.20
Rancimat hours	8	0.99	0.99	−0.01	0.31	0.98	1.01	−0.01	0.40
Simultaneous model									
ADV	9	0.98	0.98	0.02	0.18	0.95	0.98	0.32	0.23
Rancimat hours	9	0.98	0.99	0.05	0.50	0.96	1.01	−0.11	0.61

The results shown are averaged over the four training/validation folds.

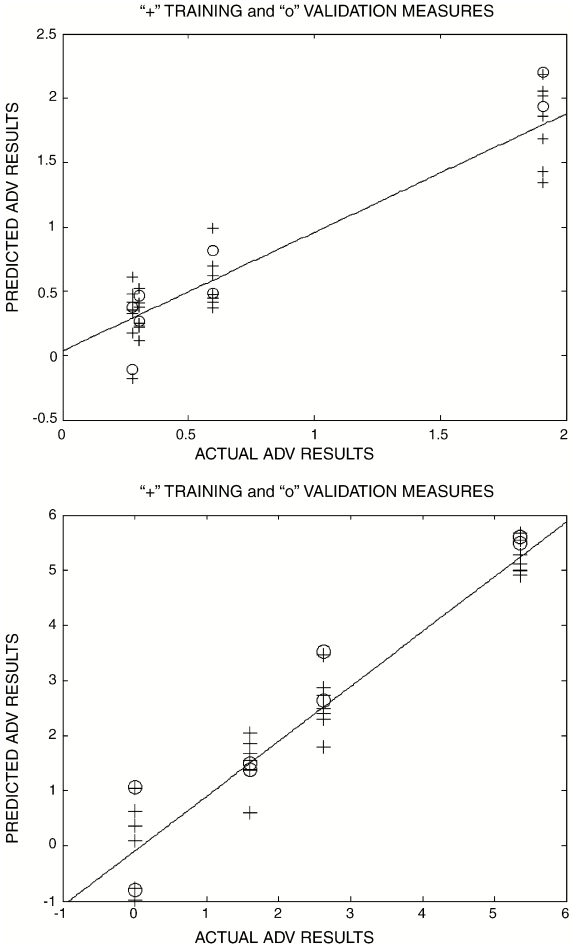

Fig. 3. Results of the PLS models. Actual vs. predicted values of the Rancimat test (top) and the ADV test (bottom).

results, ADV results and both simultaneously). The data in the training matrices were used for variable selection. Once the variable selection procedure was completed, the final PLS models were built using restricted training matrices (keeping only the *m/z* variables selected). The training and validation of these final models was conducted as before.

Table 4 summarises the averaged training and validation results for the GA-PLS models. Performing a GA-based variable selection resulted in an improved predictive ability of the PLS models. The specific models performed better than the simultaneous one. The GA selected an average of 14 out of the 78 *m/z* variables for the prediction of ADV or Rancimat results (specific models). Fifteen variables were selected in average for the simultaneous prediction of ADV and Rancimat results. The selected variables for the ADV model ranged between *m/z* 35 and *m/z* 91 and the variables selected for the Rancimat model ranged between *m/z* 36 and *m/z* 94. Finally, the simultaneous model used variables in the range *m/z* 35 to *m/z* 92. Once again, the specific models made use of a lower number of LVs than the simultaneous one (8 instead of 9). After variable selection, the results of MS e-nose show a high correlation with the results of the Rancimat and ADV tests. This correlation is higher than when no variable selection was performed, especially for the prediction of the ADV.

Fig. 3 shows the predicted ADV and Rancimat results versus the actual values using the specific models. The crosses are the results of the leave-one-out cross-validation using the training measurements and the circles correspond to the validation measurements. The predictions for validation measurements (i.e., not used for training) lie in the range of the predictions for training measurements. The PLS model was not very accurate in the prediction of the Rancimat results for samples in class D. This is because the value of the Rancimat test for these samples (i.e., 0.01 h) was extremely low. In fact, class D samples corresponded to spoilt crisps with a Rancimat value near 100 times lower than the value for crisps in class C (the previous stage of rancidity).

4. Conclusions

We have reported a new method, based on the use of an SPME MS e-nose, to assess oxidative and hydrolytic rancidity in crisps. Unlike more traditional methods such as the Rancimat and ADV tests, our system does not rely on the previous extraction of oil from the samples, but directly evaluates volatile compounds from the headspace of the crisps. This results in easier sample preparation, avoids the use of organic solvents for oil extraction and speeds up the analysis. The time needed to analyse a sample with the e-nose system is around 28 min. This includes the 20 min needed for the SPME, 3 min for acquisition of the mass spectra and 5 min for cleaning the SPME fibre. This compares favourably with the Rancimat method, which can take several hours to produce a result if the oil under analysis is of good quality.

The effectiveness of the e-nose in the assessment of the quality of crisps was demonstrated in two different applications. First, the e-nose was used to classify the crisp samples according to four pre-established rancidity stages. The system input the data from the MS (pre-processed by a PCA) into a fuzzy ARTMAP classifier. The success rate in classification was estimated to be around 93% (validation results). The system could always perfectly discriminate between fresh (class A) and rancid (B, C and D) crisps. In the second step, the e-nose was used to predict the results of the ADV and Rancimat tests by building quantitative PLS models. The correlation between our e-nose and the ADV and Rancimat tests was investigated. A good correlation existed between our instrument and the Rancimat and ADV tests (the correlation coefficients were 0.98 and 0.97, respectively).

For the particular applications developed here, a reduction in the number of input variables results in more parsimonious models with higher predictive ability. The use of a PCA to reduce the dimensionality of input variables resulted in improved performance of the fuzzy ARTMAP classifier applied to the discrimination among the four rancidity stages. The GA-based variable selection led to a significant improvement of the PLS models that were built to predict the results of the ADV and Rancimat tests using the MS e-nose.

According to these results, our SPME-MS e-nose can become an alternative tool for the assessment of oxidative and hydrolytic rancidity in crisps.

Acknowledgements

We are grateful to Carles Badia and Elena Freixes from Frit Ravich, S.L. who supplied the crisp samples, run the rancidity accelerating treatments and performed the ADV and Rancimat tests. This work was funded by INIA (project number CAL01-052-C2) and by FIS (Thematic Network C03/08).

References

[1] J.C. Allen, R.J. Hamilton, Rancidity in Foods, 2nd ed., Elsevier Applied Science, 1989.
[2] Measuring Rancidity in Fats and Oils. http://www.ralstonanalytical.com/.
[3] R.L. Heiniö, P. Lehtinen, K.M. Oksman-Caldentey, K. Poutanen, Differences between sensory profiles and development of rancidity during long-term storage of native and processed oat, Cereal Chem. 79 (3) (2002) 367–375.
[4] C.R. Hamilton, D. Kirstein, 2003. http://www.darlingii.com/.
[5] R. Marsili, Controlling the Quality of Fats and Oils, 1993. http://www.foodproductdesign.com/.
[6] Official Methods and Recommended Practices of the American Oil Chemists' Society, 5th ed., AOCS, 1998.
[7] T.C. Pearce, S.S. Schiffman, H.T. Nagle, J.W. Gardner (Eds.), Handbook of Machine Olfaction, VCH/Wiley, Weinheim, 2002.
[8] M. Vinaixa, A. Vergara, C. Duran, E. Llobet, C. Badia, J. Brezmes, X. Vilanova, X. Correig, Fast detection of rancidity in potato crisps

using e-noses based on mass spectrometry or gas sensors, Sens. Actuators B, in press.
[9] M. Adechy, V.P. Shiers, J.B. Rossell, Study of rancidity and resistance to oxidation in edible oils and fats using electronic nose technology in comparison with conventional techniques, Leatherhead Food RA Research Reports 751, 1998.
[10] R.T. Marsili, SPME-MS-MVA as an electronic nose for the study of off-flavors in milk, J. Agric. Food Chem. 47 (1999) 648–654.
[11] R. Marsili, Flavor, Fragrance and Odor Analysis, Marcel Dekker, New York, 2002.
[12] E. Schaller, S. Zenhäusern, T. Zesiger, J.O. Bosset, F. Escher, Use of preconcentration techniques applied to a MS-based electronic nose, Analusis 28 (2000) 743–749.
[13] B. Dittmann, S. Nitz, Strategies for the development of reliable QA/QC methods when working with mass spectrometry-based chemosensory systems, Sens. Actuators B 69 (2000) 253–257.
[14] G.A. Carpenter, S. Grossberg, N. Markuson, J.H. Reynolds, D.B. Rosen, Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps, IEEE Trans. Neural Networks 3 (5) (1992) 698–713.
[15] E. Llobet, E.L. Hines, J.W. Gardner, P.N. Bartlett, T.T. Mottram, Fuzzy ARTMAP based electronic nose data analysis, Sens. Actuators B 61 (1999) 183–190.
[16] P. Geladi, B.R. Kowalski, Partial least squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[17] R. Leardi, M.B. Seasholtz, R.J. Pell, Variable selection for multivariate calibration using a genetic algorithm: prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data, Anal. Chim. Acta 461 (2002) 189–200.
[18] T. Eklöv, P. Mårtensson, Selection of variables for interpreting multivariate gas sensor data, Anal. Chim. Acta 381 (1999) 221–232.
[19] H.C. Goicochea, A.C. Olivieri, A new family of genetic algorithms for wavelength interval selection in multivariate analytical spectroscopy, J. Chemometr. 17 (2003) 338–345.
[20] R. Leardi, Application of genetic algorithm-PLS for feature selection in spectral data sets, J. Chemometr. 14 (2000) 643–655.
[21] Y.B. Che Man, J.L. Liu, R. Abdul Rahman, B. Jamilah, Shelf-life of fried potato chips using RBD palm olein, soybean oil and their blends, J. Food Lipids 6 (1999) 287–298.
[22] L. Davis, The Handbook of Genetic Algorithms, Van Nostrand Reinhold, New York, 1991.

Biographies

Maria Vinaixa, graduated in Chemistry from the Universitat Rovira i Virgili of Tarragona (Catalonia, Spain) in 2002. Since 2003 she has been a PhD student in the Electronic Engineering Department at the Universitat Rovira i Virgili (Tarragona, Spain). Her work focuses on the development of MS based e-noses and their applications.

Eduard Llobet, graduated in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), (Barcelona, Spain) in 1991, and received a PhD in 1997 from the same university. His thesis dealt with the use of sensor dynamics to enhance selectivity. During 1999 he was a visiting fellow at the Sensor Research Laboratory, University of Warwick, UK. He is currently an associate professor in the Electronic Engineering Department at the Universitat Rovira i Virgili (Tarragona, Spain). His main areas of interest are in the fabrication and modelling of semiconductor chemical sensors and in the application of intelligent systems to complex odour analysis. He has co-authored over 120 technical papers in journals, conference proceedings and book chapters and holds 3 patents. Dr. Llobet is a member of the IEEE, Electron Devices Society.

Jesús Brezmes, graduated in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), (Barcelona, Spain) in 1993. Since 1993, he has been a PhD student in the Signal Processing and Communications Department at the same university. He is currently an associate professor in the Electronic Engineering Department at the Universitat Rovira i Virgili (Tarragona, Spain). His main area of interest is in the application of signal processing techniques such as neural networks to chemical sensor arrays for complex aroma analysis. He has co-authored over 100 technical papers in journals, conference proceedings and book chapters and holds 3 patents. Dr. Brezmes is a member of the IEEE.

Xavier Vilanova, graduated in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), (Barcelona, Spain) in 1991, and received his PhD in Electronic Engineering in 1998 from the same university. He is currently an associate professor in the Electronic Engineering Department at the Universitat Rovira i Virgili (Tarragona, Spain). His main areas of interest are in the fabrication, modelling and simulation semiconductor chemical sensors. He has co-authored over 100 technical papers in journals and conference proceedings and holds 3 patents.

Xavier Correig, graduated in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), (Barcelona, Spain) in 1984, and received his PhD in 1988 from the same university. He is a full professor of Electronic Technology in the Electronic Engineering Department at the Universitat Rovira i Virgila (Tarragona, Spain). His research interests include heterojunction semiconductor devices and solid-state gas sensors. He has co-authored over 150 technical papers in journals and conference proceedings and holds 3 patents. Dr. Correig is a member of the IEEE.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper IV

Llobet, E.; Gualdrón, O.; **Vinaixa, M**.; El-Barbri, N.; Brezmes, J.; Vilanova, X.; Bouchikhi, B.; Gómez, R.; Carrasco, J. A.; Correig, X., Efficient feature selection for mass spectrometry based 'electronic nose' applications. *Chemometrics and Intelligent Laboratory Systems* **2007,** 85, 253-261.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Available online at www.sciencedirect.com
ScienceDirect
Chemometrics and Intelligent Laboratory Systems 85 (2007) 253–261
Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

Efficient feature selection for mass spectrometry based electronic nose applications

E. Llobet [a,*], O. Gualdrón [a], M. Vinaixa [a], N. El-Barbri [b], J. Brezmes [a], X. Vilanova [a], B. Bouchikhi [b], R. Gómez [c], J.A. Carrasco [c], X. Correig [a]

[a] *MINOS, Microsystems and Nanotechnology for Chemical Analysis, Universitat Rovira i Virgili, Avda. Països Catalans, 26, 43007, Tarragona, Spain*
[b] *Electronic and Information Processing Laboratory, Faculty of Sciences, Physics Department, University Moulay Ismaïl, B.P. 4010, Beni M'Hamed, Meekness, Morocco*
[c] *Dept. de Ingeniería, Instituto del Frío, Consejo Superior Investigaciones Científicas, José Antonio Novais 10, 28040 Madrid, Spain*

Received 14 December 2005; received in revised form 19 June 2006; accepted 6 July 2006
Available online 8 August 2006

Abstract

High dimensionality is inherent to MS-based electronic nose applications where hundreds of variables per measurement (*m/z* fragments) – a significant number of them being highly correlated or noisy – are available. Feature selection is, therefore, an unavoidable pre-processing step if robust and parsimonious pattern classification models are to be developed. In this article, a new strategy for feature selection has been introduced and its good performance demonstrated using two MS e-nose databases. The feature selection is conducted in three steps. The first two steps are aimed at removing noisy, non-informative and highly collinear features (i.e., redundant), respectively. These two steps are computationally inexpensive and allow for dramatically reducing the number of variables (near 80% of initially available features are eliminated after the second step). The third step makes use of a stochastic variable selection method (simulated annealing) to further reduce the number of variables. For example, applying the method to an Iberian ham database has resulted in the number of features being reduced from 209 down to 14. Using the surviving *m/z* fragments, a fuzzy ARTMAP classifier was able to sort ham samples according to producer and quality (11-category classification) with a 97.24% success rate. The whole feature selection process runs in a few minutes in a Pentium IV PC platform.

Keywords: Feature selection; Mass spectrometry; Electronic nose; Simulated annealing; Neural networks; Iberian ham

1. Introduction

In the last few years, with the use of mass spectrometry (MS), a new branch within electronic nose research has developed and gained importance. Unlike in classical gas chromatography/mass spectrometry systems (GC/MS), in MS based electronic noses the sample delivery unit directly injects complex volatile mixtures (such as the ones generated in the headspace of foodstuffs or beverages) into an ionisation chamber, without a previous separation step (provided by GC). This results in very complex ionisation patterns that are recorded at the detector side. These ionisation patterns are then processed by pattern recognition engines to perform tasks associated to electronic nose systems such as classification, recognition and, to a limited extent, quantification [1–6]. Although the detector of a MS gives a signal that depends linearly, at least within a range, on the abundance of any given mass to charge ratio, the complexity of the ionisation patterns that are analysed in some particular applications justifies the need of using non-linear pattern recognition methods, including neural networks [4].

In MS-based electronic noses, every mass to charge ratio (*m/z*) in the mass spectra can be thought of as a sensor. In accordance with the electronic nose philosophy, a priori knowledge of the components present in the headspace being analysed should not been required. This is why most applications developed using this approach consider spectra consisting of a wide range of *m/z* ratios (e.g. from m/z 35 to m/z 300), which cover the fragmentation of volatile molecules. This implies that over two hundred features are going to be available for the pattern recognition analysis. Therefore, it is not

* Corresponding author.
E-mail address: eduard.llobet@urv.net (E. Llobet).

0169-7439/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2006.07.002

uncommon that the number of features exceeds the number of measurements available to train the pattern recognition methods and this is a dangerous situation because there is a high risk of overfitting [7]. Actually, a significant number of sensors (i.e., *m*/*z* ratios) can be irrelevant (i.e., noisy) for the application considered, while other sensors can show highly correlated responses. A step of dimensionality reduction seems, therefore, imperative prior to attempt the building of pattern recognition methods.

Different strategies have been reported for the reduction of dimensionality. These basically consist of either choosing directly among the variables available [8,9] (e.g., *m*/*z* ratios) or to compute new variables called factors (e.g., by performing a principal component analysis or a linear discriminant analysis, etc.) and selecting among the factors [10]. In that sense, the application of multivariate statistical techniques such as discriminant partial least squares (PLS-DA) provides the possibility of understanding the data generated by MS-e-noses based on the overall properties of the sample and perform data compression and classification without the need of additional information about its chemical composition [11]. Selecting from the full spectrum of mass to charge ratios is challenging because there is considerable overlapping among the spectra and distinctive features can be almost imperceptible. Furthermore, spectra are affected by noise. There are different sources of background noise in GC/MS systems. For example, spectral background noise is associated with contaminants present in the ionisation chamber (such as ambient air and contaminants present in the carrier gas). To a higher extent, spectral background noise is due to the counting principle and inherent noise associated with the ion multiplier. Additionally, a baseline drift may appear due to co-eluting compounds, septa and temperature induced column bleed. However, methods based on the selection of *m*/*z* ratios are interesting because the variables chosen carry direct relevant chemical information. Therefore, these methods are expected to be robust toward the experimental conditions of each specific application. Unlike in *m*/*z* selection, factor selection uses the full spectrum (e.g. including noisy or redundant m/z ratios) to compute the factors before selecting from among them. The selection of an optimal subset of factors is not necessarily straightforward because the magnitude of an eigenvalue is not always a measure of its significance for the calibration [12]. Furthermore, unlike *m*/*z* ratios, factors have no direct chemical meaning.

Once determined that selecting among *m*/*z* ratios is the more interesting approach, it should be pointed out that an exhaustive search is out of question, given the high number of variables considered for selection. Several methods that avoid being exhaustive, the so-called greedy methods, have been reported as useful. These include deterministic methods such as branch and bound, sequential forward selection, sequential backward selection or stepwise selection and, stochastic methods such as genetic algorithms or simulated annealing [13–22]. Deterministic methods can make a fair selection with relatively few operations but are prone to get trapped in a local optimum of the search space. On the other hand, stochastic methods such as genetic algorithms or simulated annealing are more likely to find a global optimum at the cost of lengthy computation. For example, a genetic algorithm for variable selection running in a Pentium IV PC platform can take as long as several days to converge to a good solution when the number of variables for selection is above two hundred. Therefore, applying a stochastic method to select among the features found in MS-based electronic nose applications can easily turn to be unpractical.

In this paper we introduce a new method for an effective feature selection especially suitable for applications where the dimension of feature space is high, a significant degree of correlation exists between features and some of them are affected by noise, such as in MS electronic nose applications. The method is efficient in the sense that after the selection process, only those features that are important for the application considered are retained to build the pattern recognition models and all the process is conducted at a very low computational cost. The usefulness of the new feature selection method is assessed using two different MS-based electronic nose databases.

2. Experimental

2.1. Solvent database

The first database consisted of measurements taken from samples with a well-characterised headspace. The samples were obtained from 4 different solutions of pure ethanol containing added impurities (trichloroethylene, 1-butanol, ethylbenzene and toluene). The composition of the different solutions is shown in Table 1. From the 4 different initial solutions, 6 different samples were obtained (6 solutions were prepared in 6 different flasks). For each one of these 6 samples, 5 different aliquots of 10 ml were taken and placed into 20-ml glass vials and sealed hermetically with silicone septa and caps. In total 120 vials were analysed (this preparation technique allows for obtaining small differences between samples, which are representative of experimental errors). Sampling based on solid-phase micro extraction (SPME) was performed with a 75-μm Carboxen/PDMS fiber purchased from Supelco (Supelco Park, Bellefonte, PA). Prior to any extraction, the fibre was conditioned following the manufacturer's recommendations. In each measurement, the fibre was pushed out of its stainless steel housing and exposed to the sample headspace for 20 min at room temperature. The SPME holder assembly scale was adjusted to 3.0 scale units to ensure that the fibre was positioned in the headspace above the sample in exactly the same way from run to run.

Table 1
Composition of the different solutions in database 1

Compounds				
Sample #	TCE	1-B	EB	TOL
S1	1	1	–	1
S2	1	1	1	1
S3	1	1	1	–
S4	1	–	1	1

Quantities are expressed in % dissolved in ethanol.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

A Shimadzu QP 5000 GC/MS (Shimadzu Corp., Tokyo, Japan) was used to implement a MS-based electronic nose. The instrument was equipped with a capillary column (Supelcowax, 30 m × 0.25 mm i.d., × 0.25 mm coating thickness). The volatile compounds trapped on the SPME fibre were subsequently desorbed for 3 min at 280 °C into the glass-lined injection port of the GC, actuated in the splitless mode. The carrier gas was helium 99.995% set to 1.0 ml/min. The temperature of the GC oven and of the GC/MS interface was held constant at 250 °C so chromatographic separation was avoided. Mass spectra were recorded at a rate of 2 scans/s over *m*/*z* ratios that ranged between 40 to 150 amu, operating the MS in the electron impact (EI) mode (70 eV). This range is known to contain all the *m*/*z* ratios with higher intensities for the compounds involved in this database. The total ion current obtained under these conditions took the form of an asymmetrical peak as shown in Fig. 1. The response matrix contained the mean abundance values of the mass fragments recorded between scan 1 and scan 70 for each sample analysed. Further data processing was performed on the relative mass spectra (i.e., normalised by the amplitude of the highest peak).

2.2. Iberian ham database

Eleven types of Spanish Iberian dry-cured hams were analysed. Samples were obtained directly from five producers and they differed in the type of food the pigs fed on during their fattening period (i.e., either acorn or fodder) and in their quality (type of pigs). Table 2 gives more details on the hams used.

Samples were prepared as follows: 3 grams of ham (taken from biceps femoris) were crushed and introduced in 10 ml glass vials, which were then sealed with a septum and an aluminium cap. For each type of ham, 10 samples were prepared (exception made of one type from Extremadura with nine only). This gave a total of 109 ham samples to be analysed. Sampling was based on static headspace. A headspace autosampler Agilent 7694 was used. Oven, loop and transfer line temperatures were set to 90, 100 and 110 °C, respectively. The times for vial equilibration, pressurisation, loop filling, loop equilibration and injection were 30, 0.4, 0.15, 0.2 and 1 min, respectively. Reproducible headspace samples were injected into the injection port of a Hewlett–Packard 6890 series II gas chromatograph coupled to a mass selective detector (Hewlett–Packard HP 5973; Wilmington, DE, USA). The injection port was used in splitless mode and maintained at 280 °C. The system was equipped with a HP 19091J-215 (50 m × 0.32 mm id, film thickness 1.05 μm) column, kept at 200 °C in isothermal conditions. In this way, chromatographic separation was avoided and the column merely acted as a transfer line delivering volatiles to the mass detector. The column flow rate was set to 1.5 ml/min. Volatile compounds were co-eluted into the mass spectrometer, where mass spectra were obtained using an electronic impact mass selective detector at 70 eV, a multiplier voltage of 2706 V, and collecting data at a rate of 1 scan s^{-1} over the *m*/*z* range 45–250 amu. In this case mass fragments lighter than 45 amu were ignored because of the

Table 2
The 11 types of Spanish Iberian dry-cured hams analysed

Ham brand	Short name	# ham types	Pig feeding on
Extremadura	EX	4	acorn
Guijuelo #1	G1	1	acorn
Huelva	HU	1	acorn
Guijuelo #2	G2	2	fodder
Guijuelo #3	G3	3	fodder

The hams differ in producer, type of pigs and pigs' feeding.

