

UPC

UNIVERSITAT POLITÈCNICA DE CATALUNYA

DEPARTAMENT DE TEORIA DEL SENYAL I COMUNICACIONS

**Diseño de una nariz electrónica
para la determinación no
destructiva del grado de
maduración de la fruta**

AUTOR:

JESUS BREZMES LLECHA

DIRECTOR:

XAVIER CORREIG BLANCHAR

A mi mujer e hija

AGRADECIMIENTOS

En primer lugar quiero agradecer a mi familia todo el apoyo recibido durante mi etapa académica, en especial a mis padres y hermano. Sin su ánimo y cariño jamás habría finalizado mi carrera universitaria y por lo tanto no estaría en condiciones de defender mi tesis doctoral.

El Dr. Xavier Correig, director de esta tesis, siempre me ha apoyado tanto a nivel profesional como personal durante la realización de este trabajo. Lejos de ser una relación director-alumno, he encontrado en sus consejos y comprensión a un buen amigo y compañero de trabajo. Gracias, por lo tanto, por esa amistad y relación profesional sin la que estoy seguro que este trabajo no hubiese llegado a buen término.

A Eduard Llobet y Xavier Vilanova, compañeros de mi mismo grupo de investigación, les tengo que agradecer todo su paciencia y apoyo durante los años que ha durado este trabajo, además de las sugerencias realizadas a los artículos y memoria que conforman esta tesis doctoral. Al resto de compañeros del DEEEA, a los que no nombro por miedo a dejarme a alguien, debo agradecerles su amistad y compañerismo que han hecho de mi vida profesional una experiencia muy positiva.

Otro de los efectos positivos de esta tesis ha sido conocer y trabajar con mis compañeras de la Universidad de Lleida. Querría recalcar, además de su profesionalidad y dedicación, su disponibilidad a ayudarme en cualquier duda que me surgiera en todos los aspectos relacionados con el estudio de la fruta, conocimientos algo lejanos a mi formación como ingeniero de telecomunicación. Gracias por la ayuda incondicional que me habeis prestado todos estos años.

A Gabriel Junyent, mi tutor en TSC, quiero agradecerle su apoyo, paciencia y disponibilidad. También me gustaría destacar la confianza inicial y financiación aportada por la empresa “Catalunya Qualitat”. Para finalizar, quiero agradecer la colaboración de Marcos Yague, Beni Vazquez, Elisenda Roca y de todos aquellos que, de alguna manera u otra, me han ayudado a que esta tesis se hiciera realidad.

Este trabajo ha sido posible gracias a la financiación del proyecto CICYT-FEDER No. 2FD-97-0436 (TIC) con financiación adicional aportada por la empresa “Catalunya Qualitat”

ÍNDICE

1. INTRODUCCIÓN	1
REFERENCIAS.....	10
2. TÉCNICAS DE CONTROL DE CALIDAD EN FRUTA.....	13
2.1 INTRODUCCIÓN	13
2.2 MEDIDAS FÍSICO-QUÍMICAS.....	15
2.2.1 Firmeza.....	16
2.2.2 Análisis de Sólidos Solubles.....	17
2.2.3 Análisis de Acidez.....	19
2.2.4 Colorimetría.....	20
2.2.5 Índice de almidón.....	21
2.3 MEDICIÓN DE VOLÁTILES	23
2.3.1 Medidas aromáticas.....	23
2.3.2 Producción de etileno	26
2.4 TECNICAS EXPERIMENTALES.....	27
2.4.1 Mediciones con ultrasonidos	28
2.4.2 Espectrometría de infrarojo cercano.....	29
REFERENCIAS.....	31
3. DISEÑO DE SISTEMAS DE OLFATO ELECTRÓNICO	35
3.1 INTRODUCCIÓN	35
3.2 PROTOTIPO DE INYECCIÓN	37
3.2.1 Descripción general.....	37
3.2.2 Sistema de muestreo.....	39
3.2.2.1 Cámara de concentración	39
3.2.2.2 Cámara de medida	39
3.2.2.3 Jeringa cromatográfica	40
3.2.3 Matriz de sensores	41
3.2.3.1 Los sensores de gases	41
3.2.3.2 Sensores Figaro	43
3.2.4 Circuitos de polarización y medida	47
3.2.5 Sistema de adquisición y procesado	49
3.3 PROTOTIPO DE FLUJO CONTINUO.....	50
3.3.1 Descripción general.....	50
3.3.2 Sistema de muestreo.....	52
3.3.2.1 Cámara de concentración	52
3.3.2.2 Cámaras de medida.....	53
3.3.3 Matrices de sensores.....	55
3.3.3.1 Sensores FIS serie SP	58
3.3.3.2 Sensores FIS serie SB.....	60
3.3.4 Circuitos de polarización y medida	62
3.3.5 Sistema de adquisición y procesado	65
REFERENCIAS.....	67

4. PROCESADO DE SEÑAL.....	69
4.1 INTRODUCCIÓN	69
4.2 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)	73
4.2.1 <i>Introducción</i>	73
4.2.2 <i>Algoritmo</i>	74
4.2.3 <i>Implementación</i>	77
4.2.3.1 <i>Programación</i>	77
4.2.3.2 <i>Mejoras del algoritmo</i>	78
4.2.3.3 <i>Evaluación de resultados</i>	79
4.2.4 <i>Otras aplicaciones</i>	80
4.3 PARTIAL LEAST SQUARES	82
4.3.1 <i>Introducción</i>	82
4.3.2 <i>Algoritmo</i>	83
4.3.3 <i>Implementación</i>	84
4.3.3.1 <i>Programación</i>	84
4.3.3.2 <i>Aplicación del método a las medidas de fruta</i>	85
4.3.3.3 <i>Evaluación de resultados</i>	86
4.4 REDES NEURONALES FEEDFORWARD.....	87
4.4.1 <i>Introducción</i>	87
4.4.2 <i>Algoritmo</i>	88
4.4.2.1 <i>Estructura</i>	88
4.4.2.2 <i>Aprendizaje</i>	90
4.4.3 <i>Implementación</i>	91
4.4.3.1 <i>Programación</i>	91
4.4.3.2 <i>Aplicación a las medidas con fruta</i>	92
4.4.3.3 <i>Evaluación de resultados</i>	92
4.5 REDES FUZZY ART	93
4.5.1 <i>Introducción</i>	93
4.5.2 <i>Algoritmo</i>	95
4.5.3 <i>Implementación</i>	96
4.5.3.1 <i>Programación</i>	96
4.5.3.2 <i>Evaluación de resultados</i>	97
4.6 RED FUZZY ARTMAP	98
4.6.1 <i>Introducción</i>	98
4.6.2 <i>Algoritmo</i>	100
4.6.3 <i>Implementación</i>	102
4.6.3.1 <i>Programación</i>	102
4.7 PREPROCESADO	104
4.7.1 <i>Introducción</i>	104
4.7.2 <i>Extracción de parámetros</i>	104
4.7.3 <i>Normalización de los datos</i>	107
4.7.4 <i>Selección de variables</i>	110
4.7.4.1 <i>Criterio de sensibilidad</i>	111
4.7.4.2 <i>Por covarianza entre variables</i>	112
4.7.4.3 <i>Algoritmos genéticos</i>	114
REFERENCIAS.....	117
5. MEDIDAS CON MANZANAS.....	121
5.1 INTRODUCCIÓN	121
5.2 ESTUDIO DE VIABILIDAD.....	122
5.2.1 <i>Objetivos</i>	122

5.2.2 Metodología experimental	123
5.2.3 Evolución de la señal	125
5.2.4 Resultados	126
5.2.5 Conclusiones	131
5.3 MEDIDAS EN FLUJO CONTINUO	132
5.3.1 Objetivos	132
5.3.2 Metodología experimental	133
5.3.3 Seguimiento del proceso de maduración	134
5.3.4 Correlación entre parámetros destructivos y señales de los sensores del prototipo	137
5.3.5 Estudio de las derivas de los sensores	139
5.3.6 Conclusiones	145
REFERENCIAS	146
6. MEDIDAS CON PERAS	147
6.1 INTRODUCCIÓN	147
6.2 ESTUDIO DE VIABILIDAD	148
6.2.1 Objetivos	148
6.2.2 Metodología experimental	148
6.2.3 Evolución de la señal	150
6.2.4 Resultados	150
6.2.5 Conclusiones	154
6.3 MEDIDAS EN FLUJO CONTINUO	154
6.3.1 Objetivos	154
6.3.2 Metodología experimental	155
6.3.3 Clasificación no supervisada	156
6.3.4 Correlación entre indicadores de calidad y señales de la matriz de sensores	161
6.3.5 Conclusiones	164
REFERENCIAS	166
7. MEDIDAS CON MELOCOTONES Y NECTARINAS	167
7.1 INTRODUCCIÓN	167
7.2 ESTUDIO DE VIABILIDAD	167
7.2.1 Objetivos	167
7.2.2 Metodología experimental	168
7.2.3 Evolución de la señal	169
7.2.4 Resultados	172
7.2.5 Conclusiones	177
7.3 MEDIDAS EN FLUJO CONTINUO	178
7.3.1 Objetivos	178
7.3.2 Metodología experimental	178
7.3.3 Clasificación no supervisada	180
7.3.4 Correlación entre indicadores de calidad y señales de la matriz de sensores	183
7.3.5 Conclusiones	189
REFERENCIAS	190
8. CONCLUSIONES	191

CAPÍTULO 1:

INTRODUCCIÓN

Estos últimos años hemos asistido a una profunda reestructuración en la industria alimentaria. Durante la mayor parte del siglo XX los esfuerzos en investigación y desarrollo fueron encaminados a optimizar la producción desde un punto de vista cuantitativo. Actualmente, la existencia de excedentes de producción y la excesiva competencia ha propiciado un cambio de mentalidad en el que se prima la calidad por encima de la cantidad. Por ese motivo, el control de calidad de los alimentos está adquiriendo una importancia capital. Además, se trata de una actividad que, lejos de ser puntual, se integra cada vez más en todo el proceso productivo, abarcando desde el inicio de la producción hasta la distribución del producto final en los mercados de consumo.

Este cambio de mentalidad, si bien es general en toda actividad relacionada con la alimentación, es crítico en las producciones frutícolas. La competencia de países con mano de obra barata es muy fuerte y los costes de producción en los países desarrollados no paran de crecer. Como consecuencia, el precio final en los mercados es elevado y la viabilidad (venta) del producto solo se justifica si va acompañado de una garantía de calidad.

En este sentido, el proceso de comercialización de la variedad “pinklady” de manzana [1.1] (una de las variedades estudiadas en este trabajo) es un buen ejemplo de la forma de entender la agricultura que a buen seguro imperará en el siglo XXI. Su precio de mercado es muy superior al que se paga por otras variedades, ofreciendo a cambio una inmejorable calidad cuidadosamente estandarizada que solo se ha conseguido a base de

monitorizar y normalizar el proceso de producción, controlando desde la plantación del manzano hasta la recolección en la fecha considerada óptima para cada campaña.

En el caso de la fruta, la calidad del producto final depende de muchos factores, algunos ambientales (y por lo tanto poco controlables) y otros directamente relacionados con la gestión del producto (fecha de recolección, tiempo y tipo de conservación en cámara, etc). De todas maneras, desde el punto de vista del consumidor, el grado de maduración es el factor que más se relaciona con la calidad de la fruta [1.2]. Por este motivo la determinación de la fecha de recolección es crucial para que la fruta llegue al consumidor en su estado óptimo de maduración. Para dar respuesta a esta necesidad, diferentes métodos para determinar el estado de maduración han sido propuestos [1.3]. De ellos, destacamos los siguientes:

- **Penetromía:** Método mecánico de medición de la consistencia de la pulpa. Es una técnica rápida, barata y sencilla, aunque inevitablemente significa la destrucción de la muestra. Su relación con la maduración es más clara en algunas variedades de fruta que en otras.
- **Reflectometría:** Método que consiste en analizar el grado de glucosa de la pulpa. Esta técnica presenta las mismas ventajas e inconvenientes que la anterior. De todas maneras, su relación con la madurez de la fruta es menos clara que en el caso de la penetromía.
- **Cromatografía (líquida o gaseosa):** Es un método de análisis muy exacto pero que presenta un conjunto de inconvenientes en su aplicación práctica. La preparación previa de las muestras a analizar, la complejidad de uso de los cromatógrafos y de interpretación de los resultados, el elevado coste del equipo y la no obtención de resultados en tiempo real hacen que esta técnica sea poco utilizada. De todas formas, es adecuada en estudios de investigación sobre valoraciones de calidad ante diferentes formas de procesar la fruta. De hecho, se puede afirmar que en la actualidad, para trabajos científicos, el estudio de producción aromática como indicador de calidad es un campo de investigación muy activo [1.4,5].
- **Colorimetría:** Técnica en la que la maduración de la fruta se asocia a cambios en su coloración. El análisis objetivo mediante el color requiere la utilización de

equipos que incluyen sensores ópticos bastante complejos, con un coste bastante elevado [1.6,7].

- Impedancia acústica: Se basa en la medida de la impedancia acústica de la fruta que se quiere analizar. Algunos autores han demostrado cómo la absorción de ultrasonidos varía de forma característica con el grado de maduración. El principal problema de este método consiste en su escasa aplicabilidad en la industria ya que es difícil conseguir un buen acople acústico entre la fuente de ultrasonidos y la fruta analizada sin estropearla. En ambientes de laboratorio se ha ensayado la utilización de tanques de inmersión con resultados desiguales [1.8,9].
- La espectrometría de infrarrojo cercano es una técnica que caracteriza la reflectancia de la fruta ante diferentes longitudes de onda. Este espectro varía a lo largo del proceso de maduración y puede ser utilizado para predecir algunos indicadores de calidad, aunque no siempre tiene suficiente resolución como para seguir el proceso de maduración [1.10]. Además, requiere de un equipo costoso que debe ser calibrado frecuentemente con un láser.

Como se puede deducir de estas breves descripciones, cada uno de estos métodos presenta algún inconveniente que dificulta su aplicación práctica. De hecho, ninguna de las técnicas anteriores puede considerarse un método de medición que caracterice completamente el grado de maduración de la fruta, por lo que la búsqueda de nuevos indicadores que midan de forma fiable la calidad del producto de acuerdo con el punto de vista del consumidor final sigue siendo una actividad de gran interés en la industria frutícola.

A pesar de su importancia, debido a las dificultades que presentan las técnicas actuales, la determinación del momento óptimo de recolección suele dejarse en manos de operarios cuya decisión está basada en la experiencia práctica o en criterios de distribución. Esta inexactitud es la causa de que cantidades importantes de fruta lleguen al consumidor con una calidad inferior a la esperada dado su inadecuado estado de maduración. Y lo que es más importante, fruta de la misma variedad cosechada en la

misma zona llega al consumidor con niveles de calidad muy dispares, lo que dificulta asociar una determinada variedad de fruta a una calidad definida que permita consolidar su valor de mercado.

En definitiva, la industria frutícola busca todavía una técnica que le permita, de forma práctica, rápida, barata y fiable, determinar el grado de maduración de la fruta y poder así servir un producto de la mejor calidad posible al consumidor. En esta búsqueda es donde una nariz electrónica podría dar respuesta a la mayoría de necesidades que presenta su comercialización.

Los sistemas de olfato electrónico, popularmente conocidos como “narices electrónicas”, tienen una historia muy reciente. Aunque sus orígenes se remontan a los primitivos sistemas montados en los sesenta [1.11], el término “nariz electrónica” y los primeros sistemas inteligentes de olfato electrónico no aparecieron hasta la segunda mitad de la década de los ochenta. Se puede afirmar, de todas maneras, que el interés por este tipo de sistemas y la investigación en olfato electrónico dio un salto espectacular en la primera mitad de los noventa. Es en esta etapa en la que se acuña la definición comúnmente aceptada hoy en día de sistema de olfato electrónico: “instrumento que comprende una matriz de sensores químicos con sensibilidades parcialmente solapadas y algoritmos inteligentes capaces de reconocer aromas simples o complejos” [1.12].

Las narices electrónicas han encontrado en la industria alimenticia uno de sus campos naturales de actuación. Numerosos estudios pusieron de manifiesto su utilidad en aplicaciones tan diversas como la clasificación de diferentes tipos de licores [1.13], cafés [1.14], cosechas de vino [1.15], etc. De todas maneras, las aplicaciones comerciales con más futuro son aquellas en las que se utiliza el sistema de olfato electrónico para monitorizar el estado de diferentes productos alimenticios, bien sea durante su producción o durante su conservación [1.16-18].

Es precisamente en el campo de la fruticultura en el que los sistemas de olfato electrónico podrían convertirse en un instrumento ideal para monitorizar de forma no destructiva el estado de la fruta ya que una de las características propias del proceso de

maduración es la emisión de compuestos volátiles orgánicos. Así, por ejemplo, el periodo en el cual la fruta climatérica madura coincide con un incremento de etileno que se difunde desde el interior del fruto hacia el exterior. Una vez la fruta ha llegado a su máximo grado de maduración, se produce un decaimiento en las emisiones del mismo. Este compuesto viene acompañado de multitud de compuestos volátiles aromáticos (más de 50 en algunas variedades), los cuales presentan cambios muy significativos durante el proceso de maduración de la fruta [1.19-20]. Podemos decir, por lo tanto, que la consecución de un equipo capaz de analizar dichos volátiles en tiempo real podría ser muy útil para la monitorización del proceso de maduración.

A priori, un sistema de olfato electrónico podría aportar importantes ventajas respecto a técnicas más convencionales como la cromatografía. Entre las ventajas que podría aportar un instrumento de estas características se pueden destacar las siguientes:

- Análisis no destructivo
- Obtención de resultados en tiempo real (en cuestión de minutos)
- Portabilidad, robustez y bajo precio del equipo
- Adaptación a diferentes cantidades y variedades de fruta
- Facilidad de uso del equipo por parte de personal no cualificado

Dentro de las diferentes tareas relacionadas con la producción de fruta, un sistema de olfato electrónico podría integrarse en las siguientes actividades:

- Determinación de fecha óptima de recolecta: Tanto si la fruta se destina al consumo directo como si se destina a cámaras de conservación siempre existe un grado óptimo dentro del proceso de maduración en el que debe recolectarse para que llegue con la mayor calidad posible al consumidor final.
- Control de calidad: La emisión de componentes aromáticos por parte de la fruta está directamente relacionada con sus características organolépticas [1.21]. Una rápida inspección de la emisión de volátiles del género que entra en un centro de almacenamiento o distribución permitiría valorar la calidad de la fruta que

suministra el productor, premiando aquellas partidas con mayor “densidad” aromática.

- Control en cámaras: En toda fruta climatérica el etileno es una sustancia íntimamente relacionada con el proceso de maduración. Es tanto un indicador de que el proceso de maduración está avanzando, como una sustancia que puede activar dicho proceso. La excesiva concentración de etileno en una cámara puede ser indicativo de un estado de maduración avanzado que aconseje distribuir inmediatamente el producto para evitar su pérdida. Además, su aparición puede iniciar una reacción en cadena, ya que su presencia acelera la maduración y ésta, a su vez, produce etileno en grandes cantidades.
- Determinación del estado de maduración en el supermercado: Un usuario podría seleccionar la fruta que compra comprobando su perfil aromático para asegurar su calidad y estado de maduración.

Dado el gran abanico de posibilidades que ofrece esta tecnología en el mundo frutícola no es de extrañar que ya se hayan presentado algunos trabajos que exploran estas posibilidades. De ellos, el de Benady [1.22] puede considerarse como uno de los pioneros. Durante la realización de esta tesis doctoral, otros trabajos se han adentrado en este tipo de aplicaciones [1.23-24], aunque ninguno es lo suficientemente completo como para evaluar objetivamente la posibilidad de utilizar narices electrónicas en el campo de la fruta.

Esta tesis doctoral pretende evaluar esta posibilidad con el estudio más amplio realizado hasta la fecha sobre la utilización de una nariz electrónica para determinar el estado de maduración de fruta climatérica. Tras realizar unos primeros estudios de viabilidad, construyendo un prototipo de laboratorio inicial que será probado en diferentes variedades de fruta, se pasará a realizar un estudio mucho más exhaustivo utilizando un segundo equipo diseñado con la experiencia de los primeros resultados. Para poder evaluar objetivamente la posibilidad de utilizar un equipo de olfato electrónico para el control de calidad de la fruta se compararán los resultados obtenidos con el prototipo

con los que se obtienen con técnicas tradicionales de control de calidad. En definitiva, los objetivos principales de este trabajo pueden ser resumidos en tres puntos:

- Construir y perfeccionar un prototipo de laboratorio para evaluar la viabilidad de utilizar un equipo de olfato electrónico como instrumento de determinación del grado de maduración de la fruta de forma no destructiva.
- Buscar si existe una buena correlación entre la señal de la nariz electrónica y la maduración de la fruta, comparando los indicadores de calidad de fruta con las señales de la matriz de sensores del prototipo.
- Realizar un estudio con variedades climatéricas de fruta. En concreto, manzanas, peras, melocotones y nectarinas, todas ellas variedades climatéricas, fueron las variedades escogidas debido a su interés comercial en la zona próxima a los laboratorios en los que se ha realizado el estudio.

La implementación de sistemas de olfato electrónico para una aplicación particular requiere adaptar cada uno de los módulos que componen una nariz electrónica genérica a dicha tarea. De todos ellos, uno de los más importantes es el módulo que procesa la señal que proviene de la matriz de sensores. La misma definición de sistema de olfato electrónico deja patente su importancia y es una de las piezas angulares de este proyecto.

Aunque existen multitud de algoritmos de reconocimiento de patrones, muchas veces es necesario adaptarlos a las características de los sistemas de olfato electrónico en general y de sus aplicaciones en particular. Esta necesidad es la razón de que el procesado de señal en sistemas de olfato electrónico sea objeto de muchos trabajos de investigación [1.25] y es de opinión general que queda un largo camino por recorrer hasta que se perfeccionen algoritmos que puedan explotar al máximo la información suministrada por la matriz de sensores y eliminar interferencias como la humedad, temperatura y derivas temporales, inconvenientes presentes en la mayoría de sensores utilizados en los sistemas de olfato electrónico actuales.

En este proyecto ha sido necesario utilizar algoritmos para clasificar, cuantificar y correlar medidas realizadas con los prototipos y con técnicas de calidad en fruta. En la

mayoría de los casos su uso directo no ha sido posible y ha sido necesario adaptarlos a cada aplicación. Las modificaciones van desde simples alteraciones de las rutinas de representación de resultados hasta modificaciones substanciales del funcionamiento interno de algún algoritmo neuronal. De hecho, la modificación, adaptación, aplicación a los objetivos perseguidos y evaluación de resultados utilizando diferentes algoritmos es una de las tareas más novedosas e importantes realizadas en el marco de este proyecto.

Todo el trabajo realizado ha sido posible gracias a la financiación por parte de dos proyectos de investigación. En concreto, el “Comité Econòmic de la fruita de Lleida” financió un primer año de estudios bajo el proyecto titulado “Diseño de una nariz electrónica para la determinación no destructiva del grado de maduración de la fruta dulce”, con el que se pudo completar el estudio de viabilidad utilizando un primer prototipo. Posteriormente, a través del ministerio de Ciencia y Tecnología, el estudio tuvo continuidad a través del proyecto CICYT- 2FD-97-0436 (TIC), de dos años de duración y en el que el “Comité Econòmic de la Fruita” intervino como entidad colaboradora.

En este segundo proyecto, el “Area de Postcollita” del centro de investigación mixto “Universitat de Lleida” (UdL) – “Institut de recerca en tècniques agraries” (IRTA) formó parte del equipo de investigación, incorporación que permitió contrastar los resultados obtenidos con el segundo prototipo con indicadores de calidad de fruta. Fue la colaboración de ambas entidades lo que permitió, desde un primer momento, identificar las necesidades del sector frutícola, escogiendo las variedades de fruta de mayor interés, los indicadores de calidad más utilizados y los objetivos con mayor interés comercial.

Con la claridad de exposición en mente, se ha creído conveniente estructurar la exposición del trabajo realizado en esta tesis doctoral en ocho capítulos. En el segundo capítulo se repasan brevemente las técnicas convencionales de control de madurez y calidad de la fruta, así como técnicas más novedosas basadas en la identificación y cuantificación de los componentes aromáticos emitidos por la fruta durante su proceso de maduración. El contenido de este capítulo expondrá las ventajas y desventajas de

estos métodos convencionales que también han sido utilizados en este trabajo para poder comparar la eficacia de la nariz electrónica como instrumento de medición de calidad y madurez de la fruta frente a técnicas convencionales.

En el capítulo tres se describen con detalle los dos prototipos de olfato electrónico diseñados en el seno de este trabajo. Un primer sistema permitió evaluar la viabilidad de la idea, mientras que el segundo fue diseñado partiendo de la experiencia anterior. Su perfeccionamiento permitió utilizarlo con una variada selección de frutos y generar resultados lo suficientemente fiables como para obtener una serie de conclusiones finales.

Como ya se ha comentado previamente, una de las características que diferencia a todo sistema de olfato electrónico frente a los analizadores de gases convencionales es la inclusión de algoritmos de inteligencia artificial que exploten la multidimensionalidad de los datos recibidos de la matriz de sensores químicos. En el capítulo 4 se describen los algoritmos utilizados, la mayoría de ellos ya considerados clásicos en el área del reconocimiento de patrones. Sin embargo, como un resultado colateral de esta tesis, se presentan algunas modificaciones interesantes a algoritmos conocidos que permiten aumentar la utilidad de los mismos en aplicaciones basadas en sistemas de olfato electrónico.

Los capítulos cinco, seis y siete presentan las diferentes experiencias realizadas con los tres tipos de fruta utilizados: melocotones, peras y manzanas respectivamente. En cada uno de ellos se describen las experiencias realizadas tanto con el primer como con el segundo prototipo, presentando los resultados obtenidos y discutiendo las dificultades y soluciones aportadas para los objetivos de cada experiencia.

Finalmente, en el capítulo 8 se presentan las conclusiones de este trabajo y las líneas de investigación que, a opinión del autor, deberían continuar el trabajo realizado hasta la fecha.

REFERENCIAS

- [1.1] V.mathieu, C.Tronel, J.Mazollier *Pink-Lady*, ref. 24620 ISBN 2-87911-101-3, Centre technique interprofessionnel des fruits et legumes (ctifl), Paris, 1998
- [1.2] R.H.H. Wills, T.H.Lee, W.B. McGlasson, E.G. Hall, D.Graham, *Fisiología y manipulación de frutas y hortalizas post-recolección*, Ed. Acribia, Zaragoza, 1990
- [1.3] Rogers E. K., *Instrumentation and Sensors for the Food Industry*, Butterworth-heinemann Ltd, Oxford, U.K. 1993.
- [1.4] H. Shiota , *Changes in the volatile composition of La France pear during maturation*, Journal of the Science of Food and Agriculture, 52 (3) (1990) 421
- [1.5] M.L. López; M.T. Lavilla, I. Recasens, J. Graell, M.Vendrell, *Changes in aroma quality of ‘Golden Delicious’ apples after storage at different oxygen and carbon dioxide concentrations*, Journal of the Science of Food and Agriculture, 80, (2000) 311
- [1.6] Christensen C.M., *Effects of color on aroma, flavor and texture judgements of foods*, J. Food Sci., 48 (1983) 787-790
- [1.7] Pilando L.S., Wrolstad R.E, Heatherbell D.A., *Influence of fruit composition, maturity and mold contamination on the color and appearance of strawberry wine*, J. Food Sci., 50 (1985) 1121-1125
- [1.8] Povey, M.J.W., *Ultrasonics in food engineering. Part II: Applications*, Journal of Food Engineering, 9, (1989) 1-20
- [1.9] Dull, G.G., *Non-destructive evaluation of quality of stored fruits and vegetables*, Food Technology, May, (1986) 106-110
- [1.10] C.Ortíz et al, *Non-destructive determination of woolly peaches using impact response and Near-Infrared spectroscopy*, J. Agric. Eng, 78 (3) (2001) 281-289
- [1.11] W.F. Wilkens, A.D. Hatman, *An electronic analog for the olfactory processes*, Ann. NY Acad. Sci., 116 (1964) 608.

- [1.12] Gardner J.W., Barlett P.N., *A brief history of electronic noses*, Sensors Actuators B 18-19 (1994) 211-220
- [1.13] Hong, H. et al, *Electronic nose system with micro gas sensor array*, Sensors and Actuators B Vol. B36, (1996) 1-3
- [1.14] Singh, S., Hines, E. and Gardner, J., *Fuzzy neural computing of coffee and trained-water data from an electronic nose*, Sensors and Actuators B, Vol. B30 (1996)
- [1.15] Natale C., Davide F., D'Amico A., *An electronic nose for the recognition of the vineyard of a red wine*, Sensors Actuators B 33 (1996) 83-88
- [1.16] Schweizer-Berberich P.M., et al. *Characterisation of food freshness with sensor arrays*, Sensors and Actuators B, 18-19 (1994) 282-290
- [1.17] di Natale, C., et al., *An electronic nose for food freshness analysis*, Eurosensors X, Leuven september (1996) 681-684
- [1.18] di Natale, C., et al., *Recognition of fish storage time by metalloporphyrins-coated QMB sensor array*, Measurement Science and Technology, Vol. 7, (1996) 8
- [1.19] Smith P., *Sense for food*, Sensor Review, vol. 9, 1 (1989) 15-20
- [1.20] C. Visai, M. Vanoli. *Volatile compound production during growth and ripening of peaches and nectarines*. Scientia Hortoculturae 70 (1997) 15-24
- [1.21] T. Lavilla, J.Puy, M.L.López, I. Recasens and M.Vendrell, *Relationships between Volatile Production, Fruit Quality, and Sensory Evaluation in Granny Smith Apples stored in different Controlled-Atmosphere Treatments by means of multivariate analysis*, J. of agri. And food chem., vol 47, no.9, (1999) 3791-3803
- [1.22] Benady(15) Benady M., Simon J.E., *Fruit ripeness determination by electronic sensing*, Transactions of the ASAE, vol. 38, 1 (1995) 251-257
- [1.23] E. L. Hines, E.Llobet, J.W.Gardner, *Neural network based electronic nose for apple ripeness determination*, Electronics Letters, 35 (1999) 821
- [1.24] E. Llobet, E. L. Hines, J.W. Gardner, S. Franco, *Non-destructive banana ripeness determination using a neural network-based electronic nose*, Meas. Sci. Technol. 10 (1999) 538-548

[1.25] E.L.Hines, E.Llobet, J.W.Gardner, *Electronic noses: A review of signal processing techniques*, IEE Proceedings, Circuits Devices and Systems, Vol. 146, No. 6, (1999) 297-310

CAPÍTULO 2:

TÉCNICAS DE CONTROL DE CALIDAD EN FRUTA

2.1 INTRODUCCIÓN

El concepto de calidad en fruta ha ido evolucionando a lo largo del tiempo. Al principio la percepción de la calidad era diferente según el interés particular de cada uno de los agentes que intervenían en el proceso de producción (productor, consumidor, comerciante o consumidor). Sin embargo, cada vez hay más coincidencia entre los sectores implicados ya que todos ellos tienden a acercar sus criterios hacia los que impone el consumidor, en los que el estado de maduración de la fruta que compra juega un papel fundamental [2.1,2]. El precio de la fruta cada vez está más ligado a la calidad del producto final y, por ese motivo, las explotaciones frutícolas planifican su proceso productivo con miras a satisfacer al máximo las exigencias del sector comercial.

El fruto pasa a lo largo de su vida por una serie de etapas, caracterizadas por una secuencia de continuos cambios metabólicos. Así, después de la polinización y cuajado, la vida de las frutas puede dividirse en tres etapas fisiológicas fundamentales: crecimiento, maduración y senescencia, sin que sea fácil establecer cuando acaba una y empieza la otra.

La etapa más importante y compleja en el desarrollo de la fruta, el proceso de maduración, puede dividirse, a su vez, en dos fases: la fase de maduración fisiológica y la de maduración organoléptica. De hecho, en la literatura especializada de habla inglesa se distingue entre ambas, denominando dichos procesos como “maturation” y “ripening” respectivamente [2.3]. La primera suele iniciarse antes de que termine el crecimiento celular y finaliza, más o menos, cuando el fruto tiene las semillas en disposición de producir nuevas plantas. El crecimiento y maduración fisiológica sólo se completan adecuadamente en el árbol.

La maduración organoléptica hace referencia al proceso por el cual las frutas adquieren las características sensoriales que las definen como comestibles. Por lo tanto, se trata de un proceso que transforma un tejido fisiológicamente maduro pero no comestible en otro visual, olfatorio y gustativamente atractivo [2.4]. Aunque el resultado difiere significativamente, la maduración organoléptica se puede completar tanto en el árbol como una vez la fruta ya se ha recolectado. En general, esta etapa es un proceso que comienza durante los últimos días de maduración fisiológica y que irreversiblemente conduce a la senescencia de la fruta.

Los cambios más palpables durante el proceso de maduración son el color, sabor, olor, textura, etc. Estos cambios son el resultado de la profunda reestructuración metabólica y química que se desencadena dentro del fruto. En los frutos climatéricos, este proceso es controlado, fundamentalmente, por el etileno y su actividad respiratoria [2.5].

Por lo tanto, a medida que el fruto se desarrolla en el árbol sufre una serie de cambios anatómicos, fisiológicos y bioquímicos que son perfectamente evaluables. Debido a la importancia de obtener frutos con unas características de madurez óptimas tanto para el consumo como para su frigoconservación, de forma que lleguen con las mejores condiciones organolépticas posibles al usuario final, se debe disponer de índices para determinar el momento óptimo de recolección.

Hay algunos índices que sirven tanto para seguir la maduración en el árbol como la evolución de la calidad organoléptica durante la frigoconservación y posterior

maduración a temperatura ambiente [2.6]. Los índices más utilizados son el color de fondo, la firmeza, el contenido en sólidos solubles, el test de almidón y la acidez valorable, siendo todos ellos de empleo muy práctico. Otros, como el número de días desde plena floración, la intensidad de respiración y la producción de etileno son más indicados para estudiar las características fisiológicas [2.7].

Recientemente se han empezado a investigar técnicas de valoración de calidad basadas en la medición de los compuestos aromáticos. La ventaja de estas técnicas es su buena correlación con las características organolépticas de la fruta [2.8]. Sin embargo, su compleja aplicación sólo las hacen aconsejables en estudios en los que se evalúen diferentes técnicas de producción y no para un uso rutinario de control de calidad de las partidas de fruta que llegan a una cooperativa.

En los sucesivos apartados se irán describiendo las técnicas comentadas, ya que todas ellas han sido utilizadas en este estudio. Se trata de técnicas que se han aplicado a la fruta paralelamente a las medidas con nariz electrónica para poder comparar objetivamente entre las evaluaciones hechas con el sistema de olfato electrónico y las técnicas convencionales.

Finalmente, el apartado 2.4 describe brevemente otras técnicas no destructivas de control de calidad en fruta que están siendo investigadas actualmente, y que pueden considerarse competidores directos de las técnicas basadas en sistemas de olfato electrónico. Como todavía están en fase de investigación, estas técnicas no se han aplicado en este estudio aunque sería interesante, en un futuro próximo, compararlas con los resultados obtenidos con narices electrónicas.

2.2 MEDIDAS FÍSICO-QUÍMICAS

Los indicadores de calidad catalogados como físico-químicos pueden ser considerados como tradicionales en el mundo de la fruta. Su aplicación suele ser sencilla y los resultados se obtienen en poco tiempo, aunque su correlación con el grado de

maduración y con la calidad según el criterio del consumidor rara vez es completamente satisfactoria. De hecho, suele ser necesario utilizar varios de ellos conjuntamente para poder garantizar un control adecuado de la calidad de la fruta analizada. Los indicadores físico-químicos utilizados en este trabajo son la firmeza, la acidez, la colorimetría, la medición de sólidos solubles y el índice de almidón. Salvo la colorimetría, todos ellos requieren la destrucción de la muestra.

2.2.1 Firmeza

La firmeza es una de las técnicas más utilizadas en el control de la maduración de la fruta. Se trata de una técnica muy sencilla cuyos resultados se obtienen en cuestión de segundos. Además, el instrumento que se utiliza para aplicar esta técnica (el penetrómetro) es una herramienta relativamente barata y de un tamaño reducido que permite hacer mediciones en campo con suma facilidad. La figura 2.1 muestra un penetrómetro profesional como el utilizado en este trabajo.

La firmeza es uno de los métodos físico-químicos que mejor se correla con el estado de maduración de la fruta, especialmente en los melocotones y nectarinas [2.9], ya que la dureza de la pulpa está directamente relacionada con la madurez de la muestra. El método que se sigue para medir la firmeza de la pulpa incluye los siguientes pasos [2.10]:

- 1) Equipar a un penetrómetro profesional con el pistón adecuado (11 mm para manzanas, 8 mm para peras, melocotones y nectarinas)
- 2) Cortar con un cuchillo la piel de la fruta en dos puntos opuestos de la parte ecuatorial del fruto.
- 3) Asegurar de que la aguja indicadora de presión del aparato se encuentre marcando un 0.
- 4) Posicionar de forma vertical el aparato, colocando la punta de éste justo donde se ha efectuado el corte con una mano, mientras que con la otra se sujeta el fruto que debe estar sobre el soporte de apoyo.

- 5) Presionar hasta que se haya introducido parcialmente el émbolo (hasta una muesca que hay dibujada).



Figura 2.1: Penetrómetro utilizado para medir la firmeza de la fruta

- 6) Anotar este primer valor, y realizar el mismo procedimiento para el otro trozo de pieza sin piel cortado con anterioridad.
- 7) Expresar los resultados en libras o kilogramos ($1 \text{ kg.} = 1 \text{ lb} \times 2,2$)
- 8) Estandarizar los calibres respecto a un fruto de tamaño 100 utilizando tablas pensadas para cada variedad de fruta.

2.2.2 Análisis de Sólidos Solubles

Como los azúcares son los componentes mayoritarios en el zumo de la fruta, el análisis de sólidos solubles puede utilizarse como un estimador del contenido en azúcares en la muestra [2.11]. La técnica más común de medición de este parámetro, basada en la refractometría, requiere de instrumentos relativamente baratos, aunque las medidas no se pueden realizar en campo cómodamente.

Para este método es necesario el siguiente material de laboratorio:

- 1 Licuadora y un cuchillo
- 1 Vaso de precipitados de 250 ml

- 1 Pipeta Pasteur
- 1 refractómetro ABBE o equivalente

El método incluye los siguientes pasos:

- 1) Corte de la fruta en gajos e introducción en la licuadora. El zumo obtenido se deposita en el vaso de precipitados de 250 ml.
- 2) Toma una muestra del zumo con la pipeta Pasteur para depositarlo, en forma de gotas, sobre el prisma del refractómetro.
- 3) Medición a través del ocular, ajustando la sombra en el punto medio de la cruz para leer en la escala numerada superior el índice de refracción. El valor leído se anota en grados Brix.
- 4) La lectura irá siempre acompañada de la temperatura a la que se ha realizado.
- 5) Conversión del índice de refracción a la medición estándar de 20°C utilizando una tabla de conversión ya estipulada.

En la figura 2.2 podemos ver el refractómetro de la casa ATAGO empleado para el análisis de sólidos solubles en las diferentes variedades analizadas.

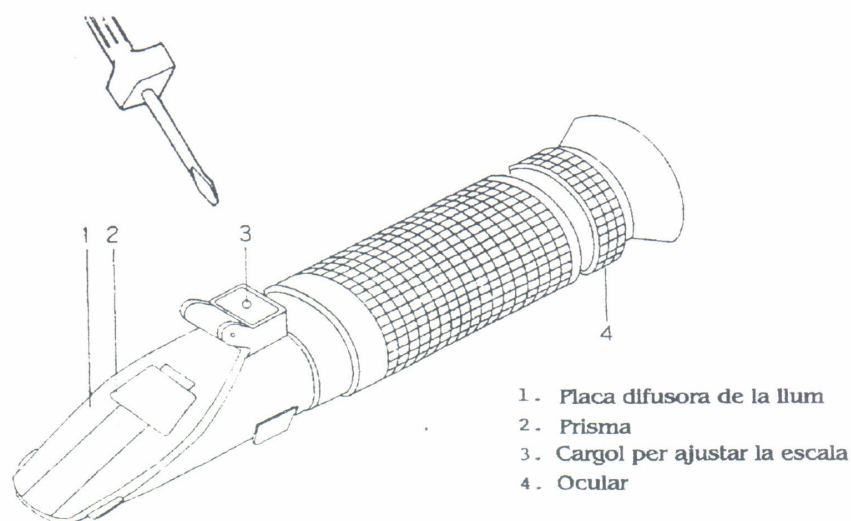


Figura 2.2: Refractómetro de masas N-1E de la casa ATAGO.

La escala de los aparatos ATAGO está hecha para que indique el valor correcto cuando estemos tomando muestras a una temperatura ambiente de unos 20 °C. Por lo tanto hay

que aplicar un factor de corrección, ya que la temperatura del laboratorio oscila entre unos 21 - 23 °C. Otra posibilidad es calibrar el aparato a la temperatura de trabajo.

2.2.3 Análisis de Acidez

Al igual que en el caso de la medición de los contenidos en sólidos solubles, para medir la acidez de la fruta se exprimen las piezas para obtener zumo. En este trabajo se ha utilizado la acidez titulable (TA). Para realizar la medición se puede aprovechar el zumo previamente extraído para obtener el índice de refractometría.

Para este segundo análisis se ha utilizado el siguiente material:

- 1 Licuadora
- Vasos de precipitados de 250 ml y 100 ml
- 2 Pipetas Pasteur
- 1 Bureta de 50 ml
- Hidróxido de sodio (NaOH) 0.1 N
- Phímetro

El procedimiento que se ha seguido es el siguiente:

- 1) Corte de la pieza en trozos pequeños, y obtención del zumo con la licuadora.
- 2) En un vaso de precipitados de 100 ml, se añaden 10 ml de zumo de un número determinado de muestras y 10 ml de agua destilada, ayudándose con la bureta y las dos pipetas de 10 ml.
- 3) Calibración del PHímetro.
- 4) Se coloca la sonda del PHímetro dentro del vaso de precipitados y se va añadiendo Hidróxido de Sodio hasta conseguir neutralizar la muestra (con un pH aproximado de 8.1, calculando el resultado como ácido málico).
- 5) Se anota los ml de NaOH gastados durante el proceso de neutralización de la muestra.
- 6) La acidez titulable se calcula según la ecuación 2.1:

$$TA = \frac{\text{ml de NaOH} \times N (\text{NaOH}) \times \text{acid meq. Factor} \times 100}{\text{ml zumo}} \quad (2.1)$$

- 7) Se limpia con agua destilada la sonda del PHímetro.