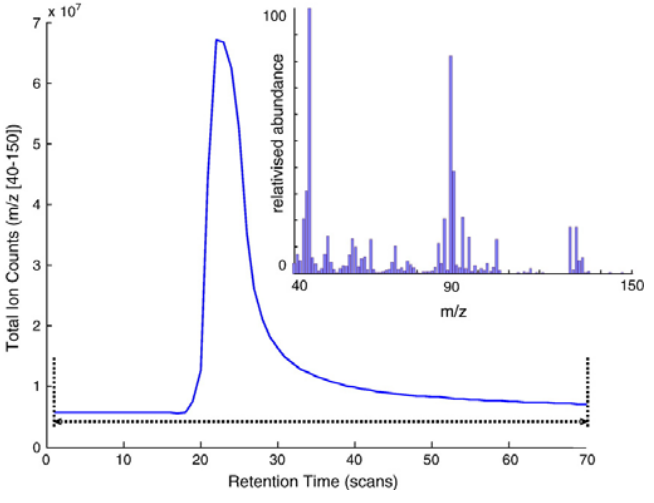

Fig. 1. Co-eluted chromatographic peak from an S2 sample (solvent database) obtained by the SPME-MS-Sensor. The insert shows a normalised mass spectrum obtained by integrating the TIC signal between scans 1 and 70.

multiple origins for such fragments. Actually, ambient air and carrier gas were far more important than ham samples for the presence of these low mass to charge fragments.

2.3. Software

The different methods for variable selection were coupled to fuzzy ARTMAP neural networks that were built using written-in-house and standard MATLAB® routines [23,24].

3. Feature selection

The feature selection introduced here consists of three steps that are run consecutively. The first step helps detecting and removing non-informative, noisy features and is conducted in a supervised way. The second step is aimed at detecting collinearity between features in an unsupervised way. As a result, highly collinear features can be removed. Finally, in the third step, a greedy search method (e.g. a stochastic one) is applied to the reduced feature set, which results from applying the first two steps. With this approach, the whole variable selection process is time efficient since the first two steps are able to dramatically reduce the number of features at a very low computational cost.

3.1. Removal of non-informative and noisy features

In electronic nose applications, the outcome sought after the system has been trained as an automated recognition or classification of new unknown samples. During the training phase, the pattern recognition ability of the system is built by using calibration samples. In the first step of feature selection, a criterion was used to rate the discrimination ability of each feature (i.e. m/z ratio). Measurements used for training were grouped in categories (e.g. measurements of the same type of ham were grouped in a category. Therefore, there were 11 ham categories). For each m/z ratio, intra-category and inter-category variances were computed. Intra-category variance was defined as the variance of an m/z ratio considered within a given category k of measurements. Therefore, the intra-category variance of the j-th m/z ratio for category k, was defined as:

$$\sigma^2_{\text{intra},jk} = \frac{\sum_{i=1}^{n_k} (m/z_{ji} - \mu_{jk})^2}{n_k - 1} \quad (1)$$

where n_k is the number of measurements within the category k, m/z_{ji} is the value of mass to charge ratio j for measurement i and μ_{jk} is the mean of mass to charge ratio j over the measurements within the category k.

The overall intra-category variance of the j-th m/z ratio was then defined as the following pooled variance:

$$\sigma^2_{\text{intra},j} = \frac{\sum_{k=1}^{d} n_k \sigma^2_{\text{intra},jk}}{\sum_{k=1}^{d} n_k} \quad (2)$$

where d is the number of different categories. In a similar way, for every mass to charge ratio, an inter-category variance was defined as the variance within the category means. Therefore, the inter-category variance was defined as:

$$\sigma^2_{\text{be},j} = \frac{\sum_{i=1}^{d} (\mu_{jk} - \bar{\mu}_j)^2}{d-1} \quad (3)$$

where $\bar{\mu}_j$ is the mean over the μ_{jk}.

The discrimination ability of the j-th m/z ratio was defined as follows:

$$\text{DA}_j = \frac{\sigma^2_{\text{be},j}}{\sigma^2_{\text{intra},j}} \quad (4)$$

The higher the discrimination ability for a given m/z ratio is, the more important is this m/z ratio to correctly discriminate between the categories. In other words, noisy or non-informative mass to charge ratios will have associated low discrimination abilities. Therefore, a set of m/z ratios, which comprises those that have the higher figure of merit, is selected for further analysis. This method would be equivalent to compute Fisher's linear discriminant if the number of categories to sort measurements within was $d=2$. This process is univariate and there is a risk of eliminating those synergetic variables that have low discrimination ability when considered individually. To minimise this problem the process described by Eqs. (1)–(4) is repeated considering all the possible combinations between two m/z ratios. Fig. 2 illustrates these criteria. As a result, a new list of figures of merit, $\text{DA}_{i,j}$, i.e., the

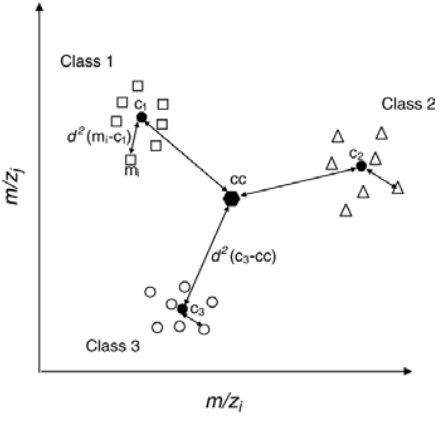

Fig. 2. Geometrical interpretation of the process used to compute the figure of merit for any two given features (case of 3 different categories). The intra-category variance is computed as the average of the squared distances between measurements within a category (e.g., m_i) and the category centroid (e.g., c_1). The between-category variance is computed as the summation of the distances between category centroids and the overall centroid (cc). $d^2(m_i - c_1)$ denotes the squared Euclidean distance between measurement i and its category centroid, c_1.

discrimination ability when m/z ratios i and j are used simultaneously, is obtained. This allows for re-selecting variables that had been removed previously, if a synergistic effect is revealed. The threshold value for the discrimination ability is heuristically set to a value that allows for retaining about 30% of the originally available m/z ratios.

However, it is important to notice that this method does not prevent redundant features (i.e., highly collinear) from being selected. This will be the task of the second step of feature selection.

3.2. Detection and removal of redundant features

Let **R** be the calibration matrix resulting from the first step of feature selection. **R** is a $(n \times p)$ matrix. Its number of columns, p, corresponds to the number of features selected in the first step and, its number of rows, n, corresponds to the number of measurements within the calibration set. If $\mathbf{R}^t$ denotes the transpose of **R**:

$$R^t = \begin{bmatrix} m/z_{11} & m/z_{12} & \dots & m/z_{1n} \\ m/z_{21} & m/z_{22} & \dots & m/z_{2n} \\ \vdots & \vdots & \dots & \vdots \\ m/z_{p1} & m/z_{p2} & \dots & m/z_{pn} \end{bmatrix} \quad (5)$$

where m/z_{ji} corresponds to the intensity of the j-th mass to charge ratio for measurement i. For any mass to charge ratio, a unity-norm response vector can be defined as follows:

$$\overrightarrow{m/z_j} = \left(\frac{m/z_{j1}}{\sqrt{\sum_{i=1}^{n} m/z_{ji}^2}}, \frac{m/z_{j2}}{\sqrt{\sum_{i=1}^{n} m/z_{ji}^2}}, \dots, \frac{m/z_{jn}}{\sqrt{\sum_{i=1}^{n} m/z_{ji}^2}} \right)$$

for $j = 1$ to p. (6)

Eq. (6) shows the unity-norm response vector for the j-th mass to charge ratio. Now, the degree of collinearity existing in the calibration set between two different mass to charge ratios can be assessed by computing the scalar product of their unity-norm response vectors as shown below:

$$\boldsymbol{P}_{j,q} = \sum_{r=1}^{n} \left(\frac{m/z_{jr}}{\sqrt{\sum_{i=1}^{n} m/z_{ji}^2}} \times \frac{m/z_{qr}}{\sqrt{\sum_{i=1}^{n} m/z_{qi}^2}} \right) \quad (7)$$

$\boldsymbol{P}_{j,q}$ is the scalar product between the normalised response vectors associated to features j and q. $\boldsymbol{P}_{j,q}$ ranges between 0 and 1. The closer to unity $\boldsymbol{P}_{j,q}$ is, the higher the collinearity between mass to charge ratios j and q is. Since p is the number of features, the collinearity of which is to be checked, the number of scalar products to be computed is $\sum_{i=1}^{p-1}(p-i)$. After these scalar products have been obtained, a collinearity threshold is set and used to determine which features are redundant and should be removed. The value of the collinearity threshold is heuristically set to a value that allows for retaining about 20% of the originally available m/z ratios (i.e., the ratios available prior to perform any variable selection). This second step of variable selection is non-supervised since, unlike in the previous step, there is no need to classify training samples according to their category.

After the removal of noisy, irrelevant and redundant features, the set of surviving features is ready for the last step of feature selection.

3.3. Stochastic feature selection

Stochastic methods such as genetic algorithms (GA) or simulated annealing (SA) are more likely to find a global optimum in the optimisation problem. A discussion on the SA algorithm is given in the Annex. These methods represent a trade off between the simple sequential methods (prone to get trapped in a local optimum) and the burden of exponential methods [18–20]. Genetic algorithms and simulated annealing solve the optimisation problem by exploring all regions of the potential solutions. Because explored points in a solution space are chosen by stochastic rather than deterministic rules, stochastic methods do not need to make assumptions about the characteristics of the problem to be solved and, therefore, apply generally. In the particular case of feature selection, these methods explore different subsets of the original set of features. Both GA and SA make use of a cost function, which in the case reported here, is an estimate of the prediction error of a neural network classifier (e.g. fuzzy ARTMAP) computed using the training measurements. This cost function is used to rank the fitness of solutions (i.e., combinations of features) during the process of stochastic feature selection. Since in most MS-based electronic nose applications a high number of variables are highly collinear or non-informative, about 80% of the original variables are eliminated by the first two steps. Therefore, the last step is aimed at fine tuning the selection process. Although stochastic feature selection methods are time-consuming, run to select among a reduced set of features that result from the two previous steps is fast.

4. Results and discussion

4.1. Analysis of the solvent database

A priori, the main challenge to correctly identify these compounds is due to the presence of ethylbenzene and toluene in the mixtures as these two species show some similarities between their mass spectra fragmentation pattern. Table 3 shows which are the most intense fragments found in the mass spectra of the different compounds used. This database is a good benchmark for the feature selection method introduced here, since looking at Table 3, one could select a set of mass to charge ratios to discriminate the different mixtures. Therefore, the main objective sought with this database is to assess whether the 3-step feature selection was able to correctly determine, out of the 111 features available, the

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Table 3
Most intense *m/z* fragments found in the mass spectra of the different compounds used in the solvent database

Compound	10 more intense *m/z* fragments
Trichloroethylene	132, 130, 95, 97, 60, 134, 47, 62, 59, 94
1-butanol	56, 41, 43, 42, 55, 45, 40, 57, 44, 53
Ethylbenzene	91, 106, 51, 65, 77, 78, 50, 92, 52, 63
Toluene	91, 92, 65, 51, 63, 45, 50, 46, 62, 89

Fragments appear sorted by decreasing intensity.

few ones would enable a classifier to correctly identify the mixtures.

Before performing variable selection, a fuzzy ARTMAP was trained and validated using the leave-one-out cross validation approach. A description of the fuzzy ARTMAP algorithm can be found elsewhere [25]. The classifier made use of 111 inputs and the number of categories was set to 4 since this was the number of different mixtures analysed. The success rate in classification was 95.83%, which corresponded to one sample S2 being identified as S4 (see Table 1).

The process of feature selection was conducted as follows. The 120 measurements available were split in 6 different feature selection datasets, which were originated from the 6 samples prepared in 6 different flasks. Each selection dataset contained 100 measurements (5 samples×5 aliquots×4 types of solutions) and their corresponding 6 validation datasets, which contained the remaining 20 measurements (i.e. 1 sample×5 aliquots×4 types of solutions) not used in the corresponding feature selection dataset. Then the process of variable selection was performed 6 times on each feature selection dataset. The first step of feature selection was applied to eliminate noisy and irrelevant features. By setting to 0.5 the threshold value of the discrimination ability, between 29 and 37 out of 111 features were initially selected, depending on the feature selection dataset used. The second step was then applied to eliminate collinear variables. By setting to 0.15 the values of the collinearity threshold, between 17 and 21 features were retained. Computing the first two steps required about 7 min in a Pentium 4 PC platform. Finally a simulated annealing feature selection was run to select among the remaining features [26]. The SA algorithms were run for 50 different annealing temperatures and the number of iterations per temperature was set to 17. More details on the simulated annealing algorithm used can be found in the Annex. In the end, only 3 features were selected (no matter what the feature selection dataset was used). These were the *m/z* ratios 46, 56 and 106. Using these three features as inputs, 6 fuzzy ARTMAP classifiers were trained employing the 6 feature selection datasets, and their performance in classification estimated using the corresponding validation datasets. The success rate in solvent mixture classification, estimated over the 6 training/validation sets was 100%. Fig. 3 shows a block diagram of the feature selection and validation processes. It is important to keep in mind that for every fuzzy ARTMAP classifier, the validation implies using measurements that have not participated in the feature selection process and are, therefore, new. Considering Table 3, it can be derived why the method has selected these specific features:

- $m/z=46$, which is one of the most relevant mass to charge ratios for toluene and not found for the other compounds in the solvent mixture.
- $m/z=56$, which is the most relevant mass to charge ratio for 1-butanol.
- $m/z=106$, the second more relevant mass to charge ratio for ethylbenzene.

It is important to notice that the different feature selection processes have disregarded using $m/z=91$. This is the most frequent ion for ethylbenzene and toluene, which would not help in discriminating between these two compounds. Finally, no mass to charge ratio that is characteristic of trichloroethylene has been selected. This is correct because trichloroethylene is present in all the different samples to be discriminated and, therefore, no information about this compound is needed for a good discrimination among the samples analysed (see Table 3). These results show that the three-step feature selection process introduced here is able to find the essential information needed to solve the discrimination problem considered. The whole process of feature selection took 15 min to complete in a Pentium 4 PC platform.

4.2. Analysis of the Iberian ham database

Initially an 11-category classification was attempted using a fuzzy ARTMAP classifier without a previous step of feature selection. Because in this database the number of measurements available was higher, a different method of cross validation was employed. A 5-fold validation was implemented, which consists of defining 5 training and validation datasets. A training dataset comprised 8 replicate measurements (out of the 10 available) per type of ham (i.e., 87 measurements in total, since one type of ham had 9 replicate measurements instead of 10). The corresponding validation dataset comprised the 2 measurements per ham sample that had been left out (i.e. 22 measurements). Therefore, the fuzzy ARTMAP classifier was trained and validated 5 times using the 5 training and validation sets and the success rate in ham classification was averaged over the 5 tests. The success rates for the 5 folds were 63.63%, 95.45%, 100%, 100% and 81.81%, which gave an overall classification success rate of 88.18% (the standard deviation was 15.61%). This corresponds, in average, to 13 samples out of 110 being misclassified. Confusions occur between samples belonging to different producers and different quality hams within a producer.

The process of feature selection was performed using the 5 training and validation sets described above. For every pair of training and validation sets, feature selection was conducted on the training set, then a fuzzy ARTMAP classifier was trained using the features selected and, finally, its success rate in ham classification was estimated using measurements in the validation dataset. The first and second steps of feature selection

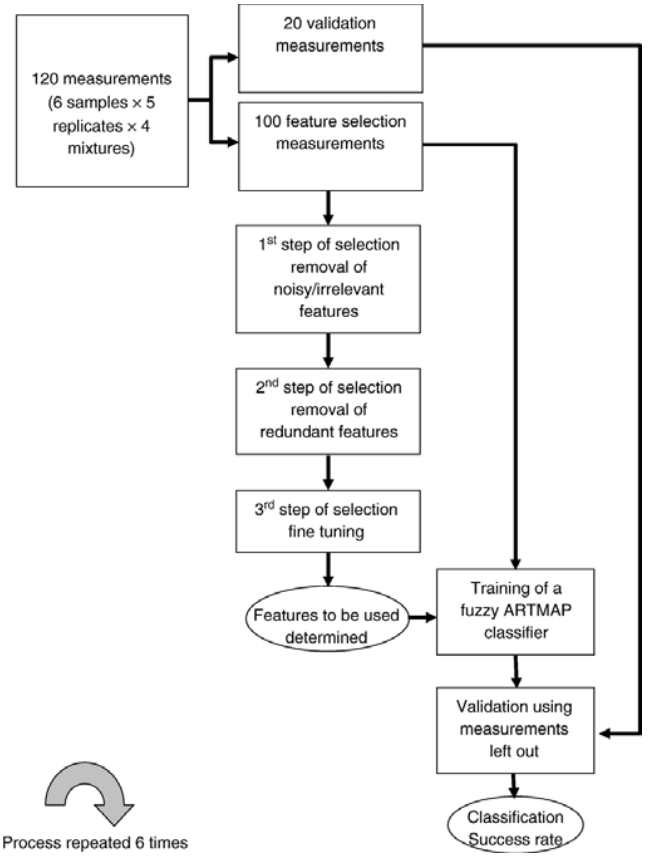

Fig. 3. Block diagram that illustrates the process of feature selection and validation with the solvent mixture database.

were applied to remove noisy, irrelevant and redundant variables in every training and validation fold. The threshold values used were 0.5 and 0.8, respectively. This resulted in 42 features being selected in average.

Then, the third step of feature selection, which consisted in selecting among the surviving features using a simulated annealing procedure, was performed. As previously, the process was conducted independently for the 5 folds available. The SA algorithms were run for 50 different annealing temperatures and the number of iterations per temperature was set to 40. Fuzzy ARTMAP classifiers (one per fold) were trained and validated using as inputs the features that remained selected after the last step. The success rates in sample classification were 81.18%, 95.45%, 90.90%, 100% and 90.90% and that gave an overall success rate of 91.68% in ham classification (the standard deviation was 6.98%). This corresponds, in average, to 9 samples out of 110 being misclassified. Confusions occur between samples belonging to different producers but never between different quality hams within a producer. The average number of features selected after the three-step variable selection was 19 (see Table 4), i.e., near 8% of the features initially available. Table 4 shows that a high number of features

Table 4
m/z fragments selected for each selection/validation fold after the three-step feature selection process

Fold #	Selected *m/z* fragments
1	45, 47, 49, 56, 58, 59, 64, 70, 71, 73, 77, 79, 80, 81, 83, 85, 94, 100, 104, 111, 114
2	45, 47, 49, 53, 56, 57, 58, 60, 61, 64, 71, 72, 77, 81, 83, 84, 94, 100, 114, 208
3	47, 48, 55, 61, 64, 67, 77, 81, 82, 83, 84, 104, 138
4	45, 49, 56, 58, 64, 71, 77, 81, 82, 83, 84, 104, 138
5	45, 47, 51, 53, 56, 57, 58, 60, 61, 64, 67, 69, 70, 71, 72, 73, 79, 81, 82, 83, 84, 85, 93, 94, 101, 105, 108, 114, 133, 138
Most frequent *m/z* fragments	45, 47, 49, 56, 58, 64, 71, 77, 81, 83, 84, 94, 114, 138

The last row shows the most frequent *m/z* fragments.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Table 5
Summary of the variable selection (VS) results for the solvent and ham databases

Database	# variables available	Classification success rate before VS	# variables retained	Classification success rate after VS
Solvent solutions	111	95.83%	4	100.00%
Dry cured hams	209	88.18%	14	94.54%

are shared by the different folds. This demonstrates the robustness of the variable selection method applied.

Finally, an eleven-category classification was envisaged using a fuzzy ARTMAP classifier using the outcome of the previous variable selection steps. Only the 14 most frequently selected *m*/*z* ratios were used as inputs of the classifier (see Table 4 for details). Its performance in the correct classification of Iberian hams was estimated to be 94.54% by using a 5-fold cross-validation approach. Only 6 out of 110 ham samples were misclassified. Furthermore, a 100% correct discrimination between hams from pigs fed on acorn or fodder was found to be possible. These results compare favourably to the 88.18% classification success rate reached with a fuzzy ARTMAP classifier that used all the features available (i.e., 209). Table 5 summarises the results of the variable selection process for both the solvent and ham databases.

A short discussion on the fragments selected by the feature selection method and used to build the ham classification models is as follows. Differences in pig feeding lead to different volatile profiles obtained from crushed samples of subcutaneous fat and meat in Iberian hams. The levels of hexanal and pentanal, which arise mainly from the oxidation of linoleic acid, are rather similar regardless of pig feeding. On the other hand, nonanal, the most important aldehyde derived from oleic acid is found in significantly larger quantities in pigs fed on acorn than in pigs fed on fodder [27,28]. The *m*/*z* fragment 114 selected in the model is present in the mass spectrum of nonanal, and therefore, helps in discriminating between acorn and fodder fed pigs. Other fragments selected such as *m*/*z* 77, may arise from aromatic volatiles, *m*/*z* 71 from esters, alkanes, propylketones and butanoate and *m*/*z* 45 could be due to the presence of carboxylic acids or alcohols. Finally, the presence of pentylketones and methylketones is revealed by *m*/*z* 56 and 58, respectively. All these compounds have been reported to be characteristic of the headspace of dry cured Iberian hams [27,28].

5. Conclusions

A new strategy for feature selection has been introduced and its good performance demonstrated using different MS e-nose databases. The feature selection consists of three steps, the first two being aimed at eliminating non-informative and highly collinear features, respectively. The removal of noisy and redundant features is computationally inexpensive and allows for dramatically reducing the number of variables (near 80% of initially available features are eliminated after the second step). This is especially interesting to solve MS-based electronic nose problems where the number of features (*m*/*z* fragments) available per measurement is high. The third step makes use of simulated annealing, which is a stochastic search method to further reduce the number of variables (fine-tuning of the feature selection process).

The strategy has been applied initially to a database consisting of synthetic mixtures of volatile compounds. This simple database has been used to show that the feature selection process is able to identify a minimal set of fragments that enables the correct discrimination between mixtures using a simple fuzzy ARTMAP classifier. Furthermore, given the simple nature of the problem envisaged, it was possible to show that the fragments selected 'made sense', that is, were characteristic ionisation fragments of the species present in the mixtures to be discriminated.

Once the correct performance of the feature selection method was demonstrated, it was applied to an additional database (Iberian hams).

Applying the method to the Iberian ham database resulted in the number of features being reduced from 209 down to 14. Using the surviving features, a fuzzy ARTMAP classifier was able to discriminate ham samples according to producer and quality (11-category classification) with a 97.24% success rate. It was also possible to identify, with a 100% success rate, whether the pigs had been fed on acorn or fodder.

For the different databases studied, performing variable selection results in a dramatic decrease in dimensionality and an increase in classification performance. The methods introduced here are useful not only to solve MS-based electronic nose problems, but are of interest for any electronic nose application suffering from high-dimensionality problems, no matter what the sensing technology is employed.

Acknowledgements

This work was funded in part by CICYT under project no. TIC2003-06301, by the Thematic Network in Metabolism and Nutrition ref. C03/08 and by AECI under project no. 39/04/P/E. O. Gualdrón acknowledges a pre-doctoral fellowship from the Universidad de Pamplona, Colombia.

Appendix A. Simulated annealing algorithm

Simulated annealing (SA) is a stochastic technique derived from statistical thermodynamics for finding near globally optimum solutions to complex optimisation problems (i.e., with a high number of degrees of freedom). The algorithm proceeds stepwise through a search space defined by all possible solutions to the optimisation problem. After each iteration (e.g. after a feature has been removed), the value of the cost function for the new step is compared to that of the previous step. If the new solution is better than the old one, the removal of the feature is confirmed. If the new solution is worse than the old one, there is still a probability, *p*, for the removal of the feature to be accepted. This offers the algorithm the possibility to jump out of a local optimum. Otherwise, the removal of the feature will be discarded and the previous step

will be the starting point for the next attempt to eliminate a variable. The probability p for accepting a worse solution depends on the difference between the new and previous solution as follows:

$$p = \exp\left(-\frac{\Delta E}{T_i}\right) \quad (a.1)$$

where ΔE=fitness (new)−fitness (old) and T_i is the annealing temperature during the i-th iteration. Since the cost function (or fitness) being optimised by the SA is the prediction error of a fuzzy ARTMAP classifier, ΔE is positive when the new solution is worse than the old one. The initial temperature is set by the user, and the whole process of feature selection is repeated a fixed number of times for monotonically decreasing annealing temperatures (as the annealing temperature is reduced, the probability for accepting a worse solution decreases significantly since T_i appears in the denominator of the exponential factor in Eq. (a.1)). After every change in the annealing temperature, the algorithm starts selecting from the complete set of features. Therefore, the role of temperature in the SA algorithm is to set the tolerance of the algorithm to accept a worse solution. For example, a high initial value would imply a high tolerance during the first iterations.