Figura 2.3: PHímetro para analizar la acidez de las muestras

2.2.4 Colorimetría

La colorimetría es el único de los métodos físico-químicos que no requiere la destrucción de la muestra. Para realizar la medición se utiliza un aparato calibrado denominado colorímetro.

En el caso de variedades rojas se realizan mediciones de color tanto en las zonas más coloreadas como en las menos coloreadas. En cambio, en las variedades verdes y amarillas se miden varios puntos y se hace la media.

La función del colorímetro es describir la coloración de la epidermis de la pieza de fruta objeto de la medición. Para ello devuelve tres parámetros, L*, a*, b*, siguiendo el

estándar C.I.E.L*a*b* (apertura de diámetro 8 mm, plato blanco de referencia, iluminación estándar D65 y observador a 10°).

La luminosidad viene descrita por L*. El color negro presenta una luminosidad de 0 mientras que el blanco presenta una luminosidad de 100. Los parámetros a* y b* se utilizan para evaluar la saturación y el tono. La saturación nos da la pureza de un color y el tono es el color propiamente dicho. Para el cálculo se utilizan las expresiones 2.2, 2.3 y 2.4:

$$\text{Saturación} = (a^{*2} + b^{*2})^{1/2} \quad (2.2)$$

$$\text{Tono en variedades rojas} = \arctg b/a \quad (2.3)$$

$$\text{Tono en variedades verdes y amarillas} = a^* + b^* \quad (2.4)$$

En este trabajo se ha utilizado un colorímetro triestímulo Minolta CR-200, cuyo precio ya no es tan asequible como el de los instrumentos utilizados en las técnicas físico-químicas descritas anteriormente. El procedimiento que se debe seguir para realizar la medida es el siguiente:

- 1) Coger el colorímetro y borrar todos los datos de medidas anteriores
- 2) Calibrar el instrumento. Para ello es necesario colocar el cabezal de medida sobre el plato de calibración e invocar a la función “Calibrate” hasta que el aparato indique que está preparado.
- 3) Poner al sistema en modo medida apretando el botón “measure”
- 4) Realizar la medida sobre la superficie de la muestra a medir
- 5) Anotar los valores de los parámetros L*, a*, b*

2.2.5 Índice de almidón

Durante el proceso de maduración, el almidón de algunas variedades de fruta se rompe en azúcares. Esta conversión empieza en el corazón del fruto y avanza por la pulpa hacia la periferia. La pauta de conversión del almidón es característica de cada variedad y para cuantificarla se pueden utilizar diferentes escalas. La figura 2.4 muestra la escala con nueve valores utilizada en este trabajo.

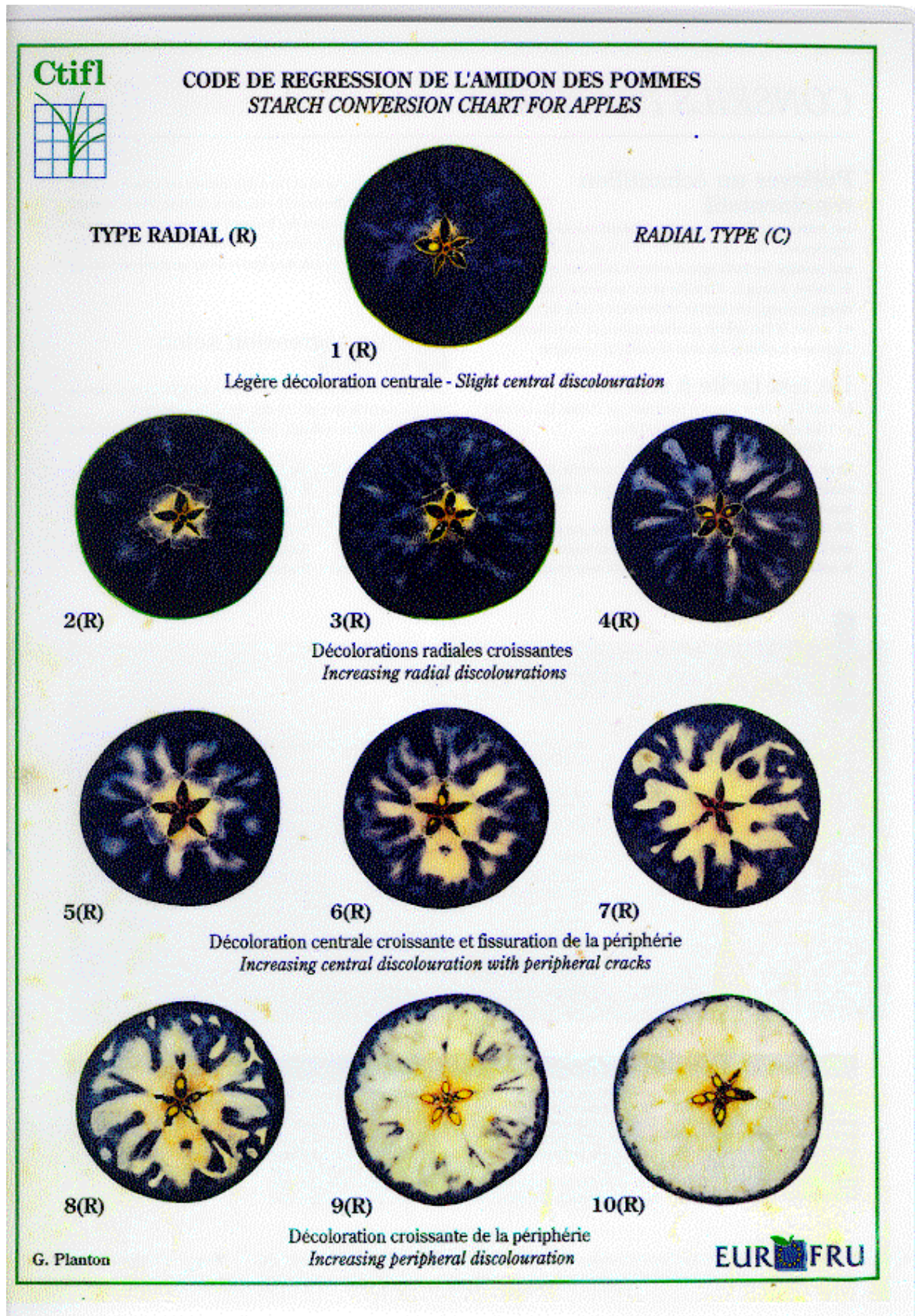


Figura 2.4: Escala utilizada para obtener el índice de almidón

La utilización de yodo, que reacciona con el almidón formando un color negro, permite visualizar las zonas en las que todavía existe almidón. El procedimiento que se debe seguir es el siguiente:

- 1) Cortar el fruto transversalmente, por la parte ecuatorial
- 2) De una de las dos mitades obtenidas, cortar una rodaja de aproximadamente 0.5 cm de espesor.
- 3) Verter con cuidado una solución 0.5% I₂-KI en la bandeja de forma que quede a una altura de 1-2 mm.
- 4) Colocar las rodajas en la bandeja con la ayuda de las pinzas; solamente debe quedar cubierta la cara inferior de la rodaja. Añadir lentamente más solución 0.5% I₂-KI si hiciese falta.
- 5) Dejar en reposo durante cinco minutos y observar la cara inferior.

2.3 MEDICIÓN DE VOLÁTILES

2.3.1 Medidas aromáticas

Las características fundamentales que determinan la calidad organoléptica del fruto son la ausencia de defectos, la textura, el “flavor” y el aspecto externo (incluyendo el tamaño, color y forma). Todas ellas se pueden correlacionar con un determinado grado de maduración.

A diferencia de los dos primeros, el “flavor “ es un atributo muy complejo, ya que está determinado por el equilibrio entre los ácidos, los azúcares y los componentes volátiles principalmente [2.12]. En definitiva, el “flavor” es el resultado de combinar tres propiedades sensoriales diferentes pero complementarias (gusto, el olor y aroma) , siendo esta última su principal componente [2.13]. De las tres, el gusto es la menos importante.

El olor es, después del color, la propiedad que nos afecta más significativamente a la hora de aceptar un alimento. El olor es la percepción, por medio del olfato, de las sustancias volátiles liberadas desde los alimentos de forma espontánea a temperatura ambiente. Aunque se han desarrollado muchas teorías que intentan explicar como se genera la percepción del olor molecular, la teoría más aceptada es la del encaje o acoplamiento [2.14].

El olor debe diferenciarse claramente del aroma, que es la percepción de las sustancias aromáticas después de introducirse los alimentos en la boca y trocearlos, llegando al sistema olfativo por vía retranasal.

El olor y el aroma característico de cualquier fruta es debido a la existencia de las sustancias aromáticas presentes en la piel y en la pulpa [2.15,16], formando una compleja mezcla de componentes orgánicos muy relacionados con el proceso de maduración, ya que la mayoría se sintetizan durante la fase climatérica [2.17].

Para interpretar el aroma de un fruto es necesario conocer la naturaleza, cualidad, cantidad e intensidad aromática de cada componente [2.18]. También es importante conocer a lo largo del desarrollo del fruto como se modifica su composición en cantidad y en tipo de sustancias [2.19].

A pesar de que los primeros intentos de identificación de compuestos aromáticos se dieron en la primera mitad del siglo XX, su éxito fue reducido debido a la instrumentación físico-química existente [2.20]. Hay que tener presente que la concentración media de las sustancias aromáticas en la fruta no suele superar los 50mg/Kgr, encontrándose valores máximos para cada sustancia de 100ug/kg [2.21].

La evolución tecnológica hace que hoy en día el análisis de los componentes aromáticos se realice a través de la cromatografía de gases combinada con la espectrometría de masas [2.22].

Para realizar un estudio cualitativo y cuantitativo de las sustancias aromáticas es necesario extraerlas. Existen muchos métodos de extracción [2.23], cada uno de los

cuales es adecuado a un tipo concreto de alimentos. En este trabajo se han combinado dos técnicas que originalmente se utilizaban por separado, el espacio de cabeza y la extracción por arrastre con gas. La utilización del espacio de cabeza (ahora denominado estático) presenta el problema de que extrae poca cantidad de volátiles y sólo permite identificar medidas de las sustancias cuantitativamente más abundantes. La combinación de esta técnica con el arrastre por gas facilita la volatilización y captura de sustancias. A esta nueva técnica se le denomina espacio de cabeza dinámico [2.24-27].

En el espacio de cabeza dinámico la fruta se deposita en un recipiente hermético. Una vez cerrado, se hace pasar una corriente de gas (N_2 o aire) a través del mismo. Esta corriente acelera la emisión de volátiles de la fruta, que son arrastrados hasta un cartucho donde son capturados por un polímero granulado poroso como el Tenax GC, el Porapak Q, o el Chromosorb 150. Posteriormente, los componentes atrapados son recuperados haciendo pasar por el polímero disolventes orgánicos, generalmente éter dietílico, o mediante desorción térmica [2.28]. La cantidad de componentes aromáticos extraídos depende del tiempo empleado en el arrastre, del flujo de gas y del polímero poroso utilizado.

Si las sustancias aromáticas son desorbidas por disolventes, la solución resultante se analiza por cromatografía de gases directamente. Si la desorción es térmica, a bajas temperaturas se eliminan las trazas de agua por elución, de forma que a medida que aumenta la temperatura se van liberando las sustancias volátiles que son arrastradas por un gas portador a un colector frío normalmente conectado a un cromatógrafo de gases. [2.29,30]

Las ventajas del espacio de cabeza dinámico son la relativa sencillez del análisis y el hecho de que los aromas que se analizan son los que en condiciones normales alcanzan la mucosa olfativa, siendo los verdaderos responsables del aroma de la fruta. Su principal desventaja radica en la dificultad de relacionar los componentes analizados con la composición interna del fruto [2.31].

El método exacto que se ha seguido en este trabajo para la monitorización de componentes aromáticas ha sido el siguiente:

- 1) Para cada medida, un 1 kg. de fruta se introduce en un recipiente de pirex de 10 litros de capacidad. Inmediatamente después se genera un flujo de nitrógeno de 50-60 ml.min⁻¹ que atraviesa la cámara con las muestras de fruta, arrastrando los componentes aromáticos generados hacia un cartucho de vidrio cuyo interior esta relleno de 1 g de Tenax-CG 60-80.
- 2) El proceso de extracción se realiza en la oscuridad, en una habitación con la temperatura controlada a 20°C. Los volátiles se recuperan de la fase absorbente con 20 ml de dietilo de éter. Las fases de éter sólido son tratadas con ultrasonidos y posteriormente se someten a un proceso de centrifugado para su separación. A continuación, se realizan dos extracciones adicionales con 20 y 10 ml de éter dietílico respectivamente bajo idénticas condiciones a las ya mencionadas. Las fases orgánicas se agrupan y evaporan hasta un volumen de 0.1 ml por burbujeo en una corriente de nitrógeno.
- 3) Finalmente, la identificación y cuantificación de los componentes volátiles se hace con un cromatógrafo de gases Hewlett-Packard 5890 equipado con un detector de ionización por llama. Para la identificación y confirmación de diferentes componentes, se utiliza la combinación de un espectrómetro de masas acoplado al cromatógrafo de gases antes mencionado. Los espectros de masas se obtienen por impacto electrónico de iones con una energía de 70 eV. Los datos de espectro obtenidos se comparan entonces con los de la librería NIST HP59943C [2.32].

2.3.2 Producción de etileno

El etileno es una de las moléculas orgánicas más sencillas con actividad biológica. Sucesivos trabajos de investigación pusieron de manifiesto que este compuesto puede ser considerado como la hormona vegetal que regula varios aspectos de la maduración y

senescencia [2.33-34]. Es precisamente el aumento de su biosíntesis hasta concentraciones estimuladoras lo que inicia la maduración en los frutos climatéricos. A partir de ese momento, la producción de etileno aumenta al mismo tiempo que se produce el incremento respiratorio [2.35].

La definición precisa de la fase climatérica hace referencia al periodo en el que aparecen una serie de cambios bioquímicos iniciados por la producción autocatalítica de etileno, comportando un aumento en la respiración y culminando con la maduración organoléptica de la fruta. Esta fase solo se presenta en los frutos denominados climatéricos, entre los que se encuentran las peras, manzanas, melocotones y nectarinas. Por lo tanto, la monitorización del etileno producido por muestras de fruta climatéricas aporta información fundamental en la determinación de su estado de maduración.

Dicha monitorización puede realizarse extrayendo una muestra del corazón de fruta o midiendo su emisión volátil. En este trabajo se ha optado por esta segunda opción, ya que ésta utiliza un proceso parecido al de la medición de volátiles aromáticos.

El proceso exacto empieza con la introducción de 1 kg. de fruta en el recipiente, que a continuación se cierra. Esta cámara continuamente es aireada con aire humidificado a 20°C con un flujo constante de 2 l.h⁻¹. Posteriormente, una muestra de 1 ml de gas se inyecta en un cromatógrafo de gases Hewlet-Packard 5890 equipado con un detector de ionización por llama y una columna de 1.5m x 3 mm empaquetada con óxido de aluminio.

El análisis se hace isotérmicamente a 100°C. Los flujos del gas portador N₂, del aire y del H₂ son, respectivamente, 45, 400 y 45 ml min⁻¹. El inyector y el detector se mantienen a 120°C y 180.

2.4 TÉCNICAS EXPERIMENTALES

Tanto la tecnología basada en ultrasonidos como la espectroscopía en el infrarojo cercano son tecnologías que están siendo consideradas como alternativas no destructivas

para monitorizar la calidad de la fruta. Su uso no se ha generalizado fundamentalmente porque son técnicas todavía inmaduras sobre las que se tiene que trabajar para eliminar sus inconvenientes. Por ese motivo existen numerosos trabajos de investigación en los que se presentan resultados obtenidos sobre diferentes variedades de fruta.

2.4.1 Mediciones con ultrasonidos

En la introducción de esta memoria ya se han referenciado algunos trabajos en los que se han utilizado los ultrasonidos para monitorizar el grado de maduración de la fruta. De hecho existen múltiples estudios con variedades diferentes de fruta, aunque los principios básicos de la técnica son similares.

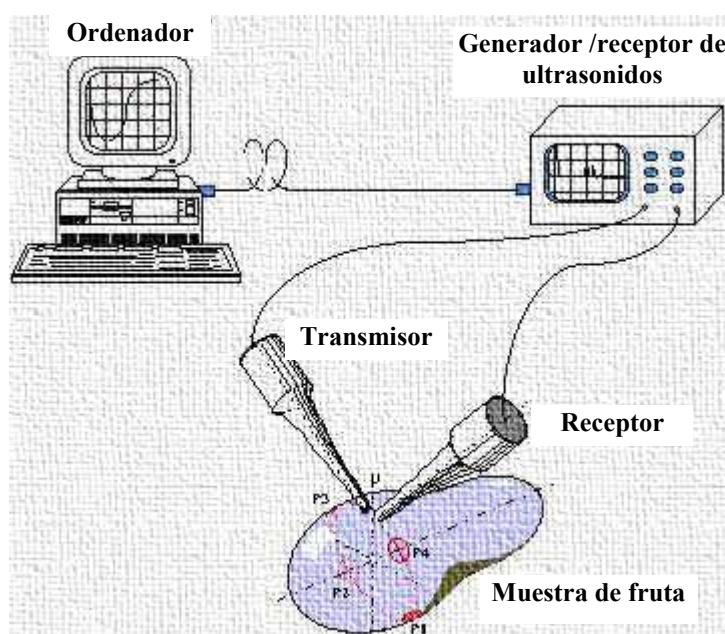


Figura 2.5: Esquema gráfico del proceso de medida mediante ultrasonidos

La figura 2.5 muestra la disposición estándar de medida. En ella se puede distinguir la pieza de fruta a caracterizar, un emisor/receptor de ultrasonidos, las sondas transmisora y receptora, y el ordenador personal que controla el proceso de medida y almacena los resultados para posteriormente procesarlos con algún paquete estadístico. Los parámetros que se extraen de las mediciones son la atenuación que sufre la onda acústica en su recorrido y la velocidad con la que atraviesa la carne del fruto. Las

mediciones se repiten variando la posición de las sondas transmisora y receptora para reducir la variabilidad debida a la posición en la que se mide la pieza.

Diversos trabajos [2.36,37] prueban que existe un grado de correlación entre los parámetros acústicos y determinados indicadores de calidad como la firmeza en el aguacate o la acidez y los contenidos de azúcar en mango. De todas formas, a los problemas ya mencionados de acoplamiento entre las sondas y las piezas de fruta se suma la escasa repetitividad de resultados, inconveniente que requiere realizar múltiples mediciones para obtener una media significativa de los parámetros acústicos deseados.

2.4.2 Espectrometría de infrarojo cercano

La espectrometría de infrarojo cercano (Near-infrared spectroscopy, NIR) es una técnica no destructiva con la que se están realizando numerosos estudios sobre fruta. En [2.38,39] se presentan dos trabajos en el que se intenta determinar el contenido en sólidos solubles, acidez y firmeza en manzanas, mientras que en [2.40] se puede encontrar un estudio para determinar la calidad interna de melocotones y nectarinas midiendo sólidos solubles y contenido total en azúcares, sorbitol y clorofila.

El principio en el que se basa la técnica es en caracterizar la reflectancia de la pieza de fruta ante diferentes longitudes de onda. Una medición típica incluye mediciones espaciadas 10 nm entre 900 y 1400 nm. Para ello se utiliza un fotospectrómetro que genera luz con una lámpara halógena de Tugsteno. Utilizando técnicas basadas en difracción se conduce la luz a través de fibra óptica hasta la superficie de la pieza a caracterizar que a su vez recoge la luz reflejada. Sobre los índices de refracción en todo el margen frecuencial se pueden extraer multitud de parámetros como la primera o segunda derivada del espectro. Estos parámetros, junto a técnicas de reconocimiento de patrones, permiten realizar predicciones o clasificaciones de la fruta en función de su estado de maduración.

De todas maneras, en muchas variedades la técnica no tiene suficiente resolución como para predecir determinados parámetros. Además, el espectrofotómetro debe ser

calibrado cuidadosamente con fuentes de luz estables como puede ser un láser de Helio-Neón.

REFERENCIAS

- [2.1] V.Urbina Vallejo, *La calidad de los frutos*, revista de fruticultura, vol. 5, núm 2, marzo-abril 1990
- [2.2] R.H.H. Wills, T.H.Lee, W.B. McGlasson, E.G. Hall, D.Graham, *Fisiología y manipulación de frutas y hortalizas post-recolección*, Ed. Acribia, Zaragoza, 1990
- [2.3] Rhodes M.J., *The climateric and ripening of fruits*, en: *The biochemistry of fruits and their products Food science and technology 1: A series of monograph*, Ed. A.C. Hulme, Academic Press London and New York (1971) 333-373
- [2.4] Leopold, A.C; Kriedemann P.E., *Plant growth and developement*. Ed.McGraw Hill, 2 Nova York (1975) 16-34 .
- [2.5] Hulme A.C; Rhodes M.J., *Pome fruits*, en: *The biochemistry of fruits and their products Food science and technology 1: A series of monograph*, Ed. A.C. Hulme, Academic Press London and New York (1971) 333-373
- [2.6] Knee M; Hatfield S; Smith S.M., *Evaluation of various indicators of maturity for harvest of apple fruits intended for long-term storage*, J. Hort. Sci, vol 64 (1989) 403-411.
- [2.7] Knee M., *Pome fruits*, en *Biochemistry of food ripening* Ed. Seymour J.D. , Taylor J.E., Tucker G.A., Chatman and hall, London U.K (1993) 325-345
- [2.8] López M.L., Lavilla M.T., Recasens I., Graell J., Vendrell M., *Changes in aroma quality of 'Golden Delicious' apples after storage at different oxygen and carbon dioxide concentrations*, Journal of the science of Food and Agriculture, vol 80 (2000) 311-324.
- [2.9] Crisosto C.H., *Optimum procedures for ripening stone fruit*, Management of ripening fruit, december 1994, 24-25
- [2.10] Recasens, I; Soria Y., *Práctica 2: Determinación de parámetros de calidad postcosecha en: Prácticas de fisiología de la poscollita*, ETSEA-Universitat de Lleida
- [2.11] Mitcham B., Kader A., *Methods for determining quality of fresh horticultural commodities*, Perishables Handling newsletter, University of California at Davis, August 1995, 1-11

- [2.12] Monin A., *Etude de la qualite gustative de la pomme golden delicious*, rev. Agric. (1970) vol 3, 471-487
- [2.13] Panasiuk D.; Talley F.D.; Sapers G.M., *Correlation between aroma and volatile composition of Macintosh apples*, J.Food Science (1980) vol 45, 989-991
- [2.14] Amore J.E., *Molecular basis of odour*, Thomas C.C. Ed. Springfield, Illinois, USA, 1970, 10-27.
- [2.15] Buttery R.G., *Vegetable and fruit flavors en: Flower research recent advances*. Ed. Teramishir;flath R.A. sugisawa H. Ed, Marsel Decker, New york (1981) 175-216
- [2.16] Fellman J.K., Mattinson D.S., Bosckit B.C., Mateis T.T.,Patterson M.E. *Esther biosynthesis in rome apples subjected to low oxigen atmospheres*, Postharvest Biology and Technology (1993) vol 3, 201-214
- [2.17] Tressl R., Drawert F., *Biogenesis of banana volatiles*, J. Agric. Food chem, (1973) vol 21, 560-565
- [2.18] Medina I., Suárez J.J, Martinez J.L., *Aromas alimentarios (1 y 2)*, Alimentación, equipos y tecnología, (1994) Junio 87-92
- [2.19] Nursten H.E., *Volatile compounds: the aroma of fruits en: The biochemistry of fruits and their products Food science and technology 1: A series of monograph*, Ed. A.C. Hulme, Academic Press London and New York (1970) 239-267.
- [2.20] Medina I., Martínez J.L, Suárez J.J., *El aroma de manzano revista: alimentación equipos y tecnología*, marzo(1996), 55-61.
- [2.21] Mattheis J.P., Fellman J.K., Chen P.M., Patterson M.E., *Change in headspace volatiles during physiological development fo bisbe delicious apple fruit* , J.Agric. Food Chem (1991), vol 39, 1902-1906
- [2.22] Dimick P.S., Hoskin J.C., *Review of apple flower-state*, Critical reviews in food science and nutrition (1983) vol 18, 387-409.
- [2.23] Medina I, Suárez J.J., Martínez J.L., *Aromas alimentarios (y 2)*, Alimentación, equipos y tecnología (1994) 99-102
- [2.24] Paillard N., *Biosinthese des produits volatiles de la pomme: formation des alcools et des esthers a partir de acids gras*, Phitocheminstry (1979) vol 18, 1165-1171
- [2.25] Knee M., Hatfield S.G.S., *The metabolism of alcohols by apple fruit tissue*, Journal of Science Food Agr. (1981),Vol 32, 593-600

- [2.26] De pooter H.L. Coolsaet B.A. Dirink T.J.; Schamp N.M.; *GLC of the headspace after concentration on tenax GC and of the esencial oils of apples, fresh celery, fresh lovage, honey suckle and ginger powder: essencial oils and aromatic plants*, Baer Heim s vendsen, A and SCHEFFER J.J. C ed. Martinus Nijhosf,dwjunk publisher doredreg the netherlands, (1985) 24-36
- [2.27] Kakiuchi N., (1986).
- [2.28] Wehmer w. Köhler T., *A simple desorption devive for gas chromatographic aroma analysis using dinamic headspace technique*, gartenbawissenschaft, (1992) vol. 57, 126-129.
- [2.29] Schamp N.; Dirink P , , *The use of headspace concentration on tennax for objective flavor quality evaluation of fresh fruit (strawberry and apple)* . Chem. Of food and beverages. Recent developements (1982) 25-47
- [2.30] Salinas M.R., Alonso G.L., Esteban-infante F.J., *Absortion-thermal desorption-gas chromatography applied to the determination of wine aroma*, J.Agric. Food Chemistry, 42, (1994) 1328-1331.
- [2.31] Feys M., Toback P. Maes E., *Volatiles of apples (var shone van boskoop): isolation and identificacion*, J. Food Technology, 15, (1980) 485-492
- [2.32] Yang S.F., Hoffman N.E., *Ethylene biosynthesis and its regulation in haisher plants annu. Rev. Plant physilologic*, (1984) vol. 35, 155-189
- [2.33] Mattoo A.K., White W.D, *Regulation of ethylene biosynthesis en:The plant hormone ethylene* Mattoo A.K. and Suttle J.C: Ed. CRC Press, Bocaraton Florida (1991) 21-42;
- [2.34] Abeles F.B., Morgan P.W., Saltveit M.E. Jr, *Ethylene in plant biology*, Academic Press, New York (1992) 213-254
- [2.35] Reid M.S., Rodhes M.J.C., Hulme A.C., *Changes in Ethylene and CO₂ during the ripening of apples*, J. science Food. Agric (1973) vol 24, 971-979.
- [2.36] Mizrach A., Flitsanov U., Fuchs Y., *An ultrasonic nondestructive method for measuring maturity of mango fruit*, Transactions of the ASAE (1997)40 (4) 1107-1111
- [2.37] Mizrach A., Galili N., Gan-mor S., Flitsanov U., Prigozin I., *Models of ultrasonic parameters to assess avocado properties and shelf life*, Journal of Agricultural Engineering Research, (1996) vol 65, 261-267

[2.38] Bellon-Maurel V., Vigneau J.L., *NIR fast spectrometer for fruit internal quality assessment: reproductibility study*, Harvest and postharvest technologies for fresh fruits and vegetables(1995), 471-476, ASAE publication

[2.39] Lammertyn J., Nicolai B,Ooms K, *Non-destructive measurement of acidity, soluble solids and firmness of Jonagold apples using NIR-spectroscopy*, Transactions of the ASA E(1998), 41 (4), 1089-1094

[2.40] Slaughter D C, *Non-destructive determination of internal quality of peaches and nectarines*, Transactions of the ASAE (1995), 38(2), 1571-1575

CAPITULO 3:
DISEÑO DE SISTEMAS DE
OLFATO ELECTRÓNICO

3.1 INTRODUCCIÓN

Como ya se ha comentado anteriormente, los primeros sistemas de olfato electrónico experimentales aparecieron a finales de los ochenta. De hecho, se lleva más de dos décadas investigando y perfeccionando este tipo de instrumentos, a pesar de lo cual las narices electrónicas actuales distan mucho de la perfección y versatilidad del sistema olfativo humano. Hoy por hoy es impensable diseñar una nariz electrónica genérica que pueda substituir a un operario en todas y cada una de las actividades en las que actualmente es necesario utilizar el olfato humano.

Desde principios de los noventa existen narices electrónicas comerciales que permiten realizar estudios sobre la utilización de esta tecnología en diferentes aplicaciones [3.1-5]. Estos equipos son de carácter genérico ya que no han sido diseñados específicamente para una aplicación concreta. Sin embargo, debido al estado actual de esta tecnología,

hay muchas aplicaciones en las que la optimización del instrumento es vital para un funcionamiento eficaz.

Por este motivo existe otra clase de sistemas de olfato electrónico que se han diseñado desde un principio para aplicaciones específicas. Esta dedicación exclusiva permite una optimización del equipo que de otra forma no puede realizarse. A este tipo de sistemas en inglés se les denomina “Application Specific Electronic Noses” o “ASEN “ (narices electrónicas para aplicaciones específicas) y solo recientemente están apareciendo en el mercado equipos comerciales que siguen esa filosofía, como el “Freshsense”, sistema expresamente diseñado para monitorizar el estado de conservación del pescado [3.6]. El futuro de estos equipos, a juicio del autor de esta memoria, es mucho más prometedor que el de las aproximaciones genéricas.

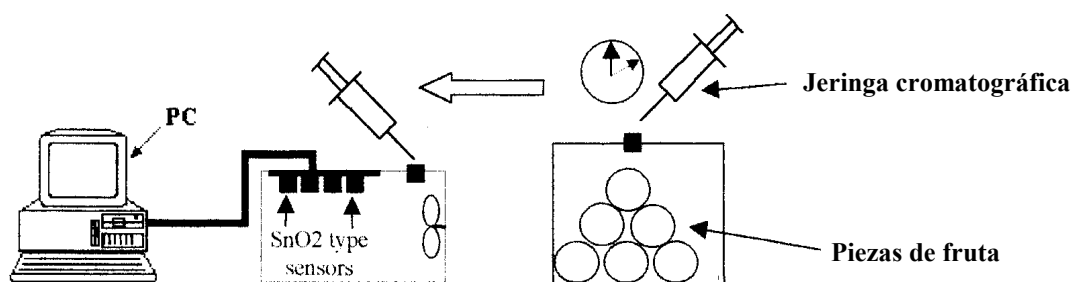
Los prototipos de nariz electrónica diseñados en este trabajo pertenecen a esta nueva categoría de sistemas, puesto que han sido expresamente concebidos para monitorizar el nivel de madurez y la calidad de la fruta. Se han diseñado dos prototipos puesto que la filosofía de trabajo que se ha seguido es la de un perfeccionamiento sucesivo. Así, el objetivo del primer equipo consistía simplemente en comprobar la posibilidad de utilizar una nariz electrónica como instrumento de seguimiento del proceso de maduración de la fruta. Posteriormente, basándose en la experiencia adquirida, se diseñó un nuevo equipo optimizado para aplicaciones concretas como son la determinación de la fecha óptima de recolección en melocotones, evaluación del mejor periodo de conservación en cámara de peras decanas, etc.

La diferencia fundamental entre ambos prototipos reside en el método de muestreo. Mientras que en el primer caso se utilizó el procedimiento de inyección, en el prototipo más avanzado se incorporó un muestreo automatizado utilizando un flujo de aire como portador de la muestra. Por ese motivo al primero lo llamaremos prototipo de inyección y al segundo prototipo de flujo continuo.

3.2 PROTOTIPO DE INYECCIÓN

3.2.1 Descripción general

Una nariz electrónica es un equipo de medida formado por una serie de módulos que trabajan de forma conjunta con el objetivo de cuantificar y/o clasificar muestras gaseosas o aromáticas. Tanto en el primer como en el segundo prototipo construidos en este trabajo se pueden distinguir tres módulos básicos con funcionalidad similar pero cuyas características constructivas son notablemente diferentes. En concreto, ambos sistemas se componen de un módulo de muestreo, de la matriz de sensores químicos y de un sistema informático.



Ordenador Personal Cámara de medida Cámara de concentración

Figura 3.1: Esquema general del prototipo de inyección

La figura 3.1 muestra el esquema general del primer prototipo. En él se pueden distinguir los diferentes elementos que conforman el equipo. El módulo de muestreo lo componen la cámara de concentración, la jeringa cromatográfica y la cámara de medida, en cuyo interior se aloja el segundo módulo, la matriz de sensores de gases. El ordenador es el tercer módulo que completa el sistema; sus funciones cubren las tareas de adquisición y procesado de las señales eléctricas que provienen de la matriz de sensores.

La figura 3.2 muestra una fotografía de este primer prototipo midiendo el grado de maduración de unas manzanas. En ella se puede distinguir perfectamente la cámara de concentración, la cámara de medida y el ordenador personal.

El proceso de medida en este prototipo empieza con la introducción de muestras de fruta en la cámara de concentración, donde permanecen el tiempo necesario para producir suficientes volátiles. Transcurrido un intervalo prefijado de tiempo se recoge una muestra de la atmósfera contaminada de la cámara de concentración y se inyecta en la cámara de medida utilizando una jeringa cromatográfica.

En ese momento la matriz de sensores reacciona ante la presencia de nuevos volátiles en la cámara de medida cambiando sus características eléctricas, variaciones que son monitorizadas y grabadas por el ordenador personal a través de una placa de adquisición de datos y un programa creado específicamente para gestionar todo el proceso de medida.

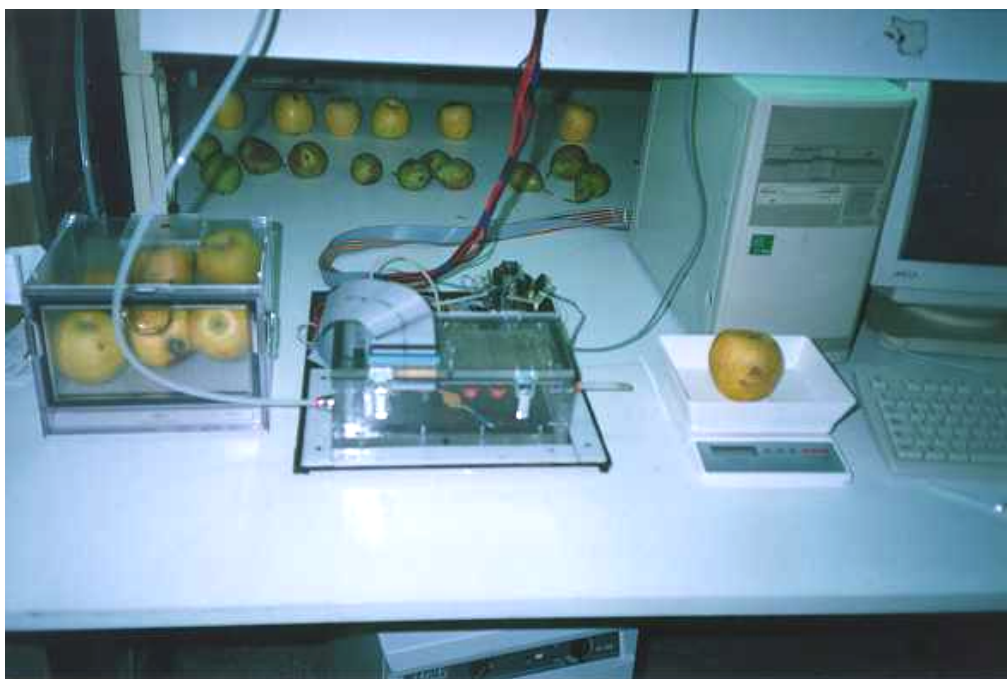


Figura 3.2: Fotografía del prototipo de inyección midiendo la madurez de manzanas

Posteriormente, el mismo ordenador procesa los datos con algoritmos diseñados a medida utilizando un paquete de programación matemática comercial.

3.2.2 Sistema de muestreo

Como ya se ha comentado, el sistema de muestreo está formado por tres módulos que pueden distinguirse perfectamente en la figura 3.1: Una cámara de concentración, una cámara de medida y una jeringa cromatográfica.

3.2.2.1 Cámara de concentración

El motivo principal por el que se ha incorporado una cámara de concentración es para obtener volátiles con concentraciones elevadas de forma sencilla. Además, su inclusión permite ajustar mucho más las dimensiones de la cámara de medida (aumentando así la sensibilidad de los sensores) al tiempo que dota al sistema de una gran flexibilidad, ya que con cámaras de concentración de diferentes dimensiones se pueden realizar medidas para diferentes cantidades de fruta o incluso para diferentes variedades (desde peras a sandías, por ejemplo).

Desde el punto de vista constructivo, se trata de un recipiente hermético de metracrilato de dimensiones 20x20x10 cm, lo que proporciona un volumen total de 5 litros. En su parte frontal dispone de una tapa con anclajes y junta de goma que permite introducir las piezas de fruta a medir. Dispone, además, de un orificio sellado con septum que puede ser atravesado por la punta de la jeringa cromatográfica para poder extraer una muestra de los volátiles generados por la fruta depositada en su interior.

3.2.2.2 Cámara de medida

La cámara de medida es donde se encuentra la matriz de sensores químicos del sistema. Se trata de una cámara estanca que no es necesario abrir para realizar cada medida ya que incorpora un septum por el que se pueden inyectar los volátiles que provienen de la cámara de concentración mediante una jeringa.

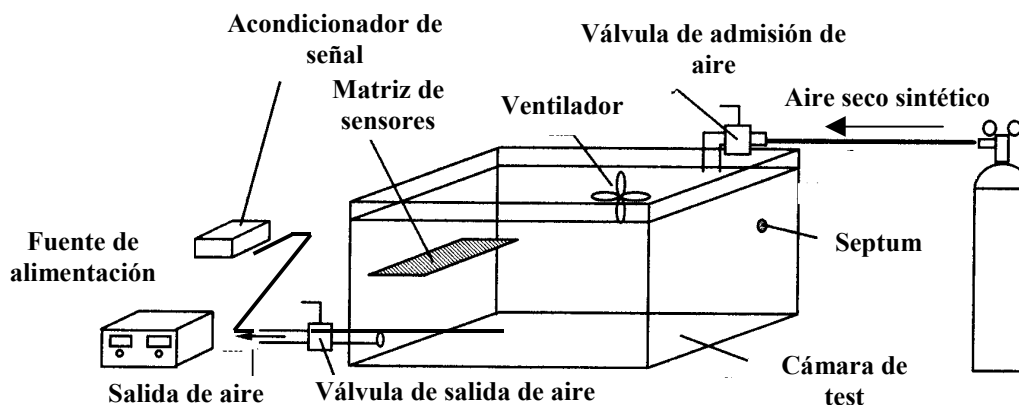


Figura 3.3: Esquema detallado de los componentes de la cámara de medida

La figura 3.3 muestra, con mayor detalle, los elementos que conforman la cámara de medida. Dentro de la misma, además de los sensores, se puede encontrar un ventilador que homogeneiza su atmósfera y un equipo comercial de medición de temperatura y humedad.

Desde el punto de vista constructivo, se trata de una cámara de metacrilato de 1.7 litros de volumen total (14.5x13x9 cm). En sus paredes se han practicado dos orificios adicionales para la entrada y salida de aire seco, de forma que después de cada medida puede ser limpiada de contaminante con aire a presión sin necesidad de abrir la tapa de la misma. Los cables de conexión, tanto de los sensores como de la sonda de temperatura y humedad, atraviesan la carcasa a través de un agujero fresado y sellado en el que se ha incorporado un conector de 54 pines.

3.2.2.3 Jeringa cromatográfica

Para realizar el transporte de la muestra gaseosa presente en la cámara de concentración hasta la cámara de medida se utiliza una jeringa cromatográfica para gases con una capacidad de 50 ml. En concreto, el modelo utilizado pertenece a la casa Hamilton y está diseñada para análisis de cromatografía líquida de alta precisión.

3.2.3 Matriz de sensores

Dentro del módulo genéricamente denominado matriz de sensores se ha incluido tanto los sensores de gases que conforman la matriz como toda la circuitería electrónica asociada.

3.2.3.1 Los sensores de gases

Los sensores de gases utilizados en ambos prototipos son de tipo semiconductor, concretamente de óxido de estaño. Este tipo de sensores son los más utilizados en los sistemas de olfato electrónico puesto que presentan una alta sensibilidad ante la presencia de diversos volátiles orgánicos, pueden ser integrados en un equipo portátil y su coste de fabricación es bajo. Además, no requieren de una instrumentación demasiado compleja para su operación y pueden ser utilizados en tiempo real [3.7-9].

Los sensores semiconductores basan su funcionamiento en la variación de la resistencia que presenta su capa activa ante diferentes compuestos volátiles. En la literatura, por lo tanto, están catalogados como sensores resistivos [3.10-13].

Cuando el sensor se calienta a una temperatura cercana a 400°C, ante la ausencia de oxígeno, los electrones libres fluyen fácilmente entre las fronteras granuladas de dióxido de estaño. En aire puro, el oxígeno, que atrapa electrones debido a su afinidad electrónica, queda adsorbido en la superficie de dióxido de estaño creando una barrera de potencial en las fronteras granulares. Esta barrera dificulta la libre circulación de electrones aumentando la resistencia eléctrica de la capa activa.

Ante una atmósfera rica en gases reductores (como pueden ser los gases combustibles), la superficie de dióxido de estaño absorbe estas moléculas gaseosas provocando su oxidación. Este proceso disminuye la barrera potencial facilitando la circulación de electrones libres, lo que reduce la resistencia del sensor.

La reacción entre los gases y el oxígeno de la superficie varía según la temperatura y composición de la capa activa. Combinando diferentes temperaturas de trabajo y dopantes en la capa activa se pueden crear sensores con sensibilidades optimizadas para detectar compuestos volátiles concretos, consiguiendo de esta manera una familia de sensores con sensibilidades solapadas entre sí.

La sensibilidad de este tipo de sensores se suele definir por la relación entre incremento de concentración de un determinado gas y el incremento de resistencia que se produce. En determinados rangos de concentración, la relación entre la resistencia del sensor y la concentración del gas desoxidado puede ser descrita mediante la ecuación experimental (3.1) :

$$R = A \cdot [C]^{-\alpha} \quad (3.1)$$

Donde: R, resistencia eléctrica del sensor

A y α son constantes características de cada gas

[C], concentración del gas

Los fabricantes suelen describir las especificaciones de cada modelo con gráficas logarítmicas en las que se representa la relación entre la resistencia final y la de reposo (R_S/R_0) frente a la concentración de diferentes volátiles. De todas formas, estas gráficas representan características típicas ya que la reproducibilidad entre sensores no es muy buena.

Además de utilizarse en sistemas de olfato electrónico, los sensores de óxido de estaño se han probado en diferentes aplicaciones comerciales. Su falta de selectividad permite considerarlos en un amplio rango de aplicaciones, aunque esa misma característica ha limitado en gran medida su aplicación industrial. La tabla 3.1 muestra algunas de las aplicaciones en las que se ha estudiado utilizar este tipo de sensores:

Categoría	Uso doméstico	Uso comercial e industrial
<u>Gases combustibles</u> - Metano - Propano - Hidrógeno	- Alarma de gases para casas, vehículos y barcas	- Sistemas de detección de gas para comercios, plantas industriales y petroquímicas - Detectores de gas portables
<u>Gases tóxicos</u> - Monóxido de carbono - Amoniacó - Otros	- Detectores de CO para casas, vehículos, barcas	- Detectores de fugas de amoniacó para refrigeradores - Detectores de amoniacó para el sector primario - Detectores de gas portables
<u>Alcohol</u>	Etilómetros	Etilómetros profesionales
<u>Vapores orgánicos</u>		Detectores de disolventes para fábricas
<u>CFCs</u> (Clorofluorocarbonatos)		Detectores CFCs para congeladores, aires acondicionados, procesos de limpieza para componentes eléctricos...
<u>Otros</u> - Oxígeno - Hidrógeno		- Detectores de oxígeno - Detectores de oxígeno para el mantenimiento de transformadores, baterías, etc.
<u>Detectores de fuego</u>	- Alarmas de fuego	- Sistemas de detección de fuego

Tabla 3.1: Aplicaciones de los sensores químicos de óxido de estaño.

3.2.3.2 Sensores Figaro

Figaro Inc. fue la primera empresa en comercializar este tipo de sensores [3.14,15] . La larga vida útil de sus unidades, su elevada sensibilidad, rapidez de respuesta, excelente durabilidad y bajo coste, son las características que permiten considerar su utilización en un gran abanico de aplicaciones dentro del sector industrial y de servicios.

En consecuencia, se creyó oportuno dotar de sensores Figaro al prototipo inicial. La línea de productos de la casa Figaro dispone de cuatro tipos de sensores elementales de los cuales se han usado modelos de las series 8 y 2000:

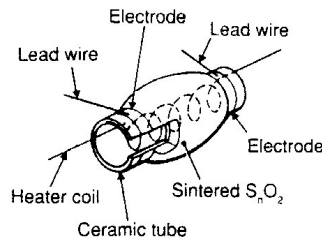


Figura 3.4: Esquema de los sensores de la serie 8

- *Sensores de la Serie 8*: Se caracterizan por tener un heater dentro de un tubo cerámico de alúmina, estando la pasta semiconductor depositada sobre la superficie cilíndrica con dos electrodos grabados en oro. Para conseguir una buena estabilidad a largo plazo, la pasta es sinterizada durante un período largo de tiempo bajo condiciones de temperatura y atmósfera controladas. Gracias a su complicada geometría, estos sensores gozan de alta sensibilidad y de buena estabilidad a largo plazo.

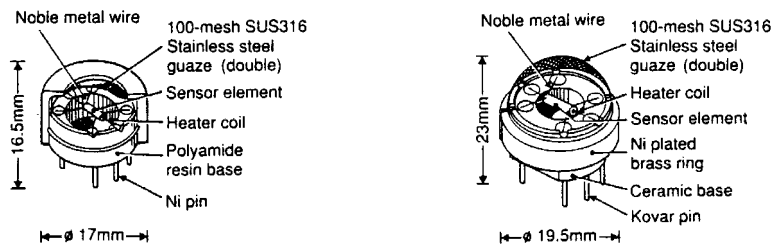


Figura 3.5: Versiones comerciales de los sensores de la serie 8

Comercialmente, los sensores de la serie 8 presentan dos configuraciones esquematizadas en la figura 3.5. La figura de la izquierda corresponde a una configuración estándar, con una base y encapsulado de resina. La segunda configuración (representada en la derecha) se caracteriza por tener una base cerámica sobre la que descansa el elemento sensor y estar recubierto por una malla metálica, empleada como medio de protección.

- *Serie 2000*: Estos sensores se caracterizan por tener una capa gruesa de óxido de metal semiconductor como elemento sensor. Tanto el *heater* como el elemento sensor están situados sobre un mismo sustrato plano. Estos sensores “thick film”

son fabricados con tecnología “screen printing” perfeccionada por la propia compañía. Además, el material sensor ya no se reduce a dióxido de estaño, incluyendo otros óxidos metálicos. En la serie 2000, existen cuatro modelos con diferentes estructuras:

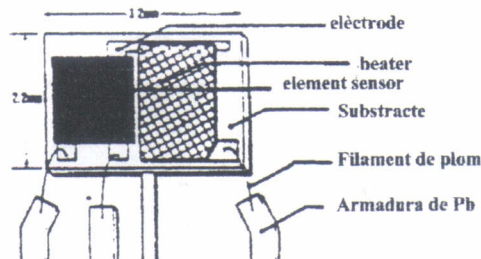


Figura 3.6: TGS Serie 2000 tipo S1

- Tipo S1: Su estructura es muy simple, dispone de una sola cara en la que se encuentran el heater y el elemento sensor.

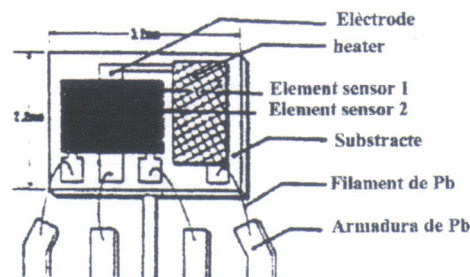


Figura 3.7: TGS Serie 2000 tipo S2

- Tipo S2: Dispone de dos elementos sensores, que están en una sola cara, lo que permite obtener información de dos sensores al mismo tiempo.

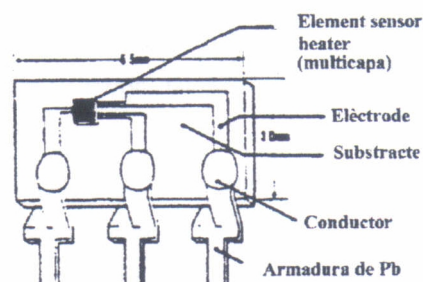


Figura 3.8: TGS Serie 2000 tipo M1

- *Tipo M1*: Su estructura es un poco más compleja, dispone de una sola cara en la que se encuentra de forma miniaturizada y multicapa el *heater* y el elemento sensor.

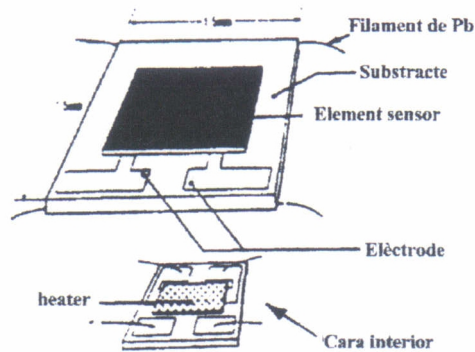


Figura 3.9: TGS Serie 2000 tipo D1

- *Tipo D1*: Se distingue del resto por tener dos caras. En una se encuentra el elemento sensor y en la otra el *heater*.

La tabla 3.2 muestra los sensores escogidos para el primer prototipo. En ella se especifica, para cada sensor, el modelo, su aplicación principal y los gases que detecta con mayor sensibilidad.

El criterio fundamental que se siguió al escoger estos sensores fue, por encima de cualquier otro, la diversidad. A pesar de que cada uno de los sensores se optimiza para sentir una determinada clase de vapores, lo cierto es que la mayoría son sensibles a un gran espectro de volátiles. Por ese motivo, sensores que en principio podían descartarse se introdujeron también en la matriz. Siguiendo con esa filosofía se optó por utilizar sensores de dos series diferentes para incrementar el espectro de tecnologías presente en el prototipo inicial, algo cada vez más habitual en las narices electrónicas comerciales, en las que, en casos extremos, se llegan a combinar matrices de sensores semiconductores con espectrómetros de masas [3.1].

Sensor	Aplicación	Detección
TGS 800	Control de calidad del aire	Aire contaminado: Tabaco, gasolina, etc...
TGS 822	Detección de vapores de alcohol	Alcohol, Tolueno, Xileno, etc...
TGS 825	Detección de gas tóxico	Sulfuro de hidrogeno
TGS 826	Detección de gas tóxico	Amoniaco
TGS 880	Control de cocinas	Gases volátiles, vapor de agua en alimentos
TGS 882	Control de cocinas	Vapor de alcohol en los alimentos
TGS 2610	Detección de gas combustible.	Hidrocarburos en general
TGS 2611	Detección de gas combustible.	Metano, Gas natural
TGS 2620	Detección de solubles orgánicos.	Alcohol, solubles orgánicos

Tabla 3.2: Sensores utilizados en el primer prototipo

3.2.4 Circuitos de polarización y medida

Los sensores de la casa Figaro, como todos los sensores de óxido de estaño, pueden catalogarse como sensores resistivos. Además, para trabajar correctamente, requieren que la capa activa esté a temperaturas elevadas, del orden de 300°C, por lo que necesitan una resistencia calefactora o “heater” que eleve la temperatura hasta su valor de trabajo.

A pesar de tener diferente patillaje, tanto los sensores de la serie 8 como los de la serie 2000 presentan un modo de operación muy parecido. La figura 3.10 muestra los patillajes de las series 8 y 2000 con sendos esquemas situados a la izquierda y derecha respectivamente.

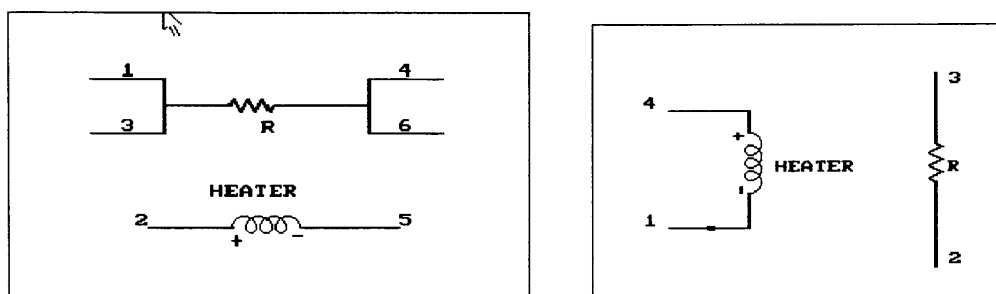


Figura 3.10: Esquema de los patillajes de los sensores Taguchi serie 8 y 2000

Como se puede observar, ambas series tienen terminales de acceso que permiten alimentar tanto la resistencia calefactora (“heater”) como medir la resistencia que presenta la capa activa del sensor ante la presencia de diferentes volátiles.

El circuito para medir la resistencia de la capa activa es fundamental ya que la variación de este parámetro físico es la base del funcionamiento de este tipo de sensores. La figura 3.11 muestra el circuito de medida de los sensores de la serie 8. El circuito de la serie 2000 es muy parecido.

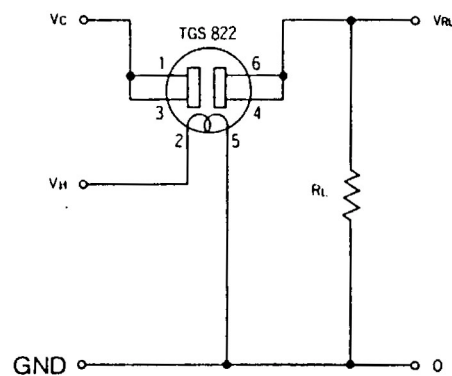


Figura 3.11: Circuito de medida de cualquier sensor de la serie 8

Como se puede observar en la figura, el circuito de medición que se utiliza se basa en el principio del divisor de tensión. Del análisis del circuito de la figura se desprende la siguiente fórmula:

$$R_S = \frac{V_C \cdot R_L}{V_{RL}} - R_L \quad (3.2)$$

Donde:

- V_C : Tensión de referencia (10 V d.c.)
- V_h : Tensión de heater (5 V d.c.)
- R_L : Resistencia de carga
- V_{RL} : Tensión de salida (variable a medir con la tarjeta de adquisición de datos)
- R_S : Resistencia medida del sensor, la cual variará según los volátiles presentes en la cámara.

El circuito de alimentación del heater y el divisor de tensión para medir la resistencia activa deben implementarse tantas veces como sensores tenga el equipo. Por ese motivo, para reducir el volumen, sólo la placa impresa que contiene los sensores está dentro de la cámara de medida. La placa que contiene la electrónica que controla todos los heaters y extrae las tensiones necesarias para medir las resistencias de todos los sensores se encuentra fuera de la misma.

3.2.5 Sistema de adquisición y procesado

Como se ha comentado con anterioridad, el módulo de procesado de señal está basado en un ordenador personal convenientemente equipado que es controlado por dos programas diferentes.

Un primer programa se encarga de la adquisición de datos y su posterior almacenamiento en un fichero de texto. Para ello debe controlar una tarjeta de adquisición de datos de 16 canales conectada al ordenador a través de su bus de expansión. El programa realiza un muestreo periódico de la tensión presente en la resistencia de carga de los circuitos de medida de cada uno de los sensores que componen la matriz. A partir de esa tensión calcula la resistencia de la capa activa mediante la ecuación 3.2. Al finalizar el proceso de medición, el conjunto de datos se archiva en un fichero de texto. Este archivo contiene la evolución temporal que han experimentado las resistencias activas de cada sensor de la matriz durante la medida.

Debido a la particularidad tanto del equipo como de la aplicación, este programa se escribió expresamente para el prototipo. En su desarrollo se utilizó un compilador de C comercial.

Posteriormente, una vez realizadas las medidas pertinentes, un segundo programa extrae las características importantes y aplica algoritmos de reconocimiento de patrones. Este segundo programa está escrito utilizando el lenguaje matemático MATLAB [3.16]. Los algoritmos de reconocimiento de patrones integrados son el análisis de componentes

principales (PCA) y las redes neuronales artificiales. En el capítulo 4, Procesado de señal, se detallará con mayor profusión la utilización de estas técnicas en el análisis de los datos generados por la matriz de sensores de la nariz electrónica.

3.3 PROTOTIPO DE FLUJO CONTINUO

3.3.1 Descripción general

En la figura 3.12 se puede contemplar una fotografía del segundo prototipo. Como ya se ha comentado anteriormente, la principal diferencia entre el primer y segundo equipo radica en el módulo de muestreo. En el nuevo prototipo este módulo está basado en la generación y control de un flujo de aire que dirige los volátiles desde la cámara de concentración hacia la de medida. Por este motivo, el sistema de muestreo consta de las tradicionales cámaras de concentración y medida más cuatro electroválvulas y una bomba de aire.

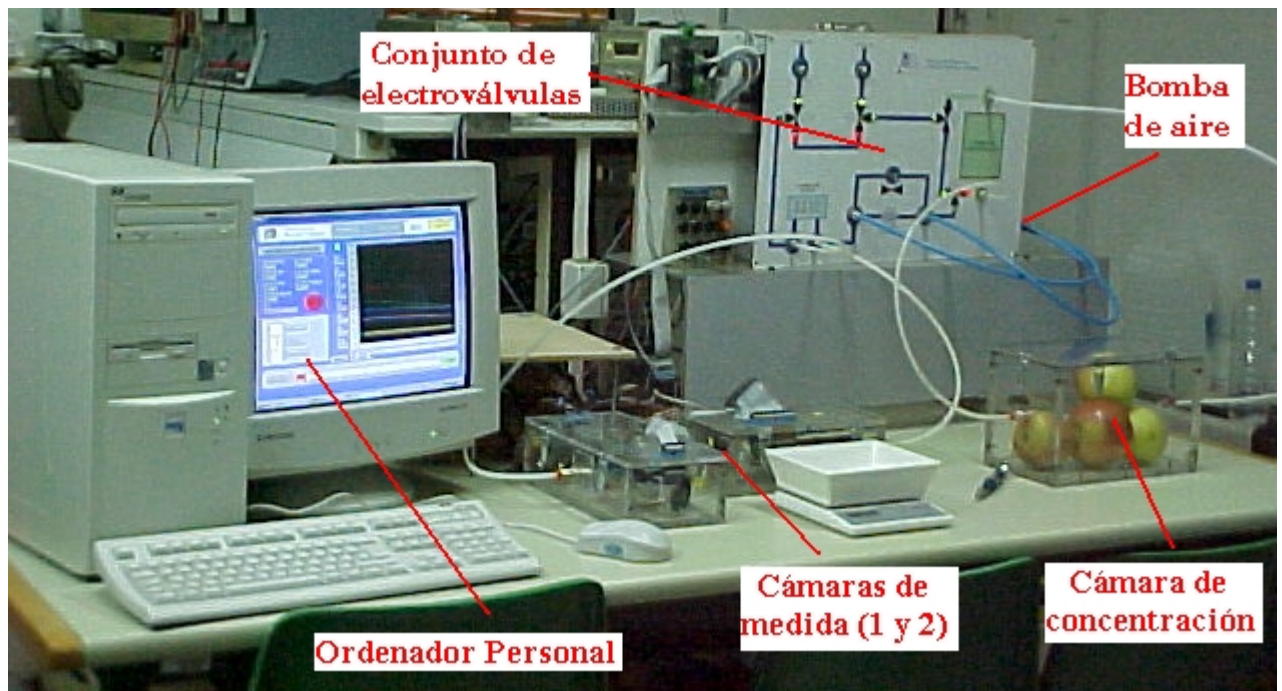


Figura 3.12: Vista general del segundo prototipo.

El procedimiento de medida en este segundo equipo consta de tres fases bien diferenciadas: La fase de concentración, la fase de medida y la fase de reposo. En todas ellas el flujo de aire que atraviesa la cámara de medida (donde están ubicados los sensores) permanece constante. La figura 3.13 muestra un esquema del módulo de muestreo en el que se pueden distinguir claramente los diferentes caminos que sigue el flujo de aire en cada una de las fases.

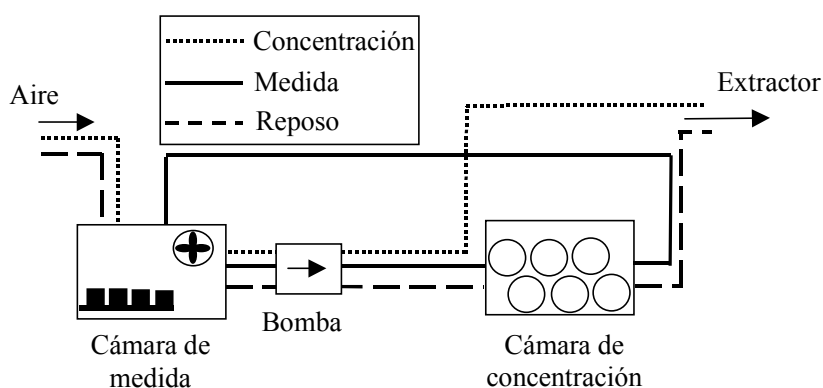


Figura 3.13: Diferentes caminos para el flujo de aire según la fase de medida

Cuando el sistema se encuentra en la fase de concentración, la bomba toma el aire del laboratorio y las electroválvulas lo conducen a través de la cámara de medida. El flujo de aire abandona el sistema y es devuelto al exterior sin circular a través de la cámara de concentración, ya que las electroválvulas han cerrado dicho camino con la finalidad de que se acumulen volátiles en su interior. Esta fase puede durar unos 90 minutos, tiempo suficiente para acumular una gran concentración de volátiles y así obtener una buena señal de los sensores.

Durante la fase de medida, la bomba de aire empuja los volátiles a través de un circuito cerrado desde la cámara de concentración a las de medida. Durante dicha etapa no entra ni sale aire del exterior. Para esta fase un tiempo total de 10 minutos es considerado adecuado, ya que este intervalo de tiempo permite que los sensores alcancen un valor estable.

Cuando la medida ha finalizado, el sistema pasa a la fase de reposo. En esta etapa se pretende limpiar todo el circuito de medida y restablecer la respuesta de los sensores a su resistencia de reposo. El aire es tomado del laboratorio, circula en primer lugar a través de las cámaras de medida, seguidamente pasa a través de la cámara de concentración, y empuja los volátiles que hayan podido quedar fuera del circuito. Una duración apropiada para esta fase puede ser de 15 minutos.

3.3.2 Sistema de muestreo

Este módulo es el que genera y controla el flujo de aire durante el proceso de medida. Dentro de él se encuentran las dos cámaras de medida, una bomba de aire con un flujo de dos litros por minuto y un conjunto de 4 electroválvulas.

3.3.2.1 Cámara de concentración

La cámara de concentración es un recipiente hermético de metacrilato en el que se introduce la fruta. Al igual que en el primer prototipo, se trata de una cámara que tiene como finalidad acumular los volátiles desprendidos por las piezas que se hallan en su interior.

La cámara dispone sobre dos de sus paredes laterales una entrada y una salida de aire que permiten trasladar los volátiles hasta las cámaras de medida en el momento adecuado. Si se cierran estas entradas mediante electroválvulas, se aísla el interior de la cámara. De esa manera, transcurrido un tiempo, se habrá creado una concentración de volátiles que posteriormente serán arrastrados por el flujo de aire hasta las cámaras de medida en las que se encuentran los sensores.

Además, esta cámara dispone de una tapa a través de la cual podemos introducir las piezas. Consta de unos anclajes y una junta de goma situados por su perímetro para poder cerrarla herméticamente. Dicha tapa tiene un *septum*, para poder extraer muestras de su atmósfera o bien inyectar sustancias de calibración con una jeringa cromatográfica. Sus dimensiones son 17.7x26.3x15.5 cm, totalizando un volumen de

7.2 litros. La figura 3.14 muestra una foto de la cámara de concentración con piezas de fruta en su interior. Como se puede observar, en su interior se pueden alojar un número máximo de muestras que dependerá de la variedad de fruta que se pretenda estudiar.



Figura 3.14: Cámara de concentración del prototipo de flujo continuo

3.3.2.2 Cámaras de medida

A diferencia del primer prototipo, en el segundo equipo se dispone de 2 cámaras de medida conectadas en serie con un volumen de 1.7 litros cada una. En ellas se alojan los diferentes sensores empleados, unos ventiladores para homogeneizar la atmósfera y unas sondas que facilitan la lectura de la humedad y temperatura en su interior, ya que estos dos factores influyen en la respuesta de los sensores de óxidos metálicos [3.17-18]. Para evitar el mayor número de interferencias posibles el laboratorio se mantuvo siempre a una temperatura de 20 °C y la humedad relativa entre el 50 y 65 %. Sin embargo, la humedad relativa en el interior de las cámaras se mantuvo siempre entorno al 30% de humedad relativa en reposo, aumentando como máximo al 45% al realizar medidas con fruta. La inclusión de una sonda de humedad permitió utilizar el

incremento de humedad relativa como información adicional para compensar su influencia utilizando algoritmos de aprendizaje supervisado [3.19-21].

El motivo de utilizar dos cámaras radica en la flexibilidad que aporta esta aproximación. Si para una aplicación se quieren utilizar menos sensores se puede conectar una de ellas, con la consecuente reducción en volumen y aumento de señal en los sensores debido a la mayor concentración de volátiles. En los estudios realizados con fruta, se alojaron sensores de las casas Figaro [3.14-15] y Fis [3.22] (series 8 y SB respectivamente) en la primera cámara, mientras que en la segunda se instalaron sensores FIS de la serie SP.

Por lo tanto, el prototipo dispone en total de 3 tipos de sensores, cuya tecnología de fabricación y respuesta ante la presencia de volátiles es diferente. Frente a una configuración en la que existieran sensores de un solo fabricante, la aproximación escogida permite monitorizar un mayor espectro de volátiles.

Las características constructivas de las dos cámaras son idénticas. En ambos casos se han construido con metacrilato transparente. En su interior, además de los sensores (colocados en la tapa que cierra la cámara) y el sistema de medición de humedad y temperatura (apoyado sobre la base de la cámara), se incluye un ventilador para homogeneizar la atmósfera del interior.

Al igual que en la cámara de concentración, las cámaras de medida disponen de una entrada y salida de aire, a través de las cuales circula un flujo de aire constante que en la fase de medida transporta los volátiles de la fruta que provienen de la cámara de concentración. Para sellar herméticamente las cámaras, se emplean 4 anclajes metálicos que ejercen una presión sobre la tapa y el cuerpo de la misma. Esta tapa descansa sobre una junta de goma que asegura el cierre hermético.

Se ha dispuesto un agujero fresado, posteriormente sellado, para colocar un conector a través del cual se alimenta eléctricamente a los sensores y al ventilador. Por ese mismo orificio se conectan las señales de los sensores hasta las placas que adaptan su señal

para ser adquiridas por las tarjetas de adquisición de datos. La figura 3.15 muestra las dos cámaras de medida conectadas en serie.

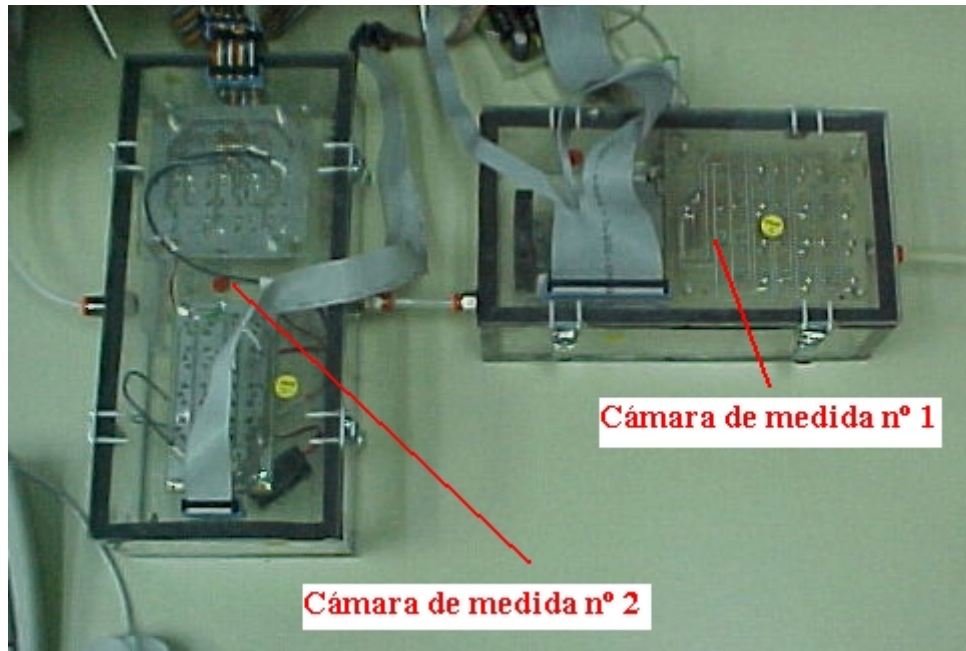


Figura 3.15: Detalle de las dos cámaras de medida del prototipo de flujo continuo

3.3.3 Matrices de sensores

Al igual que en el primer prototipo, los sensores utilizados en esta nariz electrónica fueron todos de óxido de estaño. Mientras que en el primer equipo se utilizaron sensores exclusivamente de la casa Figaro, en este segundo instrumento también se incluyeron sensores Fis de las series SB y SP.

Con la experiencia adquirida con el primer prototipo, el segundo equipo se diseñó con la intención de que fuese fácilmente reconfigurable. Esta característica se traduce, fundamentalmente, en poder cambiar de forma sencilla los sensores que constituyen la matriz de la nariz electrónica. La tabla 3.3 detalla los sensores utilizados así como las aplicaciones para las que han sido pensados.

<u>SENSOR</u>	<u>APLICACION</u>
FIS (Serie-SB)	
S1.-SB-AQ1A-00	Propósito general. Calidad del aire
S2.-SB-19-00	Hidrógeno
S3.-SB-AQ4-00	Humo de tabaco
S4.-SB-11A-00	Propósito general. Vapores inflamables
S5.-SB-31-00	Propósito general. Disolventes orgánicos
S6.-SB-30-00	Alcohol. Disolventes orgánicos
S7.-SB-15-00	Propano/Butano
Taguchi (Serie-8)	
S1.-TGS-800-1	Calidad del aire. Humo de tabaco, Vapores de gasolina
S2.-TGS-800-2	Calidad del aire. Humo de tabaco, Vapores de gasolina
S3.- TGS-880-2	Control de cocina. Olores alimenticios,vapores, humedad.
S4.- TGS-880-1	Control de cocina. Olores alimenticios,vapores, humedad.
S5.- TGS-825-1	Sulfato de hidrógeno
S6.- TGS-822	Disolventes orgánicos
S7.- TGS-882-5	Control de cocina. Vapores alcohólicos de alimentos
FIS (Series-SP)	
S1.-SP-MW0	Propósito general. Control de cocina.
S2.-SP-12A-00	Metano
S3.-SP-11-00	Propósito general. Vapores inflamables
S4.-SP-MW1-00	Humedad. Control de cocina.
S5.-SP-AQ1-00	Propósito general. Calidad del aire
S6.-SP-19-00	Hidrógeno
S7.-SP-AQ3-00	Humo de tabaco

Tabla 3.3. Sensores utilizados en el segundo prototipo

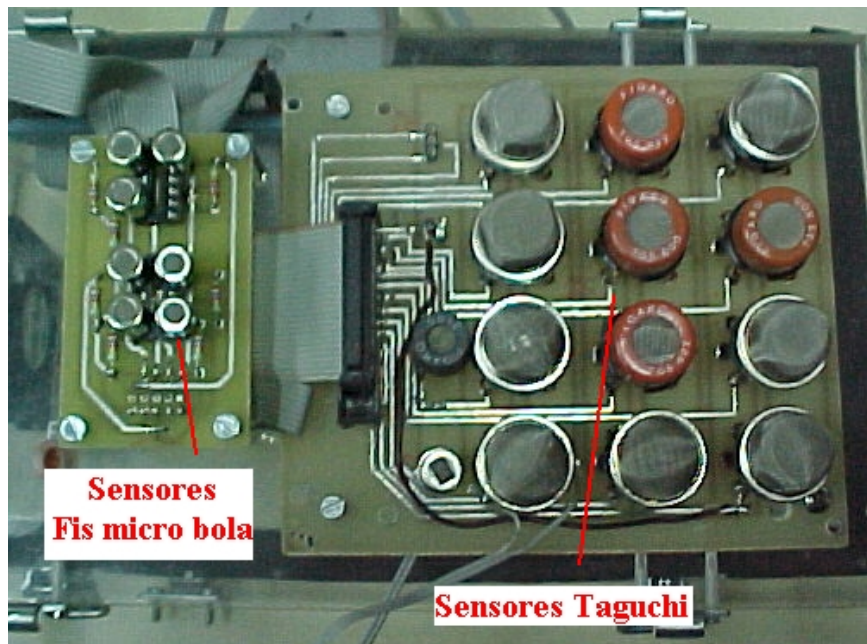


Figura 3.16: Matrices de sensores ubicadas en la cámara de medida 1

La figura 3.16 muestra con detalle las dos matrices de sensores de tecnología diferente que se alojan en la primera cámara de medida. En ella se puede llegar a incluir un total de 21 sensores repartidos en dos placas impresas. Una de ellas contiene sensores Taguchi y la otra sensores FIS microbola (SB). En la segunda cámara se ha incluido una única matriz con sensores FIS serie SP. Las dimensiones de la cámara pueden permitir, en un futuro, incluir sensores adicionales. La figura 3.17 muestra una foto de los sensores incluidos en la segunda cámara de medida.

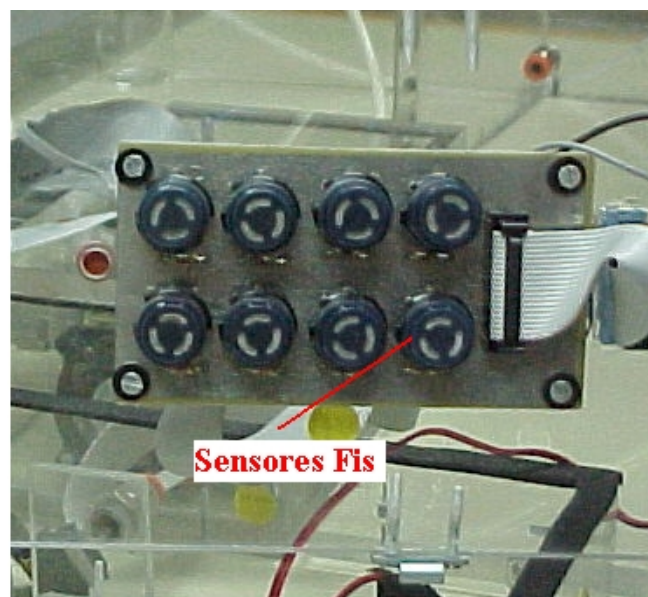


Figura 3.17: Matriz de sensores en la cámara de medida número 2

A continuación pasaremos a describir las características más relevantes de estas dos nuevas series de sensores, puesto que las características de los sensores Taguchi ya han sido expuestas en apartados anteriores.

3.3.3.1 Sensores FIS serie SP

Son sensores de Óxido de Estaño cuyo funcionamiento y tecnología es muy similar al de los Taguchi Serie 8 [3.15]. En la siguiente figura podemos ver un esquema del elemento sensor correspondiente a los sensores de la serie SP.

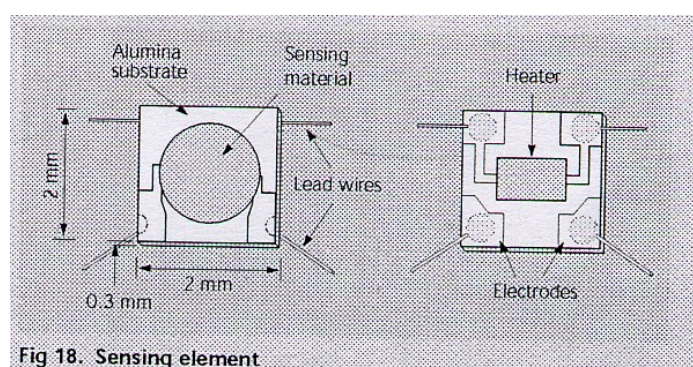


Figura 3.18: Elemento sensor de la serie SP.

Su fabricación se basa en una avanzada tecnología de impresión Thick-Film aplicada sobre un delgado substrato cerámico de alúmina en el que se encuentran los electrodos de oro. Esta configuración permite una reducción de potencia del 50% respecto a los sensores de la casa Figaro. En la cara opuesta del substrato se deposita el heater de platino u óxido de rutenio. El encapsulado que protege al elemento sensor se puede ver en la figura 3.19.

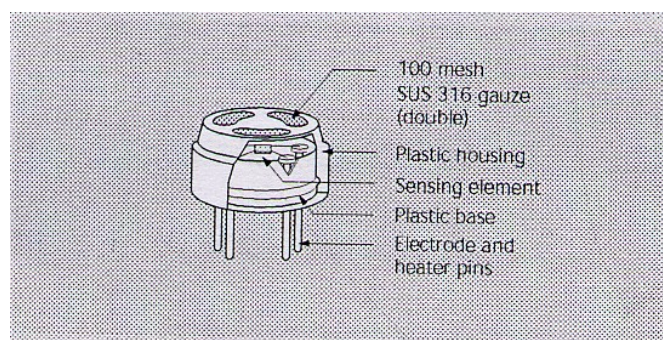


Figura 3.19: Encapsulado del sensor SP

La tabla 3.4 muestra de forma detallada los diferentes modelos de la serie SP [3.22] que comercializa Fis, incluyendo los gases para los que se ha optimizado la sensibilidad de cada sensor y el consumo de potencia de heater.

Categoría	Gas detectado	Modelo (consumo de potencia)
Mezcla de dos gases	CO y Metano	#
Gases inflamables	Hidrocarburos	SP-11-00 (Ph: 400 mW)
	Metano	SP-12A (Ph: 380Mw)
	Propano/butano	SP-15A-00 (Ph: 380 mW)
	Hidrógeno	SP-19-00 (Ph: 315 mW)
Disolventes orgánicos	De propósito general	SP-31-00 (Ph: 315 mW)
	Alcohol	SP-32-00 (Ph: 315 mW)
Gases refrigerantes	HCFCs (R-22)	SP-41-00 (Ph: 400 mW)
	HFCs (R-134a)	SP-42-00 (Ph: 400 mW)
		SP-42A sensor bajo prueba
Gases Tóxicos	Monóxido de carbono	#
	Sulfuro de Hidrógeno	SP-51 sensor bajo prueba
	Amoniaco	SP-53-00 (Ph: 315mW)
Gases oxidantes	Ozono	SP-61 sensor bajo prueba
	Óxido de nitrógeno	#
	Clorhídrico	SP-65 sensor bajo prueba
Control de alimentos	De propósito general	SP-MW0-00 (Ph: 400 mW)
	Humedad	SP-MW1-00 (Ph: 400 mW)
	Alcohol	SP-MW2-00 sensor bajo prueba
	Gases combustibles	SP-MW3-00 sensor bajo prueba
Control de calidad de aire interior	Propósito general	SP-AQ1-00 (Ph: 315mW)
		SP-AQ2-00 (Ph: 315mW)
		SP3-AQ2-00 (Ph: 315mW)
		SP3-AQ2Y-00 (Ph: 315mW)
	Humo de cigarrillos	SP-AQ3-00 (Ph: 315mW)
	CO ₂	Bajo test
Gases emitidos por los vehículos	Gasolina	SP-AD1/SP-AD3 bajo test
	Gas-oil	SP-AD2/SP-AD3 bajo test

Tabla 3.4: Campo de aplicaciones de los sensores Fis SP(Ph = Potencia heater).

3.3.3.2 Sensores FIS serie SB

Al igual que la casa Figaro con su serie 2000, FIS ha sacado al mercado sensores con tecnología diferente a la de las series 8 y SP [3.22]. En la serie SB, el objetivo primordial ha sido la disminución del consumo de potencia y tamaño del sensor gracias a la integración de todo el conjunto que constituye el sensor (elemento sensor, encapsulado, patillaje, etc.).

La serie SB [3.23] tiene un elemento sensor como el de la figura 3.20, semiconductor tipo *mini bead*, principalmente compuesto de dióxido de estaño (SnO_2). La bobina calefactora y los electrodos están incrustados en el elemento sensor.

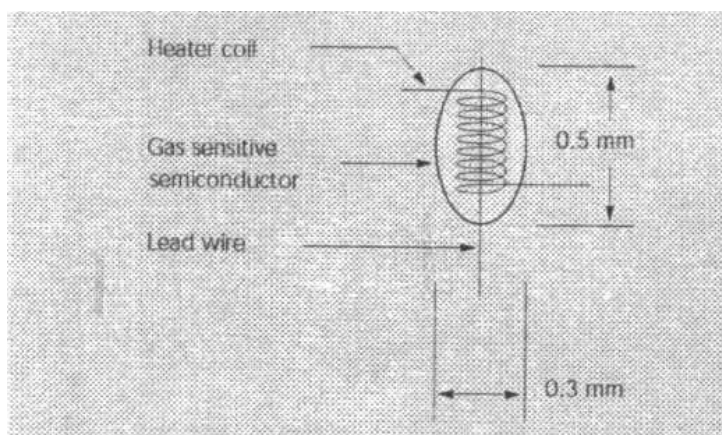


Figura 3.20: Elemento sensor

Los sensores Fis micro bola que se han empleado en la primera cámara de medida presentan un encapsulado metálico, que protege al elemento sensor. Dicho encapsulado está compuesto por dos mallas metálicas de acero inoxidable que descansan sobre una base de plástico.

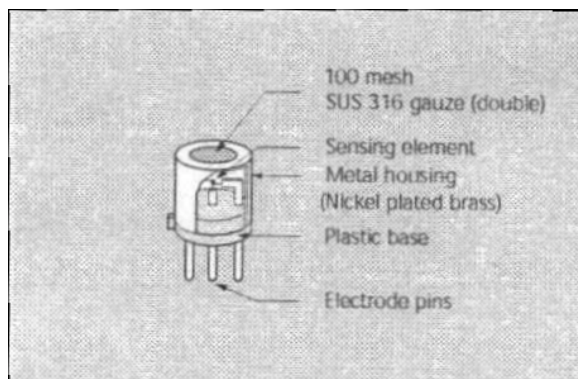


Figura 3.21: Presentación comercial de los Fis serie SB

La tabla 3.5 recoge de forma detallada los diferentes modelos FIS SB [3.23] con los gases que detectan y su consumo de potencia. Obsérvese que ésta es notablemente inferior a la que consumen los FIS SP, facilitando su incorporación en equipos portátiles.