References

[1] R. Marsili, SPME-MS-MVA as an electronic nose for the study of off-flavors in milk, J. Agric. Food Chem. 47 (1999) 648–654.
[2] E. Fallik, S. Alkali-Tuvia, B. Horev, A. Copel, V. Rodov, Y. Aharoni, D. Ulrich, H. Schulz, Characterisation of 'Galia' melon aroma by GC and mass spectrometric sensor measurements after prolonged storage, Post-harvest Biol. Technol. 22 (2001) 85–91.
[3] C. Peres, F. Begnaud, L. Eveleigh, J.L. Berdagué, Fast characterization of foodstuff by headspace mass spectrometry (HS-MS), Trends Anal. Chem. 22 (2003) 858–866.
[4] M. Vinaixa, S. Marín, J. Brezmes, E. Llobet, X. Vilanova, A. Ramos, V. Sanchís, Early detection of fungal growth in bakery products by use of an electronic nose based on mass spectrometry, J. Agric. Food Chem. 52 (2004) 6068–6074.
[5] P. Martí, R. Boqué, M. Riu, O. Busto, J. Guasch, Fast screening method for determining 2,4,6-trichloroanisole in wines using a headspace-mass spectrometry (HS-MS) system and multivariate calibration, Anal. BioAnal. Chem. 376 (2003) 497–501.
[6] M.P. Martí, J. Pino, R. Boqué, O. Busto, J. Guasch, Determination of ageing time of spirits in oak barrels using a headspace-mass spectrometry (HS-MS) electronic nose system and multivariate calibration, Anal. BioAnal. Chem. 382 (2005) 440–443.
[7] P. Jonsson, J. Gullberg, A. Nordström, M. Kusano, M. Kowalczyk, M. Sjöstrolm, T. Moritz, A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS, Anal. Chem. 76 (2004) 1738–1745.
[8] B. Dittmann, S. Nitz, Strategies for the development of reliable QA/QC methods when working with mass spectrometry-based chemosensory systems, Sens. Actuators, B 69 (2000) 253–257.
[9] C.B. Lucasius, M.L.M Beckers, G. Kateman, Genetic algorithms in wavelength selection: a comparative study, Anal. Chim. Acta 286 (1994) 135–153.
[10] U. Depczynski, V.J. Frost, K. Molt, Genetic algorithms applied to the selection of factors in principal component regression, Anal. Chim. Acta 420 (2000) 217–227.
[11] D. Cozzolino, H.E. Smyth, W. Cynkar, R.G. Dambergs, M. Gishen, Usefulness of chemometrics and mass spectrometry-based electronic nose to classify Australian white wines by their varietal origin, Talanta 68 (2005) 382–387.
[12] J. Sun, A correlation principal component regression analysis of NIR data, J. Chemom. 9 (1995) 21–29.
[13] T. Eklöv, P. Mårtensson, I. Lundström, Selection of variables for interpreting multivariable gas sensor data, Anal. Chim. Acta 381 (1999) 221–232.
[14] Lu Xu, Wen-Jun Zhang, Comparison of different methods for variable selection, Anal. Chim. Acta 446 (2001) 477–483.
[15] N. Paulsson, E. Larson, F. Winquist, Extraction and selection of parameters for evaluation of breath alcohol measurement with an electronic nose, Sens. Actuators, A 84 (2000) 187–197.
[16] J. Brezmes, P. Cabré, S. Rojo, E. Llobet, X. Vilanova, X. Correig, Discrimination between different samples of olive using variable selection techniques and modified fuzzy ARTMAP neural networks, Proceedings of the 9th International Symposium on Olfaction and Electronic Nose, ISOEN'02, Rome, Italy, vol. 1, 2002, pp. 188–190.
[17] T. Artursson, M. Holmberg, Wavelet transform of electronic tongue data, Sens. Actuators, B 87 (2002) 379–391.
[18] J.M. Sutter, J.H. Kalivas, Comparison of forward selection, backward elimination, and generalized simulated annealing for variable selection, Microchem. J. 47 (1993) 60–66.
[19] D. Broadhurst, R. Goodacre, A. Jones, Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry, Anal. Chim. Acta 348 (1997) 71–86.
[20] L. Nolle, D.A. Armstrong, A.A. Hopgood, J.A. Wware, Simulated annealing and genetic algorithms applied to finishing mill optimisation for hot rolling of wide steel strip, Int. J. Know.-based Intell. Engin. Sys. 6 (2002) 104–111.
[21] E. Llobet, J. Brezmes, O. Gualdrón, X. Vilanova, X. Correig, Building parsimonious fuzzy ARTMAP models by variable selection with a cascaded genetic algorithm: application to multisensor systems for gas analysis, Sens. Actuators, B 99 (2004) 267–272.
[22] J.W. Gardner, P. Boilot, E.L. Hines, Enhancing electronic nose performance by sensor selection using a new integer-based genetic algorithm approach, Sens. Actuators, B 106 (2005) 114–121.
[23] The Mathworks Inc., Matlab User's Guide, 2004, http://www.mathworks.com.
[24] B. Wise, N.B. Gallager, PLS ToolBox 2.0, Eigenvector Research, http://www.eigenvector.com.
[25] R. Ionescu, E. Llobet, X. Vilanova, J. Brezmes, J.E. Sueiras, J. Calderer, X. Correig, Quantitative analysis of NO_2 in the presence of CO using a single tungsten oxide semiconductor sensor and dynamic signal processing, The Analyst 127 (2002) 1237–1246.
[26] M. Shen, A. LeTiran, Y. Xiao, A. Golbraikh, H. Kohn, A. Tropsha, Quantitative structure-activity relationship analysis of functionalized amino acid anticonvulsant agents using k-nearest neighbor and simulated annealing PLS methods, J. Med. Chem. 45 (2002) 2811–2823.
[27] I. González-Martín, J.L. Pérez-Pavón, C. González-Pérez, J. Hernández-Méndez, N. Álvarez-García, Differentiation of products derived from Iberian breed swine by electronic olfactometry (electronic nose), Anal. Chim. Acta 424 (2000) 279–287.
[28] M.L. Timón, L. Martín, M.J. Petrón, A. Jurado, C. García, Composition of subcutaneous fat from dry-cured Iberian hams as influenced by pig feeding, J. Sci. Food Agric. 82 (2002) 186–191.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper V

Marin, S.; **Vinaixa, M**.; Brezmes, J.; Llobet, E.; Vilanova, X.; Correig, X.; Ramos, A. J.; Sanchis, V., Use of a MS-'electronic nose' for prediction of early fungal spoilage of bakery products. *International Journal of Food Microbiology* **2007,** 114 (1), 10-16.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Available online at www.sciencedirect.com
ScienceDirect
International Journal of Food Microbiology 114 (2007) 10–16
INTERNATIONAL JOURNAL OF Food Microbiology
www.elsevier.com/locate/ijfoodmicro

Use of a MS-electronic nose for prediction of early fungal spoilage of bakery products

S. Marín [a,*], M. Vinaixa [b], J. Brezmes [b], E. Llobet [b], X. Vilanova [b], X. Correig [b], A.J. Ramos [a], V. Sanchis [a]

[a] Food Technology Department, Lleida University, CeRTA-UTPV, Alcalde Rovira Roure, 191, 25198 Lleida, Spain
[b] Department Electronic Engineering, Universitat Rovira i Virgili, Avda. Països Catalans, 26, 43007 Tarragona, Spain

Received 11 July 2005; received in revised form 23 October 2006; accepted 3 November 2006

Abstract

A MS-based electronic nose was used to detect fungal spoilage (measured as ergosterol concentration) in samples of bakery products. Bakery products were inoculated with different *Eurotium*, *Aspergillus* and *Penicillium* species, incubated in sealed vials and their headspace sampled after 2, 4 and 7 days. Once the headspace was sampled, ergosterol content was determined in each sample. Different electronic nose signals were recorded depending on incubation time. Both the e-nose signals and ergosterol levels were used to build models for prediction of ergosterol content using e-nose measurements. Accuracy on prediction of those models was between 87 and 96%, except for samples inoculated with *Penicillium corylophilum* where the best predictions only reached 46%.

Keywords: Fungal growth; Electronic nose; Ergosterol; Bakery product

1. Introduction

Fungal spoilage is an important issue in bakery products. Although its impact from the safety point of view is not significant, the presence of visible colonies on the products detracts the companies' image for the consumer, resulting in economic losses in the medium–long term. Some companies use the measurement of the water activity of the final products as an index for fungal spoilage prediction and batches rejection. However, there is a need for more reliable alternative methods.

The investigation of the detection of fungal volatiles as a method for mould detection in food substrates started in cereals. For a long time, human olfaction has been used as the primary criterion for the acceptance of grain for human consumption in many countries (Jonsson et al., 1997). Several studies have been published. These publications cover different sampling techniques and detection instruments (Tuma et al., 1989; Börjesson et al., 1989, 1992; Jonsson et al., 1997).

Two main different approaches have been used in the past for the detection of fungal volatiles, the first one being the use of GC or GC–MS (which allows for identification of the released compounds), and more recently the use of the so-called electronic noses has been investigated. The later systems promised more rapid and simple performance, but they do not provide the identification of compounds; that's why they are coupled to multivariate analysis or artificial neural networks (ANN) algorithms to provide rapid estimation of the fungal condition of the samples. Electronic noses may be based on either gas sensors arrays or a mass spectrometer. Anyway, e-noses based on non-specific semiconductor sensors suffer from serious drawbacks such as poor sensibility, poor selectivity and long-term drift. That is why MS-based e-noses are becoming an increasingly used alternative to semiconductor gas sensor-based e-noses. This approach overcomes most of the difficulties present with the former approach, especially drift and selectivity problems. Moreover, measurements can be done in less than 5 min since no separation column is used. MS e-noses can be implemented easily in most quality control laboratories that already have GC/MS instruments (Vinaixa et al., 2004).

* Corresponding author. Tel.: +34 973 702555; fax: +34 973 702596.
E-mail address: smarin@tecal.udl.es (S. Marín).

0168-1605/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2006.11.003

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Börjesson et al. (1990) observed a weak correlation between volatile compounds and CO_2, but no correlation to CFU, when assaying *Penicillium* inoculated samples of oat grains and several cereal-based agar media. Moreover, when comparing ergosterol levels with off-odors as graded by humans, no correlation was found (Schnürer and Jonsson, 1992). The sampling system (a dynamic headspace approach) used consisted of a continuous air stream through cultivation vessels, so that concentration of volatiles occurred in an adsorbent material, and then the volatiles were desorbed and injected in a GC.

The use of a gas sensors e-nose has been shown to be useful for the prediction of ergosterol and CFU counts in wheat samples, with different levels of mould activity, by means of artificial neuronal networks (Jonsson et al., 1997). In that case, gas sensors were exposed for 60 s to the air pumped from the chamber where wheat samples were heated at 50 °C. Similarly, Börjesson et al. (1996) heated grain samples at 65 °C for 5 min, and sampling was done by pumping headspace air from the heating unit through the sensors chamber for 2 min.

Some authors have collected volatiles from fungal cultures by different systems (headspace solid-phase microextraction, diffusive sampling using Tenax tubes...) and on different substrates for chemotaxonomical purposes (Larsen and Frisvad, 1994; Wilkins et al., 2000). In those cases more complex volatile molecules were obtained as longer incubation periods were used. Moreover, volatile terpenes, which are the compounds mainly involved in the discrimination among fungal species, are produced under the more aerated conditions, their metabolism being partially inhibited by low oxygen/carbon dioxide ratios (Larsen and Frisvad, 1995; Larsen, 1997).

The objective of this work was to use a solid-phase microextraction (SPME) technique coupled to a MS-based electronic nose (SPME-MS-based nose) in order to predict mould growth (ergosterol content) of fungal contaminated samples of bakery products.

2. Material and methods

2.1. Samples

Strains of seven fungal species (*Aspergillus flavus*, AF, *Aspergillus niger*, AN, *Eurotium amstelodami*, EA, *Eurotium herbariorum*, EH, *Eurotium rubrum*, EU, *Eurotium repens*, ER, and *Penicillium corylophilum*, PE) were isolated from bakery products and one isolate of each species was used for the present study. Studies were performed on bakery product analogues prepared as described by Abellana et al. (1999), adjusted to a water activity of 0.95, and 8×8×20 mm pieces introduced in 20-ml headspace vials. Analogues were randomly needle inoculated in quadruplicate with 10^6 spores ml^{-1} suspensions of each of the seven cultures mentioned above and uninoculated analogues were used as control blanks. Once sealed, all vials were incubated at 25 °C until their measurement. Sampling was performed after 2, 4 and 7 days.

2.2. Solid-phase microextraction (SPME) for MS-based e-nose

A 75-μm Carboxen/polydimethylsiloxane fibre was used for extraction in all SPME sampling. Prior to that, the fibre was conditioned following the manufacturer's recommendations. In each measurement, the fibre was introduced through the septum into the vial and exposed to the headspace of fungal cultures for 20 min at room temperature. After that, the fibre was immediately transferred to the desorption step.

Thermal desorption of volatiles trapped on the fibre was conducted for 3 min in the chromatograph injection port (equipped with a 0.75-mm ID liner) at 300 °C. The split valve was closed during desorption. The fibre was always left for 5 additional minutes in the injector port to ensure complete cleaning.

A Shimadzu QP 5000 gas chromatograph–quadrupole mass spectrometer was used to implement a MS-based e-nose. The chromatographic column was replaced by a 5 m×0.25 mm×1 μm ID apolar pre-column that only acted as a transferline from the injector port to the mass detector. Temperature was set isothermal at 250 °C to co-elute all volatile components in one single peak. This implies that the components in the headspace of the vials passed directly to the mass detector without any chromatographic separation. For a given measurement, the resulting mass spectrum gave a fingerprint characteristic of the volatiles present in the headspace of the samples. Helium flow was set to 1.4 ml min^{-1}. The mass spectrometer operated in electron impact ionization mode (70 eV) and acquired in a scan range from m/z 35 to m/z 120 at 0.5 scan s^{-1}. Ion source temperature was set to 250 °C.

2.3. Ergosterol analysis

A modification of the method by Gourama and Bullerman (1995) was applied. Recovery rates were around 84% for the concentrations found in the study. 15 ml of 10% KOH in methanol was added to each vial and was shaken for 30 min in an orbital shaker at 220 rpm in a horizontal position. A 10-ml aliquot was transferred to a screw cap tube and placed in a hot water bath (55–60 °C) for 20 min. The tubes were then allowed to cool down to room temperature. 3 ml of water and 2 ml of hexane were added to the tubes, which were then agitated in a Vortex mixer for 1 min. After separation of layers, the upper layer (hexane) was transferred to a 10-ml vial. Hexane extraction was repeated twice using 2 ml each time. The extracts were combined and evaporated to dryness under a stream of nitrogen. The dry extracts were dissolved in 2 ml of methanol, and forced through 0.45-μm acetate filters. The HPLC equipment consisted of a Waters 515 isocratic pump (Waters Associated, Milford, MA), a Waters 717plus autoinjector, and a Waters Spherisorb ODS2 C18 column (4.6×250 mm). The Waters 2487 variable wavelength UV detector was set at 282 nm. The mobile phase was methanol at 1 ml min^{-1}. Ergosterol standard was purchased from Sigma (St. Louis, Mo).

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Table 1
Ergosterol content (ng/vial) of inoculated cake analogues along time

	48 h	96 h	168 h
A. flavus [a]	211±21	830±552	12,916±6595
A. niger [a]	<200	1121±627	12,975±4326
E. amstelodami [b]	346±125	1179±1049	7224±3095
E. herbariorum [a]	485±213	6488±294	15,933±1646
E. repens [a]	519±265	5446±1012	14,021±482
E. rubrum [a]	<200	2972±1647	13,047±1411
P. corylophilum [a]	<200	851±568	2206±1392
Control [b]	<200	<200	<200

[a] Mean±SD of 4 replicates.
[b] Mean±SD of 8 replicates.

2.4. Experimental planning

Since it was not possible to run the 32 sample vials in one day for the MS-based e-nose, experiments were carried out in two separate batches, the first one involving *E. amstelodami*, *E. repens*, *E rubrum*, *E. herbariorum* plus uninoculated controls, and the second set including *A. flavus*, *A. niger*, *P. corylophilum*, *E. amstelodami* (again) and uninoculated controls.

2.5. Multivariate analysis

Data generated by the e-nose device were collected and processed with a written-in-house software based on Matlab 6.1 (The MathWorks, Inc., Natick, Massachussets). For each measurement an unresolved single ion chromatographic peak was obtained and a response spectrum was generated by averaging mass spectra along the detected peak. Since measurements were performed in scan mode from *m/z* 35 to 120, the average intensity of each mass could be used as a variable. Therefore, each measurement was described by 85 variables and the response matrix had 85 columns and as many rows as experiments carried out. This data matrix was the *X*-matrix for further Partial Least Squares Regression (PLS) analysis, the *Y*-matrix being the ergosterol concentrations in the samples. This method performs particularly well when the various *X*-variables express common information, i.e. when there is a large amount of correlation, or even colinearity. PLS is a bilinear modeling method where information in the original *X*-data is projected onto a small number of underlying ('latent') variables called PLS components. The *Y*-data are actively used in estimating the 'latent'

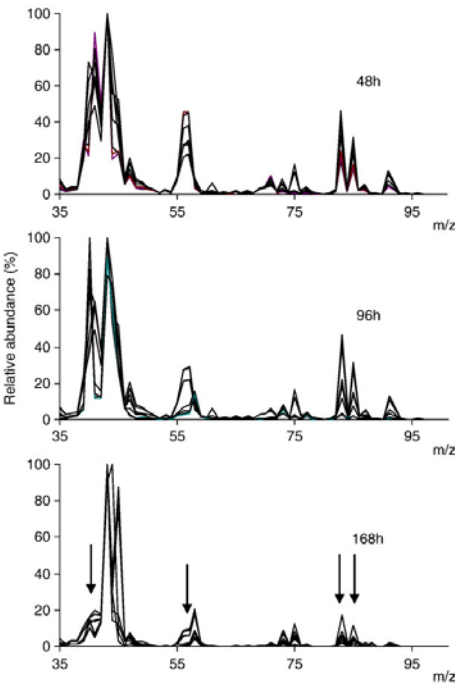

Fig. 1. Relative abundance of the *m/z* fragments obtained by SPME-MS-electronic nose from the uninoculated bakery products (controls) after incubation times of 48, 96 and 168 h. Different lines represent the different 8 replicates.

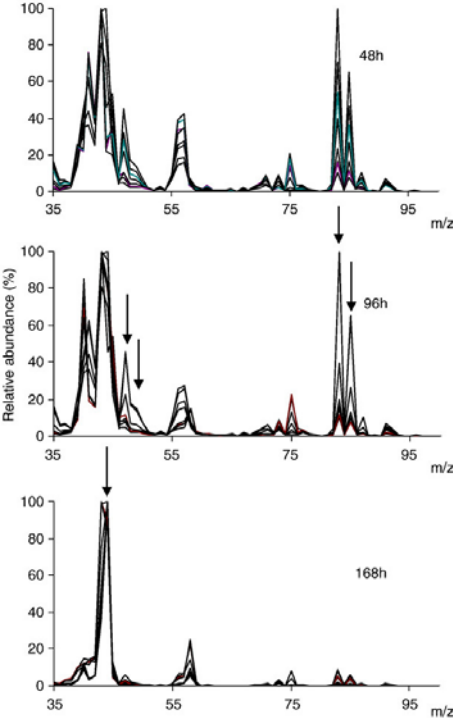

Fig. 2. Relative abundance of the *m/z* fragments obtained by SPME-MS-electronic nose from bakery products inoculated with *E. amstelodami* after incubation times of 48, 96 and 168 h. Different lines represent the different 8 replicates.

variables to ensure that the first components are those that are most relevant for predicting the *Y*-variables. Interpretation of the relationship between *X*-data and *Y*-data is then simplified, as this relationship is concentrated on the smallest possible number of components. PLS regressions were carried out by using The Unscrambler® version 7.6 (CAMO ASA, Norway).

3. Results

3.1. Ergosterol content of samples

Ergosterol content clearly increased with time for all fungal species tested (Table 1). After 48 h ergosterol levels were negligible for some species (*A. niger*, *E. rubrum*, *P. corylophilum*), while after 7 days values in the range 2200–15,900 ng were found. The highest levels of ergosterol were found for *E. herbariorum* and *E. repens*, regardless of the incubation time. Similar values were found for both *Aspergillus* species, while *P. corylophilum*, the slowest growing, showed low levels. No ergosterol was found in the control samples.

3.2. MS-based spectra of samples

First of all, mass spectra from volatiles obtained from the uninoculated bakery product samples gave a number of peaks that were also present in inoculated samples. The relative abundance of most of those peaks decreased as incubation time increased as shown in Fig. 1. In particular, *m*/*z* 39–41 decreased sharply at the last reading, as well as 56–57, 83 and 85.

After 48 h and 96 h the spectra of *A. flavus*, *A. niger*, *E. herbariorum*, *E. repens*, *E. rubrum*, *P. corylophilum* were similar to those of control samples, while *E. amstelodami* presented some differences, referring mainly to peaks 47–49, 83 and 85, whose abundance increased for some of the replicates (Fig. 2).

After 168 h, the general response was a clear predominance of the abundance of *m*/*z*=43–45, while all other peaks in the spectra dramatically decreased or disappeared — it must be taken into account that the response recorded (Figs. 1 and 2) is relative abundance of *m*/*z* fragments, thus a high increase in peaks 43–45 probably made the remaining ones disappear in the relative scale.

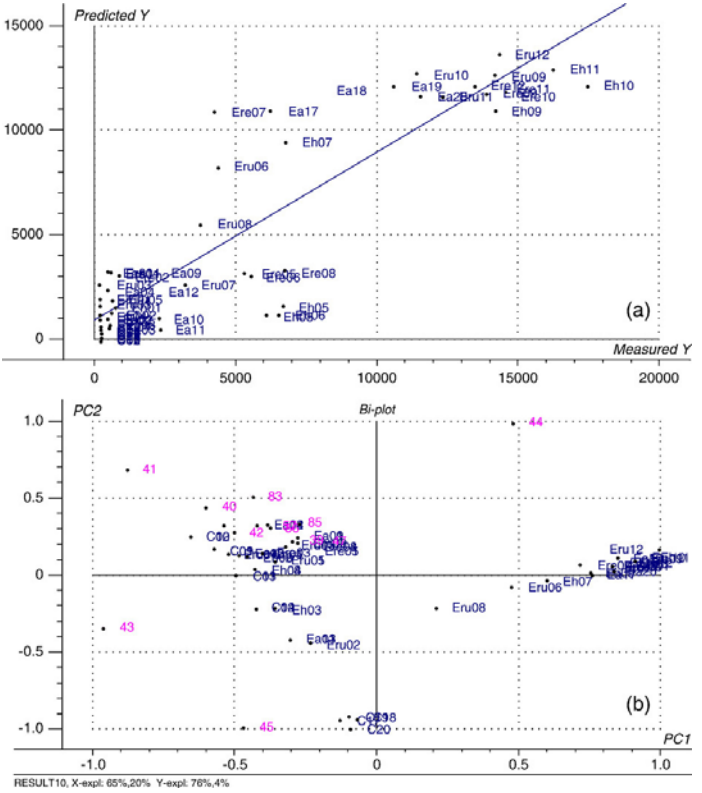

Fig. 3. PLS model for prediction of ergosterol content in the first set of samples based on MS-electronic nose spectra; a) measured ergosterol levels versus predicted values by the validated model; b) bi-plot showing the projections of the samples (coded C, Ea, Eh, Ere, Eru) and MS-variables (coded by the molecular sizes of *m*/*z* fragments) on PC1 and PC2.

3.3. Correlation among MS and ergosterol content

As there were slight deviations in the results for control and *E. amstelodami* samples, both sets of experiments were analyzed separately. For the first one, involving *E. amstelodami*, *E. herbariorum*, *E. repens*, *E. rubrum* and controls, the PLS analysis revealed a good correlation among MS measurements and ergosterol content in the samples (*R* after cross validation = 0.88, Fig. 3a). The model, after excluding the MS-variables with negligible regression coefficients, had 2 PC explaining 76 and 4% of the total variance, respectively. A bi-plot of both variables and samples (Fig. 3b) shows that the most important *m/z* variable in the model was the *m/z* 44 which had a positive projection in the PC1 axis as well as most of the 168 h samples plus some *Eurotium* inoculated samples after 96 h, suggesting that the abundance of this *m/z* fragment increased with time and greatly contributed to the explained variance by the model. The other significant *m/z* variables (43, 45, 41, 42, 40...) were negatively correlated to ergosterol content — again, it must be taken into account that the response recorded is relative abundance of *m/z* fragments, thus a high increase in peak 44 decreased the remaining ones in the relative scale, although they probably increased in absolute values.