Aplicación	Gas detectado	Modelo (consumo de potencia)
Mezcla de dos gases	CO y Metano	SB-95-00 (Ph: 120 mW - máx.)
Gases inflamables	Hidrocarburos	SB-11 ^a -00 (Ph: 120 mW)
	Metano	SB-12 ^a sensor bajo prueba
	Propano/butano	SB-15-00 (Ph: 120 mW)
	Hidrógeno	SB-19-00 (Ph: 120 mW)
Disolventes orgánicos	De propósito general	SB-31-00 (Ph: 120 mW)
	Alcohol	SB-30-00 (Ph: 120 mW)
Gases refrigerantes	HCFCs (R-22)	SB-41-00 (Ph: 120 mW)
	HFCs (R-134a)	SB-42-00 (Ph: 120 mW)
		SB-42 ^a sensor bajo prueba
Gases Tóxicos	Monóxido de carbono	SB-50-00 (Ph: 120 mW - máx.)
		SB-500-00 (Ph: 120 mW - máx.)
	Amoniaco	SB-53 sensor bajo prueba
Gases oxidantes	Óxido de nitrógeno	SB-62 sensor bajo prueba
Control de la calidad del aire (sistemas de control de ventilación, ...)	De propósito general	SB-AQ1A-00 (Ph: 120 mW)
	Humos de cigarrillos	SB-AQ4-00 (Ph: 140 mW)

Tabla 3.5: Campo de aplicaciones de los sensores Fis SB (Ph = Potencia heater)

Al igual que los sensores de la serie SP, los sensores SB son sensores resistivos cuyo funcionamiento está basado en la relación entre la resistencia del sensor (R_s), y la concentración de los gases. La resistencia del sensor se decrementa en función del gas depositado sobre el sensor según una función logarítmica (de igual forma que los TGS).

El rasgo más significativo de estos sensores es el bajo consumo del heater, garantizando, además, las características típicas de un sensor de óxido de estaño:

- Larga vida útil
- Alta fiabilidad
- Modo de empleo sencillo
- Bajo coste

3.3.4 Circuitos de polarización y medida

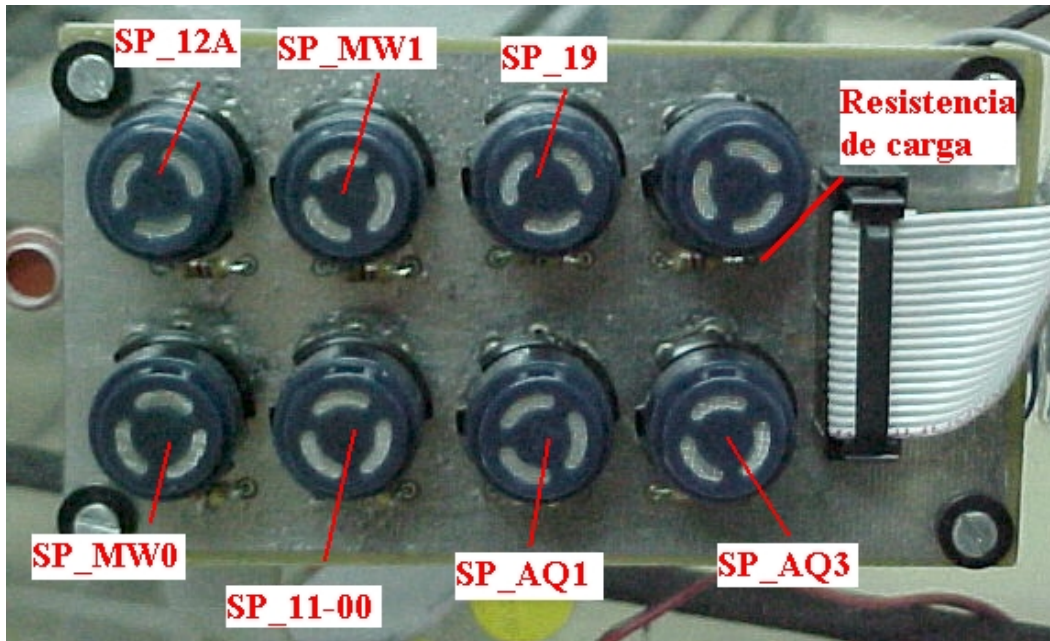


Figura 3.22: Detalle de la placa que contiene los sensores Fis SP

La figura 3.22 muestra con detalle la placa donde se aloja la matriz de sensores FIS SP. Como ya se ha comentado con anterioridad, esta placa se encuentra en la segunda cámara de medida. Los sensores Fis están montados sobre un zócalo idéntico al utilizado para los sensores Taguchi.

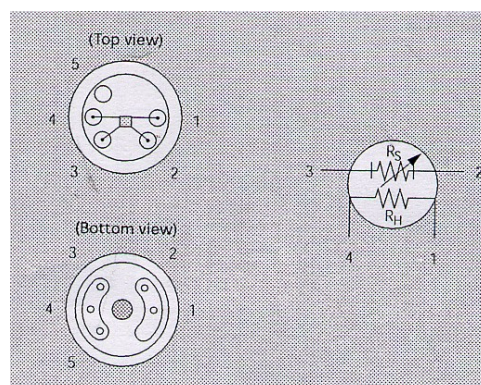


Figura 3.23: Patillaje de los sensores FIS SP

En la figura 3.23 se puede ver el layout de los pins y su equivalente eléctrico, que será de gran utilidad a la hora de diseñar el circuito de medida de la resistencia del sensor. Se trata de una disposición idéntica a la de los sensores de la casa Taguchi, lo cual implica que su circuito de medida lo forma una simple resistencia de carga que configura un divisor de tensión con la R_s del sensor, exactamente el mismo circuito que en su momento se detalló para los sensores Taguchi. (Véase figura 3.11)

Los sensores Fis micro bola, dada su reducida dimensión y patillaje, están montados sobre un zócalo "dual in line", aprovechando los pines de éste para insertar los sensores. Además, se ha utilizado esta misma placa para colocar las resistencias de carga del circuito de medida. Véase a continuación dicha placa:

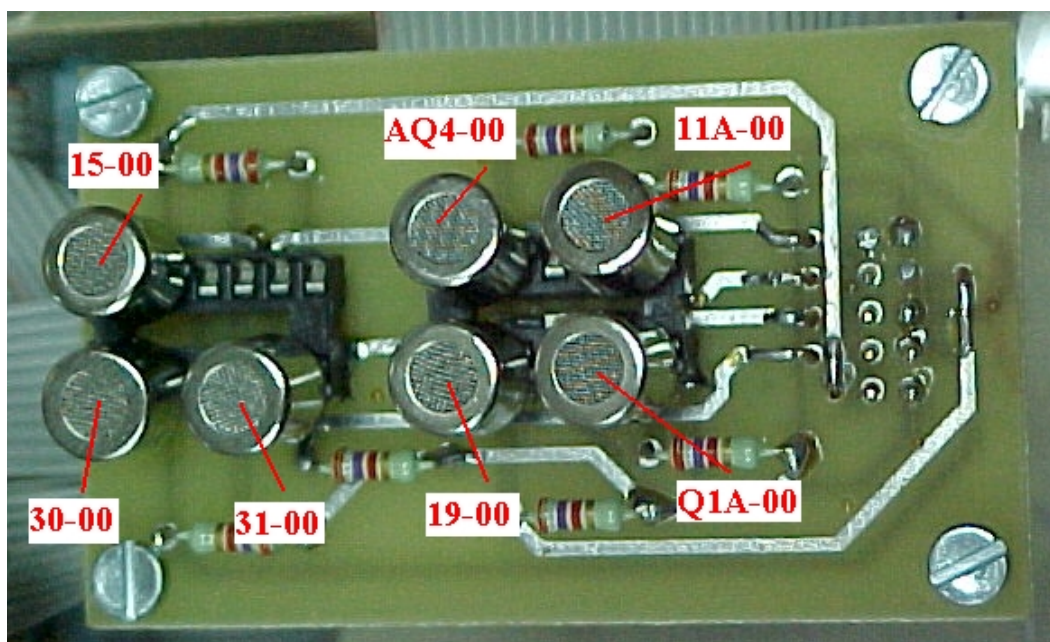


Figura 3.24: Detalle de la placa que contiene todos los sensores Fis micro bola

Los sensores Fis micro bola se encuentran ubicados en una placa situada en el interior de la primera cámara de medida. Sin embargo, para ahorrar volumen, los circuitos de medida de estos sensores se encuentran en otra placa (interficie) situada fuera de la cámara.

Para poder medir las variaciones de la resistencia del sensor, debemos ayudarnos de un circuito de medida (facilitado por el fabricante) formado básicamente por una resistencia de carga y una tensión de control:

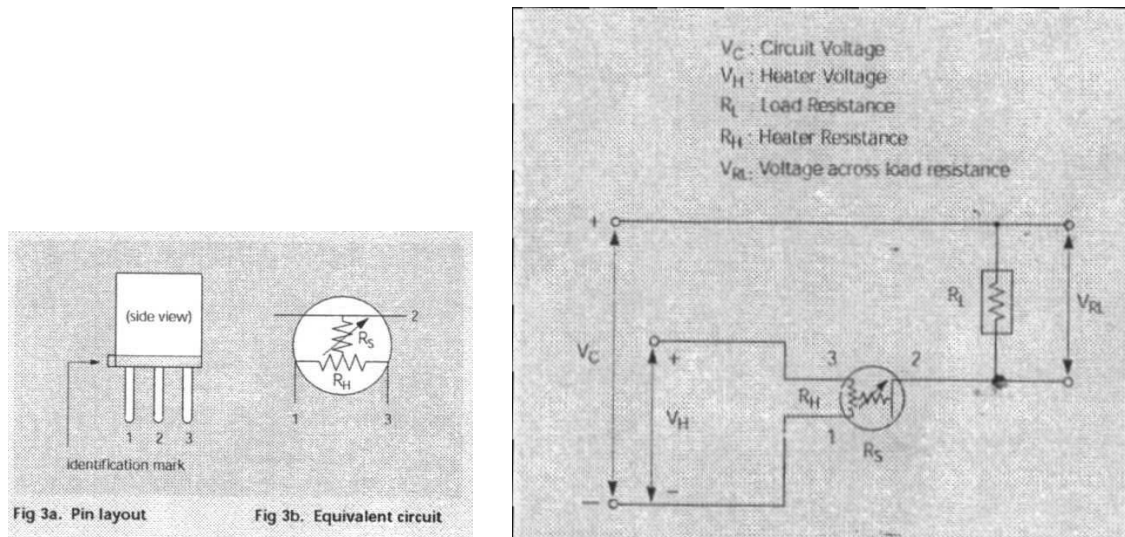


Figura 3.25: Circuito de medida D.C. de los sensores Fis micro bola, y patillaje del sensor.

Las condiciones de operación del circuito de la figura anterior son las siguientes:

- V_C : Tensión de control, que debe ser menor a 5 V d.c.
- V_H : Tensión de heater, la cual debe estar entre 900 mV y 1 V, aproximadamente
- R_L : Resistencia de carga, su valor es variable. Se recomienda que sea superior a 750Ω . En nuestro caso se optó por un valor de $R_L = 2700 \Omega$.
- V_{RL} : Tensión de salida flotante que deberemos referenciar a una masa común con el circuito que detallaremos a continuación, en el siguiente punto.
- R_S : Resistencia medida del sensor, que variará según los volátiles presentes en el sensor. Dicho valor será el que calcularemos mediante software, y lo monitorizaremos para que el usuario tenga conocimiento de la evolución de la respuesta del sensor.

La placa interficie se encarga de realizar la comunicación entre las señales recogidas procedentes de los sensores alojados en la cámara de medida número 1 y la tarjeta de adquisición de datos. Según se ha podido ver en la figura 3.25, la tensión de salida V_{RL}

no está referenciada a una masa común. Para poder fijar dicha tensión a una masa común (y evitar así incompatibilidades con los canales de entrada de la tarjeta) tanto para el sensor, resistencia de carga, fuentes de alimentación y la tarjeta de adquisición, se ha diseñado un circuito basado en un amplificador operacional que garantiza una copia referida a masa de la tensión flotante que conectamos a su entrada. La figura 3.26 detalla este esquema:

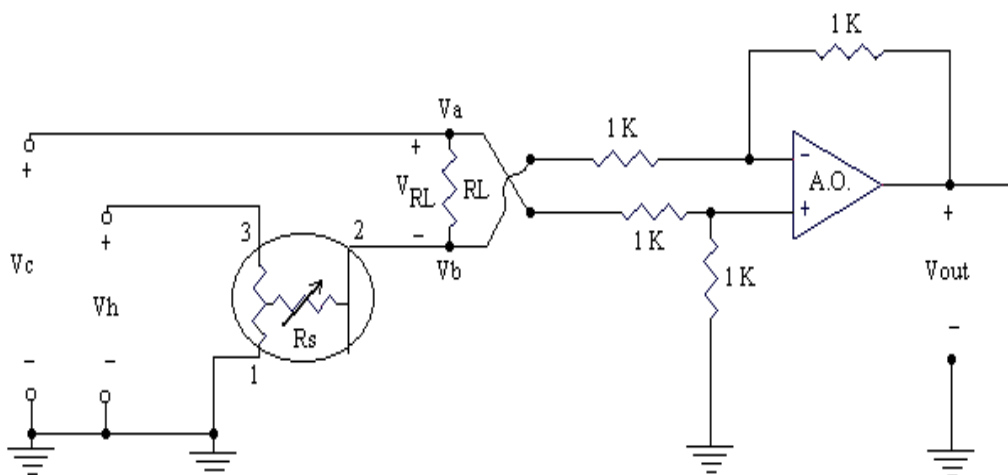


Figura 3.26: Circuito para referenciar a masa la tensión de salida de los sensores Fis micro bola

Básicamente la placa interficie está constituida por 7 circuitos como el de la figura anterior (uno para cada sensor Fis micro bola utilizado), y además dispone de una serie de conectores para cable aéreo plano, a través de los cuales se reciben las señales de la matriz de sensores Fis micro bola.

Todas estas señales, junto con las de alimentación, son canalizadas hasta un conector, y enviadas hacia la tarjeta (incluida dentro del ordenador personal PC), compartiendo una única masa común, tanto para las señales, como para la tarjeta.

3.3.5 Sistema de adquisición y procesado

Al igual que en el prototipo inicial, un ordenador personal gestiona el proceso experimental y posteriormente aplica algoritmos de reconocimiento de patrones para extraer resultados sobre las medidas realizadas. Estas dos tareas son realizadas por

programas diferentes que no es necesario que se ejecuten de forma paralela ni coordinada.

A diferencia del primer equipo en el que el muestreo se realizaba de forma manual, el nuevo equipo gestiona el proceso de forma automática. Con la ayuda de 2 tarjetas de adquisición de datos (modelo PCL-818 [3.24]), con un total de 32 canales analógicos/digitales de entrada, y un programa realizado a medida en lenguaje C, se controla todo el proceso de medida.

CÁMARA N°	SENSORES	TARJETA DE ADQUISICIÓN
1	Taguchi y Fis micro bola	PCL 818_1
2	Fis SP	PCL 818_2

Tabla 3.6: Distribución de las placas de adquisición de datos, según las cámaras de medida

Las electroválvulas son gobernadas por las salidas digitales de la tarjeta de adquisición, reconduciendo el flujo de aire según en la fase de la medida que se encuentre el sistema. Cuando los sensores están expuestos a los volátiles, durante la fase de medida, el ordenador almacena los cambios de resistencia que experimentan los sensores. Al finalizar cada medida, el programa de adquisición genera un fichero de texto que almacena en el disco duro del ordenador. Los datos que procesa este programa son:

- La resistencia de todos los sensores (Taguchi, FIS SB y FIS SP)
- La temperatura (de la cámara de medida 1 y 2).
- La humedad (de la cámara de medida 1 y 2).

Estos cálculos se presentan en pantalla para que el usuario pueda ir comprobando como varía la resistencia de los sensores en tiempo real.

Posteriormente, un nuevo programa matemático de alto nivel procesa la señal para extraer los resultados requeridos por cada aplicación. Al igual que en el primer prototipo, se trata de algoritmos programados con MATLAB. En el capítulo 4, “Procesado de señal”, se profundizará sobre los algoritmos de inteligencia artificial que utilizan los prototipos diseñados en este trabajo.

REFERENCIAS

- [3.1] Alpha Mos inc., <http://www.alpha-mos.com>
- [3.2] Lennartz Electronic, <http://www.lennartz-electronic.de>
- [3.3] Cyranno, <http://www.cyrannosciences.com>
- [3.4] Applied Sensor inc., <http://www.appliedsensor.com>
- [3.5] MicroSensor Systems, <http://www.microsensorsystems.com>
- [3.6] Element Ltd., <http://www.element.is>
- [3.7] Göpel W., *Sensors, a Comprehensive Survey*, Vol. 2, Chemical and biochemical Sensors, VCH, Weinheim, 1991, pp. 430-466.
- [3.8] Moseley P.T., Norris J.O.W. and Williams D.E., *Techniques and Mechanisms in Gas Sensing*, Adam Hilger, Bristol, 1991
- [3.9] di Benedetto L.T., Alexander P.W. and Hibbert D.B., *Portable battery-powered flow injection analyser for volatile alcohols using semiconductor gas sensors*, Anal. Chim. Acta, 321 (1996), 61-67
- [3.10] Moseley P.T., Prudenziati M., *Thick Film sensors*, Elsevier Science, Amsterdam, 1994, 261-288
- [3.11] Moseley P.T. and Tofield B.C., *Solid-State gas sensors*, Adam Hilger, Bristol, 1987, 1-16
- [3.12] Kostafad P., Non-stoichiometric, *Diffusion and Electrical Conductivity in Binary Metal Oxides*, Wiley-Interscience, New York, 1972
- [3.13] Clifford P.K. and Tuma D.T., *Characteristics of semiconductor gas sensors I. Steady-state gas response*, Sensors and Actuators,3, (1983), 233-54
- [3.14] N. Taguchi, US Patent 3 631 436, 1971
- [3.15] Figaro Engineering Inc, *Figaro Gas Sensor Technical Reference*
- [3.16] Matlab (version 5.3), The Mathworks, inc. <http://www.mathworks.com>
- [3.17] Van Geloven P., Honore M., Roggen J., Leppavuori S., Rantala T., *The influence of relative humidity on the response of tin oxide gas sensors to carbon monoxide*, Sensors and Actuators B, 4, (1991), 185-188

- [3.18] Romppainen P., Lantto V., Leppavuori S., *Effects of water vapour on the CO response behaviour of tin dioxide sensors in constant temperature and temperature-pulsed modes of operation*, Sensors and Actuators B,1, (1990), 73-78
- [3.19] J.Brezmes, E.Llobet, X.Vilanova, X.Correig, *Vapour identification and quantification under variable moisture conditions using gas sensor arrays and artificial neural networks*, Proceedings of the 11th European Conference on Solid-State Transducers, Eurosensors XI, Warsaw (Poland) 1997, 1317-1320
- [3.20] J.Brezmes, E.Llobet, X.Vilanova, X.Correig, *Quantification of volatile compounds using gas sensor arrays and artificial neural networks*, Proc. Of the International conference on Engineering applications of neural networks, EANN96, Londres (1996), pp 91-96
- [3.21] Brezmes, J.; Llobet, E.; Vilanova, X.; Correig, X., *Compensation of moisture effects in the quantification and identification of VOC's using artificial neural networks*, Proceedings of CAC'96, Chemometrics in Analytical Chemistry, Tarragona (1996), pp 109
- [3.22] FIS inc., <http://www.fisinc.co.jp>
- [3.23] Products Review, Sensor and systems technology, FIS inc., <http://www.fisinc.co.jp>
- [3.24] PC labcard, Advantech co, Ltd., <http://www.advantech.com>

CAPITULO 4:

PROCESADO DE SEÑAL

4.1 INTRODUCCIÓN

Una de las piezas angulares en toda nariz electrónica la constituye el conjunto de técnicas que tratan la información obtenida a través de la matriz de sensores químicos. De hecho, algunos fabricantes catalogan como narices electrónicas instrumentos de medida tradicionales, como los espectrómetros de masas, a los que añaden un software de análisis multivariante que funciona sobre un ordenador personal que se acopla al equipo [4.1].

El procesado de señal multivariante que se integra en las narices electrónicas actuales tiene su origen en la química analítica, también denominada “quimiometría”. A las técnicas como el Principal Component Analysis (PCA) o el Partial Least Squares (PLS), consideradas ya tradicionales, se han ido añadiendo nuevos algoritmos que se engloban dentro del campo denominado genéricamente “Reconocimiento de patrones” o “Técnicas de inteligencia artificial”. Dentro de esta última denominación englobaríamos a los diferentes algoritmos que implementan redes neuronales artificiales. Generalmente, estas técnicas se programan en un ordenador que además controla todo el proceso de medida y archiva convenientemente los datos adquiridos en cada experiencia. En el caso de las narices portátiles, los algoritmos inteligentes son implementados en microcontroladores, ya que la potencia de cálculo necesaria en la

mayoría de aplicaciones no es muy elevada debido a que no es necesario dar la respuesta inmediatamente, si no que se pueden tolerar algunos segundos o incluso minutos de retardo.

Todos estos algoritmos pueden ser sometidos a diferentes clasificaciones atendiendo a sus características. Así podemos distinguir entre algoritmos supervisados y no supervisados si nos fijamos en el proceso de aprendizaje, algoritmos de clasificación o de cuantificación en función de la naturaleza (binaria o analógica) de la respuesta que deben proporcionar a cada estímulo de entrada, lineales o no lineales según las operaciones que realicen, paramétricos o no paramétricos en función de si se hacen suposiciones iniciales sobre el proceso a modelar, etc. Debido a los diferentes objetivos que se han perseguido en cada experiencia, en este trabajo de investigación se han probado diferentes técnicas multivariantes. La tabla 4.1 resume las principales características de los algoritmos aplicados.

Técnica	Algoritmo	Aprendizaje	Paramétrica	Aplicación
PCA	Lineal	No supervisado	No	Clasificación
PLS	Lineal	Supervisado	Si	Cuantificación
Feedforward-backpropagation ANN	Neuronal	Supervisado	No	Clasificación/ cuantificación
Fuzzy Art	Neuronal	No supervisado	No	Clasificación
Fuzzy Artmap	Neuronal	Supervisado	No	Clasificación

Tabla 4.1: Principales características de los algoritmos aplicados al los sistemas de olfato diseñados

La evolución de las narices electrónicas va ligada a la de las técnicas que se han aplicado para procesar los datos obtenidos en cada medición. Desde sus comienzos diversos autores han presentado resultados aplicando diferentes técnicas de procesado. Al principio se utilizaron algoritmos lineales ya conocidos en química analítica. Por ejemplo, en [4.2] se describen los resultados obtenidos en la clasificación de diferentes tipos de café utilizando funciones discriminantes, mientras que en [4.3] se describe la clasificación cervezas y whiskies japoneses utilizando análisis de componentes principales.

De todas maneras, la detección de mezclas gaseosas (aromas, por ejemplo) va ligada a la aparición de interacciones no lineales entre los diferentes compuestos que reaccionan con los sensores. Para procesar esta información es conveniente utilizar algoritmos que no requieran linealidad en los datos y que no hagan suposiciones a priori sobre los procesos que rigen dichas interacciones. Como las redes neuronales con tres capas presentaban suficientes grados de libertad para solucionar cualquier problema de clasificación, muchos autores se dedicaron a estudiar su aplicación a los sistemas de olfato electrónico [4.4]. Otra de las ventajas que mostraban era su capacidad para cuantificar la concentración de diferentes compuestos en una mezcla gaseosa. En [4.5] se puede encontrar uno de los trabajos pioneros en los que se explotaba esta capacidad y en [4.6] uno de los primeros trabajos en los que se intenta discernir entre diferentes perfumes utilizando redes neuronales.

Lejos de ser un tema de investigación agotado, el procesado de señal en sistemas de olfato electrónico sigue siendo un activo campo de investigación impulsado por el constante abaratamiento de los sistemas informáticos que los soportan. Cada año se proponen nuevos métodos que pretenden eliminar o reducir las limitaciones que presentan los sensores químicos. Por ejemplo, en [4.7] se presenta uno de los primeros trabajos que aborda el tema de la calibración y en [4.8] el autor de esta tesis presenta un estudio de cómo una red neuronal puede neutralizar la interferencia de la humedad en el funcionamiento de los sensores de óxidos metálicos. Otras redes neuronales como las de resonancia adaptativa [4.9] y técnicas de procesado que explotan los parámetros dinámicos de la respuesta de los sensores son tendencias con gran futuro en el mundo de los sistemas de olfato electrónico. Para mayor información, se recomienda al lector que consulte la referencia [4.10] en la que se presenta un completo repaso de todas las técnicas utilizadas hasta la fecha además de valorar las diferentes líneas de trabajo que deben definir el futuro del procesado de señal en sistemas de olfato electrónico.

En los equipos de olfato electrónico, el objetivo principal del procesado de señal es el de extraer la información relevante para obtener la respuesta deseada, minimizando las interferencias (ruido) que pueda introducir la matriz de sensores. En el caso de la fruta, estas técnicas deben realizar dos funciones claramente diferenciadas:

- Clasificar muestras en función de su estado de maduración. Para ello se pueden aplicar cualquiera de las técnicas que clasifican las medidas a partir de la información multidimensional obtenida con la nariz electrónica. Las técnicas no supervisadas permiten ver la similitud entre medidas sin imponer ningún criterio, mientras que las supervisadas permiten clasificar atendiendo a un criterio previamente establecido en la fase de entrenamiento.
- Estimar el valor de diferentes indicadores de calidad de forma no destructiva a partir de la señal que se obtiene en la matriz de sensores. En este caso es necesario aplicar técnicas de cuantificación multivariantes como el PLS o las redes neuronales feed-forward que aprenden a relacionar señales en los sensores (generadas por los volátiles de la fruta) con condiciones físico-químicas en las piezas de fruta analizadas.

En este capítulo, cada subapartado describe la utilización de un algoritmo diferente, dando especial relevancia a los aspectos novedosos que se hayan podido derivar de la adaptación de técnicas conocidas a la aplicación que nos ocupa. Dependiendo del algoritmo y su utilización, los aspectos novedosos recaen en la modificación del propio algoritmo, en la evaluación de resultados, en el modo de utilización, etc. Así, por ejemplo, el PCA se ha utilizado de forma innovadora en la detección de derivas en los sensores. Otras técnicas, mucho más novedosas (como Fuzzy Artmap), han tenido que ser adaptadas internamente para un funcionamiento adecuado en el campo de las narices electrónicas.

Por otro lado, no hay que olvidar que la aplicación de técnicas de reconocimiento de patrones requieren un preprocesado previo mediante el que se extraiga y normalice la información relevante que suministra la matriz de sensores químicos. Este paso puede ser crucial en el éxito de la aplicación y debe ser tratado con sumo cuidado. Por este motivo se ha incluido un subapartado en el que se explican con detalle todos los pasos previos a la aplicación de técnicas multivariantes, que incluyen desde la extracción de parámetros relevantes hasta la selección de variables.

Todos estos aspectos han llevado a organizar este capítulo en 7 subapartados en los que se intenta cubrir todos los aspectos referentes al tratamiento de la señal que se han implementado en esta tesis. Así, en los dos siguientes subapartados se tratarán las técnicas lineales (PCA y PLS respectivamente). El subapartado 4 inicia el estudio de redes neuronales con la topología feedforward y entrenamiento backpropagation; los apartados 5 y 6 describen las redes Fuzzy Art y Fuzzy Artmap y, finalmente, en el apartado 7 se presentan las diferentes técnicas y estrategias de preprocesado que se han aplicado para optimizar los resultados.

4.2 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

4.2.1 Introducción

El análisis de componentes principales es una técnica originalmente propuesta en 1981 [4.11] que surgió como respuesta a la creciente cantidad de datos que podían ser obtenidos en cada medida gracias a los instrumentos de laboratorio de nueva generación [4.12]. Así, por ejemplo, un espectrómetro típico FTIR puede proporcionar datos característicos en 10000 longitudes de onda diferentes para cada medida. Esta nueva situación crea una saturación de información cuya consecuencia más probable es la incorrecta extracción de la información que es realmente relevante para la descripción del experimento.

En definitiva, la necesidad que surge con el nuevo instrumental de laboratorio (incluyendo la nariz electrónica) es doble: es necesario comprimir y es necesario extraer toda información relevante del voluminoso conjunto de datos obtenido ya que en muchas ocasiones la información esencial no depende de variables aisladas sino de la interrelación entre las mismas. El algoritmo PCA aborda estos problemas y por ese motivo es una de las técnicas más utilizadas por los químicos analíticos y, por extensión, por todos aquellos investigadores que trabajan con narices electrónicas ya

que permite reducir, representar y extraer información relevante al mismo tiempo [4.13].

PCA es un método que asume colinearidad entre las variables que intervienen. En otras palabras, se trata de un algoritmo lineal que puede funcionar incorrectamente en procesos altamente no lineales como pueden ser las interacciones químicas entre sensores y compuestos volátiles. De todas formas, funciona sorprendentemente bien en muchas aplicaciones con narices electrónicas, en las que las concentraciones de volátiles no son muy elevadas y el comportamiento de los sensores no es excesivamente alineal.

La compresión de datos y extracción de información relevante se hace más necesaria en aquellas situaciones en las que existe una falta de selectividad en cada uno de los sensores que componen la matriz. Por ese motivo el análisis de componentes principales es un método idóneo para explotar el concepto de sensibilidades solapadas que se aplica en la mayoría de las narices electrónicas.

El algoritmo PCA suele ser clasificado como un método no supervisado de reconocimiento de patrones, ya que su uso más extendido con las narices electrónicas se limita a representar bidimensionalmente un conjunto de medidas para ver si se pueden determinar agrupaciones (“clusters”) espontáneas entre las diferentes medidas realizadas previamente, aunque, como veremos posteriormente, también existen modificaciones que permiten aplicar el algoritmo en modo supervisado.

4.2.2 Algoritmo

Para aplicar el análisis de componentes principales se parte de un conjunto de datos dispuestos en forma matricial con una estructura de m filas por n columnas, suponiendo que se han realizado m experiencias y cada una se ha descrito con n variables. Así, cada fila corresponde a una descripción completa de un experimento y cada columna a una variable concreta que se utiliza como indicador en n experimentos. Para poder aplicar el algoritmo PCA, las columnas de esta matriz (las variables) deben estar centradas (“mean centered”).

Es lícito considerar, por lo tanto, que cada experimento o fila es un vector que pertenece a un espacio vectorial de dimensión n expresado en base canónica (en la que cada una de las coordenadas está asociada directamente a una de las variables descriptivas que proporciona el instrumento de medida). El análisis de componentes principales es un endomorfismo o cambio de base, pasando de la base canónica a una nueva base formada por los autovectores de la matriz de covarianza de los datos. Matemáticamente:

$$E_n(X) \xrightarrow{f} E_n(Y)$$
$$X = t_1 p_1^T + t_2 p_2^T + t_3 p_3^T + t_4 p_4^T \dots + t_k p_k^T \quad (4.1)$$
$$Y = [t_1 \ t_2 \ t_3 \ t_4 \ \dots \ t_k]$$

donde:

- X es la matriz de datos original (dimensión $m \times n$)
- t_k es el vector de coordenadas de las medidas sobre la componente principal k -ésima (denominado “score”)
- p_k es el vector de la componente principal k (denominado “load”), que es autovector de la matriz de covarianzas de X y ortogonal al resto de autovectores
- Y es la matriz de coordenadas de las n medidas referida a la nueva base de componentes principales (dimensión $m \times n$)

En definitiva, los “scores” son las coordenadas de cada una de las medidas respecto a la nueva base, por lo que de alguna manera especifican las relaciones existentes entre experimentos. Por el contrario, los “loads” definen la contribución de cada variable original sobre las componentes principales, aportando información sobre la interrelación entre las variables originales.

Lo que realmente interesa de esta transformación es que los autovectores de la matriz de covarianzas de X marcan los ejes de máxima varianza de las medidas. Así, el primer vector, llamado componente principal 1, contiene la mayor proporción posible de la varianza inicial en las medidas, la segunda componente principal, la mayor varianza residual (con la restricción de ser ortogonal a la componente principal 1), etc.

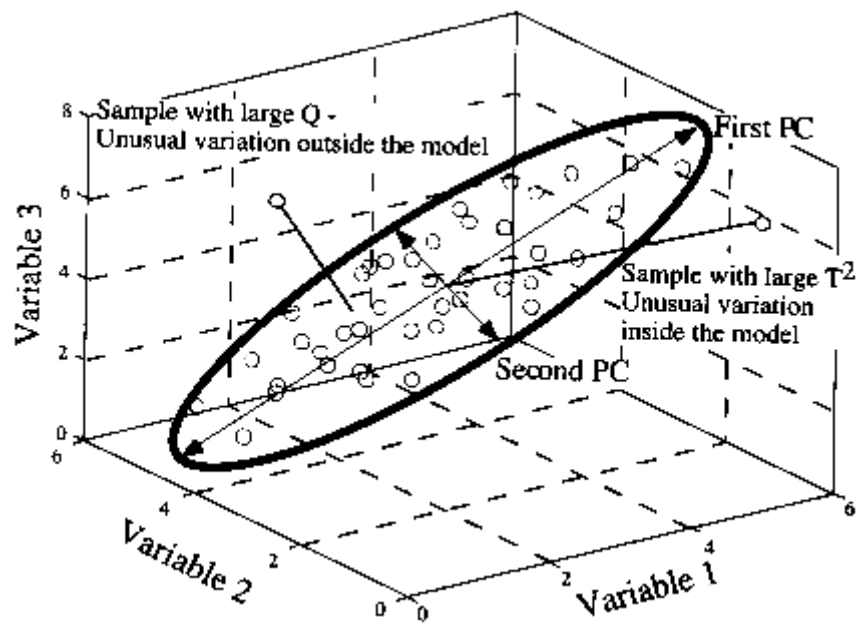


Figura 4.1: Ejemplo gráfico del algoritmo PCA

En principio, la dimensión de la nueva matriz es idéntica a la de la original, pero podemos reducir la dimensionalidad conservando únicamente las coordenadas (“scores”) respecto a las primeras componentes principales, que son las que marcan las direcciones de máxima varianza (y, por tanto, de máxima información). Eso hace posible la representación gráfica en dos dimensiones de datos multidimensionales asegurándonos de que esa representación capturará la máxima varianza posible del conjunto de datos en dos dimensiones. La figura 4.1 ilustra este concepto, en el que se puede observar que, partiendo de un conjunto de datos tridimensionales, se proyectan las medidas sobre un plano bidimensional definido por las dos primeras componentes principales. A la proyección bidimensional resultante se le denomina gráfica de “scores”, mientras que a la proyección de las variables originales sobre la nueva base se le denomina gráfica de “loads”.

En el algoritmo PCA existen una serie de parámetros y variables que permiten determinar el número óptimo de componentes principales para describir correctamente un conjunto de datos, además de algún indicador que permite identificar variaciones inusuales dentro del modelo de alguna medida, lo cual puede indicar la presencia de un “outlier” o experimento erróneo dentro de la rutina de medidas realizadas. Puesto que el

PCA es un método ampliamente conocido, no nos extenderemos mas en su descripción algorítmica.

4.2.3 Implementación

4.2.3.1 Programación

Los algoritmos PCA utilizados en los prototipos de nariz electrónica diseñados en esta tesis doctoral se han implementado en el entorno de programación que ofrece en el paquete informático MATLAB [4.14]. Las subrutinas básicas PCA se encuentran en un paquete multivariante denominado PLS-Toolbox 2.0 [4.15] que se acopla al programa principal. En concreto, existen dos subrutinas, *pca* (versión clásica) y *pcapro* (PCA en modo supervisado).

Sin embargo, estas funciones son excesivamente generales para los objetivos de nuestro trabajo. Ha sido necesario programar una serie de funciones auxiliares para facilitar la integración y evaluación de estos algoritmos en los objetivos planteados para cada variedad de fruta.

En concreto, ha sido necesario remodelar la representación gráfica de resultados para poder asignar códigos numéricos y de color para distinguir correctamente entre multitud de medidas diferentes en cada aplicación. Para ello se han programado algunas funciones gráficas adicionales para representar los datos con mayor claridad, cuyo formato se irá observando en las figuras de resultados en los capítulos 5,6 y 7.

Aunque originalmente catalogado como método no supervisado, el algoritmo PCA puede ser también utilizado como un método supervisado. En todo algoritmo supervisado puede distinguirse una etapa de aprendizaje (entrenamiento o calibración) y una etapa de evaluación o test. En este modo de trabajo, el PCA determina, en la fase de calibración, las componentes principales con un conjunto de medidas de entrenamiento. Las medidas de aprendizaje son proyectadas y se identifican las agrupaciones que corresponden a cada una de las sustancias que se han utilizado para calibrar al equipo. Posteriormente, en la fase de evaluación, las nuevas medidas se proyectan sobre las

componentes principales del modelo PCA generado anteriormente. Dependiendo de la posición de su proyección sobre el diagrama PCA se relacionará la nueva medida con alguna de las sustancias utilizadas en la fase de calibración (previsiblemente sobre la agrupación más cercana a su proyección). Este modo de utilización se ha programado integrando la función *pcapro* a un conjunto de instrucciones más amplio para acoplar su uso a diferentes objetivos del estudio.

4.2.3.2 Mejoras del algoritmo

Uno de los problemas que generalmente se achaca al análisis de componentes principales es la falta de criterio al escoger las direcciones de máxima varianza. Esto es debido a que el método no distingue entre varianza debida a la falta de repetitividad de los sensores (incluyendo posibles derivas) y la debida a información relevante. En este proyecto se ha trabajado sobre una modificación del análisis con el objetivo de reducir esta deficiencia.

La novedad radica en el cálculo de las direcciones de máxima varianza. En vez de utilizarse una serie de medidas de entrenamiento se utiliza una medida representativa de cada clase que se calcula como el valor medio de todas las medidas de entrenamiento que pertenecen a dicha clase. De esta forma el método calcula las direcciones que proporcionarán la máxima separación posible entre centroides de diferentes clases. Si el número de medidas es suficientemente elevado, factores como el ruido serán minimizados y la varianza capturada será debida, fundamentalmente, a diferencias intrínsecas a cada clase de medidas.

Se realizaron diferentes pruebas y, en todas ellas, los resultados solo mejoraron ligeramente. La gran desventaja de esta aproximación radica en que es necesario asignar correctamente las medidas de entrenamiento a una determinada agrupación, algo difícil debido a la gran variabilidad entre muestras de fruta sometidas a unas mismas condiciones. Además, limita el uso del PCA a una modo de aplicación supervisado. Por este motivo, a pesar de haber hecho diversas pruebas, no se llegó a aplicar la modificación a medidas con fruta.

4.2.3.3 Evaluación de resultados

Para valorar la resolución obtenida con la nariz se ha tenido que definir un parámetro que indicase el poder de resolución en cada PCA ateniéndonos a las clases a las que cada medida pertenece realmente. Esta función basa su cálculo en la relación entre varianza intraclase y la varianza interclase. La varianza intraclase puede definirse como la distancia media, en el gráfico de la proyección, entre cada medida de una clase concreta y la posición de la centroide de dicha categoría. Por otro lado, la varianza interclase se calcula como la distancia media entre las centroides de todas las categorías consideradas en el experimento. La figura 4.2 y las ecuaciones 4.2, 4.3 y 4.4 detallan un cálculo de resolución a modo de ejemplo.

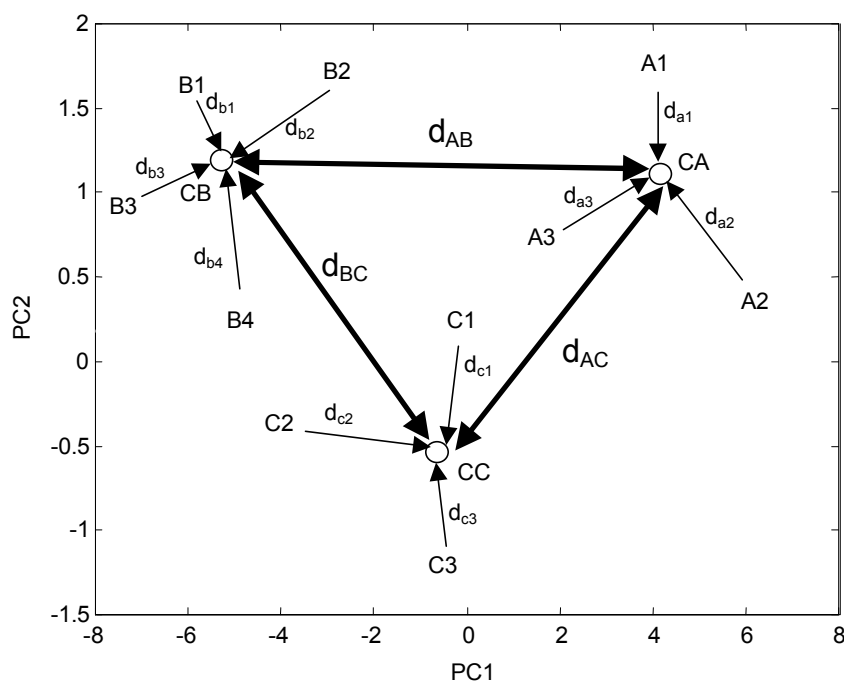


Figura 4.2: PCA con 10 medidas que pertenecen a tres clases diferentes. El siguiente recuadro ilustra el cálculo del poder de resolución

Los pasos detallados de este cálculo son los siguientes:

1. Cálculo de la centroide para cada una de las clases de medidas existentes en la proyección
2. Cálculo de la distancia media entre centroides (varianza interclase)
3. Cálculo, para cada clase, de la distancia media entre todas sus medidas y su centroide
4. Obtención del valor medio del cálculo anterior (varianza intraclase)
5. División de la varianza interclase por la varianza intraclase

Para cada clase, distancia intraclase media (varintraclase):

$$v_A = \frac{1}{3}(da1 + da2 + da3) \quad v_B = \frac{1}{4}(db1 + db2 + db3 + db4) \quad v_C = \frac{1}{3}(dc1 + dc2 + dc3) \quad (4.2)$$

$$\text{Distancia intraclase media: } vm = \frac{1}{3}(v_A + v_B + v_C) \quad (4.3)$$

$$\text{Distancia media interclases: } vmi = \frac{1}{3}(d_{AC} + d_{AB} + d_{BC}) \quad (4.4)$$

$$\text{Poder de resolución: } res = \frac{vmi}{vm} \quad (4.5)$$

4.2.4 Otras aplicaciones

Una de los problemas más acuciantes en las narices electrónicas actuales es la presencia de derivas en los sensores químicos del equipo. Aunque inicialmente pocos investigadores contemplaban su monitorización y control, lo cierto es que es de vital importancia tener en cuenta este efecto en toda aplicación en la que las medidas estén relacionadas con el paso del tiempo. El caso de la fruta es un ejemplo claro de esta situación. Durante su proceso de maduración, a medida que avanzan los días, la producción aromática crece monótonamente. Sin embargo, si los sensores aumentan su conductancia base debido a derivas temporales (envejecimiento, contaminación, etc), es

posible que su conductancia final se modifique de la misma forma, efecto que se puede atribuir erróneamente al proceso de maduración de la fruta.

Aunque existen algunos estudios en los que se intenta compensar este efecto [4.16,17] (alguno de los cuales esta relacionado con los diagramas PCA), lo cierto es que no se han utilizado de forma generalizada. En este trabajo se ha optado por monitorizar el proceso de medida y ver si existen derivas que puedan tener influencia en los resultados. Para ello se han utilizado los diagramas PCA y medidas de calibración a lo largo de las diferentes experiencias realizadas con fruta.

El método consiste en calcular el diagrama PCA con todas las medidas de fruta de una determinada experiencia, identificando las medidas con una etiqueta con el día en que se realizaron. A continuación, utilizando el método PCA en modo supervisado, se proyectan las medidas de calibración sobre las de la fruta, utilizando de esta manera las componentes principales obtenidas anteriormente con las medidas de fruta.

Idealmente, si no hubiesen derivas, las medidas de calibración deberían caer todas en un mismo punto del gráfico, ya que todas ellas se realizan en idénticas condiciones. Sin embargo, la falta de repetitividad de los sensores y las posibles derivas pueden repartir las proyecciones por diferentes zonas del gráfico.

Estudiando la evolución de las proyecciones de las medidas de calibración sobre las medidas de fruta es fácil determinar si las agrupaciones son debidas al proceso de maduración o a la deriva de los sensores de la nariz electrónica. Se remite al lector al capítulo 5 donde el apartado 5.4.3 y las figuras 5.6 y 5.7 muestran gráficamente un claro ejemplo de aplicación de cómo monitorizar derivas con este método. En ese apartado se realiza una discusión más profunda sobre las diferentes posibilidades del mismo.

4.3 PARTIAL LEAST SQUARES

4.3.1 Introducción

En esta tesis se han considerado diferentes algoritmos para correlacionar los datos de los sensores con los indicadores de calidad para la fruta. De entre los lineales se pueden destacar los algoritmos PCR (Principal Component Regression), PLS (Partial Least Squares) y MLS (Minimum Least Squares). PLS, el método finalmente escogido, puede considerarse un término medio entre PCR y MLS.

En el algoritmo PCR se buscan, al igual que en un PCA, las componentes principales que mejor describen la varianza de las variables predictoras. Con estas componentes principales es con las que se monta un modelo de correlación entre las variables predictoras y la predicha (en PCR solo se puede predecir una variable por modelo). En el espectro opuesto, MLR busca un factor simple que maximice la correlación con la variable a predecir.

El algoritmo PLS [4.18-21], tal y como ya se ha comentado, toma una aproximación intermedia. En general se puede afirmar que PLS intenta maximizar la covarianza. Para ello se vale de una serie de variables latentes (el equivalente a componentes principales) que capturan la máxima varianza posible y escoge, para montar el modelo de correlación, aquellas cuya covarianza con las variables a predecir sea máxima. Tomando este criterio PLS suele funcionar mejor que PCR o MLS en la mayoría de situaciones, motivo por el que se decidió utilizar este algoritmo para correlacionar la señal de los sensores con los indicadores de calidad de la fruta.

Como ya se ha comentado, PLS es un método lineal. Sin embargo, aplicando algunas modificaciones puede ser convertido en un algoritmo no lineal. En el PolyPLS [4.22,23], la modificación consiste en incrementar la matriz de datos con sucesivas potencias de las variables predictoras, de manera que lo que en un principio estaba restringido a relaciones lineales queda ahora ampliado a relaciones polinómicas del

orden que se decida. El único problema de este método es que aumenta el número de parámetros a ajustar y en aplicaciones con pocas medidas esto produce un sobreentrenamiento con escasa capacidad de generalización. Tras varias pruebas realizadas con medidas de fruta se comprobó que sistemáticamente se obtenían mejores resultados con el algoritmo PLS sin modificar, a pesar de ser exclusivamente lineal.

En resumen, PLS es un método de correlación multivariante entre una serie de variables de entrada (en nuestro caso las señales que provienen de los sensores de la nariz electrónica) y una serie de variables de salida (en nuestro caso indicadores de calidad de fruta). A diferencia de PCR, puede predecir varios parámetros de salida con un único modelo. Su funcionamiento es supervisado, de manera que en la fase de entrenamiento se crea un modelo que relaciona variables predictoras (variables de entrada) con variables a predecir (variables de salida) maximizando la covarianza entre ambos conjuntos. A partir de ese momento, en la fase de evaluación, PLS se convierte en una herramienta de predicción en la que se utiliza el modelo construido en la fase de entrenamiento junto con nuevas señales de entrada para predecir alguna de las variables de salida para la que ha sido entrenado.

4.3.2 Algoritmo

La manera de implementar el algoritmo de PLS ha sido objeto de numerosos estudios, por lo que no se considera apropiado describirlo con detalle en esta memoria. De todas maneras, a modo de reseña se puede comentar que en su implementación destacan dos algoritmos diferentes, el NIPALS (non-iterative Partial Least Squares), más intuitivo, y SIMPLS, algo más efectivo y fácil de calcular.

Por ese motivo, en este trabajo se ha optado por utilizar SIMPLS como algoritmo de cálculo de PLS. SIMPLS se basa en la ortogonalización de una secuencia de Krylov para calcular los pesos del modelo, estrategia poco intuitiva pero eficaz. Los resultados de este método son idénticos a los del algoritmo NIPALS en predicciones de una única variable y ligeramente mejores en predicción multivariante ya que explota mejor los criterios de covarianza.

4.3.3 Implementación

4.3.3.1 Programación

Al igual que con el algoritmo PCA, los algoritmos PLS utilizados corren bajo el entorno Matlab y son funciones ya programadas incluidas en el paquete de funciones opcional PLS Toolbox 2.0. En concreto, la función que se utiliza se denomina *simpls*. Esta función se utiliza para crear el modelo PLS a través del método SIMPLS. Requiere como entrada la matriz de medidas con las variables de entrada, la matriz con los valores correctos de las variables a predecir y el número máximo de variables latentes a considerar. La salida devuelve, básicamente, una matriz con los coeficientes de predicción del modelo y una serie de valores que informan sobre lo ajustado que ha quedado el entrenamiento.

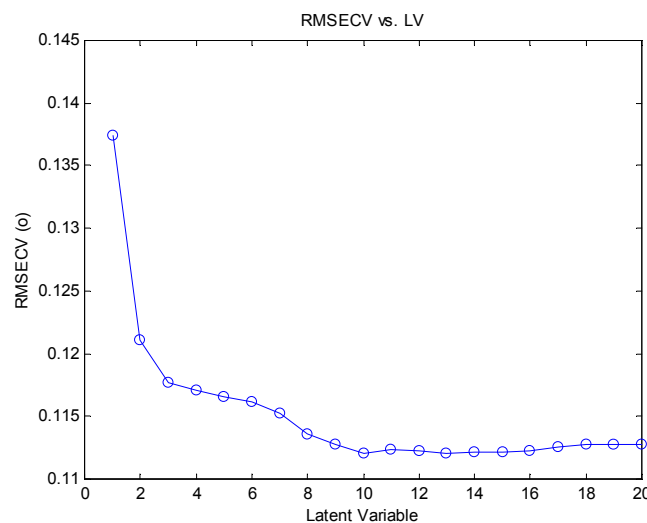


Figura 4.3: Estudio del número óptimo de variables latentes por la función *simpls*

El algoritmo determina el número óptimo de variables latentes realizando una validación cruzada en la que calcula el error cuadrático medio (RMSECV) con un número creciente de variables latentes. La figura 4.3 muestra un gráfico típico de resultados en el que se observa que, para el ejemplo en cuestión, el modelo óptimo queda creado con 10 variables latentes.

La aplicación del modelo a predicciones posteriores, en las que las medidas son diferentes a las utilizadas en entrenamiento y en las que se desconoce el valor de las variables a predecir, es realmente simple. Suponiendo que la matriz de coeficientes del modelo resultante se denomine B , la matriz con nuevos datos de medidas X_N , y la matriz de salida con las correspondientes predicciones Y_N , la forma de obtener las nuevas predicciones es tan simple como indica la ecuación 4.6:

$$Y_N = X_N * B \quad (4.6)$$

4.3.3.2 Aplicación del método a las medidas de fruta

Para determinar la validez de los modelos generados mediante PLS es necesario realizar predicciones de medidas de las que se sabe el valor correcto de salida para poder comparar la estimación PLS con el valor real.

De todas maneras, para que un modelo PLS de buenos resultados es necesario que en la fase de entrenamiento se utilicen suficientes medidas como para tener una buena representación estadística del conjunto de entrada. Al mismo tiempo, resulta interesante poder evaluar un número de medidas elevado para asegurarnos de que el sistema realiza predicciones exactas. Teniendo en cuenta que la obtención de medidas de entrenamiento suele ser una tarea costosa, resulta sumamente interesante optimizar su uso en el proceso de evaluación. Por ese motivo, en todos los estudios de correlación en esta tesis se utiliza una validación cruzada (“crossvalidation”) para determinar la capacidad de predicción de la nariz electrónica a partir de las señales de los sensores.

Suponiendo que disponemos de m medidas, en una validación cruzada de orden n se extraen n medidas del conjunto y se entrena con las $m-n$ medidas restantes. Una vez obtenido el modelo se evalúan las n medidas que se han apartado previamente del conjunto de entrenamiento. Este proceso se repite m/n veces (suponiendo m múltiplo de n), de manera que cada medida se ha utilizado para entrenamiento y para evaluación.

En este trabajo, debido a la escasez de medidas en algunos de los experimentos, se han utilizado, en la mayoría de los casos, validaciones cruzadas de orden 1 (también denominadas “leave-one-out”), lo que supone haber generado tantos modelos PLS como medidas se disponen. Esta filosofía no está pensada para proponer un modelo PLS concreto para la predicción de cada parámetro de calidad, sino para probar la validez de la aproximación.

4.3.3.3 Evaluación de resultados

La evaluación de los resultados requiere de parámetros y algunas herramientas gráficas que permitan esclarecer y comparar con qué variables y métodos se obtienen las mejores predicciones.

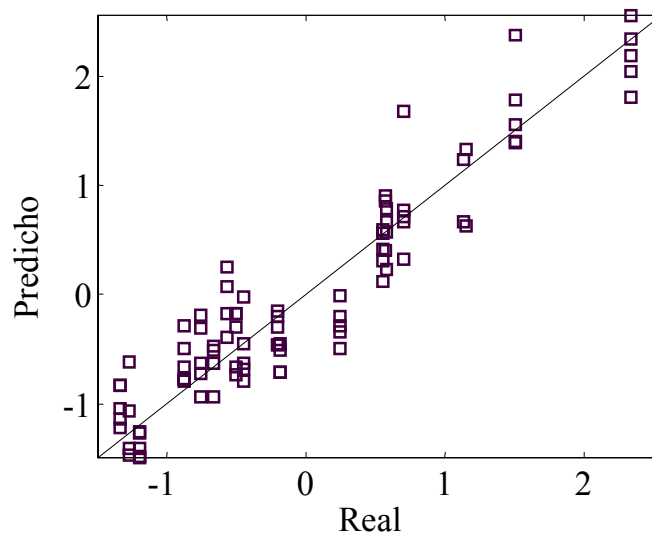


Figura 4.4: Resultado gráfico de predicción con modelos PLS

La figura 4.4 muestra una gráfica típica de resultados, en la que las medidas están representadas por recuadros. El eje horizontal muestra el valor real de la medida y el vertical el valor predicho. Una predicción perfecta pondría todos los puntos sobre la diagonal del gráfico, donde coincide valor real con valor predicho.

Mediante los puntos de la gráfica se puede hacer una regresión lineal, de la cual se puede obtener el punto de corte (parámetro “b”) y su pendiente (simbolizada por una

“m”). El coeficiente de correlación de dicha regresión nos informa de hasta que punto los resultados están alineados formando una recta. Por último, otro parámetro que nos puede dar una buena idea de la calidad de la predicción es el error cuadrático medio (“ssq”) que suma la diferencia entre el valor real y predicho de cada medida elevado al cuadrado. En una predicción perfecta, la pendiente de la recta m valdría 1, el punto de corte valdría 0, el coeficiente de correlación 1 y el error cuadrático medio 0.

4.4 REDES NEURONALES FEEDFORWARD

4.4.1 Introducción

Dentro del campo de la inteligencia artificial hace tiempo que las redes neuronales ocupan un lugar de privilegio [4.24]. Este tipo de algoritmos, cuya topología está inspirada en las redes neuronales biológicas [4.25], ajustan sus modelos de forma experimental, sin necesitar un conocimiento previo de las características de los datos que debe procesar ni de los mecanismos que los producen. Además, su forma matemática de operar no está restringida a relaciones y procesos lineales.

Dentro de los múltiples algoritmos de redes neuronales que se han propuesto hasta la fecha destaca la topología “feed-forward”, ya que fue la estructura que catapultó a las redes neuronales a su situación de privilegio actual. Además, responde perfectamente al paradigma de lo que se entiende hoy en día por una red neuronal artificial. De hecho, está demostrado que las redes neuronales feed-forward con entrenamiento backpropagation y una capa oculta son aproximadores universales [4.26]. Eso quiere decir que, con un entrenamiento adecuado, pueden interpolar cualquier relación entre el vector de entrada y el de salida, lo cual es ideal en situaciones en las que pueden haber procesos no lineales o en los que es imposible inferir una relación analítica.

Estas características hacen que este tipo de redes sean muy utilizadas como algoritmos de procesamiento de datos en las narices electrónicas [4.27]. Hay que tener en cuenta que las interacciones entre mezclas gaseosas complejas (los aromas pueden contener centenares

de componentes) y los sensores químicos son procesos extremadamente difíciles de comprender y modelar. Para mayor dificultad, la mayoría de ellos pueden presentar relaciones fuertemente no lineales, algo que no pueden digerir bien la mayoría de algoritmos clásicos de reconocimiento de patrones.

Sin embargo, para conseguir esa interpolación, el conjunto de medidas de entreno tiene que ser suficientemente representativo del conjunto de posibles vectores de entrada, lo que implica que sin un número considerable de medidas de entrenamiento no se puede realizar una interpolación con los niveles de exactitud deseables. Es este aspecto el que limita el uso de las redes neuronales cuando el número de medidas realizadas no es muy elevado. Por otro lado, el entrenamiento de las redes con el algoritmo backpropagation es computacionalmente muy costoso, por lo que los métodos de cross-validation que se pueden utilizar suelen ser de órdenes elevados, de manera que queden disminuidas el número de iteraciones. En definitiva, las dos grandes limitaciones de este tipo de redes son la necesidad de un amplio conjunto de medidas de entreno y la potencia de cálculo necesaria para ejecutar la fase de entrenamiento.

4.4.2 Algoritmo

4.4.2.1 Estructura

La figura 4.5 muestra el esquema básico de esta estructura. Toda red neuronal puede considerarse como un sistema de computación compuesto por una serie de elementos más simples (neuronas) que se interconectan entre sí. En las redes “feed-forward” las conexiones son de izquierda a derecha (“hacia adelante”). La señal de salida de una neurona sirve de señal de entrada a todas las neuronas de la columna inmediatamente a su derecha, existiendo unos pesos que ponderan esa señal.

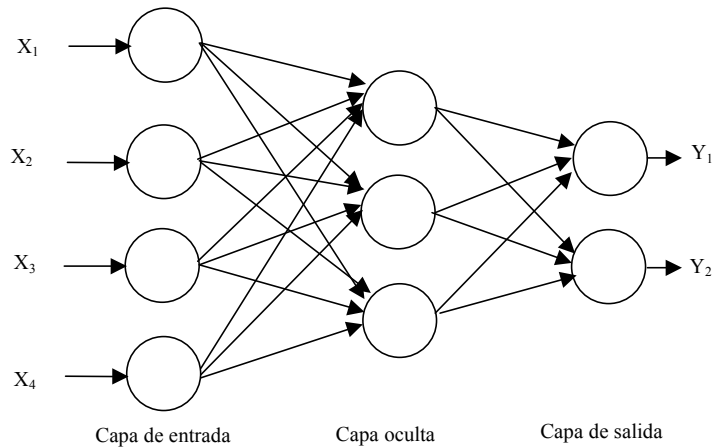


Figura 4.5: Estructura básica de una red feed-forward

Las redes que se han utilizado en este proyecto son de tres capas, considerando las señales de entrada como una capa. Las otras dos son una capa oculta y la capa de salida. La figura 4.6 muestra como trabaja una neurona. La señal total de entrada es una combinación lineal de todas las salidas de las neuronas de la capa anterior más una componente continua o “bías”, estando multiplicada cada una de las salidas por un coeficiente o peso. Una vez se ha calculado la señal de entrada total, se aplica una función de activación, generalmente no lineal, que da como resultado la salida de dicha neurona. Existen múltiples funciones de activación (la única restricción es que deben ser funciones continuamente derivables) y en este proyecto se ha escogido la función sigmoide que se detalla en la figura 4.7.

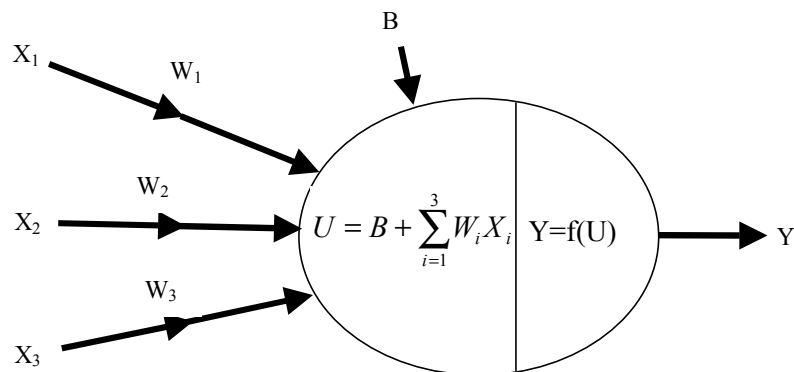


Figura 4.6: Esquema del funcionamiento de una neurona

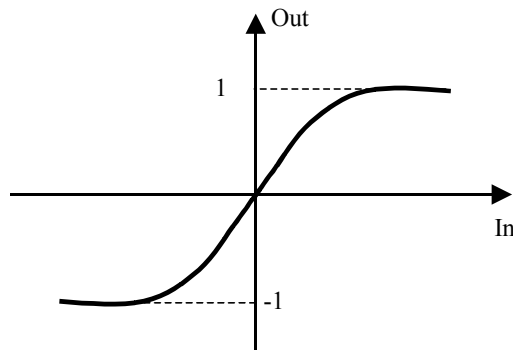


Figura 4.7: Función sigmoide de salida

4.4.2.2 Aprendizaje

El modo de funcionamiento de esta red es supervisado. Es decir, requiere de una fase de entrenamiento para poder ser utilizada posteriormente. En esta fase se le deben presentar parejas de vectores, uno de entrada y otro con la salida correcta para que la red ajuste sus pesos, aprendiendo de esta manera a mapear la relación entre el espacio de entrada y el de salida. El modo de actualizar los pesos se basa en el algoritmo backpropagation (retropropagación del error) [4.28]. Las ecuaciones 4.7, 4.8 y 4.9 muestran las fórmulas de actualización de pesos y propagación de errores:

$$\begin{aligned}\Delta w(i, j) &= lr \cdot \delta(i) \cdot P(j) \\ \Delta B(i) &= lr \cdot \delta(i)\end{aligned}\tag{4.7}$$

Donde

- lr : Ritmo de aprendizaje
- $B(i)$: Bias
- $w(i, j)$: peso de la conexión entre la neurona i y la j
- $\delta(i)$: vector de error
- $P(j)$: entradas de las neuronas

- Error en las neuronas de salida:

$$\delta_k = (\delta_k - y_k) y_k\tag{4.8}$$

- Propagación del error a las neuronas de la capa oculta:

$$\delta_i = \sum_{k=1}^m \delta_k w_{ik} f_i(Net_i) \quad (4.9)$$

A medida que el ritmo de aprendizaje crece (a través del valor lr), la modificación de los pesos es mayor y el entrenamiento más rápido, aunque por el contrario puede dar lugar a grandes oscilaciones provocando la imposibilidad de llegar al objetivo de error establecido.

Existen diferentes técnicas para mejorar el proceso del entrenamiento. Una de ellas es la del término de momento, pensada específicamente para evitar que el proceso de aprendizaje se detenga en un mínimo local del error. La ecuación (4.10) muestra este término adicional. La técnica consiste en añadir a la expresión $\Delta w(i,j)$ un nuevo término β denominado término de momento. Si la constante β vale 0, el incremento en el peso es el de siempre, en cambio, cuando vale 1, el incremento del peso es idéntico al sufrido en la iteración anterior. Típicamente el valor de β suele ser de 0.95.

$$\Delta w(i, j) = \beta \cdot \Delta w(i, j) + (1 - \beta) \cdot lr \cdot D(i) \cdot P(j) \quad (4.10)$$

Otra técnica que se utiliza es la de utilizar un ritmo de aprendizaje (lr) adaptativo que permitirá converger a la solución en el mínimo tiempo posible. Con esta técnica, si el nuevo error calculado tras la modificación de los pesos supera a la antigua en una cantidad superior definida a través de su cociente (un valor típico es 1.04), los nuevos pesos son desestimados, decrementándose el ritmo de aprendizaje por una constante (0.7 normalmente). En caso de que el error sea menor, el lr se incrementará multiplicándolo por una constante con un valor típico de 1.05.

4.4.3 Implementación

4.4.3.1 Programación

La red se ha implementado, una vez más, bajo el entorno Matlab. El conjunto de funciones utilizadas se incluyen en el Toolbox de redes neuronales “Neuralex” [4.29].

En concreto, en el entrenamiento se han utilizado las funciones *trainbpx* y *logsig* para implementar redes con una capa oculta, funciones de activación sigmoideas y aprendizaje con término de momento y ritmo de aprendizaje adaptativo.

La evaluación es más simple y para ello solo se requieren multiplicar las matrices de pesos de las sucesivas capas por las señales de entrada, activándolas a través de funciones sigmoideas.

4.4.3.2 Aplicación a las medidas con fruta

En el procesado de las medidas realizadas con fruta la red se ha utilizado tanto para clasificación como para evaluación. En el caso de la clasificación se ha utilizado una codificación *n-a-1*, de manera que en la salida se han habilitado tantas neuronas como clases existentes entre las medidas. La identificación ideal de clase se realiza con un vector en el que todos sus valores son cero excepto el de la neurona de la clase correcta, cuyo valor es la unidad.

En el caso de las medidas cuantitativas, se han habilitado tantas neuronas de salida como parámetros a cuantificar simultáneamente. Hay que destacar que en este último caso se han tenido que escalar los datos para que estuviesen comprendidos entre -1 y 1 , tal y como se detallará con mayor profusión en el apartado del preprocesado de los datos.

El proceso de entreno y evaluación se ha realizado, como en el caso del PLS, aplicando técnicas de validación cruzada. A diferencia del PLS, en el caso de las redes neuronales la validación cruzada no siempre ha sido de orden 1 debido a la carga computacional que exige su entrenamiento.

4.4.3.3 Evaluación de resultados

La evaluación de los resultados obtenidos con las redes neuronales se miden en % de aciertos en el caso de que la red se haya utilizado para clasificar. En el caso de la

cuantificación, es decir, la predicción del valor de ciertos parámetros a partir de las señales de los sensores, se han seguido los mismos métodos de evaluación que en el caso del algoritmo PLS.

4.5 REDES FUZZY ART

4.5.1 Introducción

La teoría de la resonancia adaptativa (ART) fue introducida como una teoría que intentaba emular la manera en cómo el cerebro humano procesa la información [4.30,31] . Desde entonces, esta teoría ha evolucionado hacia una serie de algoritmos neuronales para el aprendizaje no supervisado. Estos algoritmos son capaces de crear clases estables ante la presentación de secuencias de entrada arbitrarias con un ritmo de aprendizaje rápido o lento. Dentro de estos algoritmos se pueden destacar el ART1 [4.32], ART2 [4.33] y ART3 [4.34].

Fuzzy Art [4.35] es una evolución del algoritmo ART1. Éste último es capaz de categorizar de forma estable entradas arbitrarias binarias. Fuzzy Art, siguiendo el mismo esquema, generaliza esta función a vectores de entrada analógicos con coordenadas comprendidas entre 0 y 1. Para ello substituye los operadores intersección (\cap) y unión (\cup) de ART1 por los operadores MIN (\wedge) y MAX (\vee), respectivamente, de la teoría de lógica difusa [4.36]. Este cambio, con la ayuda de la codificación complementaria (“complement coding”), que preserva la información de amplitud a la vez que normaliza los vectores de entrada, permite implementar un algoritmo de clasificación no supervisada de gran rapidez de aprendizaje. En la figura 4.8 se muestra un esquema del algoritmo.

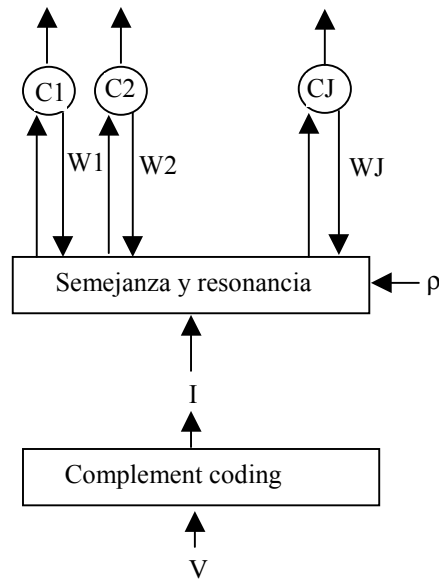


Figura 4.8: Esquema de la red Fuzzy Art

Su modo de funcionamiento es simple: cada vez que la red recibe un nuevo vector de entrada V reacciona activando uno y solo uno de los nodos de salida. Cada uno de estos nodos representa una de las diferentes clases que se han creado con las entradas anteriores. En caso de que la medida no se parezca lo suficiente a ninguno de los nodos ya asignados se crea uno nuevo que representará una nueva clase cuyo primer miembro será este vector.

Desde el punto de vista operativo, este algoritmo cuenta con dos parámetros que controlan su funcionamiento. El parámetro de vigilancia (“vigilance parameter”) determina lo riguroso que debe ser el algoritmo a la hora de agrupar medidas. Un parámetro de vigilancia muy cercano a la unidad implica una clasificación muy exigente, de manera que dos medidas deben ser muy parecidas para ser agrupadas en una misma clase. Por el contrario, un parámetro cercano a cero permite la agrupación de medidas poco parecidas, lo que, como resultado, genera una red con pocos nodos de salida, ya que el número de clases diferentes es reducido. Por su parte, el ritmo de aprendizaje queda controlado por el parámetro β , siendo su valor igual a la unidad para un aprendizaje rápido e igual a cero en caso de que no se deba aprender más.

4.5.2 Algoritmo

Incluimos, a continuación, una descripción esquemática del algoritmo:

- Vector de entrada: Cada uno de los vectores de entrada V es un vector M -dimensional donde cada una de sus componentes tiene coordenadas incluidas en el intervalo $[0,1]$.
- Codificación complementaria: A partir del vector de entrada V , se crea un nuevo vector normalizado I de dimensión $2M$ en el que la componente $I_{J+M} = 1-I_J$.
- Vector de pesos del nodo de salida j (categoría j): W_j . Inicialmente, $W_{j1}=W_{j2}=W_{j2M}=1$
- Velocidad de aprendizaje: (“learning rate”), β entre $[0 \ 1]$. Aprendizaje rápido, $\beta =1$; Aprendizaje lento, $\beta \ll 1$; Sin aprendizaje, $\beta =0$
- Parámetro de vigilancia: ρ entre $[0 \ 1]$. ρ cercano a cero implica menos categorías al agrupar con criterios de semejanza poco exigentes, ρ cercano a uno implica muchas clases, cada una con pocos miembros pero muy parecidos entre sí.
- Parámetro de selección: $\alpha >0$. Debe ser muy cercano a cero. Sirve para deshacer igualdades. Un valor típico es 0.001.
- Selección de categoría: Para cada vector de entrada V y cada categoría j se calcula la función de selección $T_j(V)$ como indica la ecuación 4.11:

$$T_j = \frac{|I \wedge W_j|}{\alpha + |W_j|} \quad (4.11)$$

Donde el operador AND (\wedge) en lógica difusa se define como:

$$A \wedge B = \min(A,B) \quad (4.12)$$

Y la norma $||$ se define como:

$$|I| = \sum_{i=1}^{2M} I_i \quad (4.13)$$

A partir de aquí inicialmente se escoge la categoría jota para la que $T_j(V)$ es máximo.

- Resonancia o reset: Se dice que aparece resonancia si se cumple la desigualdad 4.14:

$$\frac{|I\Lambda W_j|}{|I|} \geq \rho \quad (4.14)$$

En ese caso, se activa la categoría j como respuesta al vector de entrada V , lo que quiere decir que la red clasifica al vector V como de clase j . Además, se ejecuta el proceso de actualización de los pesos de dicha categoría.

En el caso de que no se cumpla la desigualdad se produce un reset: El sistema desactiva temporalmente la categoría j y vuelve a escoger una categoría siguiendo el criterio de máxima semejanza (ecuación 4.11). Si ninguna categoría “resuena”, se crea un nuevo nodo para el vector de entrada V .

- Aprendizaje: Una vez activada la categoría j debido al vector V , sus pesos son actualizados según la ecuación 4.15:

$$W_j^{NUEVO} = \beta(I\Lambda W_j^{ANTERIOR}) + (1 - \beta)W_j^{ANTERIOR} \quad (4.15)$$

Si se quiere un aprendizaje rápido, se utiliza una $\beta = 1$, un aprendizaje nulo $\beta=0$. En general, para medidas ruidosas no interesa poner $\beta=1$. Sin embargo, cuando el número de medidas es bajo y se requiere de un aprendizaje estable se puede demostrar que eso se consigue con $\beta=1$.

4.5.3 Implementación

4.5.3.1 Programación

El paquete matemático Matlab no contiene ninguna función que implemente este tipo de red, por lo que se ha tenido que programarla bajo dicho entorno. La programación se ha

hecho adaptando el uso de la red a las aplicaciones de la nariz electrónica. Desde un principio se ha implementado la codificación complementaria.

La programación se ha realizado de forma modular, con funciones que implementan cada uno de los aspectos con los que componen el algoritmo. Así, hay funciones que implementan los operadores de lógica difusa (módulo, AND, OR), una función que implementa la codificación complementaria y otra que implementa la iteración para una medida del núcleo del algoritmo. Para poder clasificar un número concreto de medidas existe una última función que invoca a todas las anteriores para formar un bucle que se repite tantas veces como medidas deban ser procesadas.

4.5.3.2 Evaluación de resultados

Para poder evaluar los resultados, se ha implementado un parámetro que sirve de figura de mérito para valorar las diferentes clasificaciones no supervisadas comparándolos con la agrupación ideal deseada. Dicho parámetro se ha denominado “separación”.

Este parámetro valora las diferencias (atendiéndonos a la categorización deseada), entre todas las medidas que activan un mismo nodo de la red. Un valor igual a cero implica que todas las medidas que activan dicho nodo pertenecen a la misma categoría real, por lo que la agrupación que realiza la red coincide con la clasificación deseada. En caso contrario, el valor del parámetro aumenta, de forma que valores de “separación” elevados indican una clasificación espontánea diferente a la deseada a priori. Hay que destacar que este parámetro penaliza la agrupación en un mismo nodo de medidas cuya categorización correcta es muy distante. Eso es debido a que cada categoría correcta es identificada por un ordinal de manera que categorías contiguas contienen medidas parecidas y categorías alejadas ordinalmente contienen medidas muy diferentes.

Inicialmente, el algoritmo calcula una “centroide” para cada una de las clases que se han formado a la salida de la red. Esta centroide se obtiene como la media aritmética de los ordinales que designan la categoría correcta de cada medida. A continuación se calcula

la distancia media absoluta entre cada una de las medidas y la centroide. Finalmente, se hace la media de estos valores y eso nos da el parámetro de separación.

Para poder realizar comparaciones adecuadas, se debe tener en cuenta que éstas solo se pueden hacer cuando el número de clases generadas es idéntico, ya que, por ejemplo, si se obtienen tantas clases como medidas (con un parámetro de vigilancia igual a la unidad) la separación media es cero, a pesar de que la clasificación realizada no aporta ninguna información.

4.6 RED FUZZY ARTMAP

4.6.1 Introducción

Las redes de tipo ARTMAP son una clase de redes neuronales que implementan un aprendizaje supervisado y una posterior clasificación de vectores multidimensionales de entrada en una serie de categorías de salida.

La red Fuzzy Artmap [4.37] proviene de la red Artmap [4.38] con las mismas transformaciones que permiten definir la red Fuzzy Art a partir de la red ART1. En definitiva, la red Fuzzy Artmap es una generalización a vectores analógicos (con componentes comprendidas entre cero y uno) de la red binaria Artmap.

La red Fuzzy Artmap presenta múltiples ventajas que la hacen muy interesante para las aplicaciones con narices electrónicas [4.39,40]. De entre todas ellas destacaremos las siguientes:

- Aprendizaje rápido (con muy poca carga computacional) de las medidas que se presentan en entrenamiento, lo que permite programar el algoritmo en dispositivos programables de bajo coste, aplicar validaciones cruzadas de orden 1 y probar con diferentes combinaciones de parámetros.

- Aprendizaje con un conjunto reducido de medidas de entrenamiento, algo muy interesante en cualquier aplicación experimental en la que sea costoso la obtención de conjuntos de medida extensos. La red presenta una habilidad particular para aprender rápidamente eventos singulares que aparecen muy pocas veces en el conjunto de entrenamiento. Por lo tanto, en dicho conjunto no es necesario que haya el mismo número de medidas de cada clase para que funcione correctamente.
- Aprendizaje continuo de nuevas características sin olvidar lo aprendido con anterioridad, algo muy útil para adaptarse a derivas producidas por sensores.
- En comparación con otros tipos de redes neuronales, Fuzzy Artmap determina automáticamente las neuronas de su capa oculta. Además maximiza el poder de generalización aprendiendo al 100% el conjunto de entrenamiento.
- Una vez entrenada, es posible extraer reglas de clasificación a partir de los pesos obtenidos tras el periodo de entrenamiento, lo que puede dar luz sobre los procesos internos y como influyen en la categorización de resultados.

Sin embargo, su implementación práctica presenta un problema que debe ser tratado con sumo cuidado. Tal y como se ha comentado con anterioridad, la red aprende el 100% de las medidas del conjunto de entrenamiento sacrificando lo mínimo posible la generalización. Sin embargo, la presencia de “outliers” o medidas erróneas en el conjunto de entrenamiento puede requerir un incremento del valor de vigilancia excesivo, lo cual perjudicará seriamente la capacidad de generalizar de la red. Este problema es una de las razones por las que este tipo de algoritmo se ha utilizado poco en narices electrónicas, ya que en este tipo de aplicaciones es sumamente difícil identificar medidas erróneas dada la baja repetitividad de las señales de los sensores. En este trabajo se han estudiado diferentes modificaciones para eliminar este problema.

Resumiendo, la red Fuzzy Artmap es una red de clasificación con aprendizaje supervisado. En una fase de entrenamiento la red necesita que se le suministre un conjunto de medidas. Cada medida debe contener un vector de entrada, que detalla los parámetros medidos en cada experiencia, y un vector de salida que codifica la categoría que se le debe asignar. Posteriormente, en la fase de evaluación solo se suministra el

vector de entrada y la red clasifica dicha medida siguiendo los criterios que ha aprendido en la fase de entrenamiento.

4.6.2 Algoritmo

Básicamente, una red fuzzy artmap está formada por dos redes Fuzzy Art conectadas entre sí por un vector de relaciones denominado “mapfield”. A una de las dos redes (la que denominaremos A) le llegan los vectores de entrada (V). A la red B le llegan, en la fase de entrenamiento, los vectores que codifican la categoría correcta de cada medida del conjunto de entrenamiento (C). La figura 4.9 esquematiza este concepto.

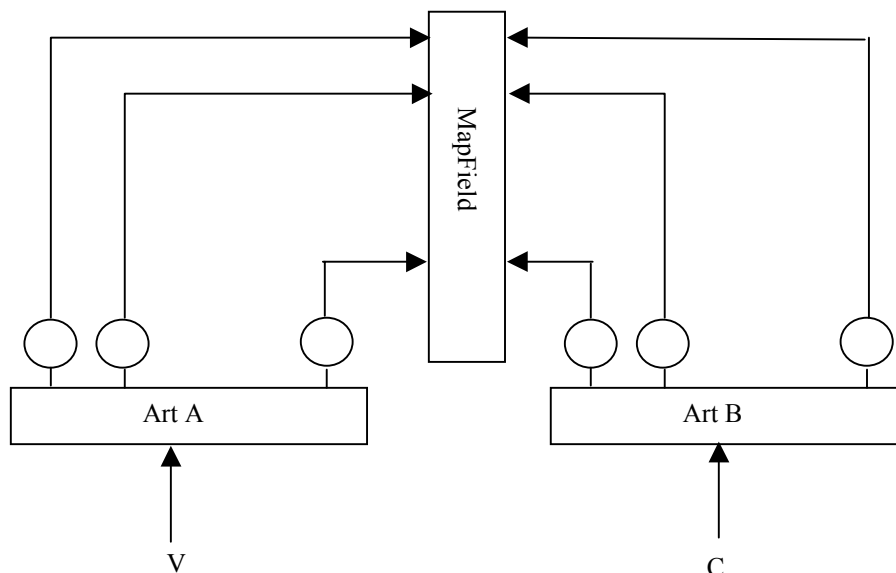


Figura 4.9: Esquema general de una red Fuzzy Artmap

Inicialmente, en la red A el vector de vigilancia es cero. En la red B se suele dar un valor igual a la unidad, ya que medidas que deban ser clasificadas conjuntamente enviarán a la red B codificaciones idénticas. Además, cualquier vector de codificación diferente, por parecido que sea al resto, debe ser detectado y debe activar una neurona de salida diferente en la red B.

Cada vez que se suministra una medida de entrenamiento, la red A activa un nodo y la red B activa otro. El mapa que las une aprende a relacionar nodos activados. De esta

forma, a cada nuevo nodo que se activa en la red A se le asocia un nodo en B. Cabe destacar que los nodos B normalmente serán imagen de varios nodos A (cada categoría contiene varias medidas), mientras que cada una de las categorías creadas en A sólo tendrá una imagen en B (cada medida solo puede pertenecer a una categoría).

Cuando una nueva medida activa un nodo en A ya existente, se comprueba si la imagen de ese nodo asignada por el mapfield coincide con el nodo que se ha activado en B paralelamente. En el caso de que no coincidan, se incrementa el valor del parámetro de vigilancia hasta que la neurona que se active en A tenga por imagen la neurona activada en B. Si no se encuentra ninguna se creará una nueva y el mapa le asignará como imagen el nodo B activado.

En definitiva, el valor de vigilancia en A solo se incrementa lo estrictamente necesario para que la red A separe en nodos diferentes las medidas que deben estar clasificadas en diferentes categorías. Suponiendo un parámetro de aprendizaje igual a la unidad para ambas redes se puede demostrar que este algoritmo aprende a clasificar correctamente el 100% de los vectores de entrenamiento. Además, ese aprendizaje es rápido y estable. A continuación se detalla el algoritmo de forma esquemática:

- ARTa: Red fuzzy art a la que llegan los vectores de entrada
- ARTb: Red fuzzy art a la que llegan los vectores de salida
- Mapfield: Módulo de mapeado que relaciona nodos de salida de ARTa con nodos de salida de ARTb
- Entradas: Supondremos que el vector de entrada de cada medida es V y el vector que codifica su clasificación correcta C.
- Normalización: Los vectores V y C se normalizan con codificación complementaria, pasando a ser los vectores I y D respectivamente.
- Pesos: El vector de pesos de la categoría k de la red ARTa se denominará W_{AK} . El vector de pesos de la categoría j de la red b se denominará W_{BJ} . El mapfield tiene un solo vector de longitud igual al número de nodos activados en A. La componente k-ésima del vector indica el nodo imagen en ARTb del nodo k de la red ARTa.

- Match tracking: Originalmente, el parámetro de vigilancia de ARTa, ρ_a , es un valor base. Si una medida a activa un nodo J en ARTa cuya imagen a través del mapfield no coincide con la activación producida en ARTb, entonces el valor de vigilancia en ARTa se incrementa según la ecuación 4.16, lo que forzará la activación de un nodo diferente en ARTa:

$$\rho_a = \frac{|I \Delta W_J^A|}{|I|} \quad (4.16)$$

- Modo de evaluación: La red ARTb se desactiva. La red ARTa recibe un vector que hace que se active su nodo k. La salida de la red es el valor de la componente k-ésima del vector de mapfield.

4.6.3 Implementación

4.6.3.1 Programación

Aunque existen varios paquetes comerciales que implementan algoritmos basados en Fuzzy Artmap no existe ninguno que implemente la red en su definición estricta. Todos estos paquetes utilizan algoritmos propietarios que no son desvelados por motivos comerciales. Los motivos de estas modificaciones ya han sido explicados con anterioridad. El problema básico reside en la posible existencia de medidas erróneas (“outliers”) en el conjunto de entrenamiento.