For the second one, involving *E. amstelodami*, *A. flavus*, *A. niger*, *P. corylophilum* and controls, the PLS analysis revealed also a good correlation among MS measurements and ergosterol content in the samples (*R* after cross validation = 0.93, Fig. 4a). The model, after excluding the MS-variables with negligible regression coefficients, had 4 PC explaining 55, 26, 8 and 2% of the total variance, respectively. A bi-plot of both variables and samples (Fig. 4b) shows that the most important *m/z* variable in the model was the *m/z* 44 which had a positive correlation coefficient, and also a high positive projection in the PC1 axis, as well as the 168 h samples. The other significant *m/z* variables (40, 43, 58...) were negatively correlated to ergosterol content.

For the individual species, the correlation obtained by regression of ergosterol content versus e-nose readings was very good, except for *P. corylophilum* which showed a poorer

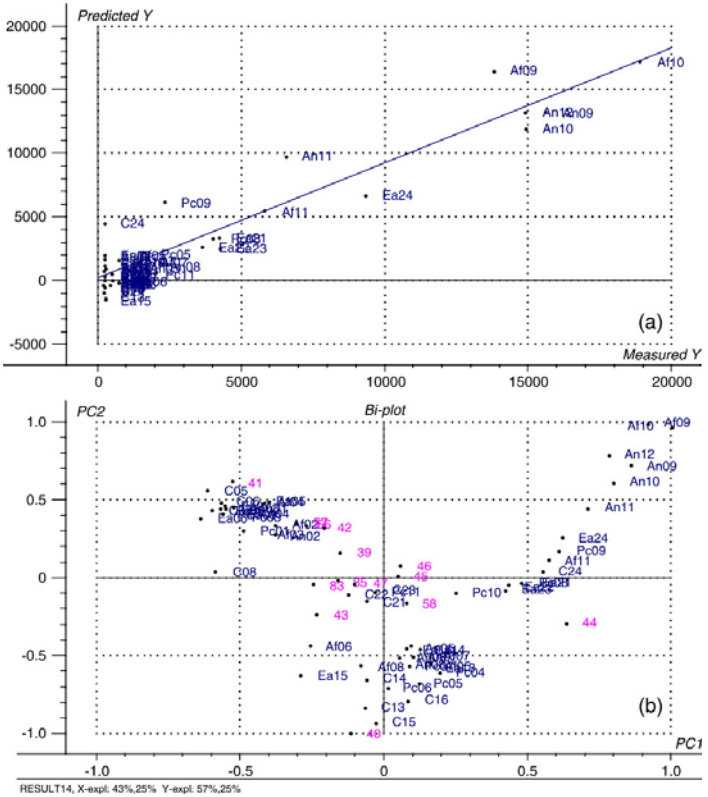

Fig. 4. PLS model for prediction of ergosterol content in the second set of samples based on MS-electronic nose spectra; a) measured ergosterol levels versus predicted values by the validated model; b) bi-plot showing the projections of the samples (coded C, Ea, Af, An, Pc) and MS-variables (coded by the molecular sizes of *m/z* fragments) on PC1 and PC2.

Table 2
PLS regression PC and correlation coefficients obtained separately for the different species

	Y-explained variance PC1 (%)	*Y*-explained variance PC2 (%)	*Y*-explained variance PC3 (%)	Calibration *R*	Validation *R*
A. flavus	79	17	2	0.99	0.95
A. niger	79	12	–	0.95	0.91
E. amstelodami	72	16	–	0.94	0.91
E. herbariorum	83	–	–	0.91	0.87
E. repens	87	–	–	0.93	0.90
E. rubrum	90	5	–	0.98	0.96
P. corylophilum	48	13	–	0.78	0.46

correlation (Table 2). Correlation coefficients ranged from 0.91 to 0.99, those of *A. flavus* and *E. amstelodami* being the highest.

4. Discussion

Different studies have been carried out to test the suitability of analysis of volatiles as a method for fungal detection. The choice of a reference method is an important issue. Most authors used CFU, ergosterol content or CO_2 released as reference methods, although ergosterol content has been shown to be a better estimator of fungal biomass of foodborne fungi than CFU (Marín et al., 2005). For this reason, in this study, ergosterol contents in the cultures were chosen to be compared with the production of volatile fungal metabolites.

Growth of the same isolates has been reported before in bakery products analogues of 0.90 a_w in terms of colony radius (Guynot et al., 2003) showing a similar growth trend: *E. repens* and *E. herbariorum* had the biggest colonies among the *Eurotium* species while *P corylophilum* led to very small colonies compared to the rest. The only difference is that *E. amstelodami* has similar or lower ergosterol levels in the present work than *Aspergillus* species, while it presented bigger colonies in the previous work; the reason is probably the difference in a_w (from 0.90 to 0.95).

Solid-phase microextraction was used as the preferred volatiles collection method; it has been pointed out as a key technique for volatile analysis. Once sampling equilibrium is established, the mass of analyte in the fibre is directly proportional to the concentration in the air. The volatile profiles change over time and differ somewhat among fungal strains: the alcohols, aldehydes, and esters always appear first and are present for each strain. Production of the phenolics lags by several days, and in some strains these compounds are barely detectable. The use of solid-phase microextraction greatly facilitates the analysis of the volatiles of lower molecular weight (Bartelt and Wicklow, 1999). Volatiles can be analyzed using different collection methods other than SPME. The simplest is to take samples with a gas-tight syringe directly from the headspace gas above a sample; alternatively, volatiles can be collected on various porous polymers (e.g. Tenax GC or Chromosorb 102). Moreover, the volatiles can be absorbed using either diffusive sampling or active purging and trapping of headspace gases (Schnürer et al., 1999).

A simple sampling method was aimed to be reproduced in the industries for monitoring of samples of the different batches of bakery products before they are sent for distribution. It should enable detection of fungal spoilage in general; no discrimination among fungal species was intended. However, this same methodology, also applied to bakery products, has been shown to be useful to discriminate among fungal genera using ANN with a 78–88% success rate after 2 days of incubation (Vinaixa et al., 2004). Vials were sealed at the beginning of the incubation period; although they were 3-g samples enclosed in 20-ml vials and fungal growth was partially restricted by a a_w of 0.95, an increased carbon dioxide/oxygen rate may be expected. Fungi emit mainly small alcohols, ketones and esters when grown in a Petri dish, while the terpenes are dominant when the same fungus is grown in a flow-through equipment. These very different results might be explained by the influence on the metabolism of the oxygen/carbon dioxide ratio or the specific flow conditions (Nilsson et al., 1996).

Sporulation could only be observed in some samples at the 7th day. Sporulation seems to coincide with an increase in the production of several volatile compounds of which some are odors, indicating that odor intensity should be positively related to the number of colony forming units. This implies that odors may not be good indicators of fungal growth measured as fungal biomass since correlations between CFU and fungal biomass may be weak (Schnürer, 1993).

Good estimates of ergosterol in the spoilt samples could be achieved by the use of the MS-electronic nose. Other authors have observed good correlation between measured CFU and ergosterol and values predicted with a trained ANN, indicating a relation between desorbed volatiles and quantitative measures of microbial growth in grains (Jonsson et al., 1997). The ergosterol contents of *P. aurantiogriseum* cultures corresponded reasonably well to fungal metabolic activity, measured as the production of volatile metabolites. At early stages of fungal growth, such as in this study, the correlation between fungal activity and ergosterol levels was high. The production of volatile metabolites on cereal grain could be detected before any visible signs of fungal growth appeared. No correlation between volatile metabolites and CFU was found, but a simultaneous increase in the levels of CO_2 and volatile metabolites was observed (Börjesson et al., 1990). Significant correlation was found between volatile metabolites and ergosterol content in cultures of *Aspergillus* and *Penicillium* species grown on wheat and oats. It was concluded that 3-methylfuran could be used as an indicator of mould growth in cereals as it was produced by the 6 fungal species tested (Börjesson et al., 1992).

The predominant m/z fragments observed in this study may correspond to CO_2 (m/z=44) plus other small fragments

frequently found in the mass spectra of low molecular mass compounds commonly produced by fungi in their early growth. Some of the most frequently reported volatile metabolites produced during fungal growth on grain or grain products include 2-methyl-1-propanol, 3-methyl-1-butanol, 3-octanone, and 1-octen-3-ol (Kaminski et al., 1974; Tuma et al., 1989; Börjesson et al., 1992). Börjesson et al. (1993) found these compounds to be produced by almost all the examined fungi (*Aspergillus* and *Penicillium* species) after 2–5 days incubation, although in different quantities. Thus, an instrumental analysis of this group of compounds should give a more reliable measure of fungal growth in cereals than would an odor assessment (Börjesson et al., 1993). A significant increase in the CO_2 content was observed in the inoculated containers after two days irrespective of the species inoculated. The production of CO_2 increased during the five days after inoculation, after which it remained constant. *A. amstelodami* produced 2-methylfuran, 2-methyl-1-propanol, 2-pentanone, 2-methyl-1-butanol, 3-octen-2-ol and 1-octen-3-ol as result of its growth on wheat, while *A. flavus* produced 2-methylfuran, 2-methyl-1-propanol and 3-methyl-1-butanol. 2-methyl-1-propanol was common to all species tested. 3-methyl-1-butanol was found to be associated with early stages of fungal growth. The change to anaerobic metabolism, may have led to an increased production of 3-methyl-1-butanol (Börjesson et al., 1989). Detection of mould growth as early as possible requires the measurement of metabolites produced at an early stage of growth. Production of 3-methyl-1-butanol reached its maximum earlier than other compounds during growth of *Penicillium cyclopium* and *A. flavus* on wheat (Börjesson et al., 1989).

To sum up, the SPME-MS-based electronic nose has been shown to be useful for prediction of spoilage of bakery products in less than 7 days. This period of time would enable sampling of this type of products (shelf life of 5–8 weeks) and rejection of the selected batches before their distribution to markets. In addition, this approach overcomes many of the difficulties encountered on traditional electronic noses, such as repetitivity, selectivity and drift issues. The instrument offers a fast alternative that can be easily automated and operated. Moreover, since many quality control departments already have GC–MS equipment, they can convert their units into a MS-based e-nose in a rather straight forward manner by just coupling the optimal sampling system and using additional pattern recognition software.

Acknowledgements

This work was supported by the MCYT (Ministerio de Ciencia y Tecnología) through the INIA (CAL-01-037-C3), AGL2004-06413/ALI and the Ramón y Cajal program, by the EC, Quality of Life Programme (QoL), KA1 on Food, Nutrition and Health (QLK1-CT-2000-01763) and by the research network 'Red de centros en metabolismo y nutrición C03/08'.

References

Abellana, M., Benedí, J., Sanchis, V., Ramos, A.J., 1999. Water activity and temperature effects on germination and growth of *Eurotium amstelodami*, *E. chevalieri* and *E. herbariorum* isolates from bakery products. Journal of Applied Microbiology 87, 371–380.

Bartelt, R.J., Wicklow, D.T., 1999. Volatiles from *Fusarium verticillioides* (Sacc.) Nirenb. and their attractiveness to nitidulid beetles. Journal of Agricultural and Food Chemistry 47, 2447–2454.

Börjesson, T., Stöllman, U., Adamek, P., Kaspersson, A., 1989. Analysis of volatile compounds for detection of molds in stored cereals. Cereal Chemistry 66, 300–304.

Börjesson, T., Stöllman, U., Schnürer, J., 1990. Volatile metabolites and other indicators of *Penicillium aurantiogriseum* growth on different substrates. Applied and Environmental Microbiology 56, 3705–3710.

Börjesson, T., Stöllman, U., Schnürer, J., 1992. Volatile metabolites produced by six fungal species compared to other indicators of fungal growth on cereal grains. Applied and Environmental Microbiology 58, 2599–2605.

Börjesson, T., Stöllman, U., Schnürer, J., 1993. Off-odours compounds produced by molds on oatmeal agar: identification and relation to other growth characteristics. Journal of Agricultural and Food Chemistry 41, 2104–2111.

Börjesson, T., Eklöv, T., Jonsson, A., Sundgren, H., Schnürer, J., 1996. Electronic nose for odor classification of grains. Cereal Chemistry 73, 457–461.

Gourama, H., Bullerman, L.B., 1995. Relationship between aflatoxin production and mold growth as measured by ergosterol and plate count. Food Science and Technology 28, 185–189.

Guynot, M.E., Marín, S., Sanchis, V., Ramos, A.J., 2003. Modified atmosphere packaging for prevention of mold spoilage of bakery products with different pH and water activity levels. Journal of Food Protection 66, 1864–1872.

Jonsson, A., Winquist, F., Schnürer, J., Sundgren, H., Lundström, I., 1997. Electronic nose for microbial quality classification of grains. International Journal of Food Microbiology 35, 187–193.

Kaminski, E., Stawicki, S., Wasowicz, E., 1974. Volatile flavor compounds produced by molds of *Aspergillus*, *Penicillium* and Fungi imperfecti. Applied and Environmental Microbiology 27, 1001–1004.

Larsen, T.O., 1997. Identification of cheese-associated fungi using selected ion monitoring of volatile terpenes. Letters in Applied Microbiology 24, 463–466.

Larsen, T.O., Frisvad, J.C., 1994. A simple method for collection of volatile metabolites from fungi based on diffusive sampling from Petri dishes. Journal of Microbiological Methods 19, 297–305.

Larsen, T.O., Frisvad, J.C., 1995. Comparison of different methods for collection of volatile chemical markers from fungi. Journal of Microbiological Methods 24, 135–144.

Marín, S., Ramos, A.J., Sanchis, V., 2005. Comparison of methods for estimating mould biomass of foodborne fungi. International Journal of Food Microbiology 99, 329–341.

Nilsson, T., Larsen, T.O., Montanarella, L., Madsen, J.O., 1996. Application of head-space solid-phase microextraction for the analysis of volatile metabolites emitted by *Penicillium* species. Journal of Microbiological Methods 25, 245–255.

Schnürer, J., 1993. Comparison of methods for estimating the biomass of three food-borne fungi differing in growth morphology. Applied and Environmental Microbiology 59, 552–555.

Schnürer, J., Jonsson, A., 1992. Ergosterol levels and mould colony forming units in Swedish grains of food and feed grade. Acta Agriculturae Scandinavica Section b-Soil and Plant Science 42, 240–245.

Schnürer, J., Olsson, J., Börjesson, T., 1999. Fungal volatiles as indicators of food and feeds spoilage. Fungal Genetics and Biology 27, 209–217.

Tuma, D., Sinha, R.N., Muir, W.E., Abramson, D., 1989. Odor volatiles associated with microflora in damp ventilated and non-ventilated bin-stored bulk wheat. International Journal of Food Microbiology 8, 103–119.

Vinaixa, M., Marín, S., Brezmes, J., Llobet, E., Vilanova, X., Correig, X., Ramos, A.J., Sanchis, V., 2004. Early detection of fungal growth in bakery products using an e-nose based on mass spectrometry. Journal of Agricultural and Food Chemistry 52, 6068–6074.

Wilkins, K., Larsen, K., Simkus, M., 2000. Volatile metabolites from mold growth on building materials and synthetic media. Chemosphere 41, 437–446.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper VI

Vinaixa, M.; Brezmes, J.; Llobet, E.; Vilanova, X.; C.Burian; Correig, X., Saint Petersburg-Russia, 1-3 May **2007**; *Proceedings of the 12th International Symposim on Olfaction and 'electronic nose'.*

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Two or three-way data analysis in MS-Sensor devices. Which is the optimal approach?

M. Vinaixa*, J. Brezmes, E. Llobet, X. Vilanova, C.Burian, X. Correig
MINOS, DEEEA, Universitat Rovira i Virgili (URV)
Avda. Països Catalans, 26, 43007 Tarragona, *maria.vinaixa@urv.cat

Abstract

The goal of this paper is to compare the performance of two-way versus three way classifier models for their application in pattern recognition of MS-Sensor devices. Three classifier models, namely bilinear PLS-DA, trilinear N-PLS-DA and PARAFAC-MLR-DA (multilinear regression discrimination using loads for first mode as X-block predictors) are applied and compared in the framework of a classical MS-Sensor application such as olive oil discrimination.

Introduction

In recent yeats it has been shown that the rapid analysis of the global volatile fraction of foods by mass spectrometry without chromatographic separation produces signals (signatures) that contain significant and useful information. The so called MS-Sensor approach has been applied to compare the volatile fingerprint profile of samples and classify these samples according to their respective signatures. The main advantages rely on the fact that sample treatment is eliminated or reduced to a minimum and the fact that since chromatographic separation is avoided, near-real time results can be obtained. In Spain, olive oil constitutes an important economic activity and much attention has been devoted to the study of such a product in recent years. Classical techniques used to assess its quality are usually time consuming and present a low throughput. MS-Sensor devices represent a novel oportunity to assess the quality of olive oils in a rapid, solvent free and easy way. Common pattern recognition algorithms applied on MS-Sensor data make use of the averaged mass spectra along the detected peak. Nevertheless, considering this averaged mass spectrum may lead to losing useful temporal information. Even when chromatographic resolution is avoided, a sort of diffusion is observed on the isothermal peak. This fact allows us to consider the possibility of computing this extra-information by considering the three way nature of the data using of trilinear algorithms such N-PLS or PARAFAC. Multi-way methods are particularly useful for treating data with more than two sources of variability like the data generated by MS-Sensor devices, where the response in ion counts arriving at the detector is measured as a function of time and mass/charge ratio. In fact, data provided from a MS-Sensor should be arranged as multi-way array where the first mode represents samples, the second one corresponds to mass spectra and the third to the elution profiles. Exploiting differences in the time response of the analytes can enhance the subtle variations in the spectra and therefore classification performance may be improved. The main goal of this paper is to see whether classifier algorithms may benefitiate from the use of second-order methods, even in the case of crude or poor chromatographic separation such direct MS-Sensor devices. To date, the application of second-order methods to classifier models in MS-Sensor devices has not been reported in the literature.

Experimental

Five different virgin olive oils were received directly from producers in 100 ml transparent glass bottles airtight sealed. They were kept frozen before any analysis

was performed. The day before the analysis, all the samples were kept out of the freezer and were exposed at room temperature in darkness. Equal amount of samples (5 g) were placed in 20-ml glass vials that were immediately sealed with silicon septum. Six different aliquots were pipetted from each of the five oils giving a total of 30 vials to be analysed. SPME extraction of volatiles was performed by introducing a 50/30 µm DVB/Carboxen/PDMS (Supelco, Bellafonte, PA.) fibre into the vial and exposing them to the headspace of oil for 20 minutes. Afterwards, thermal desorption of the volatiles trapped on the fibre was conducted for 3 minutes in the chromatograph injection port at 270ºC. Volatiles trapped on the fibre were delivered to a 5- m deactivated fused silica column. The column was kept isothermal at 250ºC and the helium flow was set to 1.4 ml/min. The split valve was closed during desorption. The quadrupole mass spectrometer acquired in scan mode, and the mass range used was m/z 35 to m/z 200 at 0.5 scan/sec. The fibre was left 5 additional minutes in the injector port to ensure its complete cleaning. MS-Sensor data was imported into ASCII format and finally loaded into MATLAB version 6.5 (Mathworks, Inc., Natick, MA) for further data processing. The MATLAB routines used for PLS-DA, NPLS-DA and PARAFAC-mlr-DA were adapted from the PLS_Toolbox, version 3.5 (Eigenvector Technologies, Inc.,Manson, WA).

Results and Discussion

Raw data provided by the MS-Sensor was arranged as a three-way array data set $\underline{\mathbf{R}}$ (30×166×205). A modification of the previous reported RAFFT aligment algorithm was applied to $\underline{\mathbf{R}}_1$ (30×166×205) before any modelling in order to overcome retention time shifts from run to run. This modification consisted in the application of this algorithm to each m/z channel instead of doing so to the reconstructed TIC signal. Finally a two-way response matrix $\mathbf{R}_2$ (30×166) was obtained by averaging m/z values along the 205 scans considered. Exploratory data analysis was conduced using PARAFAC and PCA in order to observe the main trend on the data sets. Figure 1 shows PARAFAC's loads plot for samples modelled using six components. The same trend is observed in both PARAFAC and PCA models. The five types of oils cluster well appart. Separation according origins is achieved in both cases and it seems that a gradation of scores indicating oil quality can also be deduced. Component 2 allows to ditinguish between geographical origins.

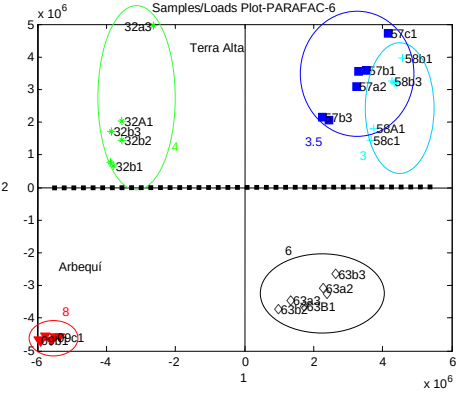

Figure 1. 6 component PARAFAC loads plot

Finally $\underline{\mathbf{R}}_1$ (30×166×205) and response matrices $\mathbf{R}_2$ (30×166) were split up into two parts: a training set and a test set. In the training phase a 5-category classification was envisaged according to each type of olive oil using either two-way PLS-DA or three way related methods as PARAFAC-mlr-DA or N-PLS-DA and four of the six measures in each category were used. The remaining two vials unseen by the model were used as a test set for further validation. Prior to any calculation, data were centered across the first mode for the three-way array and meancentered in the case of two-

way matrix. The training phase of the models was evaluated using a cross-validation method, and RMSECV was used to assess the optimal number of factors for model fitting. Once the model had been trained, predicted category for test set samples was attempted. The predicted y-value from the calibrated models results in a continuous variable that can be interpreted as a class similarity index. Each calculated class prediction value can be compared with a Bayesian distribution curve to determine for a given predicted y-value the probability that this value belongs to that original class. Furthermore a threshold of "predicted y" is determined above which a sample is considered to be a member of the class. These thresholds are also calculated for the three models and class assignment is done comparing against this value. Figure 2 shows class predicted values for the test set samples using N-PLS-DA model.

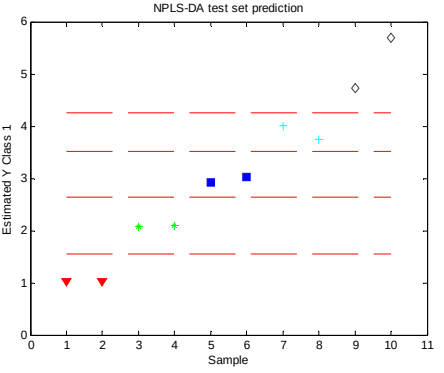

Figure 2. N-PLS-DA class predicted

The dashed red horizontal lines represents the calculated thresholds. Regression between actual and predicted Y-values were performed for the three models to better assess their accuracy. The correlation coefficient, slope and intercept were calculated. Table 1 summarizes the characteristics of the N-PLS-DA, PLS-DA and PARAFAC-MLR-DA models calibrated for the classification of the olive oil samples.

	PLS-DA	N-PLS-DA	PARAFAC-MLR-DA
#LV's	11	11	6
%Succees rate test-set	100	100	100
RMSECV	0,14	0,17	0,23
RMSEP	0,19	0,25	0,23
X val	99,98	99,53	100
Y val	98,24	99,42	99,19
b (test)	-0,05	-0,01	0,14
m (test)	1,03	1,01	0,97
R2 (test)	0,99	0,97	0,97

Table 1. models' characteristics

Several conclusions can be drawn from this table. Regardless of which classifier method employed, all of them achieved a 100% success rate in classification of test samples. Concerning the number of components needed to fit each model, PARAFAC-MLR-DA is pointed to be the simplest model and consequently the most parsimonious one. Nevertheless, taking a look to RMSECV and RMSEP, we can conclude that despite the fact that all the models are able to classify new samples in a correct way, PLS-DA seems to outperform three-way methods. Therefore, for this particular application, the introduction of the third dimension does not allow to improve the performance of the MS-sensor device. Anyway, the parsimonity observed in the PARAFAC model and the fact that the performance remains quite acceptable leads us to consider that this model is more robust and may outperform classical approaches in a broader set of applications.

References

[1] J. W. H. Wong, C. Durante and H. M. Cartwright, *Analytical Chemistry* **2005**, *77*, 5655-5661.
[2] W. P. Gardner, R. E. Shaffer, J. E. Girard and J. H. Callahan, *Analytical Chemistry* **2001**, *73*, 596-605.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper VII

El-Barbri, N.; Amari, A.; **Vinaixa, M**.; Llobet, E.; Bouchikhi, B.; Correig, X., Building of a metal oxide gas sensor-based 'electronic nose' to assess the freshness of sardines under cold storage, *Sensors and Actuators B*, **2007**, 128 235–244.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Available online at www.sciencedirect.com
ScienceDirect
Sensors and Actuators B 128 (2007) 235–244
SENSORS AND ACTUATORS B CHEMICAL
www.elsevier.com/locate/snb

Building of a metal oxide gas sensor-based electronic nose to assess the freshness of sardines under cold storage

N. El Barbri [a], A. Amari [a], M. Vinaixa [b], B. Bouchikhi [a], X. Correig [b], E. Llobet [b,*]

[a] *Sensor Electronic & Instrumentation Group, Faculty of Sciences, Physics Department, Moulay Ismaïl University, B.P. 11201, Zitoune, Meekness, Morocco*
[b] *MINOS, Microsystems and Nanotechnologies for Chemical Analysis, Universitat Rovira i Virgili, Avda. Països Catalans 26, 43007 Tarragona, Spain*

Received 27 September 2006; received in revised form 7 June 2007; accepted 8 June 2007
Available online 15 June 2007

Abstract

We report on the building of a simple and reproducible electronic nose based on commercially available metal oxide gas sensors aimed at monitoring the freshness of sardines stored at 4 °C. Sample delivery is based on the dynamic headspace method and four features are extracted from the transient response of each sensor. By using an unsupervised method, namely principal component analysis (PCA), we found that sardine samples could be grouped into three categories (fresh, medium and aged), which corresponded to an increasing number of days that sardines had spent under cold storage. Then, supervised linear or non-linear pattern recognition methods (PARC) such as discriminant factor analysis (DFA) or fuzzy ARTMAP neural networks (FANN) were successfully applied to build classification models to sort sardine samples according to these three states of freshness. The success rate in classification was 96.88% for the neural network classifier. Additionally, 10 volatile species that indicated the evolution of sardines with the number of days of cold storage were identified by SPME/MS/GC.

Keywords: Fish freshness; Electronic nose; Metal oxides; Pre-processing; Pattern recognition

1. Introduction

Freshness, defined as the number of storage days at a certain temperature [1], is the single most important attribute when assessing fish quality. Microbiological, biochemical, sensory changes and electronic noses are associated with deterioration of fish quality during handling and storage [2,3]. Although a variety of biochemical, physical, microbiological methods and sensory evaluation [4] have been used to assess fish freshness, an electronic nose is still the most promising method for achieving such a goal [5,6]. The methods for fish freshness control using electronic noses are based on the headspace sampling method [5] that requires injection of a volatile phase above fish into a chamber containing the sensor array. Then, this procedure requires two chambers, one for fish samples and another for gas sensors.

Chemical investigations using gas chromatographic techniques have shown that there are a lot of volatile gases emitted from degrading fish, which give rise to the overall odour of fish as a combined action [7]. The concentration of some of these volatiles increases with time as the fish spoils; indeed some of these are often used as indicators of spoilage [7]. Due to the high number of volatile compounds involved in the process, and to the fact that they also dynamically change, the measure of fish freshness over a long period of storage can be achieved with a multicomponent approach. This is a typical electronic nose application, where a number of non-selective and partially cross-correlated sensors are used to get a qualitative analysis of samples [1]. The main advantage of this approach is that the technique is non-destructive, since it involves sampling the headspace of fish [5].

Nowadays, electronic noses are being developed for many application fields like food technology, pharmaceutical industry, medical, etc. Such systems are devoted to replace analytical instruments, which are bulky and very expensive. These new intelligent gas sensor systems are mainly composed of a sensor array coupled with powerful statistical or neural network-based pattern recognition methods [8]. Electronic nose systems do not give any information about the identity of compounds causing an aroma nor about their concentration. However, with the aid

* Corresponding author. Tel.: +34 977558502; fax: +34 977559605.
E-mail address: eduard.llobet@urv.cat (E. Llobet).

0925-4005/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2007.06.007

of appropriate mathematical techniques, the electronic noses are capable of recognising the volatile pattern from a particular sample and distinguishing it from other samples [9]. Metal oxide semiconductor (MOS) sensors represent the most widely used technology in electronic nose applications. MOS sensors behave as chemo-resistors and their response signal is often measured as the relative change in their resistance when exposed to complex aromas.

In this paper, we report on the building and characterisation of an electronic nose system based on MOS gas sensors. There are two main objectives for this e-nose tool. The first one is making the system able to sort sardines that have undergone cold storage in at least two categories, namely acceptable or non-acceptable by consumers. The second one is to keep the cost of the system to a minimum (e.g., by using commercially available sensors and by defining a simple and cost-effective sample delivery method that leads to reproducible results). Reduced cost and simplicity are key factors for a widespread use of the e-nose in the logistics chain of sardines at both sides of the Mediterranean sea. We present results on the performance of the electronic nose at identifying the state of freshness of sardines, employing a combination of different pattern recognition techniques such as PCA, DFA and neural networks. Additionally, the results obtained with the gas sensor-based electronic nose are compared to reference methods such as SPME/GC/MS or HS/MS in the framework of fish freshness assessment.

2. Experimental

2.1. Sample preparation

The sardines selected for testing the system were *Sardinia pilchardus*. The different samples (mass of 40 ± 1 g) used in the experiments were obtained from the same local fish market (Meekness, Morocco) in October 2005 for the first set of experiments and in December 2005 for the second set of experiments. The sardines used for our experiments were initially as fresh as possible; not more than 3 h after being caught. The samples were immediately placed in plastic bags (bags for freezing food) (one fish per bag) and introduced in a refrigerator and kept at a constant temperature of 4 ± 1 °C. No cleaning or other sample manipulations were performed. Samples were analysed at Days 1 (i.e., no storage), 3, 5, 7, 9, 11, 13 and 15. For each measurement, a whole sardine was taken from the refrigerator, weighed and placed in the sampling vessel of the electronic nose.

2.2. Electronic nose set-up

Spoilage odours that develop as a result of microbial growth and oxidation leading to the degradation of sardines are sensed by a simple and cheaper electronic nose prototype, developed in our laboratories. The experimental system is mainly composed of four parts: sensor array, sensor cell, sampling vessel and measurement rig, and data acquisition system.

The sensor array comprises six metal oxide gas sensors Figaro TGS 8XX (with XX = 23, 25, 26, 31, 32 and 82), a temperature sensor (National Semiconductors LM35DZ), and a humidity sensor (Phylips H1). Following the manufacturer's recommendations, every TGS sensor was heated by applying 5 V to its heating resistor and its conductance was measured in a half-bridge configuration using 10 V as a supply voltage. The TGS 8XX sensors group have been used by several authors in fish freshness monitoring [5,10].

The sensor cell comprises an insulated cylindrical vessel made of plastic and a printed circuit of circular shape on which a sensor array is placed together with the sensor conditioning electronics. The temperature inside this cell is kept constant at 33 °C using a heating coil and a fan. In addition, we have placed in the sensor cell a very simple cooling system based on the principle of the water circulation in a copper tube. The cell has been especially designed to provide all sensors with the same experimental conditions (e.g., the same gas flow, ambient temperature, etc.).

For each measurement and without any additional manipulation, sardine samples were placed into standard glass sampling vessels of 1 l volume and sealed with septum caps. Each sample was kept in the sealed sampling vessel for 40 min at 25 ± 2 °C, for the headspace to develop. Each time that a new set of sardines was analysed, new glass vessels were employed. After waiting for 40 min, the vessel was connected to the e-nose system (valves and airtight tubing were used for this purpose). Pure nitrogen at a constant flow of 500 sccm was employed to allow the volatile species from the dynamic headspace to reach the sensor cell (the measurement procedure is described below). Fig. 1 shows the experimental set-up.

Sensor conductance was acquired and then digitised using a data acquisition board (PCL 812PG, Advantech). A sampling rate of 1 sample/s was used. A program in visual C^{++} 6.0 was developed to control the data acquisition. Additionally, this program was in charge of keeping constant the temperature inside the sensor cell via the temperature sensor and some digital outputs of the data acquisition card, which actuated the heating coil or the cooling system and the fan when needed.

2.3. Measurement protocol

Every day, the sensor array was powered 1 h before the measurement process started. This allowed the sensitive layers to stabilise and so to improve measurement reproducibility [11]. The response of the sensors in a flow of pure nitrogen was acquired during 50 min. This was essential because it allowed for the gas sensors to reach a stable and reproducible resistance, which was considered their baseline resistance (or their initial state) [12]. Each measurement comprised two phases. In the first phase that lasted for 20 min, the flow of pure nitrogen was interrupted. Finally, in the second phase, the response of the sensors when a flow of nitrogen carried the volatiles from the headspace of the sardine sample was acquired for 50 min. At the end of phase 2, a new measurement process could be initiated immediately by restarting at phase 1. The duration of phase 2 could be reduced down to 20 min by increasing the flow of nitrogen to 2500 sccm.

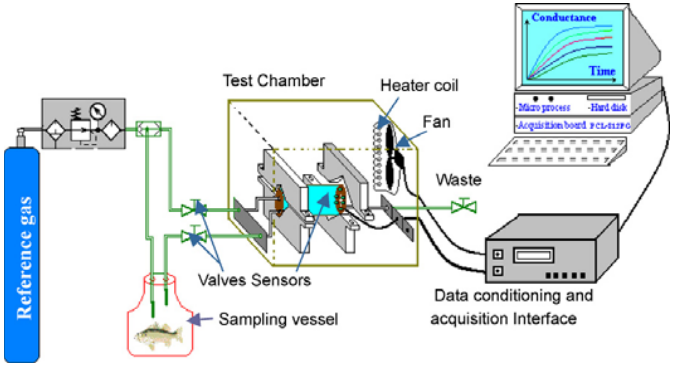

Fig. 1. Schematic representation of the experimental set-up used in the measurements.

2.4. Feature extraction and data pre-processing

For every sensor within the array and measurement performed, four representative features from the response signal were extracted. These were:

- G_0: the initial conductance of a sensor calculated as the average value of its conductance during the first 15 min of a measurement.
- G_s: the steady-state conductance calculated as the average value of its conductance during the last 5 min of a measurement.
- dG/dt: the dynamic slope of the conductance calculated between 15 and 35 min of a measurement. This corresponds to a phase where a fast increase of sensor conductance is observed.
- A: the area below the conductance curve in a time interval defined between 15 and 40 min of a measurement. This area is estimated by the trapeze method.

Following the procedures described above, 64 measurements were performed, which corresponded to eight replicate measurements at Days of storage 1, 3, 5, 7, 9, 11, 13 and 15. Therefore, the original data matrix had 64 rows (i.e., measurements) and 24 columns (i.e., 6 sensors × 4 response features/sensor).

The dataset was pre-processed using standard procedures such as mean centring, standardisation or matrix normalisation, depending on the different pattern recognition methods employed. Extracting the four aforementioned features from each sensor response and pre-processing the resulting data matrix were an automated process via a written-in-house MATLAB 6.5 program. Once pre-processed, the data were then exploited by different pattern recognition methods such as principal component analysis (PCA), discriminant factor analysis (DFA), fuzzy ARTMAP neural network and genetic algorithms (GA). In all data analysis, the dataset was normalised to set their range to (0, 1) when PCA, DFA, fuzzy ARTMAP or GA was used.

2.5. Data analysis

The main objective of using pattern recognition methods in this particular application is to estimate the performance of the electronic nose at identifying the number of days that sardine samples have undergone cold storage. Performance is assessed by employing both statistical and neural PARC methods.

Initially, a linear, unsupervised method (PCA) was used to investigate whether data clusters related to days of cold storage appeared or not. Being unsupervised, PCA groups together (separates) samples according to similarities (differences) in input data (i.e., features extracted from sensor response). On the other hand, supervised methods make use of a calibration database in an attempt of building a model to classify calibration samples according to sensor responses (input data) and the information supplied by the supervisor (the class to which any sample in the calibration database belongs to). In the second step, supervised procedures such as discriminant factor analysis and neural networks (fuzzy ARTMAP) were used to build classification models on the basis of input/output relationships.

2.5.1. Principal component analysis (PCA)

This pattern recognition technique is a powerful unsupervised method often employed with tin oxide gas sensor arrays [13,14]. To apply this method, sensor response features (i.e., response vectors) were grouped into a response matrix **X**. Since MOS sensors show highly overlapping sensitivity, this matrix is expected to contain highly collinear variables. This collinearity means that the matrix **X** will have some dominating types of variability which carry most of the information. The main objective of PCA [14,15] consists of expressing this information by a lower number of variables called principal components. These principal components are linear combinations of the original response vectors. The principal components are chosen to contain the maximum data variance and to be orthogonal.

2.5.2. Discriminant factor analysis (DFA)

As in PCA, this technique is a factorial method [16]. In fact, using this method, the data are separated in k a priori defined classes. The objective sought using DFA is to investigate if the variables are sufficient or not to allow a good a posteriori classification of the data in their a priori groups. For this aim, the discriminating procedure consists of calculating factors that maximise the differences (e.g., the variance) among all the classes and minimise these differences inside each class. Factors are linear combinations of the variables in each group. Afterwards, a decision law or discriminant function is generated. In fact, this law corresponds to a multiple regression equation using a linear combination of variables. Using such a law, new samples can be identified and classified [16].

2.5.3. Fuzzy ARTMAP neural network

Fuzzy ARTMAP is a self-organising and self-stabilising supervised classifier that shows generally superior performance in training compared with the multilayer perceptron (MLP) [17,18]. This is especially true when there are an uneven number of samples per category [19]. Therefore, a fuzzy ARTMAP network was used for sardine freshness identification.

2.6. Genetic algorithm-based variable selection

Genetic algorithms (GA) are inspired by the process of natural selection and perform a global random search on a population of solutions [20–22]. GA have been shown to solve the optimisation problem by exploring all regions of the potential solutions and exponentially searching promising areas through mutation, cross-over, and selection operations applied to individuals (i.e., chromosomes) in a population. The population is maintained and manipulated by implementing a "survival of the fittest" strategy in the search for the optimal solution. Because the next explored point in a solution space is chosen by stochastic rather than deterministic rules, GA do not need to make assumptions about the characteristics of the problem to be solved and, therefore, apply generally. When this algorithm is applied to perform variable selection, a population of n subsets or chromosomes is created; each of them contains a random combination of variables. Chromosomes are binary strings where the occurrence of a bit equal to 1 (or 0) in position i implies that ith variable is present (or absent). The cost function for each subset is then evaluated in turn and, using techniques loosely based on biological genetics and evolution, a new population is created.

Since up to four features were extracted per sensor, the dimension of input space in the application studied was 24. However, employing a high number of features does not necessarily help to build robust classification models. Therefore, a GA algorithm was considered to investigate whether using a subset of the original features available could lead to better classification models. During the variable selection process implemented here, the cost function being optimised by the GA was the prediction error of a fuzzy ARTMAP classifier. A cascaded GA was applied to prevent the selection of irrelevant variables. More details on this algorithm can be found elsewhere [20].

2.7. Characterisation of the headspace of sardines using SPME/GC/MS

A solid phase micro-extraction (SPME)/GC/MS analysis was carried out at Days 4, 5, 7 and 11 of cold storage. The sardine samples used in this experiment were obtained from a local retailer in Tarragona (Spain) the same day they had been caught. Samples were conditioned and stored in the same way as described before. The SPME fiber, 75 μm Carboxen/polydimethylsiloxane (Supelco), was introduced into the plastic bag containing the sardine sample and exposed to its headspace for 20 min at a controlled temperature (22–23 °C). Afterwards, the fiber was retracted and immediately introduced into the injection port of the gas chromatograph for 30 min in order to desorb the volatile compounds and to ensure the complete cleaning of the fiber. The gas chromatograph was a GC-17A (Shimadzu, Kyoto, Japan) equipped with a Carbowax-10 column (30 m × 0.25 mm i.d., 0.25 μm film thickness) from Supelco coupled to a Shimadzu QP-5000 mass spectrometer. The injection port was equipped with a glass liner (0.75 μm i.d.) and kept at 250 °C in splitless mode (sampling time 1 min). The helium (99.9999%) flow-rate through the column was 1.0 ml min^{-1} (constant flow). The column temperature was increased from 40 to 200 °C at 2 °C min^{-1}, and held there for 10 min. The mass detector was operated in the electron ionization mode (electron energy, 70 eV). The ion source temperature was maintained at 250 °C. A continuous scan mode at 2 scans/s was used with a mass range m/z 35–150. Volatile compounds were identified by comparing their mass spectra with those contained in the Willey 229 mass spectral database.

2.8. HS/MS analysis

The results of the electronic nose system were compared against the ones obtained with a static headspace (HS)/MS. The set-up was similar to the one described above, the only difference being that oven temperature was kept constant at 250 °C to minimise chromatographic separation. In this way, for any given measurement, the resulting mass spectrum gives a fingerprint that is characteristic of the volatiles present in the headspace of samples. Four aliquots (500 μl volume) were extracted from the headspace of the sample and input into the injector using a chromatographic syringe. HS/MS analysis was carried out with fresh sardines and with sardines at Days 4, 5, 7, and 11 of cold storage.

3. Results and discussion

3.1. SPME/GC/MS compound identification

The composition of the headspace of fish is a source of information about the freshness degree of a sample. Spoilage in fish can be detected through the measure of the amount of amines, such as trimethylamine (TMA) [1]. The identification and quantification of characteristic compounds in the headspace of sardines were determined by SPME/GC/MS. Table 1 shows the 10 components that were identified. The more intense ions

Table 1
The 10 compounds in the headspace of sardine samples identified by SPME/GC/MS and their three more intense ions

Trimethylamine	58, 42, 59
Acetone	43, 58, 42
2-Butanone	43, 72, 57
Ethanol	43, 45, 46
2-Butanol	45, 59, 41
3-Methyl-1-butanol	55, 41, 42
Ethylacetate	43, 45, 70
Dimethylsulfide	47, 62, 45
Dimethyldisulfide	94, 79, 45
Dimethyltrisulfide	126, 45, 79

for each one of these components are also shown in this table. The variation in the headspace content was evaluated by comparing the peak area of each identified compound integrated using its peak base ion (for example *m/z* 94 was used to integrate dimethyldisulfide and so on). The absolute peak area was divided by its height to be converted to arbitrary unit area (aua). In Fig. 2, an histogram shows the evolution (in aua) of the 10 identified compounds for sardine samples stored during 4, 5, 7 and 11 days. The composition and concentration of volatile compounds emanating from sardines change depending on their freshness. Spoilage odours develop as a result of microbial growth and oxidation leading to the degradation of the fish. Compounds such as trimethylamine, 2-butanone, ethanol, 3-methyl-1-butanol, dimethyldisulfide and dimethyltrisulfide are among the most volatile compounds being produced during sardine spoilage and therefore are easily recognisable in the headspace. This is in agreement with the results of Olafsdottir et al. [23] and Haugen et al. [3].

3.2. HS/MS results

An HS/MS method was implemented and a PCA was performed using as inputs the intensities of the fragments presented in Table 1. This multivariate approach was considered to evaluate whether some correlation existed between the spoilage of sardines and the change in headspace composition. As it can be appreciated in Fig. 3 (scores plot), sample clusters are well apart according to the number of days of storage. The first three principal components accounted for 82% of the total variance. The fact that most of the volatile molecules identified by SPME/GC/MS have numerous mass fragments in common increases the difficulty to scrutinize which volatiles allow for cluster separation according to different spoilage stages. The poor specificity of *m/z* fragments can be explained by the low molecular mass of volatiles sampled by means of static headspace and the high degree of fragmentation obtained with electron beam ionization. Despite this poor specificity, several conclusions can be derived taking into account the scores plot shown in Fig. 3. PC2 seems to be the main responsible for differentiating Day 11 samples from the other ones (i.e., PC2 helps to separate samples at evolved degrees of spoilage). On the other hand, PC1 is responsible for temporal separation at early and medium stages of spoilage (i.e., Days 1–7). Considering now the loadings histogram shown in

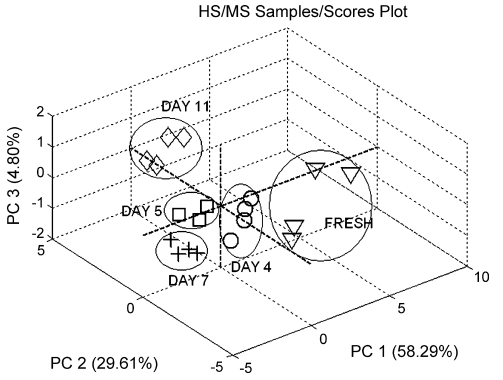

Fig. 3. Scores plot of a PCA performed on sardine data obtained with the HS/MS system.

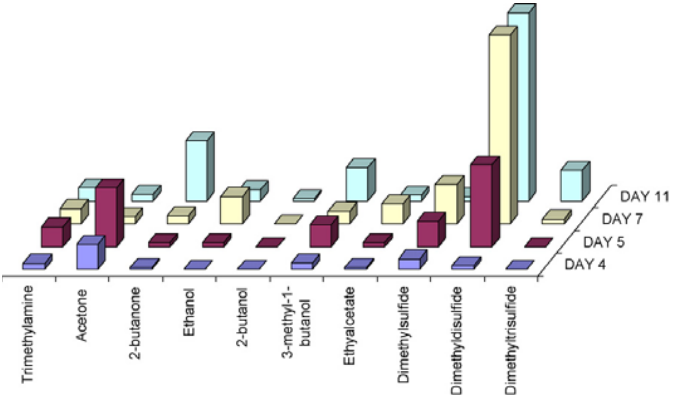

Fig. 2. Comparative histogram showing the evolution with the number of storage days for the 10 volatiles identified in the headspace of sardine samples by SPME/GC/MS.

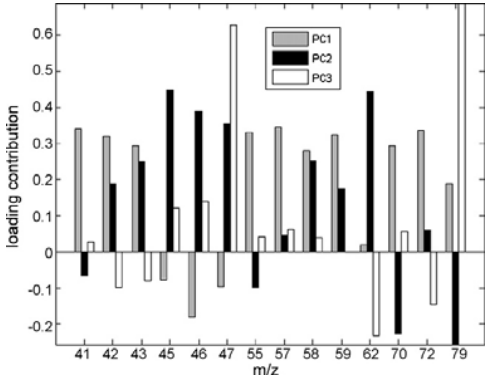

Fig. 4. Contribution (loadings) of the *m/z* fragments to the PCs of the PCA model performed on sardine data obtained with the HS/MS system.

Fig. 4, *m/z* ratios 45–47 heavily contribute to PC2. These fragments stem mainly from sulfur compounds and alcohols such as 2-butanol and ethanol. All these compounds seem to be responsible for clearly separating the most spoilt samples. That makes sense, since in accordance with Fig. 2, these are the compounds that experience the highest increase at Day 11. The most relevant *m/z* fragments on PC1 (41–43, 55, 58, 59, 70 and 72) are related to substances that are released at the beginning of fish spoilage such as trimethylamine, acetone, 3-methyl-1-butanol and ethylacetate. Finally, PC3 that explains only 5% of variance presents positive and strong correlations with ions 47 and 79, which are related to the presence of dimethylsulfide, dimethyldisulfide and dimethyltrisulfide.