Por razones de flexibilidad, al igual que la red Fuzzy Art, esta red se ha programado bajo el entorno Matlab para poderla adaptar a las medidas realizadas con nariz electrónica. Al igual que al implementar Fuzzy Art, Fuzzy Artmap se ha programado a partir de funciones más simples que son anidadas en una función con un bucle principal. Como la red Fuzzy Artmap está formada por dos redes Fuzzy Art, sus funciones principales incluyen llamadas a la función Fuzzy Art implementada anteriormente.

En la programación del algoritmo Fuzzy Artmap se ha distinguido entre una función de entreno (en la que a partir de unas medidas de entrada y sus salidas correspondientes se calculan los pesos) y una de evaluación (en la que a partir de unas medidas de entrada y los pesos se devuelven las clases a las que pertenecen según la red entrenada previamente).

Para evitar los problemas anteriormente comentados y conseguir que la red Fuzzy Artmap funcione correctamente en aplicaciones con nariz electrónica, se han implementado diferentes versiones que modifican ligeramente el algoritmo original con el objetivo de inmunizar a la red ante la presencia de outliers en el conjunto de entrenamiento. Explicar las modificaciones en cada una de estas versiones alargaría innecesariamente el capítulo, por lo que se indicará la filosofía de cada una de ellas y se indicará la versión finalmente escogida para procesar las medidas realizadas con fruta.

- Una primera versión imponía un tope al parámetro de vigilancia. En el caso de que una medida exija un aumento de dicho parámetro superior al estipulado, se concede momentáneamente dicho valor para regresar al anterior en la siguiente medida. La ventaja de esta versión radica en que se puede preservar la capacidad de generalización aprendiendo medidas extrañas que pueden ser “outliers”. Su desventaja radica en saber encontrar el valor óptimo para ese tope que permita un aprendizaje adecuado sin incurrir en un sobreentrenamiento.
- En una segunda versión se asignaba un parámetro de vigilancia particular para cada nodo de la red ARTa. Esta estrategia permite tener diferentes niveles de generalización para cada categoría, algo muy útil si la varianza de medidas de diferentes categorías es muy dispar. El inconveniente más importante de esta versión radica en la falta de convergencia en la fase de entrenamiento en multitud de situaciones prácticas.
- Finalmente, la solución que mejor resultados prácticos ha propiciado modifica el algoritmo original sólo en la etapa de evaluación. En dicha fase el algoritmo modificado ignora el factor de vigilancia alcanzado en la fase de entrenamiento

y busca, de forma iterativa, el valor más elevado que permita clasificar la nueva medida en alguna de las categorías existentes.

Junto a las funciones necesarias para implementar las diferentes versiones de la red fuzzy artmap se han diseñado funciones de validación cruzada de orden 1 (“leave-one-out”) que permiten evaluar el método dando como resultado un porcentaje de aciertos en la clasificación de las medidas, tal y como se hizo con los algoritmos PLS y de redes neuronales backpropagation.

4.7 PREPROCESADO

4.7.1 Introducción

El paso previo a la aplicación de técnicas de reconocimiento de patrones es lo que se denomina preprocesado. Su importancia es capital para el éxito de cada aplicación. Es en esta fase en la que se extrae la información relevante de los sensores para procesar luego su respuesta. Dentro del preprocesado se pueden distinguir tres etapas claramente diferenciadas, cada una de las cuales repercute de una manera esencial sobre los resultados que se obtendrán posteriormente con los algoritmos de reconocimiento de patrones.

En una primera etapa se extraen los parámetros que aportan información relevante del conjunto de cambios físico-químicos que experimentan los sensores. Posteriormente, se acondicionan estos resultados en función de los algoritmos que se vayan a aplicar. Finalmente, se pueden aplicar diferentes tipos de criterios para seleccionar las variables que aportan más información y eliminar aquellas que solo aportan ruido.

4.7.2 Extracción de parámetros

Los sensores de estaño basan su funcionamiento en el cambio de conductividad que experimenta su capa activa ante la presencia de gases reductores. Este aspecto ya ha

sido tratado con mayor detenimiento en el Capítulo 3, por lo que se aconseja al lector el volver a dicho capítulo si requiere revisar su funcionamiento.

Lo que sí es relevante es el hecho de que ese cambio de conductividad conlleva unos transitorios que llevan a la capa activa del sensor desde una situación de reposo a una conductancia final que depende del tipo de volátil y de su concentración.

La información que se puede extraer del comportamiento del sensor es muy variada. Si se obtienen los valores de conductancia inicial y/o final se dice que se están utilizando parámetros estáticos. Con ellos se pueden hacer varias combinaciones que se detallan en la tabla 1:

G_i	Conductancia inicial
G_f	Conductancia final
Δg	Incremento de conductancia ($G_f - G_i$)
Δg_n	Incremento de conductancia normalizada $\Delta g / G_i$

Tabla 4.2: Parámetros estáticos extraíbles en los sensores de óxidos metálicos

Una nueva forma de obtener información adicional es modulando alguno de los factores que afectan a la respuesta del sensor. Los parámetros que se obtienen al aplicar este tipo de aproximación se clasifican como dinámicos y existen de diferentes tipos.

La modulación de concentración es una de las maneras para obtener información adicional de los sensores. En [4.41] se demuestra como la utilización de parámetros asociados a la respuesta transitoria a un aumento brusco de concentración incrementa notablemente el poder de resolución ante diferentes gases de los sensores de óxido de estaño. Esto es debido a que el tiempo de respuesta del sensor es característico del gas o mezcla gaseosa cuya concentración se modula [4.42].

De esta modulación, el parámetro que más se utiliza es el tiempo de subida, medido como el tiempo que tarda el sensor en recorrer un determinado porcentaje del

incremento de conductancia (normalmente del 10 al 90% de ΔG). De hecho, este tipo de modulación se utiliza en los dos prototipos de nariz electrónica desarrollados, ya que la llegada de volátiles de la fruta es un fenómeno rápido en comparación con la inercia de los sensores. En el primer prototipo la concentración aumenta bruscamente a través de varias inyecciones sucesivas, algo que se ha estudiado y aporta una buena repetitividad ante medidas con gases simples [4.43]. En el segundo equipo, la concentración previa de volátiles alcanza de forma rápida la cámara de sensores ante una conmutación de las electroválvulas pertinentes. La figura 9 muestra el transitorio típico de un sensor ante la rápida llegada de aromas provenientes de la fruta en el segundo prototipo. Se puede observar como el transitorio tarda unos minutos en estabilizarse y como de él se pueden extraer parámetros estáticos y parámetros dinámicos.

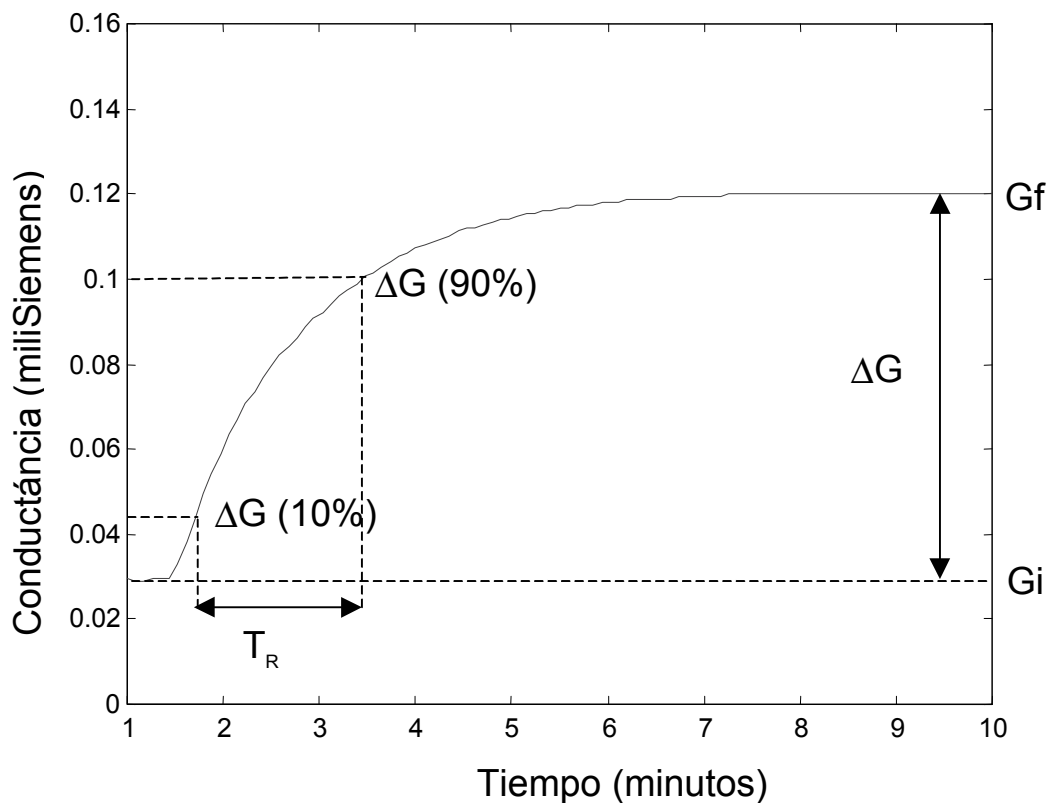


Figura 4.10: Extracción de parámetros temporales ante un típico transitorio típico de los sensores de óxido de estaño ante un brusco cambio de la concentración de volátiles emitidos por la fruta.

Otra forma de obtener parámetros dinámicos de los sensores es mediante la modulación de la temperatura de su capa activa. Aunque ya se han realizado varios trabajos al respecto, en alguno de los cuales el autor de esta tesis ha tenido una contribución

significativa [4.44,45], no se ha aplicado el concepto al prototipo de nariz diseñada para monitorizar la fruta dado que los resultados obtenidos con sensores comerciales no han sido alentadores.

4.7.3 Normalización de los datos

Una vez decididos los parámetros que deben ser extraídos de cada sensor es necesario escalarlos para que puedan ser utilizados por el sistema de reconocimiento de patrones que se considere adecuado para cada aplicación. Existen multitud de escalados diferentes: escalados lineales o logarítmicos, de variables o de medidas, etc. En este proyecto se han considerado cuatro tipos diferentes que pasaremos a detallar a continuación. Para ello cabe recordar que el conjunto de m medidas se resume en una matriz de m filas por n columnas, donde cada una de las m filas describe una medida a través de n variables (una por columna).

- Centrado de datos (“mean centring”): En esta normalización a cada una de las columnas (variables) se le resta su valor medio. Es decir, toda medida es descrita por variables de media nula. Este tipo de centrado es fundamental para los métodos lineales aplicados en este estudio (PCA, PLS), ya que sin él no funcionan adecuadamente debido a la matriz de covarianzas. De todas formas, como veremos a continuación, este tipo de normalización no suele ser suficiente si se pretenden obtener buenos resultados.
- Autoescalado (“Autoscaling”): En este método de normalización, en primer lugar se aplica un “mean centring” a las columnas, para que todas las variables tengan media cero. A continuación, se calcula la varianza de cada una de las variables (columnas) y se divide cada uno de los m valores de cada variable por dicho resultado, de forma que al finalizar la normalización cada una de las n variables que describen las m experiencias tiene media cero y varianza unidad.

Este tipo de escalado es muy útil cuando los parámetros que describen cada experiencia son de naturaleza (y por tanto de unidades) diferentes. Incluso en el

caso de que cada variable represente el mismo parámetro para un sensor hay que recordar que diferentes sensores trabajan en diferentes valores de resistencia, por lo que es conveniente que todos ellos trabajen dentro del mismo rango de valores. El éxito de este escalado radica en que, a priori, asigna la misma importancia a cada una de las variables que describen una medida, independientemente de su naturaleza o sensor del que provengan. De todas formas, hay que tener cuidado ya que este proceso puede incrementar notablemente el ruido existente en señales “pequeñas”.

- Escalado por el valor máximo: La filosofía detrás de este tipo de escalado es similar a la del autoescalado en cuanto a que su objetivo es dar la misma importancia numérica a todas las variables. En este método se divide cada columna de la matriz de experimentos por el valor máximo de cada una de las variables, de manera que todas ellas tienen un valor absoluto comprendido entre 0 y 1. A diferencia del método anterior, en este caso no se resta el valor medio de cada variable.

Este tipo de escalado es fundamental al aplicar redes neuronales de tipo backpropagation y fuzzy artmap. En el primer caso, si se utilizan neuronas con funciones de activación tipo sigmoide, su salida siempre estará en el margen comprendido entre -1 y 1 , por lo que es fundamental escalar las variables de salida entre esos dos valores. Por otro lado, la red fuzzy artmap trabaja con vectores de lógica difusa (fuzzy logic), en los que todos los valores deben estar comprendidos entre 0 y 1 para funcionar correctamente.

- Escalado de “cantidad”: Un último tipo de escalado es el que normaliza el módulo del vector que describe cada medida. Su objetivo es eliminar la información de cantidad (concentración) para facilitar la tarea de identificación y clasificación de las diferentes experiencias que se realicen con la nariz. La ecuación 4.17 muestra la fórmula que permite normalizar cada una de las filas siguiendo este criterio, suponiendo que cada medida queda descrita por N variables.

$$S'_{IM} = \frac{S_{IM}}{\sum_{I=1}^N S_{IM}} \quad (4.17)$$

donde : S'_{IM} : Componente i-ésima de la medida m normalizada

S_{IM} : Componente i-ésima de la medida m sin normalizar

La figura 4.11 intenta esquematizar, de forma gráfica, las diferencias conceptuales entre estas dos últimas normalizaciones partiendo de un conjunto de 4 medidas (dos de cada tipo de sustancias pero con concentraciones diferentes) representadas mediante gráficos polares. Como se puede observar, en la normalización de valor máximo (gráfico superior) existe alguna medida para toda variable en la que el valor máximo es la unidad. Sin embargo, la forma de las medidas de la misma sustancia es siempre la misma, aunque cambiando el tamaño.

En la normalización por medida (gráfico inferior) se puede observar como la forma de las medidas no varía respecto al gráfico original pero se pierde la información del tamaño, por lo que medidas de la misma sustancia se superponen.

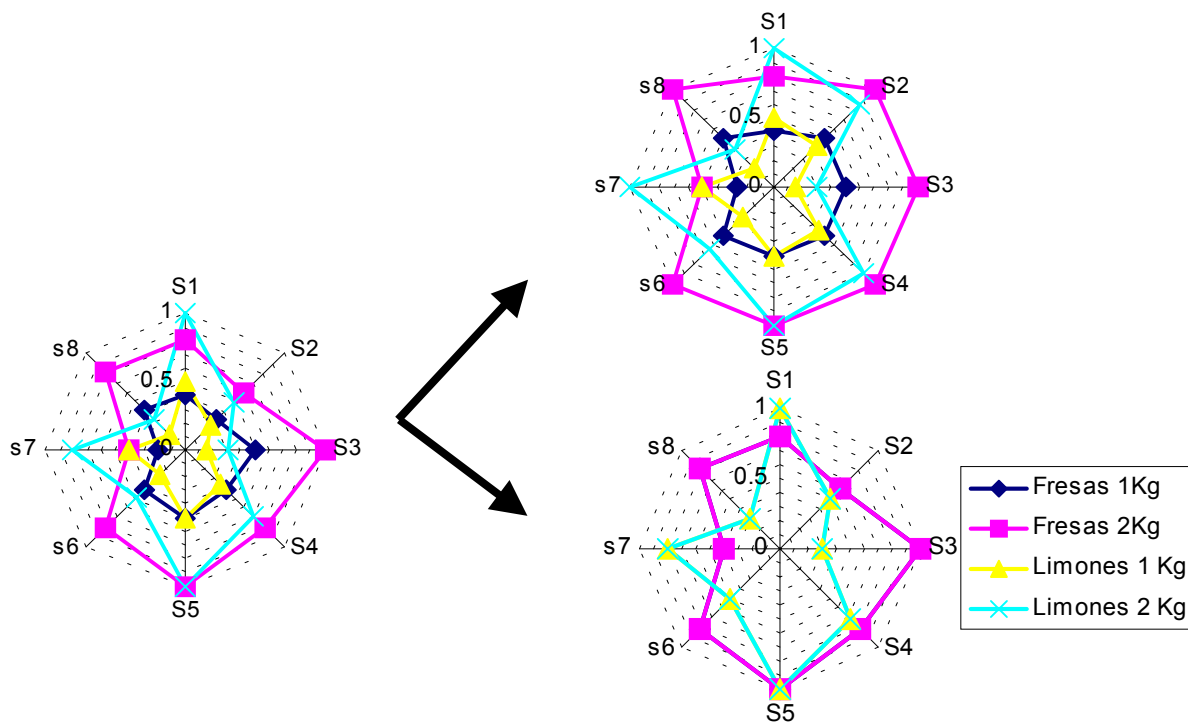


Figura 4.11: Normalizaciones por sensor (arriba) y por medida (abajo)

4.7.4 Selección de variables

En todo estudio de viabilidad en el que se pretende determinar si una nariz electrónica puede realizar satisfactoriamente una determinada función se debe realizar una cuidadosa selección de los sensores que deben formar la matriz química del instrumento. Una práctica común es empezar realizando medidas con un prototipo que contenga una matriz con muchos sensores, de manera que se obtenga la máxima información posible de cada medida. Esta aproximación suele ser muy útil ya que los recursos dedicados al realizar la medida son prácticamente los mismos con pocos o muchos sensores (solo es necesario aumentar la capacidad de adquisición de datos del ordenador). Esta práctica se ha seguido en este trabajo, ya que la realización de medidas con fruta es una práctica muy laboriosa y conviene extraer tanta información como sea posible de cada medida.

De todas formas, el aumentar la cantidad de información que se recoge no es garantía de obtener mejores resultados. Hay variables que aportan información útil y otras que solo aportan ruido. Por ese motivo es necesario escoger cuidadosamente las variables que los diferentes algoritmos de reconocimiento de patrones van a utilizar.

Si tenemos en cuenta que, además, de cada sensor se pueden extraer varios parámetros (estáticos y dinámicos) podemos encontrarnos en una situación en la que incluso haya más variables que medidas. Por ese motivo resulta interesante el utilizar algún criterio de selección de variables que permita reducir la dimensionalidad de los datos sin eliminar información útil y minimizando, a la vez, las interferencias que aporten las variables con ruido.

Dentro de los diferentes métodos existentes podemos destacar tres filosofías diferentes que se han considerado en este proyecto: Criterios de sensibilidad, de correlación entre variables y de optimización mediante la utilización de un algoritmos genéticos.

Cabe aclarar que estas estrategias no siempre han sido aplicadas, ya que en algunos casos se ha optado por incluir familias enteras de parámetros dada la dificultad para obtener criterios y resultados fiables sobre que variables escoger para obtener resultados óptimos. Por lo tanto, en algunos casos se han programado funciones iterativas de prueba y error en las que las combinaciones testadas se han escogido por intuición, sentido común o análisis gráfico de resultados.

4.7.4.1 Criterio de sensibilidad

Este método se ha aplicado, fundamentalmente, para seleccionar sensores más que para escoger variables. Su objetivo es seleccionar los sensores que presenten una mayor sensibilidad media a las sustancias que deben ser analizadas. El parámetro que mejor describe este concepto es el incremento de conductancia normalizada (Δgn), ya que nos da información sobre el cambio de conductancia relativo al valor de reposo. Otros parámetros, como la conductancia inicial o el incremento de conductancia, pueden ser engañosos ya que sensores menos sensibles pueden tener valores más elevados si su conductancia base es superior. La figura 4.12 muestra esta idea con una respuesta real al aroma de la fruta producida por dos sensores diferentes. En ella se puede observar como el sensor 1 tiene valores de ΔG y G_f mayores que el sensor 2, y, sin embargo, este último tiene una mayor sensibilidad ya que triplica su conductancia inicial mientras que el sensor 1 solo la incrementa en un 50%.

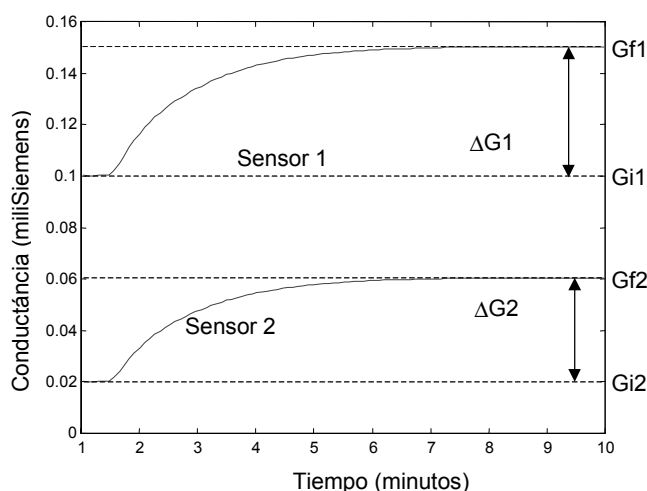


Figura 4.12: Comparación de sensibilidades entre dos sensores

Por otro lado, el escoger los sensores con mayor sensibilidad lo único que garantiza es una capacidad de detección superior, pero no garantiza un mayor poder de resolución que identifique y distinga diferentes compuestos volátiles, ya que, a lo mejor, los sensores con mayor señal reaccionan de manera muy parecida y no cubren correctamente el espectro de aromas a medir. Por lo tanto, este criterio se debe utilizar simplemente para asegurarnos de que la señal que producen los sensores es útil y no es única y exclusivamente ruido.

4.7.4.2 Por covarianza entre variables

Tanto PCA como PLS son algoritmos lineales que basan su funcionamiento en la correlación entre variables. Ambos análisis buscan unas componentes principales (denominadas variables latentes en PLS) sobre las que se pueden proyectar las contribuciones de cada una de las variables (los “Loads” en el argot quimiométrico).

Proyectando las variables sobre las dos primeras componentes principales se puede obtener información sobre la relación entre las mismas. La cercanía entre variables suele presuponer una buena correlación entre ellas. En el caso de que estén en situación completamente opuesta indica que están fuertemente anticorreladas. Estas situaciones tienen diferente interpretación según se esté trabajando con PCA o PLS.

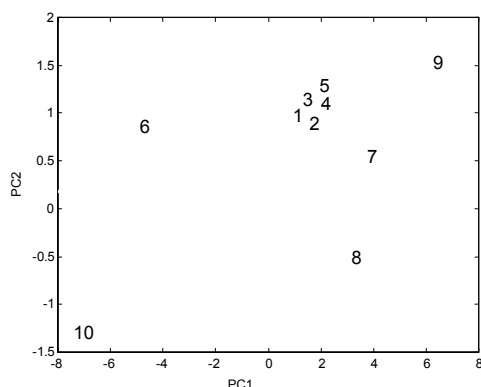


Figura 4.13: Ejemplo de interpretación de loads en un diagrama PCA

Si se está trabajando con componentes principales, variables que estén fuertemente correladas o anticorreladas aportan información similar, lo que indica que son redundantes. El eliminar todas menos una puede permitir reducir la dimensionalidad del vector que define la medida sin perder información importante. Sin embargo, cuando las variables son bastante ruidosas puede ser interesante conservar más de una para intentar promediar su efecto. La figura 4.13 muestra un ejemplo simulado en el que se observa como las variables 1 a 5 aportan información similar (están fuertemente correladas), mientras que la variable 10 está claramente anticorrelada a la variable 9. Así pues, una posible elección para reducir el número de variables podría incluir la 1, 6, 7, 8, 9, representando la variable 1 a todas las de su agrupación y la 9 a su variable anticorrelada numero 10.

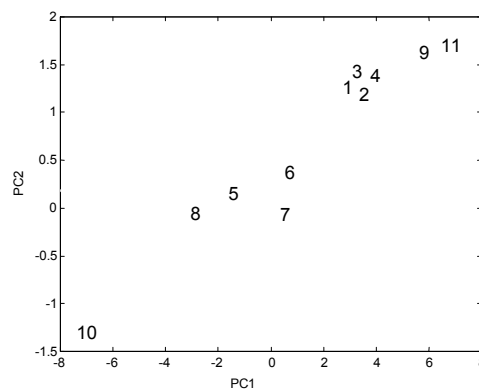


Figura 4.14: Ejemplo de interpretación de loads en un diagrama PLS

En el caso del algoritmo PLS, la cercanía entre variables predictoras y a predecir indica una buena correlación entre ellas, por lo que para construir el modelo PLS interesa incluir las variables cercanas y aquellas en posición opuesta a la variable a predecir. La figura 4.14 muestra un ejemplo ficticio para explicar este concepto, donde las variables 1 a 10 son las predictoras y la 11 la variable a predecir. A partir de la proyección bidimensional el conjunto mínimo de variables que se podrían escoger serían la 1 (como “representante” de la agrupación 1-4), la 8 y la 9. La agrupación 5-7 no está cerca de la variable a predecir y, además, su aportación a las dos primeras componentes principales es prácticamente nula.

De todas maneras, este tipo de selección de variables, a nivel práctico, no tiene mucho éxito. B.Wise, experto en el campo del análisis multivariante, cree que la existencia de variables perturbadoras puede anular por completo la eficacia de este método. Además, es un método lineal que no funciona adecuadamente en procesos no lineales.

4.7.4.3 Algoritmos genéticos

Los algoritmos genéticos son procesos de búsqueda basados en los principios de la selección y la evolución natural [4.46-48]. Las posibles soluciones a un problema son codificadas en forma de cadenas binarias y la búsqueda se inicia con una población aleatoria de posibles soluciones.

En el problema de la selección de variables, cada posible combinación es codificada con una cadena binaria tan larga como parámetros se consideren para encontrar la combinación óptima de variables. En dicha cadena, cada variable tiene asignada una posición o bit, de manera que una posible solución vendrá descrita por una sucesión de unos y ceros indicando la presencia (con un 1) o la ausencia (con un cero) de cada una de las variables en esa combinación particular.

En este tipo de algoritmos, cada miembro de la población, que representa una posible solución, es testada con algún criterio objetivo, de manera que cada uno de los miembros de la población se valora en función de su “fitness” (valor del criterio).

A las soluciones mejor valoradas se les permite sobrevivir y pasar a la siguiente iteración (“generación”), mientras que las soluciones de peor “fitness” desaparecen en las sucesivas generaciones. El algoritmo genético mantiene las iteraciones hasta que iguala o supera el “fitness” establecido como meta o hasta que existe una convergencia en la población, de manera que un determinado porcentaje de sus miembros acaben siendo idénticos.

La creación de los miembros de la población de una nueva iteración se realiza a partir de combinaciones y mutaciones entre los miembros supervivientes de la anterior

población. La combinación consiste en cruzar dos miembros de la antigua población, creando dos nuevos individuos en los que los primeros N bits son de uno de los “padres” y el resto del otro. N, denominado “crossover point” o punto de cruce es un valor aleatorio. La figura 4.15 muestra este concepto. Por otro lado, la mutación de un miembro consiste en el cambio aleatorio de algún bit de su cadena.

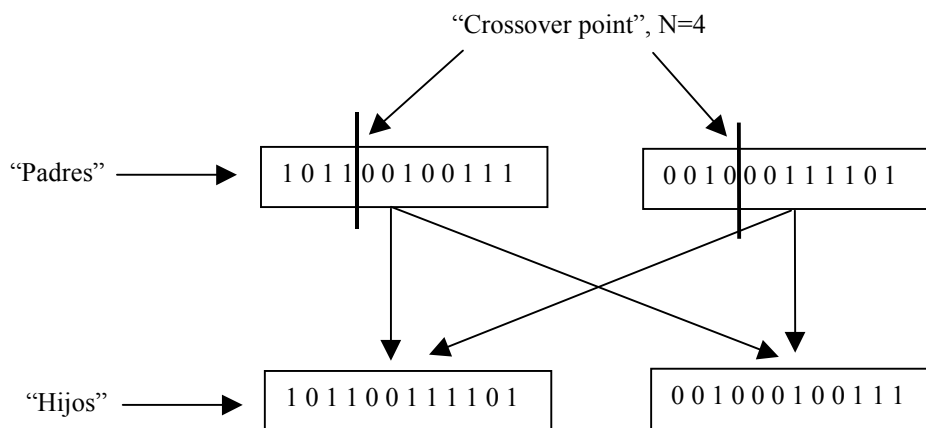


Figura 4.15: Esquema explicativo del cruce entre “padres” para la generación de “hijos”

Existe un teorema [4.49] en el que se demuestra que la iteración sucesiva permitiendo sobrevivir a las combinaciones que mejor cumplen con el criterio preestablecido tras aplicarles combinaciones y mutaciones, permite llegar a una serie de patrones (denominados “schematas”) que convergen a la solución óptima del problema.

El algoritmo genético que se ha aplicado en esta tesis tiene por objetivo encontrar la mejor combinación de variables que debe utilizar un modelo PLS para predecir cada uno de los indicadores de calidad medidos en las piezas de fruta que a su vez fueron utilizadas para obtener resultados con la nariz electrónica. Una vez más, dicha función pertenece al paquete de análisis multivariante “PLS Toolbox” [4.50] que corre bajo el entorno Matlab.

Entre las características de la implementación particular que hace su autor sobre los algoritmos genéticos destacaremos las siguientes:

- Implementa técnicas de combinación (“crossovers”) simples o dobles
- En cada iteración descarta la mitad de la población y permite sobrevivir a la mitad de los miembros con un mejor “fitness”
- El criterio o “fitness” que se utiliza se basa en una validación cruzada PLS, escogiendo de forma aleatoria grupos de medidas de las que obtiene un error cuadrático medio entre la predicción y el valor real. Por su puesto, a menor error, mejor “fitness”.

Los parámetros que utiliza dicha función son los siguientes:

- Número de miembros de la población: Indica el número de individuos que deben considerarse y evaluarse en cada iteración
- Número de términos iniciales: Número de miembros en la población inicial
- Número máximo de generaciones: Número máximo de iteraciones antes de parar el proceso
- Criterio de convergencia: Porcentaje de miembros de la población idénticos para considerar que se ha convergido a una solución y que por lo tanto se deben finalizar las iteraciones
- Probabilidad de mutación: Fracción de bits que deben ser cambiados en cada generación
- Tipo de combinación: “Crossover simple o doble”
- Número de subconjuntos en validación cruzada PLS: Como su nombre indica, numero de subconjuntos que serán iterativamente utilizados para entrenar y evaluar el “fitness” de cada miembro de la población
- Numero de iteraciones: Numero de veces que se deben formar los subconjuntos para obtener un “fitness” medio con menor “rizado” debido a la aleatoriedad.

REFERENCIAS

- [4.1] Agilent Technologies, <http://www.chem.agilent.com>
- [4.2] Gardner J.W., Shurmer H.V., Tan T.T., *Application of an electronic nose to the discrimination of coffees*. Sensors and Actuators B, (1992) vol 6, 71
- [4.3] Abe H., Yoshimura T.Y., Kanaya S., Takahashi Y., Miyashita Y., Sasaki S., *Automated odor-sensing system based on plural semiconductor gas sensors and computerised pattern recognition techniques*. Anal. Chim. Acta, (1987) 194, 1-9.
- [4.4] Bartlett P.N., Blair N., Gardner, J.W., *Electronic noses, principles, applications and outlook*. in 15e Colloque Assoc. Sci. Int. Cafe. 1993. Montpellier, France.
- [4.5] Sundgren, H., Winqvist, F., Lukkari, I., Lundström, I., *Artificial neural networks and gas sensor arrays: quantification of individual components in a gas mixture*. Meas. Sci. Technol.(1991) vol 2, 464-469.
- [4.6] Nakamoto T., Fukuda A., Moriizumi T., *Perfume and flavor identification by odor sensing system using quartz resonator sensor array and neural network pattern recognition*. Sensors and Actuators B, (1993) 10, 85-90.
- [4.7] Carey W.P., Yee S.S., *Calibration of nonlinear solid-state sensor arrays using multivariate regression techniques*. Sensors and Actuators B, (1992). 9, 113-122.
- [4.8] J.Brezmes, E.Llobet, X.Vilanova, X.Correig, *Vapour identification and quantification under variable moisture conditions using gas sensor arrays and artificial neural networks*, Proceedings of the 11th European Conference on Solid-State Transducers, Eurosensors XI, Warsaw (Poland) 1997, 1317-1320
- [4.9] Gardner J.W., Hines E.L., Pang C., *Detection of vapours and odours from a multisensor array using pattern recognition: self-organising adaptive resonant techniques*. Measurement + Control (1996) vol 29
- [4.10] E.L.Hines, E.Llobet, J.W.Gardner, *Electronic noses: A review of signal processing techniques*, IEE Proceedings, Circuits Devices and Systems, Vol. 146, No. 6, (1999) 297-310
- [4.11] Jackson J.E., *Principal component and factor analysis: Part 1-Principal components*, J.Qual. Tech (1981) vol 13,1

- [4.12] Kresta J.V., MacGregor J.F, Marlin T.E., *Multivariate statistical monitoring of process operating performance*, Can. J. of Chem. Eng., (1991) vol 69, 35-47
- [4.13] Gardner, J.W., *Detection of vapours and odours from a multisensor array using pattern recognition. Part 1: principal components and cluster analyses*. Sensors and Actuators B (1991) vol 4, 108-116.
- [4.14] The Mathworks, inc, Matlab (version 5.3), The Mathworks, inc. <http://www.mathworks.com>
- [4.15] B.Wise, N.B. Gallager, *PLS ToolBox 2.0*, Eigenvector research , <http://www.eigenvector.com>
- [4.16] Artursson T., Sjostrom M., Holmberg M., *A comparison between different drift counteraction methods*, proceedings of ISOEN 99, Tubingen, Germany september 1999, 241-242
- [4.17] Holmberg M., Winquist F., Lundstrom I., Davide F., Di Natale C., D'amico A., *Drift counteraction for an electronic nose*, Sensors and actuators B 35-36 (1996), 528-535
- [4.18] Lorber A., Wangen L.E., Kowalski A., *A theoretical foundation for the PLS algorithm*. J. Chemometrics, 1987. Vol 1, 19.
- [4.19] Geladi P., Kowalski B.R., *PLS tutorial*, Anal. Chim. Acta, 185 (1), 1986
- [4.20] Martens H., Naes T., *Multivariate calibration*, John Wiley & Sons, New York, 1989
- [4.21] Höskuldsson A., *Prediction methods in Science and Technology*, Thor publishing, Denmark 1996.
- [4.22] Wold S., Esbensen K., Geladi P., *Principal component analysis*, Chemo. And Intell. Lab. Sys., (1989) vol 7, 53-65
- [4.23] Wise B.M., *Adapting multivariate analysis for monitoring and modeling dynamic systems*, Ph D. Dissertation, University of Washington, Seattle, 1991
- [4.24] Lippman R.P., *An introduction to computing with neural nets*, IEEE ASSP Magazine , abril 1997, 4-22
- [4.25] Grossberg S., *Studies of the mind and brain*, Drodrecht, Holland: Reidel Press, 1982
- [4.26] Hornik K. et al., *Multilayer Feedforward networks are universal aproximators*, Neural Networks, (1989), vol 2, 359-366

- [4.27] Gardner J.W., Hines E.L., Wilkinson M., *The application of artificial neural networks in an electronic nose*. Meas. Sci. Technol., (1990). Vol 1, 446-451
- [4.28] Rummelhart D., McClelland J. and the PDP research group, *Parallel distributed processing*, Vol 1, chapter 8, the M.I.T. press, 1986
- [4.29] Demuth H., Beale M., *Neural Network Toolbox v1.0* , The Mathworks Inc., <http://www.mathworks.com>
- [4.30] Grossberg S., *Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction and illusions*, Biological cybernetics (1976) vol 23, 187-202
- [4.31] Grossberg S., *How does a brain build a cognitive code?*, Psychological Review, (1980) vol 1, 1-51
- [4.32] Carpenter G.A., Grossberg S., *A massively parallel architecture for a self-organizing neural pattern recognition machine*, Computer vision, graphics, and image processing, (1987) vol 37, 54-115
- [4.33] Carpenter G.A., Grossberg S., *ART2: Stable self-organization of pattern recognition codes for analog input patterns*, Applied Optics (1987), vol 26 4919-4930
- [4.34] Carpenter G.A., Grossberg S., *ART 3 hierarchical search: Chemical transmitters in self-organizing pattern recognition architectures*, International joint conference on neural networks (washington DC), 30-33 Hillsdale, NJ: Erlbaum Associates
- [4.35] Carpenter G.A., Grossberg S., Rosen D., *Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system*, Neural Networks, (1991) vol 4, 759-771
- [4.36] Zadeh L., *Fuzzy sets*, Information and control (1965), vol 8, 338-353
- [4.37] Carpenter G.A., Grossberg S., Markuzon N., Reynolds J., Rosen D., *Fuzzy Artmap: A Neural Network architecture for incremental supervised learning of analog multidimensional maps*, IEEE Transactions on neural networks, (1992) vol 3, No 5, 698-713
- [4.38] Carpenter G.A., Grossberg S., Reynolds J., *Artmap: Supervised realtime learning and classification of nonstationary data by a self-organizing neural network*, Neural Networks, (1991) vol 4, 565-588

- [4.39] Gardner J.W., Hines E.L., Pang C., *Detection of vapours and odours from a multisensor array using pattern recognition: self-organising adaptive resonant techniques*. Measurement + Control, (1996) vol 29, 1996.
- [4.40] E. Llobet, E. L. Hines, J.W. Gardner, S. Franco, *Non-destructive banana ripeness determination using a neural network-based electronic nose*, Meas. Sci. Technol. 10 (1999), 538–548
- [4.41] Llobet E., Brezmes J., Vilanova X., Fondevila Ll., Correig X., *Quantitative vapor analysis using the transient response of non-selective thick-film tin oxide gas sensors*, Transducers'97, Chicago, Digest of Technical Papers, Vol 1, ISBN: 0-7803-3829-4, 971-974
- [4.42] Vilanova X., Llobet E., Alcubilla R., Sueiras J.E., Correig, X., *Analysis of the conductance transient in thick-film tin oxide gas sensors*, Sensors and Actuators B31 (1996), 175-180
- [4.43] Llobet E., Vilanova X., Brezmes J., Alcubilla R., Calderer J., Sueiras J.E., Correig, X., *Analisis of the conductance transient of thick-film tin dioxide gas sensors under successive gas injection steps*, Meas. Sci. Technol., (1997) Vol. 8, Iss 10, 1133-38
- [4.44] Nakata S., Akakabe S., Nakasuji M., Yoshikawa K., *Gas sensing based on a nonlinear response: Discrimination between Hydrocarbons and quantification of individual components in a Gas mixture*, Anal. Chem., (1996) vol 68, 2067-2072
- [4.45] Llobet E., Brezmes J., Ionescu R., Vilanova X., Al-Khalifa S., Gardner J.W., Bârsan N., Correig X., *Wavelet Transform and Fuzzy ARTMAP Based Pattern Recognition for Fast Gas Identification Using a Micro-Hotplate Gas Sensor*, Sensors and Actuators B *en prensa* (2001)
- [4.46] Holland J.H., *Genetic algorithms*, Scientific american (1992), 267(1),66
- [4.47] Lucasius C.B., Kateman G., *Understanding and using genetic algorithms Part 1. Concepts, properties and context*, Chemom. Intell. Lab. Syst.,(1993) 19, 1-33
- [4.48] Lucasius C.B., Kateman G., *Understanding and using genetic algorithms Part 2. Representation, configuration and hybridization*, Chemom. Intell. Lab. Syst.,(1994) 25, 99-145
- [4.49] Holland J., *Adaptation in natural and artificial systems*, (1975), University of Michigan Press

CAPITULO 5:

MEDIDAS CON MANZANAS

5.1 INTRODUCCIÓN

En los capítulos 5, 6 y 7 se discuten los resultados obtenidos con las variedades de fruta que se han estudiado en este trabajo. Todas ellas (manzanas, peras, melocotones y nectarinas) son frutas climatéricas caracterizadas por una producción creciente de etileno y otros componentes aromáticos a medida que su maduración avanza. Dicha producción alcanza un máximo y después decrece con la aparición de la senescencia. De entre todas las variedades climatéricas se escogieron estas cuatro debido a su disponibilidad, a su interés comercial y al hecho de que se producen en zonas relativamente cercanas a nuestro laboratorio.

Las cuatro variedades escogidas han sido estudiadas con los dos sistemas de olfato electrónico diseñados a lo largo de este proyecto. Así, en una primera temporada se hicieron estudios de viabilidad con el prototipo inicial, mientras que en el segundo año se realizaron experimentos con objetivos concretos. Por este motivo, la estructura de los capítulos de resultados ha sido organizada de forma similar, con un primer apartado que se ocupa de las experiencias iniciales y otro que detalla los resultados obtenidos con el segundo prototipo. En [5.1] se pueden consultar los resultados obtenidos la primera temporada, mientras que en [5.2-4] se detallan los experimentos realizados con el prototipo de flujo continuo.

La metodología que se ha seguido en cada temporada es muy similar para todas las variedades de fruta estudiadas, de manera que la descripción experimental que se realiza en este capítulo sirve para ilustrar los estudios realizados con el resto de variedades que se explicarán en los capítulos 6 y 7.

5.2 ESTUDIO DE VIABILIDAD

5.2.1 Objetivos

El objetivo primordial del estudio preliminar con manzanas era comprobar si una nariz electrónica podía identificar el estado de maduración de las manzanas a lo largo de su proceso de maduración en condiciones ambientales (maduración de “shelf life”). Para ello es imprescindible que la concentración de volátiles de la fruta provoque en la matriz de sensores una respuesta que evolucione medida que avanza el estado de maduración de las muestras medidas con el instrumento.

Aplicando técnicas de reconocimiento de patrones se intentará clasificar la fruta entre tres estados de maduración: verde, madura y sobremadura. Es importante resaltar que en el estudio preliminar las mediciones destructivas se realizaron para orientar al operario en su decisión sobre como clasificar la fruta tras las medidas con nariz electrónica. Por lo tanto, la decisión última sobre si la fruta esta verde, madura o sobremadura corresponde, en este estudio preliminar, a un criterio subjetivo apoyado en técnicas destructivas como la firmeza, acidez y sólidos solubles.

Además, teniendo en cuenta que en la aplicación de redes neuronales existe una fase de aprendizaje y otra de evaluación, y que en la de aprendizaje se consideran salidas ideales a las clasificaciones hechas por el operario, es necesario aclarar que el índice de aciertos en la clasificación nos indica el grado de coincidencia entre la clasificación hecha por la nariz electrónica y el operario, y no el grado de acierto sobre el estado exacto de maduración de la fruta. Esta limitación es inevitable ya que,

desgraciadamente, no existe un criterio único y objetivo para clasificar manzanas en diferentes estados de maduración. En definitiva, en la fase de entrenamiento la red aprende a relacionar el criterio del operario con las señales que recibe de los sensores y que son provocadas por la producción aromática de la fruta en cada medida.

5.2.2 Metodología experimental

El estudio de viabilidad para manzanas se realizó con muestras de la variedad “Golden” obtenidas de una cooperativa de Mollerusa, en la provincia de Lérida. Se cogieron 56 muestras que habían sido recolectadas el 24 de setiembre. Estas muestras se dividieron en 4 grupos de 14 piezas cada uno. Los dos primeros grupos se empezaron a medir con el primer prototipo el 28 de setiembre, mientras que los grupos 3 y 4 se conservaron en la nevera y fueron sacados el 25 de octubre para empezar su seguimiento al día siguiente. En total se midieron 56 muestras y se realizaron un total de 48 medidas con el sistema de olfato electrónico.

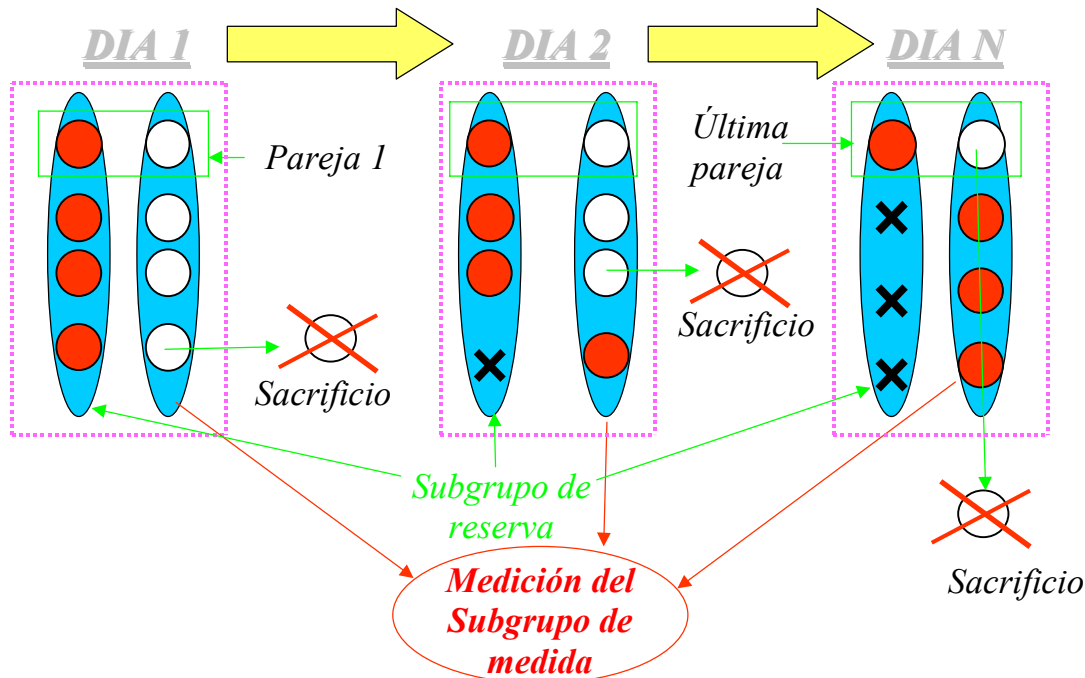


Figura 5.1: Proceso de sustitución y medida a lo largo del shelf life de la fruta

La figura 5.1 esquematiza la metodología experimental que se diseñó para poder realizar el seguimiento del proceso de maduración con técnicas destructivas y con la nariz electrónica de forma paralela. Este proceso se utilizó en todas las variedades de fruta medidas con el primer prototipo.

La base del método radica en saber emparejar correctamente las piezas de un grupo de medida, de forma que cada pareja esté formada por dos piezas de tamaño, color y tacto muy parecidos para poder afirmar que su estado de maduración es similar. Así, en el caso de las manzanas Golden, se formaron 7 parejas para cada uno de los cuatro grupos de fruta recolectada.

Una vez las piezas han sido emparejadas, se forman dos subgrupos, el de medida y el de reserva, de forma que de cada pareja una pieza se destina al subgrupo de medida y otra al de reserva. La primera medida con el sistema de olfato electrónico se realiza con el subgrupo de medida. Tras finalizar las medidas con nariz electrónica, se escoge una pieza al azar de dicho subgrupo y se aplican sobre ella mediciones destructivas.

Finalmente, se coge su pareja (que estaba en el grupo de reserva) y se incorpora al grupo de medida, de manera que el nuevo grupo de medida sigue teniendo el mismo número de piezas y la destruida queda substituida por una muestra con un estado de maduración similar.

Con las manzanas Golden se realizaron, salvo fines de semana u otros imprevistos, dos medidas diarias con la nariz electrónica. Primero se midieron los grupos 1 y 2 (desde el 24 de setiembre hasta el 19 de octubre) y posteriormente los grupos 3 y 4 que se habían conservado en la nevera (desde el 26 de octubre hasta el 12 de noviembre). En cada medida se pesaba todo el grupo y se obtenía el perímetro medio de las piezas para tener información con la que los algoritmos de reconocimiento de patrones pudiesen realizar operaciones de normalización. Además, las medidas sucesivas con nariz electrónica se espaciaron una hora para dar tiempo a los sensores a recuperar su estado inicial. La tabla 5.1 muestra el número total de medidas que se hicieron para cada grupo.

Grupo	Número de medidas
1	13
2	13
3	11
4	11

Tabla 5.1: Número de medidas para cada grupo

Las técnicas destructivas aplicadas a las manzanas en este estudio preliminar incluyeron medidas de firmeza, de acidez y de sólidos solubles. Para poder monitorizar el proceso de maduración durante el “shelf life” de la fruta (maduración a 20° y humedad relativa ambiente) por un período de unos 20 días, las medidas destructivas se practicaron aproximadamente cada tres días. Solo en estas jornadas se realizó la sustitución de la muestra destruida por su pareja de reserva.

5.2.3 Evolución de la señal

Los sensores incluidos en la nariz son los que se incluyeron en el primer prototipo, sensores ya descritos en el capítulo 3. Lo primero que se hizo fue determinar el tiempo de concentración y el número de inyecciones necesarias para provocar una respuesta adecuada en la matriz de sensores del equipo. Tras varias pruebas, se estableció que las piezas debían permanecer en la cámara de concentración una hora, y que una sola inyección de 50 ml de muestra de la atmósfera de la cámara de concentración era suficiente para generar una señal clara.

Como la respuesta de los sensores era suficientemente clara, lo importante era comprobar si la evolución de la señal permitiría seguir el proceso de maduración de las muestras. La figura 5.2 muestra la evolución de cada uno de los cuatro grupos de manzanas a los que se siguió el proceso de maduración con el primer prototipo. El eje horizontal representa los días transcurridos desde el inicio del shelf life, mientras que el vertical representa el incremento de conductancia tras la inyección de la muestra de atmósfera de la cámara de concentración.

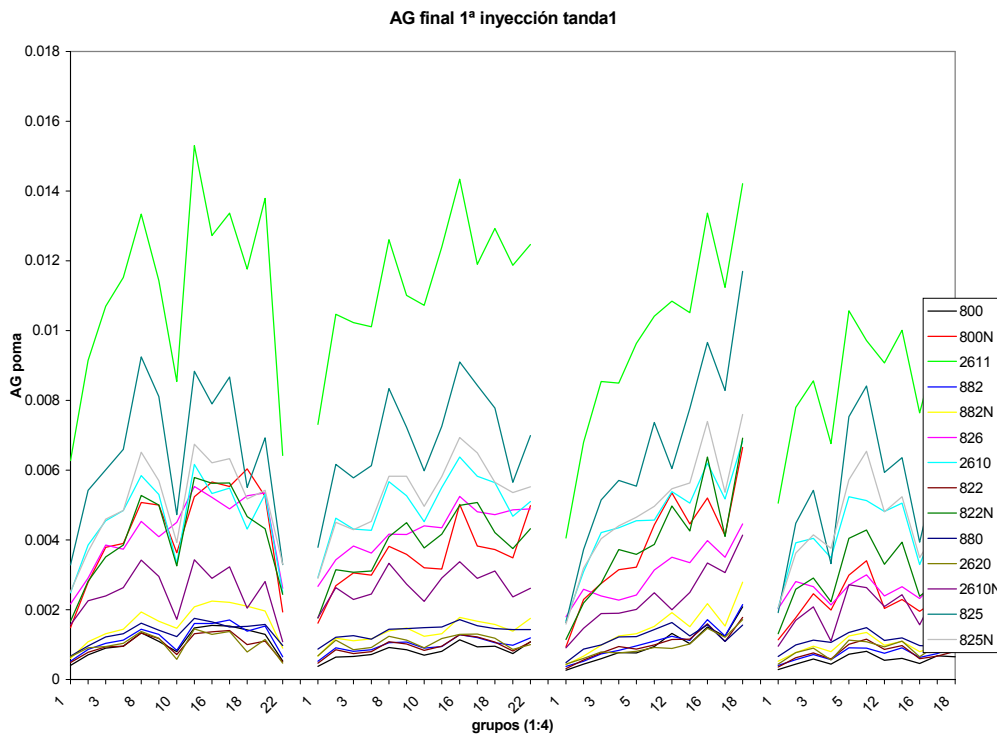


Figura 5.2: Evolución de la señal para los 4 grupos de manzanas

En la figura se puede observar, en líneas generales, que existe una cierta tendencia a aumentar el valor del incremento de conductancia (ΔG) a medida que pasan los días de “shelf life”. De todas maneras, esta tendencia es poco clara salvo para el grupo 3. De hecho, la poca repetitividad de los sensores junto a la variabilidad inherente a las piezas de fruta hace que el “rizado” entre medidas sea de una amplitud claramente superior a la evolución que se pueda derivar de la señal.

Por otro lado, también se observa que prácticamente todos los sensores siguen una tendencia similar y en lo único que se diferencian es en la sensibilidad que muestran ante los volátiles generados por la fruta.

5.2.4 Resultados

La figura 5.2 ya nos permite intuir que los resultados no van a ser muy prometedores ya que, aunque se observa una cierta tendencia, también se ve que la variabilidad en las mediciones supera con creces este comportamiento. Una primera aproximación

consistió en plasmar en un diagrama PCA las 48 medidas y observar si en alguna de las dos primeras componentes principales se observaba alguna tendencia a destacar. La figura 5.3 muestra dicho diagrama en el que cada medida ha sido etiquetada con su número de orden. Cada color representa un grupo diferente, siendo el grupo 1 de color rojo, el grupo 2 de color azul, el 3 de color añil y el cuatro de color amarillo.

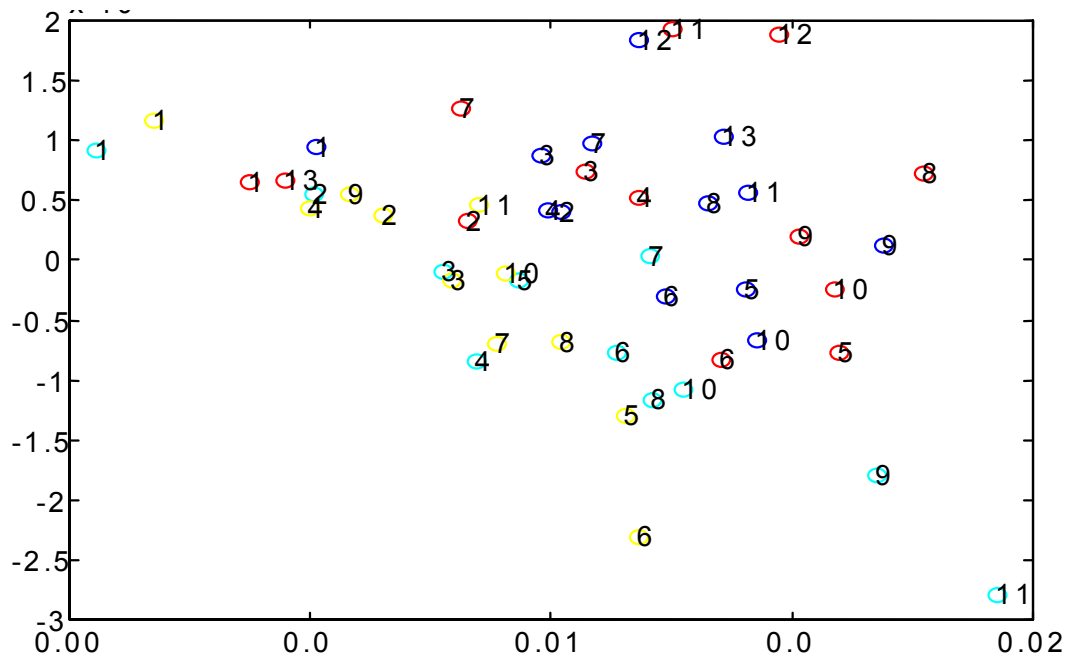


Figura 5.3: Diagrama PCA de las dos primeras componentes principales de los 4 grupos de manzanas (PC1 eje horizontal, PC2 eje vertical).

Los datos de entrada para realizar el PCA fueron el incremento de conductancia de todos los sensores. En esta figura no se pueden deducir ningún tipo de comportamiento respecto a las dos primeras componentes principales, lo que permite intuir que los resultados de clasificación no serán excesivamente buenos.

Para comprobar si era posible realizar una clasificación del estado de maduración de las piezas se utilizó una red neuronal. El objetivo era clasificar cada una de las medidas como fruta verde, madura o sobremadura.

Como se puede observar en la figura 5.4, la red utilizada es de tipo “feed-forward” con una capa oculta. Las entradas son las combinaciones de diferentes parámetros de los

sensores, mientras que la salida ideal es una codificación “uno a tres” en la que la clasificación correcta se denota por un uno y las dos clases inactivas por un cero. Así, por ejemplo, una salida ideal “0-1-0” indicaría que la clasificación correcta para esa medida sería manzanas “maduras”.

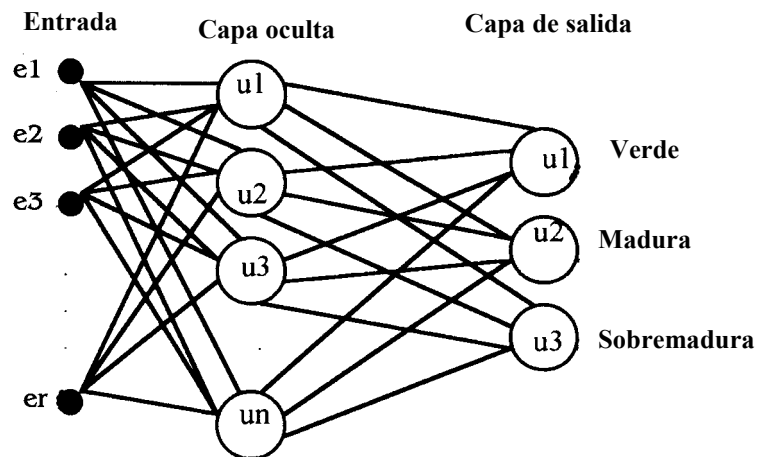


Figura 5.4: Estructura de la red neuronal para clasificar manzanas

Como se ha comentado previamente, toda red neuronal requiere de una fase de entreno y otra de evaluación. En nuestro caso se utilizó una aproximación cruzada de orden 4 para evaluar la eficacia del método. En concreto, al tener cuatro grupos diferentes de manzanas, se iteró cuatro veces el proceso de entreno y evaluación. En cada una de las iteraciones, se entrenó con tres grupos y se evaluó con el cuarto. Esta aproximación se utilizó descartando la validación cruzada de orden uno (el método “leave-one-out”) por dos razones:

- Por un lado, el entreno de una red neuronal es un proceso computacionalmente costoso, por lo que una validación cruzada de orden uno exigiría la repetición del aprendizaje 48 veces, ya que tenemos un conjunto de 48 medidas con nariz electrónica. Esto impediría probar diferentes combinaciones de entradas, número de neuronas en la capa oculta, etc., dado que el tiempo de computación necesario para cubrir todo el abanico de posibilidades sería prohibitivo.
- En segundo lugar, la utilización de grupos de fruta diferentes para el proceso de entreno y evaluación permite comprobar el comportamiento del prototipo

ante una situación más realista en la que el entrenamiento y la evaluación se hacen con fruta totalmente diferente. De utilizarse un “leave-one-out”, al evaluar la medida n (que corresponde a uno de los cuatro grupos) se entrena con el resto de medidas, entre las que se incluyen las medidas realizadas con las piezas del mismo grupo días antes y después a la medida n .

La tabla 5.2 muestra los resultados obtenidos para las mejores combinaciones de parámetros y para diferente número de neuronas en la capa oculta. En ella se han utilizado las 48 medidas disponibles, que corresponden a 4 grupos diferentes de manzanas. Como se puede observar, en las mejores combinaciones se utiliza el parámetro incremento de conductancia (ΔG) de un número reducido de sensores. Los sensores se han escogido aplicando el criterio de sensibilidad descrito en el capítulo 4.

Como se puede comprobar en la tabla, los parámetros de peso (p) y superficie media (s) de las piezas están presentes en todas las combinaciones con un índice de aciertos elevado. Esto se explica por su utilización como señales de normalización. Hay que tener en cuenta que piezas de fruta más verdes pero con un gran tamaño pueden generar la misma cantidad de volátiles que piezas más pequeñas en un estado de maduración más avanzado. El peso es, por lo tanto, un parámetro de normalización a tener en cuenta. Asimismo, la superficie media de las piezas (calculadas a partir del perímetro medio) es un dato importante si se quieren normalizar correctamente los aromas emitidos.

Como la mayoría de los sensores tiene un comportamiento similar (aunque a diferentes escalas), tal y como se dedujo de la figura 5.2, no es de extrañar que el mejor resultado corresponda a la utilización de un solo sensor (en este caso el 2610), aunque la combinación de dos sensores (2610 y 800N) da resultados muy parecidos. Debido a la variabilidad tanto de los sensores como de la fruta, así como de la subjetividad del operario, estos resultados no deben ser interpretados de forma categórica, si no más bien como un estudio indicativo.

Entre las conclusiones que se pueden extraer destaca el hecho de que, aproximadamente, el sistema está de acuerdo con el operario en 3 de cada 4 casos (73%), mientras que si la clasificación de las muestras en tres clases se realizase de forma aleatoria los aciertos se reducirían a un 33%, estadísticamente hablando. Por otro lado, se observa que con pocos sensores se obtienen buenos resultados y que las señales de normalización ayudan considerablemente a incrementar el grado de éxito en la clasificación, ya que como se observa en la tabla resumen, en las mejores combinaciones siempre aparecen el peso y superficie como parámetros incluidos. En el capítulo 7, al medir melocotones, se presenta un estudio comparativo en el que queda patente la importancia de estas dos señales.

Variables	n° sensores	Sensores	N° inputs	Neuronas ocultas	Aciertos/ Medidas	% ok
Go,G1,p,s	4	800N, 2610, 825, 825N	10	6	33/48	68.75
ΔG1,p,s	4	800N, 2610, 825, 825N	6	2,4,6	32,32,33/48	68.75
ΔG1,p,s	1	2610	3	2,4,6	34,33,35/48	<u>72.91</u>
ΔG1,p,s	2	800N,2610	4	2,4,6	33,33,34/48	70.8
ΔG1,p,s	2	800N,825	4	2,4,6	33,34,30/48	70.8
ΔG1,p,s	2	800N,825N	4	2,4,6	32,33,32/48	68.75
ΔG1,p,s	2	2610,825	4	2,4,6	32,30,32/48	66.6
ΔG1,p,s	2	2610,825N	4	2,4,6	32,32,34/48	70.8

Tabla 5.2: Resultados de clasificación con redes neuronales para manzanas

En este proceso de clasificación cruzada se han utilizado los cuatro grupos de manzanas recolectados. Sin embargo, dos de ellos fueron almacenados en un frigorífico y, por lo tanto, los grupos 1 y 2 pueden haber madurado de forma diferente a los grupos 3 y 4. De hecho, la evolución de los valores de firmeza muestra diferencias importantes entre las muestras medidas inmediatamente después de la recolecta y las almacenadas un mes en el frigorífico. Por ese motivo, para ver si se obtenían mejores resultados, se realizó una validación cruzada utilizando sólo los grupos que no pasaron por frigorífico (1 y 2).

La tabla 5.3 muestra los resultados en los que se puede observar que se utilizaron un total de 26 medidas, con dos iteraciones, de forma que en la primera se utilizó el grupo

1 como conjunto de entrenamiento y el grupo dos como grupo de evaluación, mientras que en la segunda iteración cambiaron los papeles.

Variables	n° sensores	Sensores	n° inputs	Neuronas ocultas	aciertos/medidas	% ok
Go,G1,p,s	4	800N, 2610, 825, 825N	10	6	14/26	53.8
AG1,p,s	4	800N, 2610, 825, 825N	6	2,4,6	12,12,13/26	50
AG1,p,s	1	2610	3	2,4,6	13,16,13/26	<u>61.53</u>
AG1,p,s	2	800N,2610	4	2,4,6	14,13,12/26	53.84
AG1,p,s	2	800N,825	4	2,4,6	16,14,11/26	<u>61.53</u>
AG1,p,s	2	800N,825N	4	2,4,6	15,13,13/26	57.7
AG1,p,s	2	2610,825	4	2,4,6	16,14,15/26	<u>61.53</u>
AG1,p,s	2	2610,825N	4	2,4,6	16,13,14/26	<u>61.53</u>

Tabla 5.3: Resultados de clasificación con los grupos 1 y 2

Como se puede ver, los resultados no mejoran respecto al estudio anterior. Esto puede ser debido a dos razones. En primer lugar, el número de medidas de entrenamiento baja drásticamente (se reduce a 13 frente a un mínimo de 35 en el caso anterior). Por otro lado, aunque el ritmo de maduración de los grupos 3 y 4 hayan sido diferentes eso ya ha sido tenido en cuenta por el operario, que asignó más medidas verdes y menos medidas sobremaduras a estos grupos ya que dicha clasificación no se hacía por días de “shelflife” sino por color, apariencia y por los parámetros destructivos aplicados.

5.2.5 Conclusiones

Con estos estudios preliminares sobre manzanas se puede concluir que, aunque los volátiles generados por las manzanas son detectados fácilmente por los sensores del prototipo, el proceso de maduración se traduce en una señal cuya tendencia es débil frente a otras interferencias como son la falta de repetitividad de los sensores o la variabilidad de las piezas de fruta.

Por ese motivo era fácil prever que un análisis PCA fracasaría en el intento de agrupar las piezas en función de su grado de maduración, ya que, al ser un método no supervisado, busca las direcciones de máxima varianza independientemente del tipo de información que aporten. Como la variabilidad de las muestras y la escasa repetitividad de la respuesta de los sensores aportan mucha mayor varianza que la débil tendencia del proceso de maduración, las componentes principales no reflejan la evolución de dicho proceso.

Por otro lado, un algoritmo de aprendizaje supervisado como una red neuronal es capaz de, al menos en parte, extraer la tendencia subyacente en el proceso y filtrar las interferencias debidas a sensores y fruta. Es por ese motivo que los resultados de clasificación son bastante más buenos de lo que cabría esperar con las tendencias mostradas en la figura 5.1.

En definitiva, el primer prototipo es capaz de detectar los volátiles que emiten las manzanas durante el proceso de maduración y realizar un seguimiento de dicho proceso con ciertas dificultades. Los algoritmos de reconocimiento de patrones que pueden explotar con éxito las señales obtenidas con el equipo deben ser instruidos a distinguir entre información (varianza) que debe ser eliminada e información relevante para clasificar correctamente la fruta, por lo que, a priori, siempre tendrán más éxito los algoritmos de aprendizaje supervisado como las redes neuronales backpropagation.

5.3 MEDIDAS EN FLUJO CONTINUO

5.3.1 Objetivos

Los objetivos planteados para la segunda temporada de mediciones con manzanas fueron múltiples y se centraron tanto en probar el prototipo de flujo continuo como en completar el estudio preliminar en aquellos aspectos que limitaban la validez de los resultados, como la falta de objetividad en la determinación del estado de maduración de la fruta.

Al mismo tiempo, ya que se iban a utilizar técnicas clásicas de control de calidad de forma sistemática, se creyó interesante estudiar el proceso de maduración de un tipo de manzana poco estudiada y de gran valor comercial, como es la variedad de manzana denominada “pinklady”.

En concreto, los aspectos que se trataron con detalle en este estudio fueron el seguimiento del proceso de maduración en función de los días de shelf-life, la relación existente entre las señales de los sensores y los parámetros destructivos y las derivas que experimentan los sensores durante el intervalo temporal del estudio.

5.3.2 Metodología experimental

Como se ha comentado, se utilizaron manzanas Pinklady para el estudio. 200 muestras fueron cosechadas en una finca de la provincia de Lleida en la fecha óptima de recolecta según profesionales del sector. Se formaron cuatro grupos cogiendo seis manzanas para cada uno de forma aleatoria. Las muestras sobrantes (176 en total) se reservaron para ir substituyendo a las piezas destruidas al aplicar técnicas clásicas de medida de calidad.

En cada sesión, el sistema de olfato electrónico mide de forma separada los cuatro grupos de manzanas. Al final de la jornada, se escogían al azar nueve manzanas (1 del grupo 2, 2 del grupo 3 y las 6 del grupo 4) para realizar medidas de calidad. Cada una de las piezas destruidas se substituía por otra con un color, peso y tamaño similares de las restantes piezas recolectadas. Las piezas del grupo 1 fueron la excepción al ser las mismas durante toda la experiencia.

Las medidas se prolongaron desde la recolección hasta 29 días después, manteniendo la fruta bajo condiciones estándar de “shelf life” (20°C, 50-60% de humedad relativa). Debido a la limitación en recursos (sobre todo humanos), las jornadas de medición no se realizaron todos los días (en los fines de semana, por ejemplo, no se realizó ninguna actividad). En cada sesión se median una vez los grupos 2,3 y 4, mientras que el grupo 1 se medía dos veces. En total, se realizaron 88 medidas con el sistema de olfato electrónico.

Al final de cada sesión se realizó una medida de calibración. En cada una de estas medidas, un microlitro de etanol se inyectaba en la cámara de concentración a través de una jeringa cromatográfica, de forma que el líquido se evaporaba casi instantáneamente en la cámara de concentración. Tras un minuto de tiempo de concentración el prototipo pasa a la fase de medida para seguir el proceso habitual de medida. Todas las medidas de calibración se realizaron siguiendo exactamente el mismo proceso y fueron almacenadas en el ordenador para poder monitorizar las derivas que sufren los sensores del prototipo a lo largo del período que duró el estudio.

5.3.3 Seguimiento del proceso de maduración

En primer lugar se estudió la capacidad del sistema para seguir el proceso de maduración de las manzanas a lo largo de su shelf life. Para ello se utilizaron dos métodos no supervisados como son el PCA y la red neuronal Fuzzy Art.

El análisis PCA se aplicó a las 88 medidas realizadas con nariz electrónica. Se realizaron proyecciones utilizando las diferentes familias de sensores por separado (Taguchi serie 8 y FIS series SP y SB). Se realizaron proyecciones utilizando diferentes combinaciones de parámetros y las tres primeras componentes principales en cada análisis. Para determinar las mejores agrupaciones (suponiendo que deberían agruparse por días desde la recolecta) se utilizó el criterio definido en el capítulo 4 para representaciones PCA. Puesto que la maduración es un proceso dinámico, una agrupación por días indicaría, en principio, que la nariz es capaz de seguir el proceso de maduración de las manzanas.

Tras múltiples pruebas, se comprobó que las proyecciones que presentaban agrupaciones significativas correspondían siempre a las realizadas con parámetros de los sensores FIS-SB. Desgraciadamente, el estudio realizado sobre las derivas en los sensores que se presenta en un apartado posterior atribuye estas agrupaciones a la deriva de los sensores y no al proceso de maduración. Por otro lado, ninguna combinación de parámetros utilizando los sensores Taguchi y FIS-SP generó proyecciones con

agrupaciones significativas, resultado decepcionante que, sin embargo, permite concluir que en estas dos familias de sensores no existe deriva apreciable.

Para completar el análisis no supervisado se aplicaron algoritmos Fuzzy Art a las medidas para ver como se agrupaban por semejanza . En primer lugar, se probaron múltiples combinaciones y se mantuvieron las que mejor clasificaban las medidas en función del día en que fueron realizadas. Para ello se utilizó el parámetro “separación” descrito en el capítulo 4 de esta memoria. En todas las pruebas se ajustó el parámetro vigilancia para que el algoritmo siempre clasificase las medidas de fruta en cinco clases (ya que se observaron cinco agrupaciones en el diagrama PCA).

Clase	Medidas
1	g1d01 - g1d02 - g1d02 - g2d02 - g3d02 - g4d02 - g1d03 - g1d03 - g2d03 - g3d03 - g4d03 - g1d04 - g1d04 - g2d04 - g3d04 - g4d04 - g1d07 - g1d07 - g2d07
2	g1d08 - g1d08 - g3d08 - g4d08 - g1d09 - g1d09 - g2d09 - g4d09 - g1d10 - g1d10 - g2d10 - g3d10 - g4d10 - g1d13 - g1d13 - g2d13 - g3d13 - g4d13 - g1d14 - g1d14 - g2d14 - g3d14 - g1d15 - g1d15 - g2d15 - g3d15 - g4d15 - g1d16 - g2d16 - g3d16 - g4d16
3	g1d16 - g1d17 - g1d17 - g3d17 - g4d17 - g1d19 - g1d19 - g2d19 - g3d19 - g4d19 - g1d21 - g1d21 - g2d21 - g3d21 - g4d21 - g1d23 - g1d23 - g4d23 - g1d24 - g1d24 - g2d24 - g3d24 - g4d24 - g1d27 - g4d27
4	g2d23 - g1d27 - g2d27 - g3d27 - g1d28 - g1d28 - g2d28 - g3d28 - g1d29 - g1d29 - g2d29 - g3d29
5	g4d29

Tabla 5.4: Clasificación Fuzzy Art usando sensores FIS-SB

La tabla 5.4 muestra los mejores resultados obtenidos utilizando los sensores FIS-SB. En ella, cada medida está definida por una etiqueta que especifica el grupo al que pertenece (el número después de la “g”) y el número de días entre la recolecta y la medida (el número después de la “d”). Teniendo en cuenta que los resultados con el PCA agrupaban las medidas en función del día de la medida, no es de extrañar que los resultados con el algoritmo Fuzzy Art den resultados similares (desde un punto de vista operativo, Fuzzy Art puede considerarse como una versión automática de la clasificación realizada por un PCA). Desgraciadamente, tal y como se discutirá en el apartado sobre derivas de los sensores, este resultado se debe atribuir a la deriva

temporal que sufren estos sensores a lo largo del tiempo y no al proceso de maduración de las manzanas.

La tabla 5.5 muestra los mejores resultados obtenidos para los sensores Taguchi. Como se puede observar, no hay grandes errores aunque la resolución no es muy buena, ya que la mayoría de medidas están clasificadas bajo la categoría 4. Por último, la tabla 5.6 muestra la mejor clasificación obtenida con los sensores FIS-SP, en donde se pueden detectar clasificaciones claramente erróneas, lo que nos da una idea de la eficacia de esta clase de sensores en el seguimiento de la maduración en manzanas.

Clase	Medidas
1	g1d01 - g1d02 - g1d02 - g2d02 - g3d02 - g4d02 - g1d03 - g2d03
2	g1d03 - g3d03 - g1d04 - g1d04 - g2d04 - g4d04
3	g4d03 - g3d04
4	g1d07 - g1d07 - g2d07 - g1d08 - g1d08 - g3d08 - g4d08 - g1d09 - g1d09 - g2d09 - g4d09 - g1d10 - g1d10 - g2d10 - g3d10 - g4d10 - g1d13 - g1d13 - g2d13 - g3d13 - g4d13 - g1d14 - g1d14 - g2d14 - g3d14 - g1d15 - g1d15 - g2d15 - g3d15 - g4d15 - g1d16 - g1d16 - g2d16 - g4d16 - g1d17 - g1d17 - g3d17 - g4d17 - g1d19 - g2d19 - g3d19 - g4d19 - g2d21 - g3d21 - g4d21 - g1d23 - g1d23 - g2d23 - g4d23 - g2d24 - g3d24 - g2d27 - g3d27 - g4d27
5	g3d16 - g1d19 - g1d21 - g1d21 - g1d24 - g1d24 - g4d24 - g1d27 - g1d27 - g1d28 - g1d28 - g2d28 - g3d28 - g1d29 - g1d29 - g2d29 - g3d29 - g4d29

Tabla 5.5: Clasificación Fuzzy Art utilizando sensores Taguchi

Cla se	Medidas
1	g1d01 - g1d02 - g1d02 - g2d02 - g3d02 - g4d02 - g1d03 - g1d03 - g2d03 - g4d03 - g1d04 - g1d04 - g4d04 - g1d07 - g1d07 - g2d07 - g1d08 - g3d08 - g4d08 - g1d09 - g1d09 - g2d09 - g4d09 - g1d10 - g1d10 - g2d10 - g3d10 - g4d10 - g2d13 - g4d13 - g4d16 - g3d17 - g3d19
2	g3d03 - g2d04 - g3d04 - g1d08 - g1d13 - g1d13 - g1d14 - g1d14 - g2d14 - g3d14 - g1d15 - g1d15 - g2d15 - g1d16 - g1d16 - g2d16 - g1d17 - g2d21 - g3d27 - g4d29
3	g3d13 - g3d15 - g4d15 - g3d16 - g4d17 - g4d19 - g4d21 - g4d24 - g4d27
4	g1d17 - g1d19 - g1d19 - g2d19 - g1d21 - g1d21 - g3d21 - g1d23 - g1d24 - g1d24 - g2d24 - g3d24 - g1d27 - g1d27 - g2d27 - g1d28 - g1d28 - g2d28 - g3d28 - g1d29 - g2d29 - g3d29
5	g1d23 - g2d23 - g4d23 - g1d29

Tabla 5.6: Clasificación Fuzzy Art utilizando sensores FIS-SP

Como ya se ha comentado, los resultados obtenidos para los sensores FIS-SB se deben a derivas en los sensores. En cambio, los resultados para los sensores Taguchi y FIS-SP están libres de derivas, por lo que estas dos últimas familias de sensores fueron las únicas que se utilizaron para correlar las señales medidas con el sistema de olfato electrónico frente a los parámetros obtenidos aplicando técnicas destructivas a las muestras seleccionadas en cada jornada.

5.3.4 Correlación entre parámetros destructivos y señales de los sensores del prototipo

Como se ha comentado en la sección 5.3.2, un total de nueve piezas de los grupos medidos en cada jornada fueron destruidas para obtener parámetros de calidad. Para comparar objetivamente la eficacia de la nariz electrónica frente a técnicas clásicas, los valores obtenidos con las técnicas destructivas se emparejaron con las obtenidas con la nariz electrónica en la misma jornada. De esta manera se formaron 88 parejas de valores.

Aunque inicialmente se consideró la utilización de redes neuronales tipo “feed-forward” con entrenamiento “backpropagation”, finalmente se optó por no utilizarlas debido a su carga computacional durante la fase de aprendizaje y por no disponer de un número suficientemente elevado de medidas para poder entrenar y validar la red correctamente. El algoritmo PLS (Partial Least Squares) fue escogido debido a las limitaciones de la aplicación. Durante una fase de entrenamiento, el algoritmo PLS construye un modelo que describe la relación entre las señales de los sensores y los parámetros de la fruta que deben ser predichos. En la fase de evaluación, dicho modelo predice el parámetro de calidad de una nueva medida no utilizada durante el entrenamiento a partir de las señales de los sensores obtenidas con la nariz electrónica.

Para evaluar la capacidad de la nariz electrónica en la predicción de parámetros de calidad se aplicó una validación cruzada de orden 1 (un “leave one out”). Bajo esta metodología, se utilizan 87 medidas para crear el modelo PLS y la restante sirve para testear dicho modelo. La utilización de este proceso, que se repite 88 veces (tantas como

medidas) permite explotar al máximo la información que se puede obtener de un conjunto de medidas reducido. Hay que recalcar que para cada modelo PLS creado los datos del conjunto de entrenamiento son autoescalados, mientras que la medida restante es centrada y escalada con el valor medio y la varianza del conjunto de entrenamiento. Junto a los parámetros típicos que se pueden extraer de los sensores, se suministro al modelo información de humedad relativa en el interior de la cámara así como el peso y superficie medias de las piezas utilizadas en cada medida.

Indicador de calidad	Ssq	Coef. Corr	m	b	Lv
Firmeza	0.1262	0.93	0.91	0	18
Índice de almidón	0.6	0.68	0.63	0.02	27
PH	0.29	0.84	0.78	0	12

Tabla 5.7: Resultados de predicción en indicadores de calidad

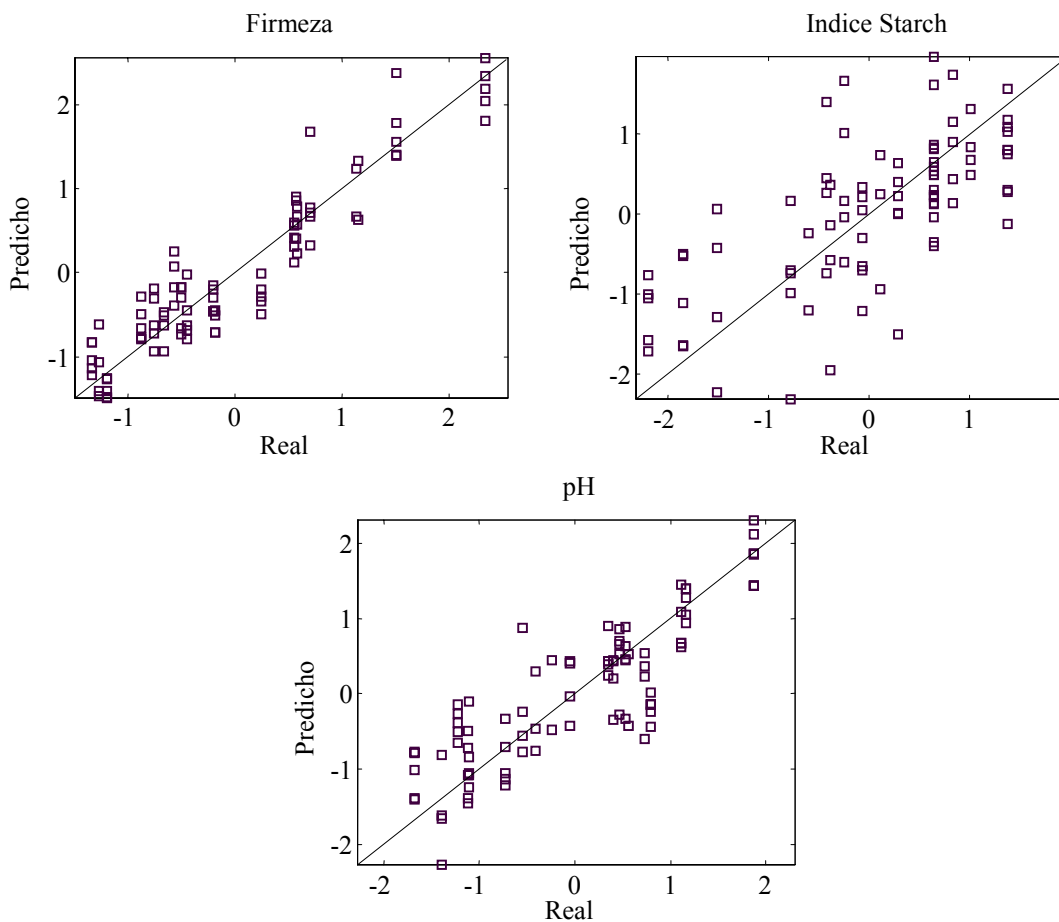


Figura 5.5: Gráficas de predicciones para los parámetros destructivos

La tabla 5.7 muestra el error cuadrático medio (S_{sq}), el coeficiente de correlación, el número óptimo de variables latentes (lv), la pendiente (m) y la constante de corte (b) para cada uno de los tres parámetros destructivos predichos por el sistema. Se puede observar como la firmeza es el parámetro destructivo mejor predicho. La figura 5.5 muestra las predicciones realizadas por la nariz comparando el valor real de cada medida con el predicho por el modelo. Cabe recordar que una predicción ideal dejaría todos los puntos sobre la diagonal, donde el valor real y el predicho son exactamente iguales. Es importante aclarar que tanto los valores predichos como los reales han sido autoescalados previamente, por lo que no debe extrañar encontrar, por ejemplo, valores negativos para la firmeza.

Como la mayoría de los indicadores de calidad presentan una evolución monótona (creciente o decreciente) a lo largo del tiempo, es importante comprobar si la deriva de algún sensor es utilizada por el modelo PLS para realizar predicciones ajustadas de alguno de los parámetros. Estas comprobaciones (tal y como se detallan en el siguiente subapartado) confirmaron que la deriva de los sensores no alteró el proceso de correlación a lo largo de la duración de todo el estudio.

5.3.5 Estudio de las derivas de los sensores

Tal y como se ha comentado previamente, las derivas en los sensores pueden ser un factor decisivo que debe ser monitorizado cuidadosamente, ya que el proceso de maduración está íntimamente ligado al paso del tiempo. En la rutina de medida de las manzanas se tomó la precaución de realizar medidas de calibración de forma sistemática para determinar qué sensores presentan derivas en el intervalo de tiempo que abarca el estudio. Solo de esta forma se puede estar seguro de si los resultados obtenidos son fruto de la detección del proceso de maduración de la fruta o son provocados por una deriva intrínseca en la señal de los sensores.