3.3. Sensor response analysis

This initial analysis consists in studying the influence of the volatile gases present in the headspace of sardine samples on the conductance of the sensors. Since measurements were gathered in 15 days only, the evolution in sensor conductance can be directly attributed to the increasing number of storage days of fish (D1, D3, D4, D5, D7, D9, D11, D13 and D15), and the effect of sensor drift can be neglected. Fig. 5 shows for the six sensors used the evolution in their response to sardine samples D1–D15. More specifically, the different sub-plots in Fig. 5 show the second step of the measurement phase as described in Section 2.3. It can be derived that the intensities of the conductance of the six sensors increase according to the number of days that sardines have undergone cold storage. This behaviour can be justified by an increase in the concentration of volatile gases given out by sardines as a function of storage time, or the occurrence of new species in the headspace of sardines (as revealed by the SPME/GC/MS analysis). Since a humidity sensor was present and no significant humidity changes were observed during the whole measurement process, it can be ruled out that water activity was responsible for the evolution in sensor response shown in Fig. 5. The sensor array used is highly sensitive to

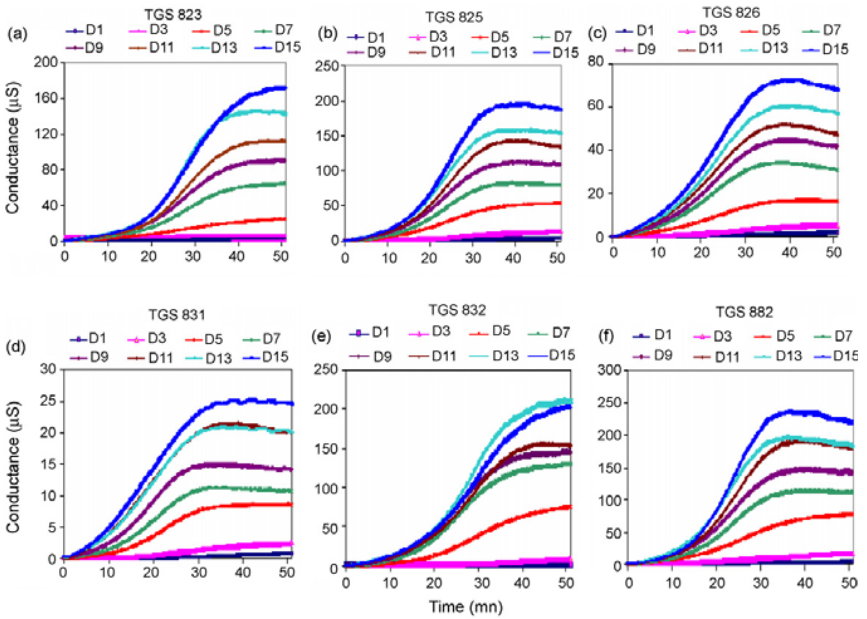

Fig. 5. Dynamic response of the sensors within the array to sardine samples that have undergone cold storage a different number of days: (a) TGS 823; (b) TGS 825; (c) TGS 826; (d) TGS 831; (e) TGS 832; (f) TGS 882. D*i*: A sample having undergone cold storage for *i* days.

these volatile species. It can also be noticed that a saturation of the dynamic response of the sensors appears more distinctly for the last days of conservation studied. The response of the sensors to the dynamic headspace of sardines is quite similar. A very small variation in conductance is observed during the first 15 min of exposure. Then, a sharp increase in conductance occurs in the time interval between 15 and 40 min. Finally, the responses show a tendency to stabilise. Sensors TGS 882, TGS 832, TGS 825 and TGS 823 showed high responsiveness for all the samples measured. Sensor TGS 882 showed the highest responsiveness, which is in agreement with the results of O'Connell et al. [10]. On the other hand, the responses of sensors TGS 826 and TGS 831 were lower and did not show a large difference between fresh samples and samples that underwent long conservation. The changes in sensor response according to the number of days the sardines were kept under cold storage are clearly shown in Fig. 6. This figure shows in histograms the evolution of the stabilised conductance of the sensors at 3, 7 and 13 days of storage. Even thought all sensors show an increase in conductance when the number of storage days is increased, this change is very large for sensors TGS 882, TGS 832, TGS 825 and TGS 823 only.

3.4. Electronic nose data analysis

3.4.1. Principal component analysis (PCA)

A mean centring pre-processing technique was applied to the response matrix. The results of the PCA show that the sensors are strongly correlated, since 91.34% of the variance of the data was explained by the first principal component, PC1. The first two components, PC1 and PC2, captured 96.22% of data variance. Fig. 7 shows the projections of the experimental results on a two-dimensional (2D) plane formed by the first two principal components. Samples can be grouped together in three different clusters. Each of these groups corresponds to fresh, medium and aged sardine samples, respectively. The first group corresponds to the samples having undergone up to 3 days of storage. The second group corresponds to the samples having undergone from 5 up to 7 days of storage. Finally, the third group corresponds to the samples that underwent from 9 up to 15 days of storage. The

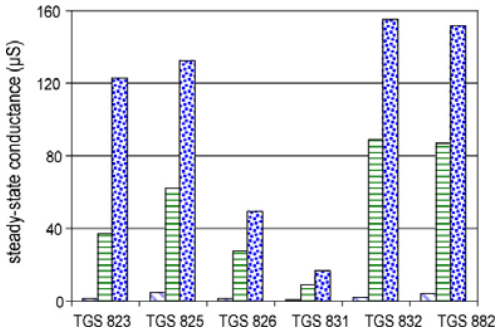

Fig. 6. Stabilised conductance (*Gs*) of the different sensors in the presence of sardine samples stored at 4 °C during 3 days (▨), 7 days (▭) and 13 days (▩).

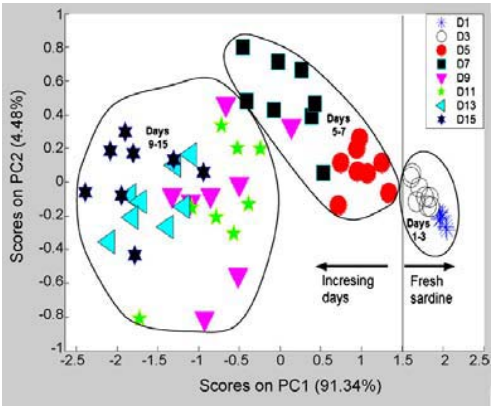

Fig. 7. Scores plot of a PCA on sardine data using the six-element gas sensor array. A vertical line separating fresh samples from medium and aged ones and ellipses grouping fresh, medium and aged samples had been added for easy identification.

fact that samples appear ordered along the first principal component according to the number of storage days is a good result, since the sensor array employed seems appropriate to envisage an application where the main goal would be to predict the number of days of cold storage undergone by sardines (or at least to classify sardines samples into fresh, medium or aged). The first principal component explains the main variance in the data (i.e., in the response of the sensors). The fact that the samples increased their scattering along the second principal component when the number of storage days increased can be due to two effects. The first one (and more important) is due to the occurrence of new volatile species in the headspace of medium and aged sardine samples. The second effect may be due to small fluctuations of ambient variables such as temperature or humidity. These results are very alike to those obtained when a PCA was computed using the HS/MS system; in that case the samples are also ordered along the first principal component according to the number of cold storage days. In particular, it can be derived that a very good agreement exists in the discrimination between fresh and aged sardines. Fresh sardines would be those that had been stored up to 3 days.

Fig. 8 shows the evolution of the scores on PC1 as a function of storage time. For each day in which measurements were performed, the average value of the scores over eight replicate measurements is shown in Fig. 8. These scores show an almost monotonic decrease during the period of storage. A polynomial fit of the scores suggests that the first PC could be used to predict the period of storage and, therefore, the decay in sardine freshness. Additionally, this figure shows that after Day 3, the scores on PC1 experience a sharp decrease. This suggests that after the third day of cold storage sardines are degrading (e.g., the threshold of acceptability should be located between Days 3 and 5). For Days 9, 11, 13 and 15, the average value of the scores on PC1 remains almost unchanged, which suggest that all these samples belong to a single category (i.e., aged fish). Once again,

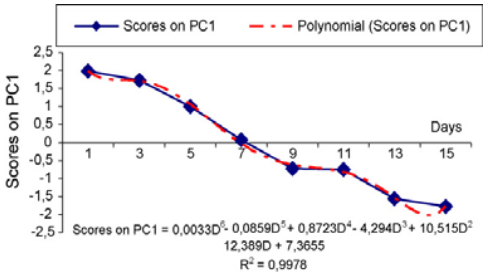

Fig. 8. Evolution of the scores on the first principal component with the period of storage and polynomial fitting.

this is in good agreement with the PCA results obtained when an HS/MS system was employed.

3.4.2. Discriminant factor analysis (DFA)

The performance of the designed electronic nose system in the classification of sardine samples into three groups, fresh, medium and aged, was evaluated applying DFA. These groups corresponded to those discovered by PCA.

For validating purposes, six samples, two from each group, were selected at random and removed from the database. This left 58 samples for training. With the remaining samples, the DFA technique was used first to build a decision law for discriminating between the three a priori defined groups as illustrated in Fig. 9. With the DFA method and employing the 24 features extracted, 96.6% of the samples were correctly classified in their a priori groups. In fact only two measurements were misclassified (Table 2). Additionally, a leave-one-out cross-validation procedure was implemented and 58 DFA models were built using 57 measurements. This led again to the same success rate in classification since only two measurements were misclassified.

Table 2
DFA discrimination results

Predicted	Actual		
	Class 1	Class 2	Class 3
Class 1	14	0	0
Class 2	0	13	1
Class 3	0	1	29

Table 3
DFA identification results for the six samples tests

Test samples	Measurement label	Identification results
Case-1	D1S3	Fresh
Case-2	D3S2	Fresh
Case-3	D5S1	Medium
Case-4	D7S4	Medium
Case-5	D9S1	Aged
Case-6	D13S3	Aged

Di: *i*th-day; Si: *i*th-sample

In the second step, the decision law obtained using DFA and the 58 remaining measurements were employed to identify the six samples that had been left out. Table 3 and Fig. 10 summarise the results of this process. These six unknown samples could be correctly identified. In conclusion, the response features of the electronic nose allow for a good a posteriori classification, and the decision law enables identifying and classifying unknown samples.

3.4.3. Fuzzy ARTMAP neural networks (FANN)

A non-linear classifier such as the Fuzzy ARTMAP was applied to the three-category classification of sardine samples. The performance of the fuzzy ARTMAP classifier was evaluated

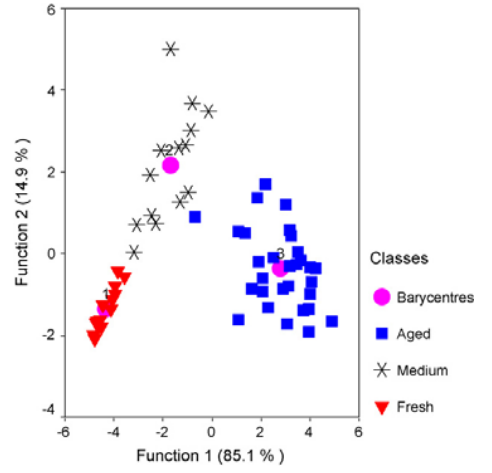

Fig. 9. DFA results in the discrimination of the three states of freshness of sardine samples.

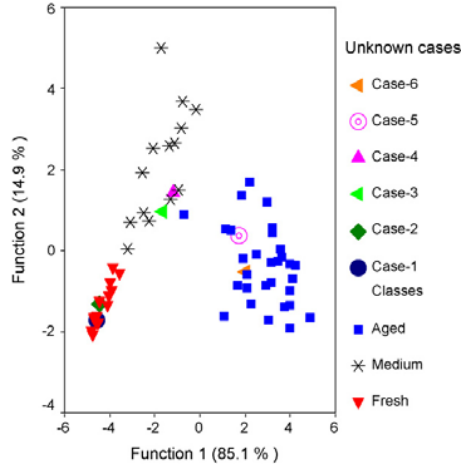

Fig. 10. Example of identifying six unknown samples using DFA.

using a leave-one-out cross-validation method. The process was as follows: given n measurements ($n = 64$ measurements within each training matrix), the model was trained 64 times using 63 vectors. The vector left out was then used for testing the model. Performance in training was estimated as the averaged performance over the 64 tests. The fuzzy ARTMAP neural network reached a 96.88% success rate in the recognition of the three states of freshness of sardine samples. Table 4 summarises the results of the fuzzy ARTMAP classifier. The two misclassified samples are as follows. One measurement belonging to class 2 (medium) was misclassified as belonging to class 3 (aged) and one measurement belonging to class 3 was misclassified as belonging to class 2. The results show that the fresh samples could be perfectly discriminated from medium and aged samples.

3.4.4. GA-based variable selection

In an attempt to prevent irrelevant input variables from being used in the classification models, a cascaded GA for variable selection was implemented. The 64 measurements available were used. The process was as follows. First a GA coupled to a fuzzy ARTMAP classifier was applied to select between the 24 variables extracted from the sensor responses. Twelve out of 24 variables were selected. A second GA was applied to select from the 12 variables that had been selected by the first GA. Seven out of 12 variables were selected. The variables that survived the process of selection were G_0, dG/dt and G_s of TGS 823, G_s of TGS 826, TGS 831 and TGS 832, and G_0 of TGS 882. Once again the performance of a fuzzy ARTMAP classifier for determining the freshness of sardine samples was estimated using the leave-one-out approach. The success rate remains constant 96.88%, when the surviving variables were used. However, the speed of classification was slightly improved.

The GAs allowed the size of the database to be reduced, which resulted in an increased speed of sample identification and in the development of more robust models for classification.

3.4.5. Electronic nose performance validation

Generally, the performance of an electronic nose is studied by measuring the repeatability and reproducibility. The first one represents the short-term precision and is evaluated by measuring replicate samples on the same day. On the other hand, the reproducibility evaluates the long-term precision of the instrument and is determined by measuring different samples on different days.

In order to evaluate the reproducibility of our electronic nose, a validation analysis for a new set of sardines was performed.

Table 4
FANN classification results

Predicted	Actual		
	Class 1	Class 2	Class 3
Class 1	16	0	0
Class 2	0	15	1
Class 3	0	1	31

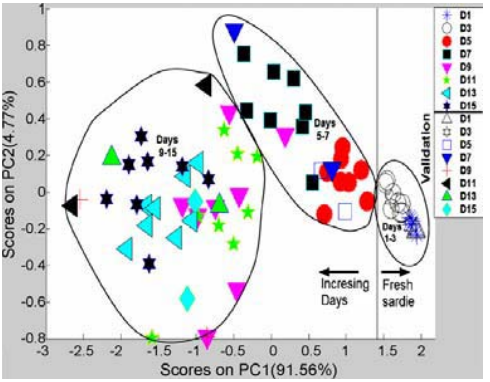

Fig. 11. Projection of the second set of measurements (second catch of sardines) onto the PCA model built using the first set of measurements.

This analysis was performed using the same experimental conditions as in the first analysis but 2 months later. Sardines from another catch were obtained from the same local fish market in December 2005 and stored under the same conditions. The new measurements (16) were projected onto the already existing PCA and DFA models. In other words, the classification models that had been built using the first sardine samples were used (without any retraining) to classify the second set of sardines. Figs. 11 and 12 show the validation results obtained with the PCA and DFA models, respectively. The new samples appear correctly distributed according to their stage of conservation. These results prove that the electronic nose system leads to reproducible results and that it can be used as a screening tool to assess the freshness of sardines.

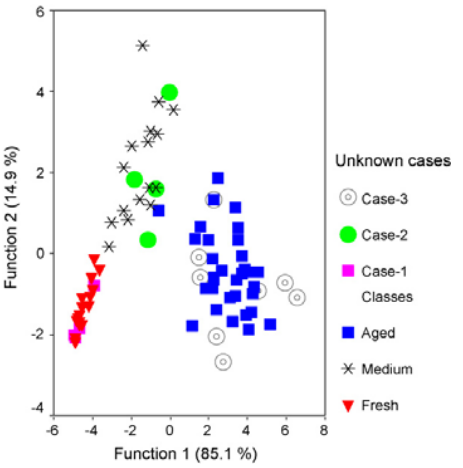

Fig. 12. Results of the DFA classification for the second set of measurements. The DFA classifier was built using the first set of measurements.

4. Conclusion

In this paper we have reported on the building of a simple electronic nose system aimed at monitoring fresh sardines stored at 4 °C. A dynamic headspace sampling method is employed together with an array of commercially available metal oxide gas sensors.

A principal component analysis has shown that the sensor signals are well correlated with the number of days that sardines have undergone cold storage. More specifically, PCA results show that the classification of sardines in three groups, namely, fresh, medium and aged, which correspond to up to 3, 7 and over 9 days of storage, respectively, seems to be possible. Therefore, discriminant factor analysis and the fuzzy ARTMAP neural networks have been used in an attempt to build classification models. These models have reached a high success rate in the classification of sardine samples (96.60% for DFA and 96.88% for fuzzy ARTMAP).

Finally, we have shown that these classification performances can be slightly improved (speed of classification) by selecting appropriate features from the sensor response signals. This process has been conducted by using a genetic algorithm. Additionally, by analysing a new batch of sardines caught 2 months later than the ones used to build the e-nose pattern recognition, we have shown that accurate reproducible results are obtained. Therefore, these results clearly demonstrate that our system is a simple, affordable and efficient tool for sensing the degradation of fresh sardines stored at 4 °C. For example, it clearly identifies fresh sardines (up to 3 days of storage) from aged sardines (5 or more days of storage). In fact, 3 or 4 days of storage could be proposed as a threshold for consumers' acceptability.

Moreover, the changes in the headspace composition according to the number of conservation days have been demonstrated by identifying the main compounds and their evolution via an SPME/GC/MS study. The results obtained with the electronic nose are supported by those obtained with an HS/MS headspace screening technique (i.e., they are quite similar). This electronic nose would be useful for testing the quality of fresh sardines along the logistics chain.

Acknowledgement

Financial support from AECI under project no. A/4349/05 is gratefully acknowledged.

References

[1] C.D. Natale, G. Olafsdóttir, S. Einarsson, E. Martinelli, R. Paolesse, A. D'Amico, Comparison and integration of different electronic noses for freshness evaluation of cod-fish fillets, Sens. Actuators B: Chem. 77 (2001) 572–578.

[2] J. Hammond, B. Marquis, R. Michaels, B. Oickle, B. Segee, J. Vetelino, A. Bushway, M.E. Camire, K.D. Dentici, A semiconducting metal-oxide array for monitoring fish freshness, Sens. Actuators B: Chem. 84 (2002) 113–122.

[3] J.E. Haugen, E. Chanie, F. Westad, R. Jonsdottir, S. Bazzo, S. Labreche, P. Marcq, F. Lundby, G. Olafsdottir, Rapid control of smoked Atlantic salmon (*Salmo salar*) quality by electronic nose: correlation with classical evaluation methods, Sens. Actuators B: Chem. 116 (2006) 72–77.

[4] B. Kilinc, S. Cakli, Chemical, microbiological and sensory changes in thawed frozen fillets of sardine (*Sardina pilchardus*) during marination, Food Chem. 88 (2004) 275–280.

[5] R. Olafsson, E. Martinsdottir, G. Olafsdóttir, T.I. Sigfusson, J.W. Gardner, Monitoring of fish freshness using tin oxide sensors, in: J.W. Gardner, P.N. Bartlett (Eds.), Sensors and Sensory Systems for an Electronic Nose, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992.

[6] G.E. Chanie, F. Westad, R. Jonsdottir, C.R. Thalmann, S. Bazzo, S. Labreche, P. Marcq, F. Lundby, J.E. Haugen, Prediction of microbial and sensory quality of cold smoked Atlantic salmon (*Salmo salar*) by electronic nose, J. Food Sci. 70 (2005) 563–574.

[7] D.B. Josephson, R.C. Lindsay, G. Olafsdottir, Measurement of volatile aroma constituents as a means for following sensory deterioration of fresh fish and fishery products, in: D.F. Kramer, L. Liston (Eds.), Proceedings of the Seafood Quality Determination Symposium, Elsevier, Amsterdam, 1986.

[8] W. Göpel, Chemical imaging: concepts and visions for electronic and bio electronic noses, Sens. Actuators B: Chem. 52 (1998) 125–142.

[9] A.K. Srivastava, Detection of volatile organic compounds (VOCs) using SnO_2 gas-sensor array and artificial neural network, Sens. Actuators B: Chem. 96 (2003) 24–37.

[10] M. O'Connell, G. Valdora, G. Peltzer, R.M. Negri, A practical approach for fish freshness determinations using a portable electronic nose, Sens. Actuators B: Chem. 80 (2001) 149–154.

[11] J. Zhu, C. Ren, G. Chen, C. Yn, J. Wu, H. Mu, A new automotive air/fuel sensor based on TiO_2-doped Nb_2O_5 thin film by ion-beam-enhanced deposition, Sens. Actuators B: Chem. 32 (1996) 209–213.

[12] A. Amari, N. El Barbri, E. Llobet, N. El Bari, X. Correig, B. Bouchikhi, Monitoring the freshness of Moroccan sardines with a neural-network based electronic nose, Sensors 6 (2006) 1209–1223.

[13] H.T. Nagle, R. Gutierez-Osuna, S.S. Schiffman, The how and why of electronic noses, IEEE Spectrum 35–39 (1998) 22–34.

[14] E. Llobet, J. Brezmes, X. Vilanova, J.E. Sueiras, X. Correig, Qualitative, quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick film tin oxide gas sensor array, Sens. Actuators B: Chem. 41 (1997) 13–21.

[15] J.W. Gardner, Detection of vapours and odours from a multisensor array using pattern recognition: Part 1. Principal component and cluster analysis, Sens. Actuators B: Chem. 4 (1991) 109–115.

[16] D.F. Morrisson, Multivariate Statistical Methods, 2nd ed., Mc-Graw-Hill, Singapore, 1988, p. 415.

[17] G. Carpenter, S. Grossberg, N. Markuzon, J. Reynolds, D. Rosen, Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps, IEEE Trans. Neural Netw. 3 (1992) 698–713.

[18] G. Carpenter, S. Grossberg, J. Reynolds, A fuzzy ARTMAP non parametric probability estimator for non-stationary pattern recognition problems, IEEE Trans. Neural Netw. 6 (1995) 1330–1336.

[19] E. Llobet, E. Hines, J. Gardner, Fuzzy ARTMAP-based electronic nose data analysis, Sens. Actuators B: Chem. 61 (1999) 183–190.

[20] E. Llobet, J. Brezmes, O. Gualdrón, X. Vilanova, X. Correig, Building parsimonious fuzzy ARTMAP models by variable selection with a cascaded genetic algorithm: application to multisensor systems for gas analysis, Sens. Actuators B: Chem. 99 (2004) 267–272.

[21] J.W. Gardner, P. Boilot, E.L. Hines, Enhancing electronic nose performance by sensor selection using a new integer-based genetic algorithm approach, Sens. Actuators B: Chem. 106 (2005) 114–121.

[22] D. Broadhurst, R. Goodacre, A. Jones, Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry, Anal. Chim. Acta 348 (1997) 71–86.

[23] G. Olafsdottir, P. Nesvadba, C.D. Natale, M. Careche, J. Oehlenschläger, S.V. Tryggvadóttir, R. Schubring, M. Kroeger, K. Heia, M. Esaiassen, A. Macagnano, B.M. Jorgenseng, Multisensor for fish quality determination, Trends Food Sci. Technol. 15 (2004) 86–93.

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Paper VIII

Vinaixa, M.; Llobet, E.; Brezmes, J.; Vilanova, X.; Correig, X. Taking advantage of the time dimension on MS-Sensor approaches: Evaluation of Iberian Ham quality using SH-MS and multi-way data analysis, *In preparation.*

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

UNIVERSITAT ROVIRA I VIRGILI
IMPROVING MS-SENSOR TECHNOLOGIES FOR FOOD QUALITY ASSESSMENT
Maria Vinaixa Crevillent
ISBN:978-84-691-9752-3/DL:T-126-2009

Taking advantage of the time dimension on MS-Sensor approaches: Evaluation of Iberian Ham quality using SH-MS and multi-way data analysis

M. Vinaixa*[a]; J. Brezmes[a], E. LLobet[a], J.A. Carrasco[b], X. Correig[a]

[a]MINOS, Dept. Enginyeria Electrònica, Universitat Rovira i Virgili, Avda. Països Catalans, 26, 43007, Tarragona, Spain.
[c]Dept. Ingeniería, Instituto del Frío, Consejo Superior Investigaciones Científicas, José Antonio Novais 10, 28040 Madrid, Spain.

Abstract

Data provided from second-order instruments such as MS-Sensor devices intrinsically present more than two sources of variability. The response in ion counts arriving at the detector is measured as a function of retention time and mass to charge ratio. Thus, these data can be arranged in a three way array layout. Common pattern recognitions algorithms traditionally used in MS-Sensor devices make use of data vectors (i.e. a single mass spectrum per sample) obtained by averaging mass spectra along the time dimension. Thus, this approach might potentially leak significant temporal information. Arranging and processing MS-Sensor data in a three-way array layout can solve this problem. This paper demonstrates for the firs time that the use of multi-way methods with MS-Sensor data compares favourably against traditional two-way pattern recognition algorithms. This assertion is illustrated in a typical food-related MS-Sensor application such as Iberian Ham quality assessment. Performance of PLS-DA (Partial Least Square Discrimination Analysis) and models containing temporal information such as unf-PLS-DA (unfolded Partial Least Square Discrimination Analysis) and N-PLS-DA (three-way Partial Least Square Discrimination Analysis) are compared proving the advantage of the three-way approach against classical two-way algorithms in the classification of two different hams qualities in accordance to the type of feeding the pigs receive. Additionally, composition parameters such as water content (χ), water activity (a_w) and salt content (NaCl) are correlated to MS-Sensor measurements using PLS, N-PLS and unf-PLS regression models. Again, N-PLS compares favorably respect to two-way methods in this quantification problem. Since multi-way data analysis implies the use of the time dimension, any variation along retention time has to be addressed. In order to achieve alignment of MS-Sensor

ISBN:978-84-691-9752-3/DL:T-126-2009

signals along time dimension a recent published algorithm (Recursive Aligment Fast Fourier Transform,RAFFT) has been adapted and applied.

Keywords: MS-Sensor, electronic nose, Iberian ham, N-PLS-DA, PLS-DA, unfolding, alignment, multi-way calibration.

1. Introduction

The characterization of raw materials and foods is of prime strategic importance to the food industry. Recent research has shown that the rapid analysis of the volatile fraction of food products by mass spectrometry can be an effective tool for assessing food quality. In this framework, and during the last few years, applications on MS-Sensor technology (also called MS-based e-noses) have been growing as an alternative to gas sensor based electronic noses[1,2]. The main advantages of using a mass analyzer instead of gas sensor arrays are widely described in the literature[3,4].

The basic working principle of MS-Sensor systems is based on the introduction of volatile components extracted from the headspace of a sample into the ionization chamber of a mass spectrometer. The mass spectra resulting from the simultaneous ionization and fragmentation of these volatiles constitutes a 'fingerprint' that is characteristic of the product being analyzed. Exploitation of this information allows the classification of samples according to their volatile pattern profile or even the correlation with the main quality parameters. The TIC (Total Ion Chromatogram) plot determines the time interval with meaningful signals, and the averaged m/z values of the mass spectrum is obtained by summing up the ion counts in these time intervals. This plot usually takes a form of an asymmetrical peak of width equal to this range[5]. Common pattern recognition algorithms applied on MS-Sensor data make use of data matrices in which columns represent each one of the mass to charge ratio (m/z) scanned and rows hold their intensities. Even when chromatographic resolution is avoided, some kind of diffusion can be observed on this asymmetrical peak which shows some retention potentially able to report additional information that may help for further modelling of data in classification or prediction tasks. Time averaging of the mass spectrum may lead to a loss of this temporal information so that this potentially useful data is not used in the pattern recognition step. Actually,

averaging mass spectra along the detected peak allows converting real three-way data to two-way by eliminating the time dimension. Therefore, the real nature of data is not respected and is changed just to adapt the response of the instrument to current available pattern recognition two-way algorithms.

This paper explores the possibility of computing this extra information using multi-way analysis [6]. Multi-way methods are particularly useful for the analysis of batch process data and analytical data that intrinsically present more than two sources of variation [7] (i. e. MS-Sensor data, where a response is being measured as a function of two parameters: m/z and time). In fact, data provided from the MS-Sensor should be arranged as multi-way array where the first mode represents samples, the second corresponds to mass spectra and the third to the elution profiles.

Our case study is related to Iberian ham quality assessment. The products derived from Iberian pigs constitute an important economic activity in Spain and much attention has been devoted to the study of dry cured ham in recent years. Volatile compound composition of ham is markedly affected by pig feeding. Moreover, remarkable sensory differences according to pig feeding have been reported[8, 9, 10]. According to the rearing system, Iberian hams can be classified into two basic categories: acorn hams (from pigs fattened outdoors, feeding based on acorn and pasture land) and fodder hams (from pigs fattened indoors, feeding based on concentrated feed). The feeding that pigs receive contributes in a remarkable way to the sensorial characteristics of hams such as flavour, those most appreciated being that ones coming from pigs fattened with acorn. Currently, characterisation of Iberian ham as a function of their diet is an issue to which many analytical efforts have been addressed [11, 12, 13, 14]. Nowadays, manufacturer producers need for fast and reliable methods either to distinguish the feeding that the animal has received or to assess quality parameters in general. Such methods would help regulatory authorities and even final consumers to avoid frauds in the Iberian ham commercialization.

Taking advantage of the time dimension in MS-Sensor devices involves the unavoidable evaluation of co-eluted peak shape in the multi-way analysis step. In order to perform direct chemometric analysis of the entire chromatographic data matrices, co-eluted chromatographic profiles must be properly aligned to compensate for minor drifts in retention time, either in global or in small sections of the chromatograms[15]. Although MS-

ISBN:978-84-691-9752-3/DL:T-126-2009

Sensor data is considered to be reproducible under optimized injection conditions, small differences in retention time between different injections will always occur. These differences may arise from small variations in the carrier flow, changes in the columns during operation, temperature variations, drift in the detectors and other unknown factors that influence retention time reproducibility. Many of chemometric techniques available for multi-way modelling rely on trillinearity, a prerequisite seldom met due to the variations in the chromatographic conditions affecting peak position and peak width. One way to tackle this problem is to pre-process the data by some kind of time alignment procedure[16]. Several methods have been proposed for alignment of second order data where the spectral information is used to guide the alignment procedure[16,15,17,18]. Nevertheless all these approaches require user intervention to set-up the optimal parameters for the alignment algorithms. This is a serious drawback from a MS-Sensor point of view because it does not allow automatic data processing. Recently, a fully automated algorithm called RAFFT (Recursive Alignment Fast Fourier Transform) has been presented by Jason W. H. Wong et al.[19]. This algorithm makes use of the Fast Fourier transform for rapid computation of the cross-correlation function that enables alignments between a target sample and samples to be optimized. It is based on spectra segmentation models and is developed to offset the need for operating parameters. Minimal segment size is determined automatically by recursive alignment from the full spectrum (global scale) to progressively smaller segments (local scale) until no further alignment is required. In this study, a modification of the RAFFT algorithm has been applied in order to overcome retention time shifts prior to modelling both three-way and unfolded data.

The main goal of this paper is to demonstrate that considering the time dimension in the signal provided by a MS-Sensor device translates into the addition of extra valuable information related to the real three-way structure of the data. To do so, a case study such as the quality assessment of Iberian ham has been performed to test this hypothesis.

The first part of this study deals with the fitting of qualitative models such as PLS-DA to classify Iberian ham samples according to the type of feeding received during the pig's fattening period. Its performance is compared to classification models containing temporal information either using three way data and N-PLS-DA (discriminant multi-way partial least squares regression, the generalization of PLS-DA to three-way data) or using a

classical PLS-DA but this time on the matrix resulting from unfolding the three-dimensional data array. Unfolding is simply a way to rearrange multi-way data arrays to a two-way matrix by concatenating the third dimension as new columns.

Despite the fact that the Partial Least Squares (PLS) algorithm was not originally designed as a tool for statistical discrimination, there have been many studies applying PLS for classification and according to Barker and Rayens [20] there is enough evidence to suggest that it performs well in that role. PLS-DA consists in a classical PLS regression where the response variable is a categorical one (Y-block response is replaced by a set of dummy variables describing the categories). It is based on a bilinear decomposition of the so-called calibration matrix so that PLS components are calculated by finding a proper compromise between the purpose of maximising covariance within X-block or explanatory variables and categorized Y-block. Therefore, PLS-DA allows for variables that define the group of individuals. In practice, a threshold between classes is determined and class assignment is done according to this threshold. Probability of a sample being inside or outside a category can be also calculated.

The N-PLS algorithm is the normal extension of PLS to three-way data [21]. NPLS regression provides a platform for developing calibration models with the ability to classify new data and provide the degree of certainty. This regression method, combined with discriminant analysis (DA), performs supervised pattern recognition for the separation of classes with a high degree of similarity [22]. To our knowledge, there are only a few works reported in the literature where N-PLS-DA has been used [22,23]. To date there are no related works where N-PLS-DA is used for classification purposes on MS-Sensor data.

In the first part of this work, the feasibility of using N-PLS-DA as a classifier in a typical MS-Sensor application is studied for the first time. In the second part of this study PLS predictive models are calculated for parameters such as *NaCl*, a_w, and *X*, and the results are compared with the results obtained by models using the temporal dimension information. The three parameters mentioned above are composition parameters currently used to control the curing process and to assess ham quality in general[24, 25]. These parameters are highly correlated with the risk of bacterial spoilage and recently it has been pointed out their correlation with the texture of hams[26, 27]. Because of their interest, prediction of these parameters from MS-Sensor measurements is also attempted. Once again, two different

ISBN:978-84-691-9752-3/DL:T-126-2009

approaches are used in order to incorporate the time dimension. The first approach uses N-PLS predictive models fed with the three way data array and the second one uses a PLS model fed with two-dimension unfolded data. Even some studies where MS-Sensor measurements are used to perform multivariate calibration have been reported in the literature[28, 29, 30]. To date there is only one work claiming for advantages when using differences in the temporal profiles of the analytes using direct sampling mass spectrometry[31]. Gardner and co-workers[31] demonstrated the superiority of three-way calibration models against multivariate and univariate classical methods in the prediction of the concentration of three analytes with similar mass spectra in mixtures of these compounds. Nevertheless, our application represents a more challenging task because the number of volatile compounds is much higher and the resulting mass spectra are highly overlapped. Therefore, the exploitation of subtle differences becomes much more difficult in our case.

2. Experimental section

2.1 MS-Sensor ham measurements

Eleven types of Spanish Iberian dry-cured hams were used in the study. They were directly marketed from five manufacturer producers and they differed in pig feeding during their fattening period (i.e., either acorn or fodder) and in their geographical origin (Extremadura, Guijuelo or Huelva).

Samples were prepared as follows: three grams of ham (taken from the biceps femoris) were cut and introduced in 10-mL headspace vials. Six different vials were prepared from each of the eleven ham samples giving a total of 66 vials to be analysed. Sampling was based on Static Headspace. Each one of the 66 vials was analysed once, no replicates of the same vial were allowed because the headspace of the samples degrades at the sampling temperatures used in the oven, so that repeatability in the extract injections can not be ensured.

A headspace autosampler Agilent 7694 was used. Oven, loop and transfer line temperatures were set to 90, 100 and 110ºC, respectively. Times for vial equilibration, pressurisation, loop filling, loop equilibration and injection were 30, 0.4, 0.15, 0.2 and 1 minutes,

respectively. Reproducible headspace samples were injected into the injection port of a Hewlett-Packard 6890 series II gas chromatograph coupled to a mass selective detector (Hewlett-Packard HP 5973; Wilmington, DE, USA). The injection port was used in splitless mode and maintained at 280ºC. The system was equipped with a HP 19091J-215 (50m × 0.32mm id, film thickness 1.05 μm) column, kept at 200ºC in isothermal conditions. In this way, chromatographic separation was avoided and the column merely acted as a transfer line delivering volatiles to the mass detector. The column flow rate was set to 1.5 ml/min. Volatile compounds were co-eluted into the mass spectrometer, where mass spectra were obtained using an electronic impact mass selective detector at 70 eV, a multiplier voltage of 2706 V, and collecting data at a rate of 1 scan s^{-1} over the m/z range 45–250 amu.

2.2 Determination of water activity (a_w), water content (X) and sodium chloride content (NaCl)

Water activity (a_w) measurements were carried out at 25ºC with a Novasina AW-SPTRINT-TH 500 instrument (Axair Ltd, Pfäffikon, Switzerland) that allows temperature-controlled measurements of a_w. The measurements were performed for each of the eleven ham samples used in this experiment. After measuring a_w, the percentage water content (X=kg H_2O / kg dry matter) of the samples was determined by drying at 103±2 ºC until reaching a constant weight (ISO 1442:1997). NaCl content was determined through the use of chloride analyser according Bernman [32].

3. Data handling

3.1. Response matrices.

Raw data provided by the MS-Sensor was arranged as a three-way array data set **R** (66×205×101), where the first index corresponds to the number of vial samples (and, therefore, measurements), the second to the mass spectrum width (number of m/z relationships considered) and the third to the number of scans performed (retention time).

Some of the non-informative part of the raw data was rejected; only m/z ratios laying in the range 45-120 amu were considered. Mass fragments lighter than 45 amu were not taken

ISBN:978-84-691-9752-3/DL:T-126-2009

into account because of their multiple origins (mainly ambient air and carrier gas). Mass fragments higher than 120 amu were also rejected because of their poor sensitivity and because the main volatiles in the headspace of ham samples reported in the literature have no significant ions above this value. At this point, it has to be kept in mind that the static headspace technique only extracts the highest volatile fraction of the headspace. Thus, this fraction probably is not going to contain compounds that could account mass spectra with significant fragments whose intensity is above 120 amu.

Finally, only 101 scans ranging from scan 50 to 150 were considered as elution profiles leading to three different data matrices: $\underline{\mathbf{R}_1}$ (66×76×101) which is going to be considered for further multi-way analysis, $\mathbf{R}_2$ (66×76) which was obtained by averaging m/z values along 101 scans considered and it was the matrix considered to apply two-way PLS approaches (Figure 1), and $\mathbf{R}_3$ (66×7676), the last matrix response considered in this paper which includes temporal information but this time in an unfolded way.

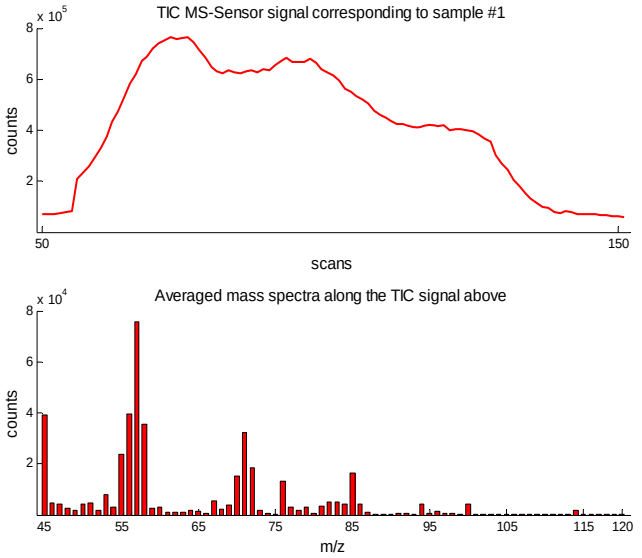

Figure 1: Obtaining T2 by averaging mass spectra along TIC signal

Unfolding is a way to convert a three-way array into a two-way matrix preserving all the data points. In our particular case, unfolding is accomplished by concatenating the third mode of $\underline{\mathbf{R}_1}$ (66×76×101). In this way, the slab conformed by the intensity of 66 measurements for each of 76 m/z channels scanned on the first scan is placed next to the

same slab for scan number two and consecutively for all the 101 scans. In the new 2-D matrix, each measurement is represented by a row, while columns represent m/z relationships at each scan, leading to the **R_3 matrix** (66×7676), where the number of columns is the product of the number of m/z relationships and number of scans (Figure 2).

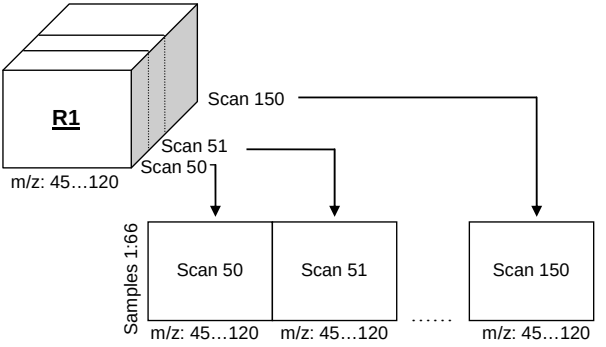

Figure 2: Illustration on how unfolding is performed in this particular case.

3.2. Software data analysis

MS-Sensor data was imported into ASCII format and finally loaded into MATLAB version 6.5 (Mathworks, Inc., Natick, MA) for further data processing. The MATLAB routines used for PLS, PLS-DA and NPLS regression were provided by the PLS_Toolbox, version 3.5 (Eigenvector Technologies, Inc.,Manson, WA). The MATLAB functions for performing N-PLS-DA were adapted from this toolbox. Joint confidence interval of the intercept and the slope tests were performed with a MATLAB routine script downloaded from [33]. Finally the routine used below in the alignment of three-way data was a modification of the RAFFT algorithm published by Wong J. W. H. et al [19].

3.3 Data pre-processing: time alignment

A modification of the RAFFT algorithm was applied to the original data contained in the **$\underline{R_1}$** (66×76×101) matrix before using any PLS model. The modification consisted in the application of this algorithm to each m/z channel instead of doing so to the reconstructed TIC signal. Afterwards, the TIC signal was reconstructed summarizing the intensities of all m/z channels for each scan of the aligned response. Fig 3 shows how the aligment modifies the m/z response on the time axis. The TIC for the 20th first measurements is shown on the

ISBN:978-84-691-9752-3/DL:T-126-2009

upper part of the figure, while the result of applying the RAFFT algorithm is shown at the bottom.

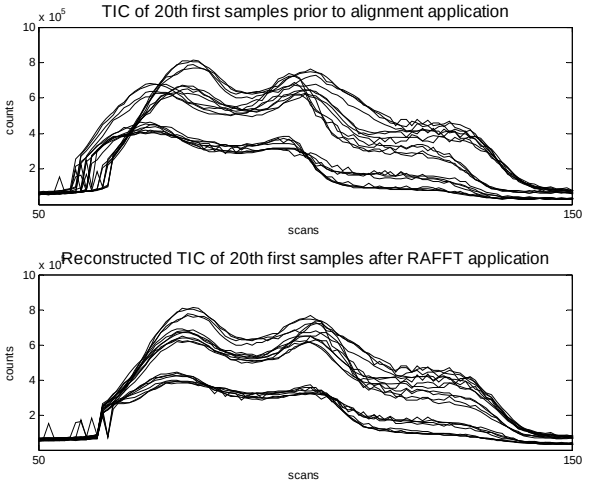

Figure 3: Comparison of raw and aligned 20th first TIC signals after applying the modified RAFFT algorithm.

3. 4. Training and Validation Datasets

The aligned response array $\underline{\mathbf{R_1}}$ (66×76×101) and response matrices $\mathbf{R_2}$ (66×76) and $\mathbf{R_3}$ (66×7676) were split up into two parts: a training set and a test set. From the total of 66 measurements performed (six different replicate flasks for each one of the eleven different hams), four of the six replicates were used in the calibration process of the models, leading to a training array $\underline{\mathbf{T_1}}$ (44×76×101) and training matrices $\mathbf{T_2}$ (44×76) and $\mathbf{T_3}$ (44×7676). The remaining 22 samples conformed the validation array $\underline{\mathbf{V_1}}$ (22×76×101) and validation matrices $\mathbf{V_2}$ (22×76) and $\mathbf{V_3}$ (22×7676) respectively.

4. Results and discussion

4.1. Exploratory data analysis

In order to determine trends on the data provided by the MS-Sensor, a previous exploratory analysis step was performed on the training set. A PCA model was used to assess tendencies in two-way training data matrices $\mathbf{T_2}$ (44×76) and $\mathbf{T_3}$ (44×7676). With a previous autoscaling of the data, a clear clustering according to pig feeding could be

observed in both cases. The main differences between both matrices consisted on the percentage of captured variance that decreased from 78% on the PCA performed on $\mathbf{T_2}$ (44×76) to 50% in the case of the unf-PCA on $\mathbf{T_3}$ (44×7676). The fact that autoscaling becomes necessary to obtain meaningful results gives the idea that the substances involved in the distinction of the feeding classes are present in low level concentrations compared to other common volatiles to both categories. Autoscaling of the data allows to all m/z variables to have the same opportunity of influencing the model. This approach is advantageous because low intensity m/z fragments involved in the underlying phenomenon we want to model are emphasized. On the other hand, using this approach, the noisiest m/z fragments are considered as important as that ones containing meaningful information.

In order to perform exploratory data analysis on the $\underline{\mathbf{T}_1}$ (44×76×101) array, the data matrix was decomposed using different number of factors using a PARAFAC model. Scaling within m/z mode was found to be necessary. Centering and scaling in three-way arrays is widely described by Bro and Smilde [21]. In three-way arrays scaling i.e. within the second mode is equivalent to multiply by the same scalar each one of the vertical slices of a three-way array. Non-negativity constraints were imposed in mode 2 and three in order to obtain a realistic solution because the spectra and time dimension should be positive. Residual analysis indicated that the optimal number of factors was three (78% of explained variance). The three component model shows much better clustering of samples according to the type of feeding than two-way models do as it is showed in scores plot comparison between the three models fitted (Figure 4). In all cases secondary effects such as ham origin or production process affect clustering since data is markedly influenced by these effects. This is much more pronounced in the case of the PLS-DA model and unf-PLS-DA where the data seems to be clustering according to producers instead of feeding type. Anyway, in all cases the third component allows to model the underlying phenomena we use for classification.

4.2 Classification according to pig's feeding

A 2-category classification was envisaged according to the type of feeding the pigs received. Prior to any calculation, data were scaled within the second mode (m/z) for the three-way array and autoscaled in the case of two-way matrices. A N-PLS-DA model was

ISBN:978-84-691-9752-3/DL:T-126-2009

used to calibrate data from the three way array $\underline{\mathbf{T_1}}$ (44×76×101), and PLS-DA models were constructed with the $\mathbf{T_2}$ data matrix (44×76) or the unfolded data matrix $\mathbf{T_3}$ (44×7676).

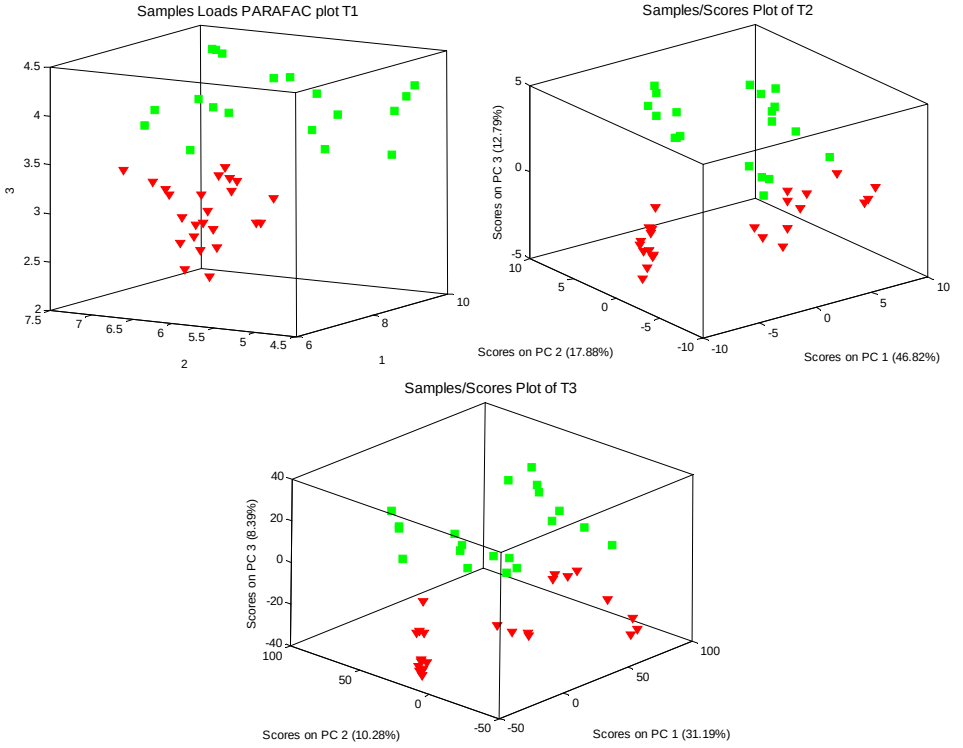

Figure 4: Scores plot of T1, T2 and T3 comparison: acorn (red spots), fodder (green spots)

The training phase of the models was evaluated using a cross-validation method. Performance during training was estimated using the root mean square error of the cross-validation (RMSECV). RMSECV tells us how well a given model fits the data in the calibration step and it is defined as in Eq.1:

$$\mathrm{RMSECV} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_{i,ref})^2}{n}} \qquad \text{[Eq. 1]}$$

where $\hat{\mathbf{y}}_\mathbf{i}$ are the values of the predicted classes with the samples left-out in each iteration of the cross-validation process, $\mathbf{y_{i,ref}}$ are scalars of vector composed by dummy variables defining current classes and *n* is the number of calibration samples (n=44). RMSECV was used to assess the optimal number of factors for model fitting (Figure 5).

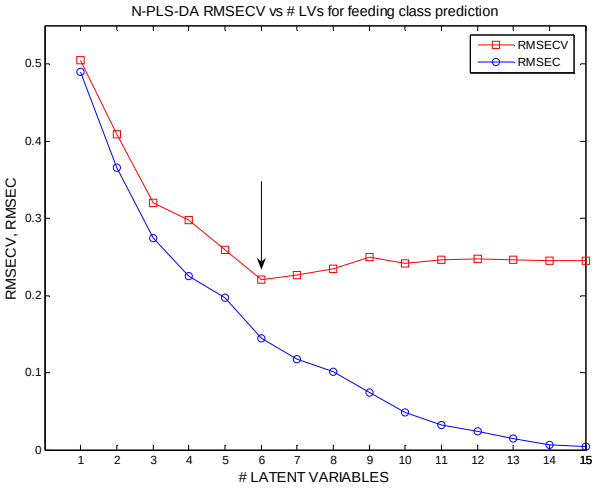

Figure 5: Assessing the optimal number of LVs from RMSECV vs #LV's plot in N-PLS-DA classifier model.

The calibration variance for X and Y-blocks is a measure of how well has the model been fitted to the training data set. This model fit was defined as it is shown in Eq 2:

$$\text{Var}\,(X_{cal}, X_{val}) = 100 \times \left(1 - \frac{SS_{mod}}{SS_{tot}}\right) \quad \text{.....[Eq. 2]}$$

where SS_{mod} corresponds to cross-validated residuals and SS_{tot} refers to the sum of squares of treated data. Cross-validated residuals were determined as the discrepancy between the fitted value and the real value left out in the full cross-validation routine.

Once calibration has been performed, a validation step is attempted using $\underline{\mathbf{V_1}}$ (22×76×101), $\mathbf{V_2}$ (22×76) and $\mathbf{V_3}$ (22×7676), which contain measurement data from the samples that belong to the test set (and therefore unseen by the models during training). Then, the prediction error (RMSEP) was estimated as stated in equation 3:

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_{i,ref})^2}{n}} \quad \text{[Eq. 3]}$$

where, this time, $\hat{\mathbf{y}}_\mathbf{i}$ represents the values of the predicted class of samples in test set, $\mathbf{y_{i,ref}}$ is the current class for these samples and n represents number of samples in the test set (n=22). RMSEP is an expression of the error that can be expected when using the calibrated model in future predictions. i.e., a measure of predictability of model. The percentage of

ISBN:978-84-691-9752-3/DL:T-126-2009

variance in validation is also accounted as in equation 4 where SS_{pred} now corresponds to the sum of the squared predicted residuals. The validation variance for the Y-block expresses how well the model will perform with new data. For the X-block it indicates how well the validation data have been projected onto NPLS-DA or PLS-DA models already fitted during training.

$$\text{Var}\,(X_{val}, X_{val}) = 100 \times \left(1 - \frac{SS_{pred}}{SS_{tot}}\right) \cdots \qquad \text{[Eq. 4]}$$

Table 1 shows the results obtained for each one of the models. Several conclusions can be drawn from this table. The most important conclusion is that N-PLS-DA is the simplest model (lowest number of LVs) and at the same time it shows the highest predictability in terms of RMSEP (since it has the lowest value). Although PLS-DA is able to fit with the highest percentage of variance in X-block, Y-validation variance decreases dramatically and that is translated to an increase of RMSEP. Thus, the PLS-DA model shows the lowest predictability . Fit in the X-block is better in the case of PLS-DA because of the higher flexibility of this model against to its trilinear generalization. The fit of a trilinear model will be lower per definition that the fit of the corresponding bilinear model because in N-PLS-DA any variation must be consistent over all scans. The PLS-DA model is often overly flexible and the increased fit to a large extent attributable to fitting the noise of the data and it is not directly translated to an increase on predictability as it can be seen in this particular case study. The unf-PLS-DA model even being a bilinear model shows the lowest fit in the X-block. The point is that, by unfolding, no relationship between different scans is imposed and in some way the structure of data is destroyed. It has to be kept in mind that an unfolded matrix holds all the scans acquired, even the ones that do not carry usable chemical information. Performing an autoscale on the entire matrix leads to put on the same level meaningful m/z variables from scans holding signal and from scans just holding noise. Therefore, $\mathbf{T_3}$ contains a lot of irrelevant information not related to the feeding type and that is why it is just fitting 7,83% of the variance in the X-block, which allows to describe 66,41% of the variance in Y as it is shown in Table 1. Even with the poor percentage of variance accounted in X-block at prediction, this model it is able to predict the feeding category of the samples better than a PLS-DA does in terms of RMSEP,

proving that the incorporation of the time dimension helps to improve the prediction ability of the models.

		% Explained variance				RMSEC		Success Rate
	# LV's	Xcal (Fit)	Xval (Val)	Ycal (Fit)	Yval (val)	RMSECV	RMSEP	
N-PLS-DA	6	72,03	71,61	91,52	77,93	0,220	0,289	100
PLS-DA	13	92,39	86,80	99,23	25,72	0,117	0,429	100
unf-PLS-DA	11	64,39	7,83	99,99	66,41	0,337	0,384	95

Table 1: Characteristics of the N-PLS-DA, PLS-DA and unf-PLS-DA models calibrated for the classification of ham samples according to the type of feeding.

The predicted y-value from the calibrated models results in a continuous variable that can be interpreted as a class similarity index. Each calculated class prediction value can be compared with a Bayesian distribution curve to determine for a given predicted y-value the probability that this value belongs to that original class. For each sample, either in prediction or in the test set, a table of probabilities is obtained (Figure 6).

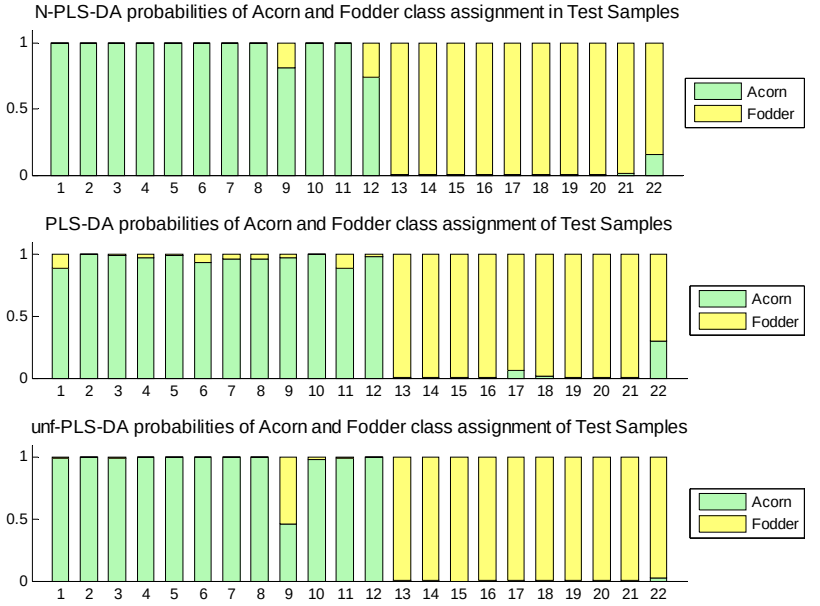

Figure 6: Table of probabilities from test-set samples in class assignment

ISBN:978-84-691-9752-3/DL:T-126-2009

The "predicted y-values" of the PLS-DA model are actually values around and between one (value assigned to acorn class) and two (value assigned to fodder class). Then, ideally, each sample which is acorn class would predict as a value of 1 (one) and each sample which is fodder class would predict as 2 (two) and the discrimination between the two would be simple. As this is often not the case, a threshold of "predicted y" must be determined above which a sample is considered to be a member of the class. These thresholds are also provided for the three models and class assignment is done comparing against this value. Figure 7 shows class predicted values for all the test set samples using the three models.

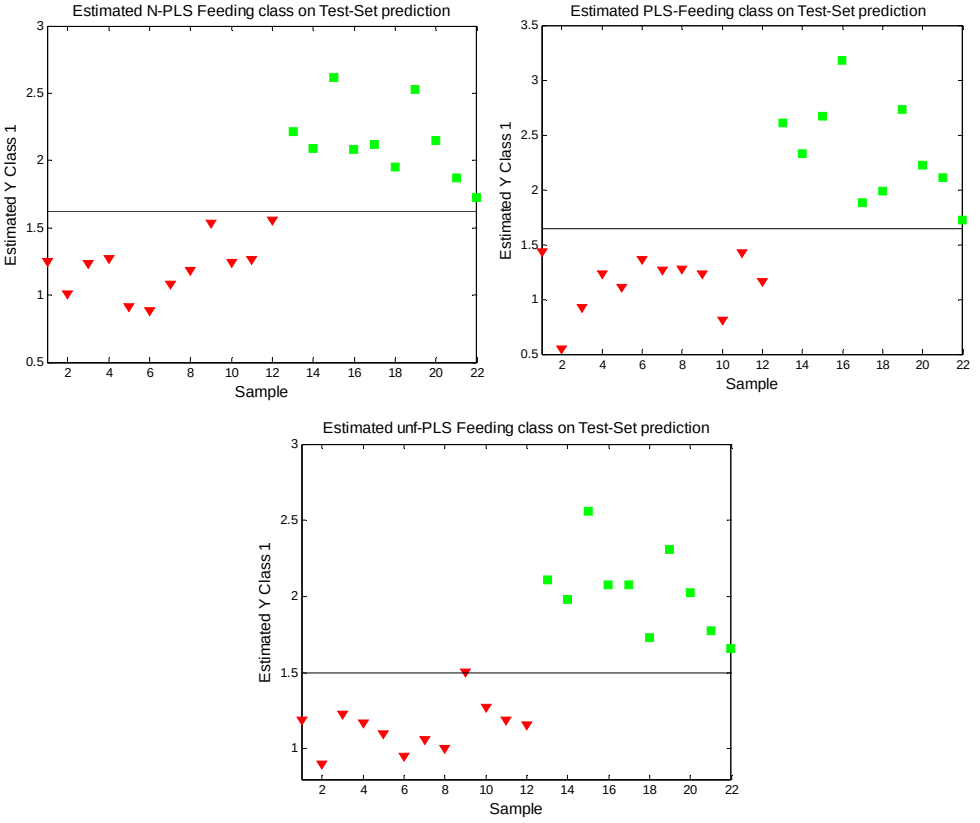

Figure 7: Acorn (red spots) or fodder (green spots) class prediction comparison between N-PLS-DA, PLSDA and unf-PLS-DA approaches for the test set samples.

The black horizontal line is the calculated threshold. As it can be seen, in the case of the unf-PLS-DA model there is a misclassified sample corresponding to sample 9. In fact, the classification of this sample has been done with a low certainity, as seen in figure S-6. Even

though the predictability of test set samples (i.e., the RMSEP) looks better for unf-PLS-DA than for PLS-DA, the former model presents a misclassified sample while the latter is able to classify all test samples in a correct way. That can be explained with the certainty of predictions. In order to compute the certainty of each model, values for reduced residuals Q and T^2 (Hottelling's) statistics are used. The larger the values, the less certain is the class assignement. Q statistics are calculated for the residual part and they simply express the goodness fit of the test data set. Fig 8 shows that 11 of the 22 samples of the test set prediction have very high values for Q in the case of unf-PLS-DA data, leading to the conclusion that this model is not modelling with enough confidence at least half of the samples on the test set; therefore, we can conclude that this approach is not able to generalize in a reliable manner. At the same time, and according to the explained variance in prediction, N-PLS-DA shows the best generalisation ability with the test set. T^2(Hottelling) expresses unusual variance of the data within the model, it is an indication of how far the projection of samples are from the multivariate mean. It can be interpreted as a measure of the leverage of samples, i.e.,the influence of samples inside the model. Trend on leverage is similar for all samples and for all models and extraordinary variations were not produced in leverage.

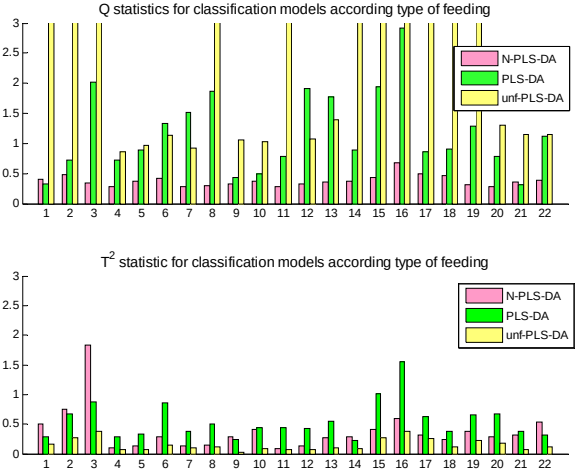

Figure 8: Value comparison between N-PLS-DA, PLS-DA and unf-PLS-DA for reduced residuals Q and T2 (Hottelling's) statistics

In general N-PLS-DA compares favourably to other models for several reasons. It is the simplest one, it is more stable than the rest since it fits data correctly, and it is able to

ISBN:978-84-691-9752-3/DL:T-126-2009

generalize to new samples. Moreover with this model the risk of overfitting is minimized and it shows the highest predictability.

4.3 Quantification of a_w, X and NaCl

This section deals with the prediction of a_w, *X* and NaCl content, benchmarking N-PLS, PLS and unf-PLS fitting and validation results. Fitting and validation were done as explained above. Table 2 shows comparative results between the three models for prediction of composition parameters. As stated before, RMSECV was used to assess the optimal number of LVs.

		a_w prediction		NaCl prediction		*X* prediction	
		Cross-validation	Test Set Validation	Cross-validation	Test Set Validation	Cross-validation	Test Set Validation
# LVs	N-PLS	8	-	9	-	7	-
	PLS	15	-	15	-	15	-
	unf-PLS	15	-	10	-	14	-
Correlation	N-PLS	0,71	0,63	0,92	0,88	0,89	0,88
	PLS	0,98	0,88	0,98	0,84	0,97	0,80
	unf-PLS	0,91	0,85	0,84	0,84	0,86	0,81
SECV, SEP	N-PLS	0,02	0,02	0,57	0,64	2,58	2,63
	PLS	0,01	0,03	0,30	1,01	1,48	3,74
	unf-PLS	0,02	0,02	0,80	0,91	2,69	2,67
%Var X-block	N-PLS	75,32	75,09	74,14	74,10	75,64	75,54
	PLS	94,10	88,09	94,52	86,87	94,01	87,85
	unf-PLS	70,06	7,96	62,95	7,08	54,14	7,43
%Var Y-block	N-PLS	99,94	96,43	98,86	98,29	99,88	99,56
	PLS	99,39	40,92	99,48	52,77	99,48	55,86
	unf-PLS	100,00	68,87	63,42	62,95	99,69	76,50

Table 2: Characteristics of the N-PLS, PLS and unf-PLS models used for the prediction of composition parameters such as aw, NaCl and X.

The trend observed in the classification models is also seen in their quantification counterparts. Trilinear models are constructed using fewer factors and, hence, they are simpler and more parsimonious than their corresponding bilinear models. For this reason, better performance is expected using three-way models. Bias in cross-validation prediction

and in test-set prediction was checked by building a joint confidence interval of slope and intercept of the predicted values vs actual values. The bias was calculated statistically significant at a 95% confidence level (Figure 9). Bias is a measure of the accuracy of the model. In the absence of bias, RMSEP equals to SEP (Standard Error Performance) that expresses the precision of the result corrected for the bias, so we continue to focus in SEP to compare the predictability of the methods. The same holds for the calibration step, where SECV (Standard Error of Cross-validation) is equal to RMSECV. In any case, bias does not exist in cross-validation steps, but when predicted measures are compared with actual values in the validation set, accuracy of un-PLS-DA models decreases in prediction leading to a presence of bias in the case of NaCl and *X* predictions (values 0.02 and 0.80 respectively).

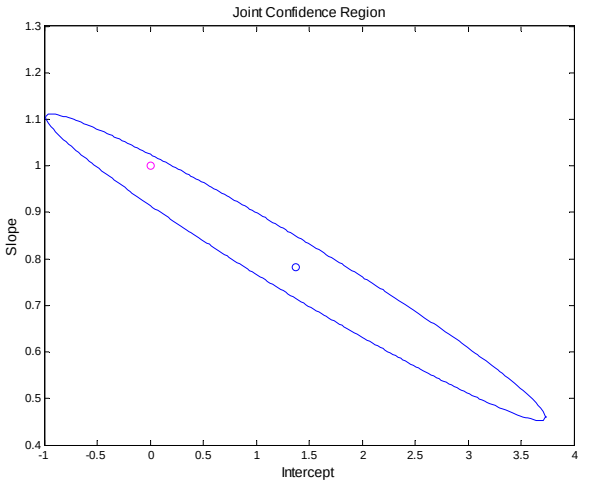

Figure 9: Joint confident interval of slope and intercept for aw prediction of test-set samples that was showed not being statistically significant at 95% confidence level

The PLS model shows the lowest SECV value and at the same time for the same model SEP gets the highest values. These facts indicate that PLS models tend to overfit data during the calibration process and afterwards they are not able to generalize the prediction as N-PLS does. PLS models present the lowest predictability in terms of Y-block variance explained in the test set. Unf-PLS models follow in the ranking. Finally, N-PLS is shown to be the model with highest predictability. These results are similar to the ones obtained with the classifier models, since the trilinear model is much more difficult to overfit since any

ISBN:978-84-691-9752-3/DL:T-126-2009

variation incorporated in the model must be consistent over all scans. For the trilinear model it can be seen that the trend is to preserve the variance for both the X-block and Y-block in the prediction process. Again, N-PLS is more stable and the value of the explained variance in the X-block is more or less the same both in the calibration and prediction steps. Referring to unf-PLS models, an important point to take in to account related to NaCl prediction, for example, is that just a 7,08% of variation in the X-block is needed to predict 62,95% of salt percentage. These results give clear evidence that a lot of irrelevant and redundant information exists in the X-variable descriptors with respect to Y. This is a common issue when using multivariate methods in unfolding matrices and it points to the fact that interpretation of unfolded models should be executed with great care. The same trend is observed for the remaining two predicted parameters.

To better assess the accuracy of the different PLS models, regression between actual and predicted values for a_w, salt and *X* content was performed. The correlation coefficient, slope and intercept were calculated (Table 2 shows only the correlation coefficient). A perfect regression would yield 1 for the correlation coefficient and slope and 0 for the intercept. Acceptable correlation coefficients have been achieved for NaCl and X. In both cases better regression coefficient for tests measurements are achieved through the use of N-PLS-DA models. Only in the case of a_w prediction, other methods overcome N-PLS in terms of the correlation coefficient. Nevertheless, the RMSEP value is the lowest. Note that RMSEP is the averaged error composed of large and small errors altogether. The reason for this discrepancy might be that some of the badly predicted test samples, which show a big difference between predicted and actual values, might be those responsible from introducing artefacts in the linearity behaviour of that correlation. These effects are not reflected in RMSEP because it is an averaged error figure.

5. Conclusions

From the results derived from this study we can conclude that, in this particular case, pattern recognition on MS-Sensor data clearly beneficiates from including temporal information either for quantification or classification purposes, since improved quantification and classification can be observed by using three-way methods.

Nevertheless, the way in which temporal information is introduced into the model is an important issue. In this context and in the particular case under study it has been demonstrated that the data generated by the MS-Sensor is better arranged in three way arrays rather than in an unfolded way. The main reason for this is the lack of trilinear constraints in the unf-PLS model, since the information across scans (temporal information) is not used to stabilize the solution. Unfolding methods lead to more unfavourable results for several reasons such as producing complex models, with many parameters which increase the risk of poor predictive capability, and the generation of less robust, interpretable and parsimonious models.

Therefore, it is reasonable to use multi-way methods as they have the potential to simplify the interpretation of the results and provide more adequate and robust models using fewer parameters.

The specific results about ham quality classification or the prediction of composition parameters obtained by the different models built can be a little optimistic, since replicates of the same ham have been always used in the models. So, to some extend, we have been evaluating how well a sample predicts its own replicate. However, since the same strategy was employed to assess the performance of N-PLS-DA, unf-PLS-DA and PLS-DA, the results showing that N-PLS-DA leads to more parsimonious models with higher prediction ability remains fully consistent.

Acknowledgements

This work has been supported by Redes de Centros en Metabolismo y Nutrición C03/08.

REFERENCES

[1] Pérez, J. L.; Nogal, M. d.; García, C.; Fernández, M. E.; Moreno, B.; Guerrero, A. *Trends in Analytical Chemistry* **2006**, *25*, 257-266.
[2] Pérès, C.; Begnaud, F.; Eveleigh, L.; Berdagué, J.-L. *Trends in Analytical Chemistry* **2003**, *22* 858-866.
[3] Marsili, R. In *Flavor, Fragance and Odor Analysis*; Marsili, R., Ed.; Marcel Dekker, Inc.: New York, 2002, pp 349-374.
[4] Dittmann, B.; b, S. N. **2000**, *69*, 253-257.
[5] Boudaoud, N.; Eveleigh, L. *J. Agric. Food Chem.* **2003**, *51*, 1530-1533.
[6] Smilde, A.; Bro, R.; Geladi, P. *Multi-way Analysis*; John Wiley & Sons Ltd.: Chichester, 2004.
[7] Wong, J. W. H.; Durante, C.; Cartwright, H. M. *Anal. Chem.* **2005**, *77*, 5655-5661.

ISBN:978-84-691-9752-3/DL:T-126-2009

[8] Timón, M. L.; Martín, L.; Petrón, M.; Jurado, Á.; García, C. *J. Sci. Food Agric* **2001**, *82*, 191.
[9] Pastorellia, G.; Magnia, S.; Rossia, R.; Pagliarinib, E.; Baldinic, P.; P. Dirinckd; Opstaelee, F. V.; Corinoa, C. *Meat Science* **2003**, *65* 571–580.
[10] Carrapiso, A. I.; Bonilla, F.; García, C. *Meat Science* **2003**, *65* 623–629.
[11] García, M.; Horrillo, M. C.; Santos, J. P.; Aleixandre, M.; Sayago, I.; Fernández, M. J.; Arés, L.; Gutiérrez, J. *Sensors and Actuators B* **2003**, *96*, 621–629.
[12] Santos, J. P.; García, M.; Aleixandre, M.; Horrillo, M. C.; Gutiérrez, J.; I. Sayago; Fernández, M. J.; Arés, L. *Meat Science* **2004**, *66* 727–732.
[13] Otero, L.; Horrillo, M. C.; García, M.; Sayago, I.; Manuel Aleixandre; Fernández, M. J.; Arés, L.; Gutiérrez, J. *Meat Science* **2003**, *65*, 1175–1185.
[14] González-Martín, I.; Pérez-Pavón, J. L.; Claudio González-Pérez; Hernández-Méndez, J.; Álvarez-García, N. *Analytica Chimica Acta* **2000**, *424*, 279–287.
[15] Nielsen, N.-P. V.; Carstensen, J. M.; Smedsgaarda, J. *J. Chromatography A* **1998**, *805* 17–35.
[16] Bylund, D.; Danielsson, R.; Malmquist, G.; Markides, K. E. *J. Chromatography A* **2002**, *961* 237–244.
[17] Johnson, K. J.; Wright, B. W.; Jarman, K. H.; Synoveca, R. E. *J. Chromatography A* **2003**, *996*, 141–155.
[18] Tomasi, G.; Berg, F. v. d.; Andersson, C. *J. Chemometrics* **2004**, *18*, 231–241.
[19] Wong, J. W. H.; Durante, C.; Cartwright, H. M. *Anal. Chem.* **2005**, *77*, 5655-5661.
[20] Barker, M.; Rayens, W. *Journal of Chemometrics* **2003**, *17*, 166-173.
[21] Bro, R. *Journal of chemometrics* **1996**, *10*, 47-67.
[22] Hall, G. J.; Clow, K. E.; Kenny, J. E. *Environ. Sci. Technol.* **2005**, *39*, 7560-7567.
[23] Guimet, F.; Boqué, R.; Ferré, J. *Chemom. Intell. Lab. Syst.* **2006**, *81*, 94 – 106.
[24] Fenaille, F.; Visani, P.; Fumeaux, R.; Milo, C.; Guy, P. A. *J. Agric. Food Chem.* **2003**, *51*, 2790-2796.
[25] Andrés, A. I.; Cava, R.; Ventanas, J.; Thovar, V.; Ruiz, J. *Meat Science* **2004**, *68*, 45–51.
[26] Ramírez, J. L. R., Universitat Autónoma de Barcelona, Bellaterra, 2005.
[27] Serra, X.; Ruiz-Ramírez, J.; Arnau, J.; Gou, P. *Meat Science* **2005**, *69*, 249-254.
[28] Martí, M. P.; Pino, J.; Boqué, R.; Busto, O.; Guasch, J. *Anal. Bioanal. Chem.* **2005**, *382*, 440-443.
[29] S, S.; J, L.; AZ, B.; EA, V.; C, D. N.; BM, N. *Postharvest biology and Technology* **2004**, *31*, 9-19.
[30] Martí, M. P.; Boqué, R.; Riu, M.; Busto, O.; Guasch, J. *Anal. Bioanal. Chem.* **2003**, *376*, 497-501.
[31] Gardner, W. P.; Shaffer, R. E.; Girard, J. E.; Callahan, J. H. *Anal. Chem.* **2001**, *73*, 596-605.
[32] Berman, N. C., New York 1976.
[33] www.quimica.urv.es/quimio/ang/maincat.html.