Al final de cada jornada, durante todo el tiempo que duró el estudio, se inyectó un microlitro de etanol en la cámara de concentración (la misma cámara en la que se depositaba la fruta) a través de un septum habilitado para este fin. Como la inyección se realizaba con una jeringa cromatográfica, el etanol líquido se volatilizaba prácticamente

al instante al ser inyectado. A partir de ese momento se activaba el prototipo para realizar una medida estándar en la que únicamente cambiaba el tiempo de concentración, que quedaba reducido a un minuto, ya que no era necesario esperar a que la fruta concentrara suficientes volátiles. En definitiva, desde el día de la recolecta al final del estudio se realizaron medidas de calibración siguiendo siempre un procedimiento idéntico. Después, junto con las medidas de fruta, se aplicaron diversos algoritmos de reconocimiento de patrones para evaluar las derivas de los sensores y decidir si influían o no en los resultados obtenidos.

Si recordamos el análisis PCA realizado a las medidas de fruta, sólo las proyecciones realizadas con parámetros de los sensores FIS-SB presentaban agrupaciones de las medidas acorde con el día de la medición. Para comprobar si las derivas habían influenciado en la proyección, las medidas de calibración se incluyeron en el análisis PCA de dos formas diferentes. Las figuras 5.6 y 5.7 muestran las proyecciones cuando se fusionaron todas las medidas siguiendo dos métodos diferentes. En estas figuras, cada medida está representada por una leyenda que distingue entre medidas de fruta y de calibración (cadena alfanumérica que empieza por una “d” o “cal” respectivamente) y que especifica el número de días transcurridos desde la cosecha. Con algunas excepciones (debidas a contratiempos en alguna medida) , la mayoría de jornadas incluyen 5 medidas de fruta (2 para el grupo 1, una para los grupos 2, 3 y 4) y una adicional de calibración.

La figura 5.6 muestra la proyección resultante en la que las medidas de calibración y de fruta se utilizan conjuntamente desde el principio para calcular las componentes principales del análisis. Como se puede observar, las medidas de fruta y calibración están claramente separadas, pero una inspección más detallada revela que la distribución de clusters dentro de el área en la que están situadas las medidas de calibración es muy parecida a la distribución por días de las medidas de fruta, tal y como se indica en la misma figura. En definitiva, este resultado no aclara si las derivas son o no el motivo de las agrupaciones observadas entre las medidas de fruta. Por ese motivo se decidió aplicar el algoritmo PCA de forma supervisada.

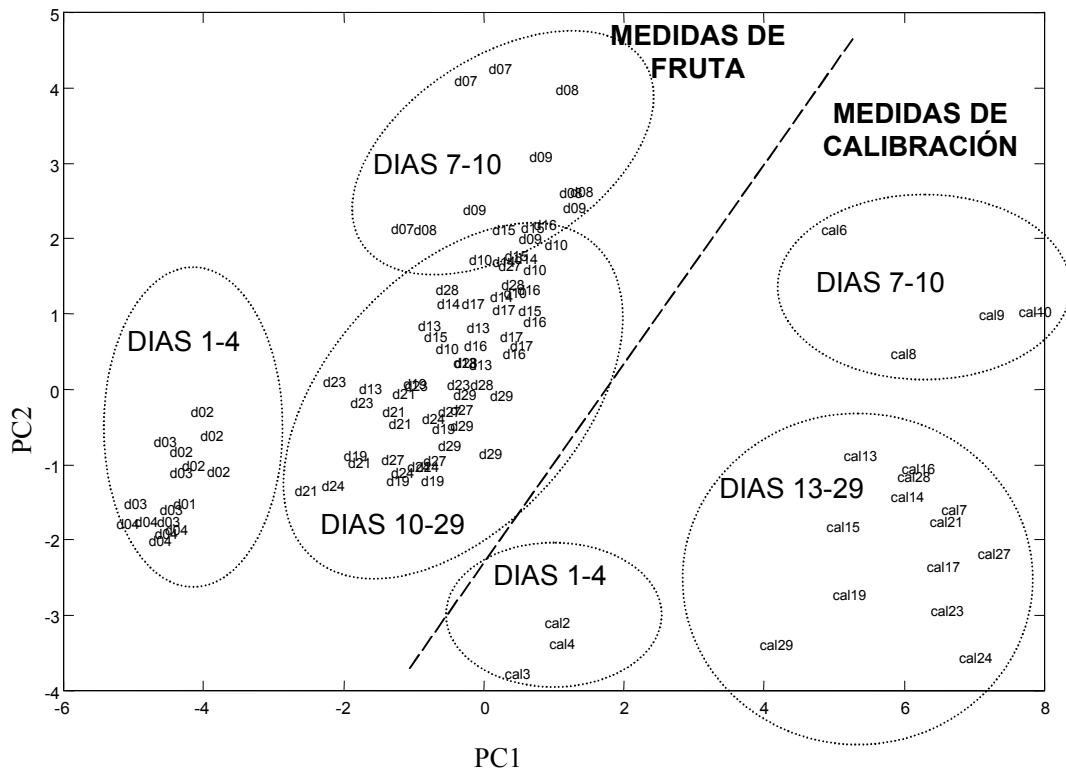


Figura 5.6: PCA con las medidas de fruta y calibración en modo no supervisado

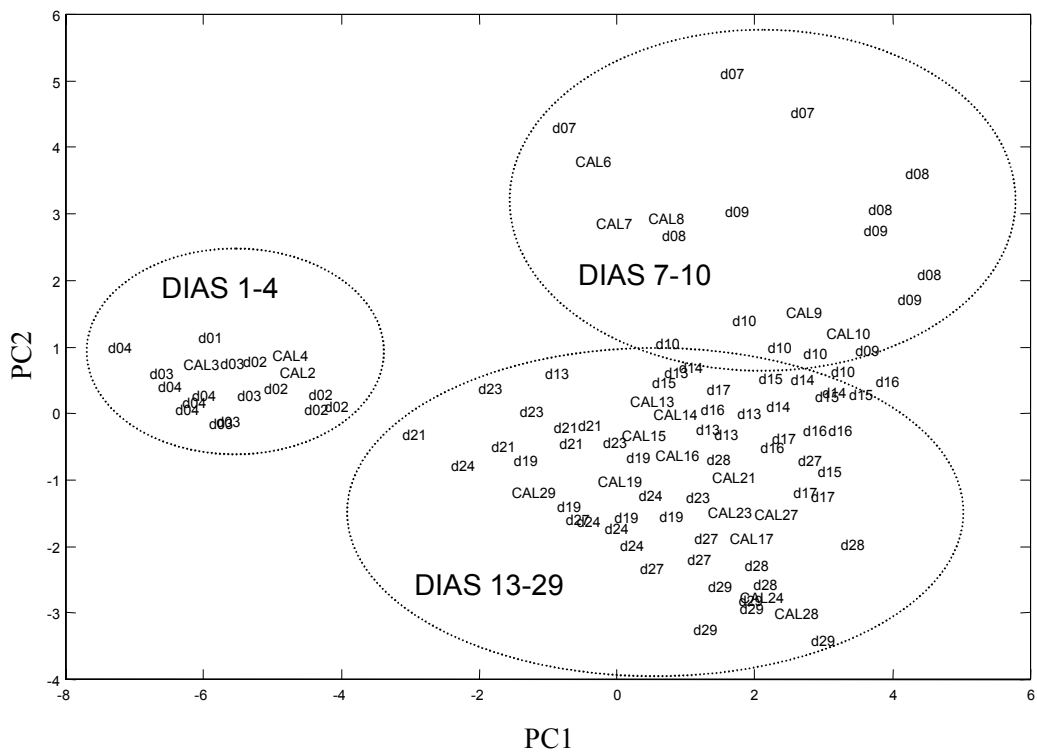


Figura 5.7: PCA de las medidas de fruta y calibración en modo supervisado

La figura 5.7 muestra las proyecciones resultantes de aplicar el PCA según la segunda opción. Este método se basa en determinar las componentes principales utilizando única y exclusivamente las medidas de fruta. Una vez determinadas las direcciones de máxima varianza, se proyectan tanto las medidas con fruta como las de calibración bajo las dos primeras componentes principales. Como se puede observar, las medidas de calibración realizadas el mismo día que las de fruta comparten una misma agrupación.

Este resultado es una clara indicación de que las derivas de los sensores son las responsables de que aparezcan agrupaciones de medidas según los días transcurridos desde la cosecha. Por lo tanto, se puede concluir que los sensores FIS-SB experimentan derivas que falsean los resultados obtenidos con el algoritmo PCA.

Clase	Medidas
1	g1d01 - g1d02 - g1d02 - g2d02 - g3d02 - g4d02 - g1d03 - g1d03 - g2d03 - g3d03 - g4d03 - g1d04 - g1d04 - g2d04 - g3d04 - g4d04 - g1d07 - g1d07 - g2d07 - g1d08 - g1d13
2	CAL2- - CAL3- - CAL4- - - CAL6- - - CAL7
3	g1d08 - g3d08 - g4d08 - g1d09 - g1d09 - g2d09 - g4d09 - g1d10 - g1d10 - g2d10 - g3d10 - g4d10 - g1d13 - g2d13 - g3d13 - g4d13 - g1d14 - g1d14 - g2d14 - g3d14 - g1d15 - g1d15 - g2d15 - g3d15 - g4d15 - g1d16 - g1d16 - g2d16 - g3d16 - g4d16 - g1d17 - g1d17 - g4d17 - g1d19 - g1d19 - g3d19 - g4d19 - g2d21 - g1d23
4	CAL8- - CAL9- - CAL10 - CAL13 - CAL14 - CAL15 - CAL16
5	g3d17 - g2d19 - g1d21 - g1d21 - g3d21 - g4d21 - g1d23 - g2d23 - g4d23 - g1d24 - g1d24 - g2d24 - g3d24 - g4d24 - g1d27 - g1d27 - g2d27 - g3d27 - g4d27 - g1d28 - g1d28 - g2d28 - g3d28 - g1d29 - g2d29
6	CAL17 - CAL19 - CAL21 - CAL23 - CAL24 - CAL27 - CAL28 - CAL29
7	g1d29 - g3d29 - g4d29

Tabla 5.8: Clasificación Fuzzy Art utilizando sensores FIS-SB incluyendo medidas de calibración

Para comprobar si los resultados obtenidos con Fuzzy Art también estaban influenciados por las derivas, el algoritmo de clasificación neuronal se utilizó con las medidas de fruta y calibración de forma conjunta. Las tablas 5.8, 5.9 y 5.10 muestran los resultados obtenidos para los sensores FIS-SB, Taguchi y FIS-SP respectivamente. Una vez más, las medidas de calibración están identificadas por una etiqueta que comienza con las siglas CAL y cuyo número identifica el día de la experiencia. La

identificación de las medidas de fruta sigue el mismo criterio ya presentado en la tabla 5.4.

Clase	Medidas
1	g1d01 - g1d02 - g1d02 - g2d02 - g3d02 - g4d02 - g1d03 - g1d03 - g2d03 - g1d04 - g1d04 - g2d07 - g4d09 - g1d10 - g2d10 - g4d10 - g1d13 - g2d13 - g4d13 - g4d15 - g1d16 - g1d19 - g2d19 - g3d19
2	CAL2- - CAL3- - CAL4- - CAL6- - CAL9- - CAL21 - CAL23
3	g3d03 - g4d03 - g2d04 - g3d04 - g4d04 - g1d27 - g1d29 - g1d29
4	g1d07 - g1d07 - g1d08 - g1d08 - g3d08 - g4d08 - g1d09 - g1d09 - g2d09 - g1d10 - g3d10 - g1d13 - g3d13 - g1d14 - g1d14 - g2d14 - g3d14 - g1d15 - g1d15 - g2d15 - g3d15 - g1d16 - g2d16 - g3d16 - g4d16 - g1d17 - g1d17 - g3d17 - g4d17 - g1d19 - g4d19 - g1d21 - g1d21 - g2d21 - g3d21 - g4d21 - g1d23 - g1d23 - g2d23 - g4d23 - g1d24 - g1d24 - g2d24 - g3d24 - g4d24 - g1d27 - g2d27 - g3d27 - g4d27 - g1d28 - g1d28 - g2d28 - g3d28 - g3d29
5	CAL7- - CAL8- - CAL10 - CAL13 - CAL14 - CAL15 - CAL16 - CAL17 - CAL19 - CAL27 - CAL28
6	CAL24
7	g2d29 - g4d29
8	CAL29

Tabla 5.9: Clasificación Fuzzy Art utilizando sensores Taguchi incluyendo medidas de calibración

Clase	Medidas
1	G1d01 - g1d02 - g1d02 - g2d02 - g3d02 - g4d02 - g1d03 - g1d03 - g2d03 - g3d03 - g4d03 - g1d04 - g1d04 - g4d04 - g1d07 - g1d07 - g2d07 - g1d08 - g1d08 - g3d08 - g4d08 - g2d09 - g1d13 - g1d13 - g4d13 - g1d14 - g3d14 - g1d19 - g1d24 - g1d28
2	CAL2- - CAL3- - CAL4- - CAL6- - CAL7- - CAL8- - CAL9- - CAL10 - CAL13 - CAL14 - CAL15 - CAL16 - CAL17 - CAL21 - CAL23 - CAL24 - CAL27 - CAL29
3	G2d04 - g3d04 - g1d09 - g1d09 - g4d09 - g1d10 - g1d10 - g2d10 - g3d10 - g4d10 - g2d13 - g1d14 - g2d14 - g1d15 - g1d15 - g2d15 - g1d16 - g1d16 - g2d16 - g4d16 - g1d17 - g1d17 - g3d17 - g1d19 - g1d21 - g2d21 - g3d21 - g1d23 - g4d23 - g1d24 - g2d24 - g3d24 - g4d24 - g1d27 - g2d27 - g3d27 - g1d28 - g1d29 - g1d29 - g4d29
4	G3d13 - g3d15 - g4d15 - g4d17 - g2d19 - g3d19 - g4d19 - g4d21 - g1d23 - g2d23 - g1d27 - g4d27 - g3d28
5	G3d16 - g1d21 - g2d28 - g2d29 - g3d29
6	CAL19 - CAL28

Tabla 5.10: Clasificación Fuzzy Art utilizando sensores FIS-SP incluyendo medidas de calibración

De las tablas se puede concluir que el algoritmo Fuzzy Art separa perfectamente las medidas de calibración de las de fruta para las tres familias de sensores. Por otro lado, fijándonos en la clasificación de las medidas de calibración, se puede ver claramente que los sensores FIS-SB las distribuyen en tres categorías según el día en que se realizaron. Por el contrario, tanto en el caso de los sensores Taguchi como de los FIS-SP se puede comprobar que no se agrupan siguiendo el orden de medida, una clara indicación de que no existe una deriva temporal en estas dos familias de sensores. Por ese motivo sólo se utilizaron estos sensores en el proceso de correlación.

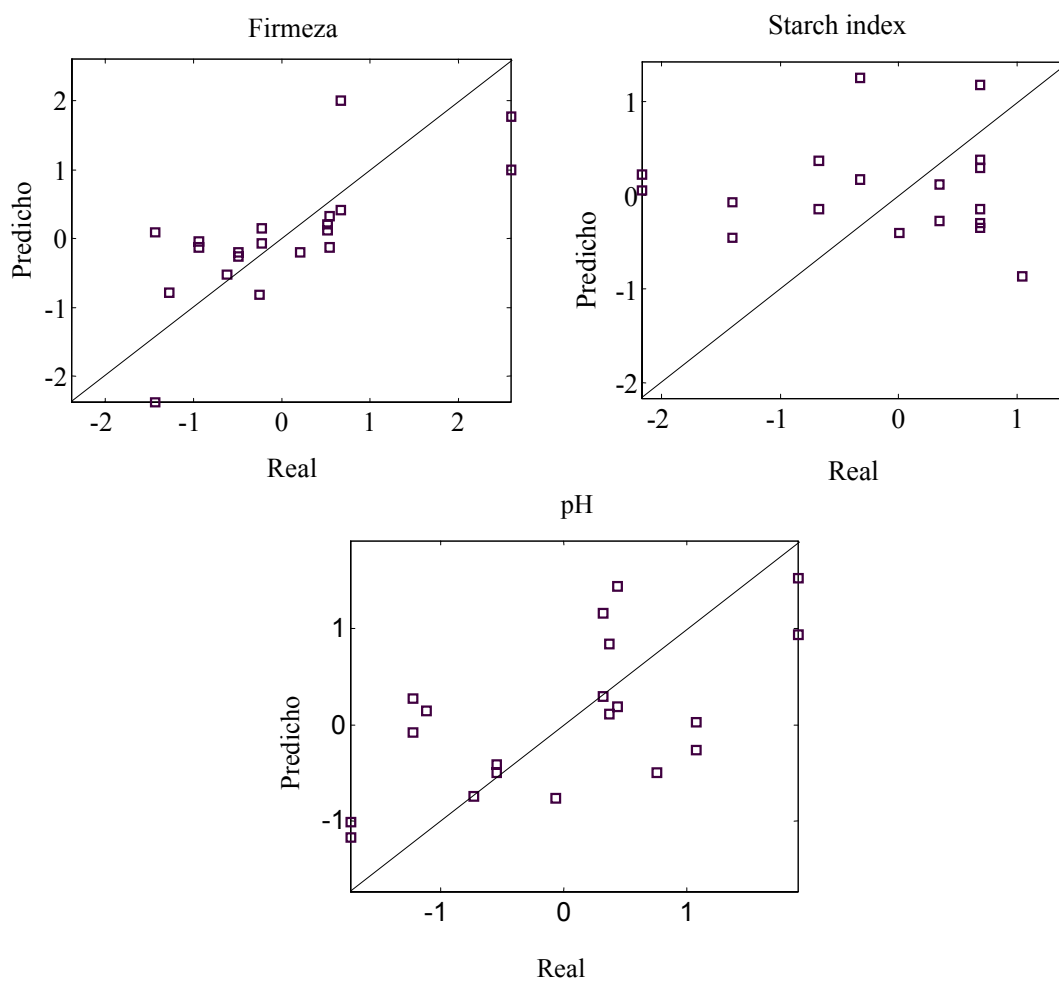


Figura 5.8: Predicciones utilizando las medidas de calibración

Finalmente, para comprobar si las predicciones realizadas con estas dos familias de sensores podían estar influenciadas por derivas, se realizaron predicciones con modelos PLS utilizando únicamente las medidas de calibración. Si las predicciones hubiesen

resultado acertadas, eso sería una clara indicación de que las predicciones que realiza el sistema de olfato electrónico son debidas a las derivas de los sensores y no a la interacción de éstos con la sucesiva evolución de los vapores emitidos por la fruta a lo largo de su proceso de maduración. Si las predicciones son hechas de forma aleatoria significaría que el éxito obtenido previamente en las predicciones de los parámetros destructivos son debidas exclusivamente a la información captada por los sensores en las sucesivas medidas realizadas con la fruta.

La figura 5.8 muestra los resultados de predicción obtenidos utilizando las medidas de calibración como la información de entrada para el algoritmo PLS. Como se puede observar, el modelo PLS realiza predicciones de forma totalmente aleatoria y por lo tanto inexacta, prueba de que la exactitud en las predicciones realizadas con las medidas de fruta son debidas a la evolución de la producción aromática a lo largo del proceso de maduración y no a la deriva de los sensores.

5.3.6 Conclusiones

Una vez más, las señales obtenidas con el segundo prototipo son suficientemente claras para que el sistema detecte la presencia de manzanas en su cámara de concentración. Sin embargo, la evolución de la señal a lo largo de dicho proceso sigue siendo débil y no es aprovechada por los algoritmos no supervisados como el PCA o Fuzzy Art para monitorizar el proceso de maduración de forma clara. Se tiene que recurrir de nuevo a un método supervisado (en este caso un algoritmo PLS) para poder obtener pruebas de que de una nariz electrónica puede monitorizar el proceso de maduración de las manzanas. La buena predicción de parámetros destructivos con las señales recogidas con la nariz electrónica (a través de medidas no destructivas) confirma la viabilidad de utilizar la nariz electrónica como instrumento de monitorización del proceso de maduración en manzanas.

Sin embargo, las interferencias como las derivas en los sensores y la variabilidad intrínseca a la fruta hace necesario un perfeccionamiento tanto de la metodología

experimental como de la tecnología en los sensores para que el sistema sea plenamente comercial.

Por otro lado , la detección de derivas importantes en los sensores FIS-SB ha permitido tomar la decisión de eliminarlos de los estudios con las variedades de pera y melocotón. Por contra, los sensores Taguchi y FIS-SP se seguirán utilizando tras confirmar que no presentan derivas apreciables que puedan falsear los resultados.

REFERENCIAS

[5.1] Brezmes J.; Llobet E.; Vilanova X.; Saiz G.; Correig X., *Fruit ripeness monitoring using an electronic nose, Sensors and Actuators B-Chemical*, (2000), vol 69, num 3, 223-229

[5.2] Brezmes J. ; Llobet E.; Vilanova X.;Saiz G.; Correig X.; *Correlation between electronic nose signals and fruit quality indicators on shelf-life measurements with pinklady apples*, Sensors and Actuators B-Chemical (2001) vol 80, 41-50

[5.3] Brezmes J; Llobet E; Vilanova X ; Saiz G; Correig X., *Evaluation of the electronic nose as a novel instrument to assess fruit ripeness*, Eurodeur-airodeur, Paris, Junio 2001

[5.4] Brezmes J; Llobet E; Vilanova; López M.L.; Recasens I, X ; Saiz G; Correig X., *Design and evaluation of an electronic nose to assess fruit ripeness*, IEEE Sensors (enviado)

CAPITULO 6: **MEDIDAS CON PERAS**

6.1 INTRODUCCIÓN

La segunda variedad de fruta que se estudió con los prototipos de olfato electrónico fueron peras. Una vez más, la primera temporada se realizó un estudio de viabilidad con el primer prototipo, mientras que en el segundo año se realizó un estudio más completo con el prototipo de flujo continuo. Tanto la metodología como los objetivos perseguidos en el estudio preliminar son similares a los expuestos para las manzanas. Como complemento a esta memoria, el lector puede consultar en [6.1] la descripción de las experiencias de la primera temporada y en [6.2,3] se pueden encontrar los resultados obtenidos con el sistema de flujo continuo.

Los estudios del primer año se centraron en comprobar el nivel de señal que generaban los volátiles emitidos por las peras así como si se detectaba una evolución a lo largo del proceso de maduración. El segundo año se realizó un estudio de la madurez de las piezas tras diferentes intervalos de frigoconservación. El estudio con parámetros de calidad fue mucho más completo que con las manzanas, ya que, por primera vez en este estudio, se identificaron y cuantificaron volátiles producidos por la fruta mediante cromatografía de gases. Este estudio paralelo permitió comparar la eficacia de la nariz electrónica frente a un método con tanta resolución como la cromatografía de gases.

6.2 ESTUDIO DE VIABILIDAD

6.2.1 Objetivos

El objetivo primordial del estudio preliminar era comprobar si una nariz electrónica podía identificar el estado de maduración de las peras a lo largo de su proceso de maduración en condiciones ambientales (maduración de “shelf life”). Para ello es imprescindible que la concentración de volátiles de la fruta provoque en la matriz de sensores una respuesta que evolucione medida que avanza el estado de maduración de las muestras medidas con el instrumento.

Al igual que en el caso de las manzanas, tras la medición con la nariz electrónica un operario clasifica las muestras en tres estados posibles de maduración y se compara dicha clasificación con la que realiza el prototipo de olfato electrónico. Una vez más, es importante recalcar que la determinación del estado de maduración por parte del operario sigue un criterio subjetivo, por lo que los resultados indican la coincidencia entre el operario y el instrumento.

6.2.2 Metodología experimental

El estudio de viabilidad con peras se realizó con muestras de la variedad “Blanquilla” obtenidas de una cooperativa de Mollerusa, en la provincia de Lleida. Se cogieron 60 muestras que habían sido recolectadas el 22 de agosto. Estas muestras se dividieron en 3 grupos de 20 piezas cada uno. Los dos primeros grupos se empezaron a medir con el prototipo el 28 de setiembre, mientras que el grupo 3 se conservó en una cámara frigorífica a una temperatura de 0.5°C aproximadamente desde el 28 de setiembre hasta el 26 de octubre, día en el que se empezaron a medir. En total se midieron 60 peras y se realizaron 39 experimentos con la nariz electrónica.

Las peras estudiadas llevaban un mes en cámara frigorífica cuando se seleccionaron para el estudio. Las condiciones de conservación fueron una temperatura de 0.5°C y una

atmósfera controlada con unos porcentajes de oxígeno y de anhídrido carbónico del 2.5% en ambos casos.

El proceso que se siguió para realizar medidas con la nariz electrónica y con técnicas destructivas de forma paralela fue el mismo que para las manzanas. En este caso, cada grupo de 20 peras fue subdividido en dos subgrupos (el de medida y el de reserva) con diez piezas cada uno de ellos.

Tras unas primeras pruebas para comprobar el nivel de señal, se constató que los aromas producidos por las peras provocaban señales muy claras en los sensores, mucho más fuertes que las de las manzanas, por lo que se consideró suficiente concentrar la fruta una hora e inyectar 50 ml de la atmósfera de la cámara de concentración en la de medida de una sola vez.

Salvo restricciones debidas a los recursos disponibles, se intentó realizar una medida diaria a cada uno de los grupos de peras. Así, al principio, en cada jornada de medición se realizaron dos medidas diarias (una para el grupo 1 y otra para el grupo 2). Cuando se acabaron estas medidas se procedió a medir el grupo 3 en solitario. Al igual que con las manzanas, en cada medida se pesaba todo el grupo y se obtenía el perímetro medio de las piezas. Además, las medidas sucesivas con nariz electrónica se espaciaron una hora para dar tiempo a los sensores a recuperar su estado inicial. La tabla 6.1 muestra el número total de medidas que se hicieron para cada grupo.

Grupo	Número de medidas
1	14
2	14
3	11

Tabla 6.1: Número de medidas realizadas con cada grupo

Las técnicas destructivas aplicadas fueron las mismas que para las manzanas: la penetromía, acidez y sólidos solubles. Para poder monitorizar el proceso de maduración durante el “shelf life” de la fruta (maduración a 20° y humedad relativa ambiente) por

un período de unos 20 días, las medidas destructivas se practicaron aproximadamente cada dos días. Solo en estas jornadas se realizó la sustitución de la muestra destruida por su pareja de reserva.

6.2.3 Evolución de la señal

La figura 6.1 muestra la evolución de la señal de los sensores para cada uno de los tres grupos de medida. En ella se muestra el parámetro ΔG para cada sensor en cada una de las medidas realizadas con nariz electrónica. En esta gráfica se puede comprobar que la evolución es clara para todos los grupos. Además, recordando que el tercero fue medido con posterioridad (el día 1 de medida de este tercer grupo es posterior a la última medida de los grupos 1 y 2) se puede observar que esta evolución no puede ser imputada a derivas temporales en los sensores.

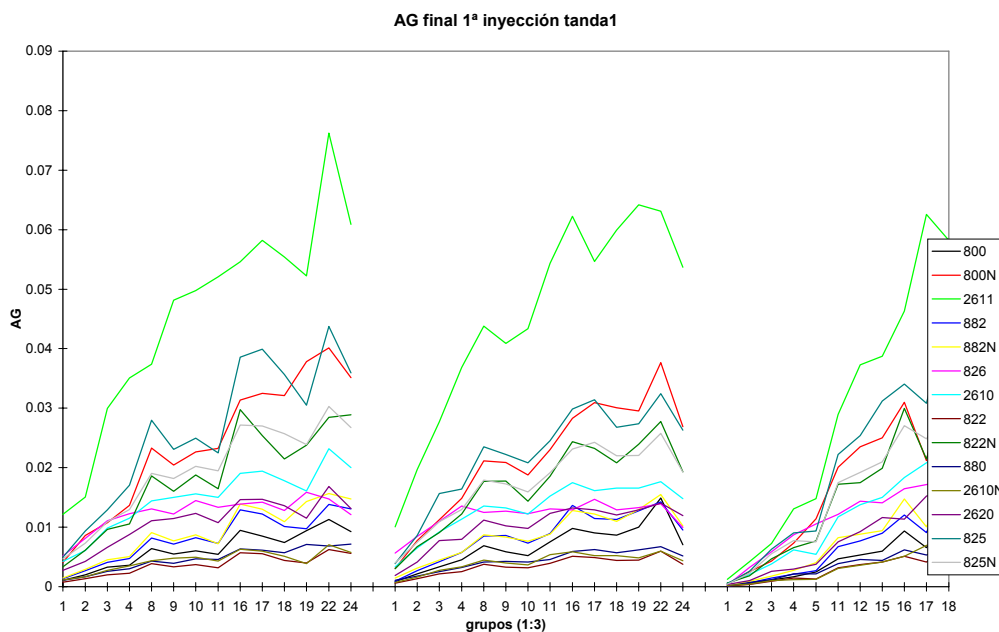


Figura 6.1: Evolución de la señal para los 3 grupos de peras

6.2.4 Resultados

La figura 6.1 ya nos permite intuir que los resultados pueden ser buenos ya que la matriz de sensores parece seguir la evolución del proceso de maduración de las piezas medidas. Al igual que con las manzanas, la primera aproximación consistió en plasmar en un diagrama PCA las 48 medidas y observar si en alguna de las componentes

principales se observaba alguna tendencia a destacar. La figura 6.2 muestra dicho diagrama en el que cada medida ha sido etiquetada con su número de orden. Cada color representa un grupo diferente, siendo el grupo 1 de color rojo, el grupo 2 de color azul y el 3 de color añil.

Los datos de entrada para realizar el PCA fueron el incremento de conductancia de todos los sensores. En esta figura se puede observar una clara evolución de las medidas de izquierda a derecha, de forma que la variación de la señal queda concentrada sobre la primera componente principal. Esto, una vez más, nos indicará que el seguimiento del proceso de maduración se puede hacer con pocos sensores. La proyección de componentes principales de menor importancia no aportan información interesante.

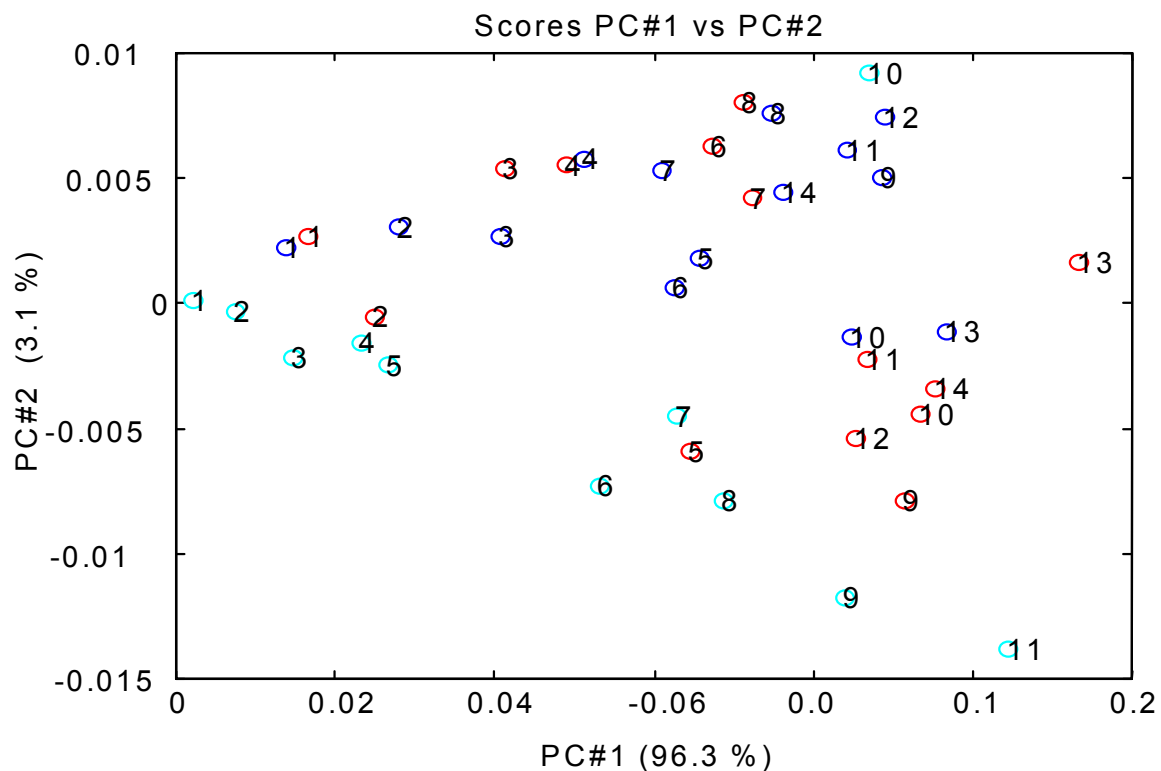


Figura 6.2: Diagrama PCA de las dos primeras componentes principales de los 3 grupos de peras.

Para comprobar si era posible realizar una clasificación del estado de maduración de las piezas se utilizó una red neuronal. El objetivo era el mismo que para las manzanas, clasificar cada una de las medidas como fruta verde, madura o sobremadura. El tipo de

red neuronal fue el mismo que el descrito en la figura 5.4. Siguiendo el criterio de sensibilidad descrito en el capítulo 4, se escogieron los sensores que detalla la tabla 6.2 .

Sensores
800N
2610
825
825N

Tabla 6.2: Sensores escogidos para la red neuronal

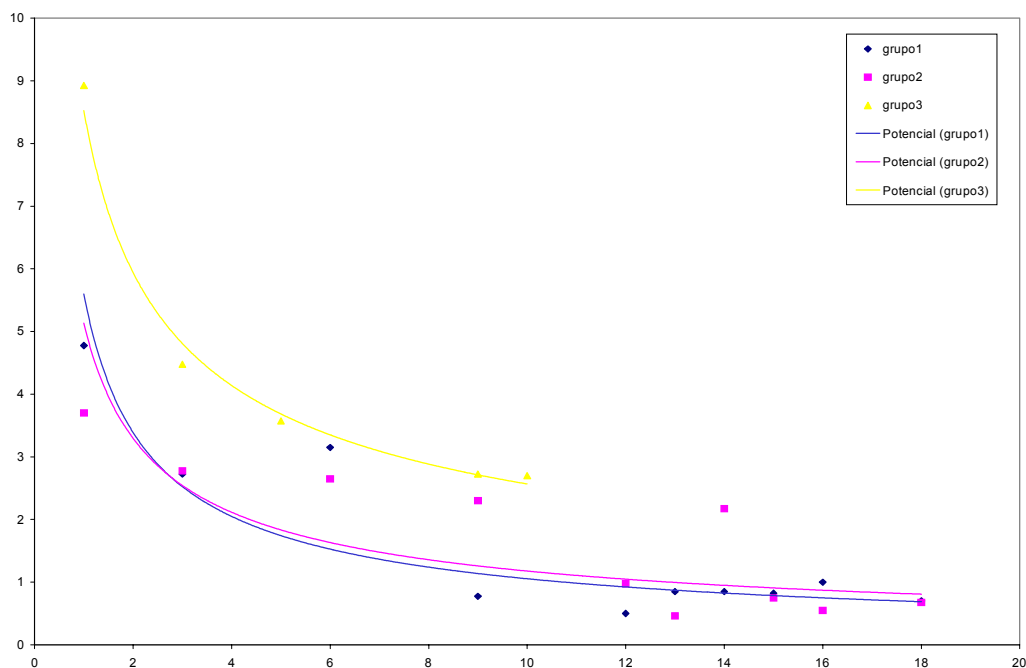


Figura 6.3: Evolución de la firmeza para los tres grupos de peras

La figura 6.3 muestra los valores de penetromía interpolados para los tres grupos frente a los días de shelf life, en los que se puede comprobar una evolución similar para los grupos 1 y 2 y otra ligeramente diferente para el grupo 3. Como la penetromía se realizó cada dos jornadas de mediciones, los valores originales están representados por puntos y la interpolación de la tendencia por una línea sólida.

La determinación de la clasificación correcta volvió a realizarse de forma subjetiva por un operario, aunque en este segundo intento la firmeza fue el factor determinante para

esta clasificación. La tabla 6.3 muestra los valores que el operario finalmente fijó arbitrariamente como frontera entre los tres estados de maduración. Los sólidos solubles y la acidez no aportaron información útil ya que las tendencias eran muy erráticas. Los casos dudosos fueron clasificados a través del aspecto externo de la fruta y la experiencia del operario.

	Verde	Madura	Sobremadura
Peras	$p > 2 \text{ kg.}$	$1 \text{ kg.} < p < 2 \text{ kg.}$	$p < 1 \text{ kg.}$

Tabla 6.3: Valores frontera de penetromía para cada grado de maduración

El proceso de entreno y evaluación se realizó a través de una validación cruzada de orden 3, de forma que se hicieron tres iteraciones en las que se entrenaba con todas las medidas de dos de los tres grupos de peras y se evaluaba con las medidas del grupo restante. La tabla 6.4 muestra los resultados en los que se puede comprobar el alto nivel de éxito conseguido. Las diferentes combinaciones de parámetros de entrada consideradas incluyeron los valores del peso medio y superficie y los parámetros G_i , G_f y ΔG de los sensores mostrados en la tabla 6.2.

Variabes	n° sensores	Sensores	n° inputs	Neuronas ocultas	Aciertos/ Medidas	% ok
G_o, G_f, p, s	4	800N, 2610, 825, 825N	10	6	34/39	87.17
$\Delta G, p, s$	4	800N, 2610, 825, 825N	6	2, 4, 6	35, 32, 31/39	89.74
$\Delta G, p, s$	1	2610	3	2, 4, 6	35, 36, 35/39	92.3
$\Delta G, p, s$	2	800N, 2610	4	2, 4, 6	33, 35, 34/39	89.74
$\Delta G, p, s$	2	800N, 825	4	2, 4, 6	34, 37, 36/39	<u>94.87</u>
$\Delta G, p, s$	2	800N, 825N	4	2, 4, 6	33, 32, 30/39	84.61
$\Delta G, p, s$	2	2610, 825	4	2, 4, 6	36, 32, 37/39	<u>94.87</u>
$\Delta G, p, s$	2	2610, 825N	4	2, 4, 6	34, 36, 34/39	92.3
$\Delta G, p, s$	2	825, 825N	4	2, 4, 6	36, 37, 34/39	<u>94.87</u>

Tabla 6.4: Aciertos con la red neuronal para las 39 medidas con peras

Con un acierto del 94% ya no se hicieron pruebas utilizando solamente los grupos 1 y 2 de peras. Además, el acertar 37 de 39 medidas se da para diferentes combinaciones de sensores, lo que indica que no es crítico el utilizar un modelo concreto de sensor, ya que hay varios de ellos que siguen correctamente el proceso de maduración de esta variedad de fruta. Sin embargo, en todas las combinaciones exitosas aparecen los parámetros de peso y superficie medios.

6.2.5 Conclusiones

Con unos resultados tan claros se puede afirmar rotundamente que la nariz electrónica puede seguir con facilidad el proceso de maduración de esta variedad de peras. Los resultados óptimos se alcanzan con diferentes combinaciones de sensores en las que el peso y superficie son parámetros siempre presentes. Por todo ello, con el segundo prototipo se perseguirán objetivos con una segunda variedad de peras que permitan corroborar los resultados obtenidos en este estudio preliminar.

6.3 MEDIDAS EN FLUJO CONTINUO

6.3.1 Objetivos

Debido a los buenos resultados obtenidos con el primer prototipo, los objetivos para la segunda temporada fueron más ambiciosos, ya que se utilizó la nariz electrónica para monitorizar los efectos de diversos factores en el proceso de maduración de las peras, comparando la eficacia del segundo prototipo con la de otras técnicas de control de calidad en fruta.

Por lo tanto, el objetivo primordial de esta segunda experiencia con peras fue el comparar la eficacia de las mediciones con nariz electrónica frente a la mayoría de las técnicas que se utilizan para monitorizar la calidad de las peras. De entre ellas, la más novedosa y completa consiste en caracterizar su perfil aromático a lo largo de su

maduración mediante cromatografía de gases y muestreo por espacio de cabeza dinámico.

Las comparaciones se realizaron al monitorizar el proceso de maduración de las peras tras diferentes intervalos de frigoconservación, diferentes fechas de recolecta y número de días de maduración “en estantería”. Mediante técnicas de reconocimiento de patrones se realizaron clasificaciones no supervisadas para ver como afectaban estos tres factores y después se realizaron correlaciones entre las señales de los sensores y los parámetros extraídos con las técnicas de calidad.

6.3.2 Metodología experimental

Para el estudio se utilizaron peras de la variedad 'Doyenne du Comice' (*Pyrus communis* L). Estas peras se cosecharon en una finca de la provincia de Lleida. En cada recolecta se cogieron 500 muestras, la mitad de las cuales se destinaron al laboratorio de postcosecha en la Universidad de Lleida y la otra mitad se transportaron al Laboratorio de Sensores de Gases y Nariz Electrónica en Tarragona. Se realizaron cinco recolectas espaciadas una semana entre ellas, desde el 23 de agosto hasta el 8 de setiembre.

Las muestras de cada recolecta fueron divididas en cinco grupos en cada uno de los laboratorios. El primer grupo se utilizó para realizar un seguimiento de “shelf life” tras cada recolecta, mientras que los cuatro grupos restantes se almacenaron en una cámara de frigoconservación. Tanto en Lleida como en Tarragona las piezas se almacenaron a la misma temperatura (1°C) y humedad relativa (90%). Cada uno de los grupos se almacenó un período de tiempo diferente, desde una semana para el grupo 2 hasta 4 para el grupo 5. El objetivo era estudiar el proceso de maduración de la fruta a temperatura ambiente en función del número de semanas de frigoconservación.

Las medidas del proceso de maduración a temperatura ambiente (“shelf life”) fueron realizadas tras uno, cuatro y siete días después de la recolección (para el grupo 1) o salida de la cámara de frigoconservación (para los grupos 2 a 5). Como se planificaron 5

recolectas, de cada una se separaron cinco grupos y cada uno de ellos fue medido en tres ocasiones (1, 4 y 7 días de “shelf life”), se realizaron un total de 75 medidas con nariz electrónica en Tarragona y 75 medidas con técnicas de control de calidad de fruta en Lleida. En el proceso de correlación, las medidas realizadas con nariz electrónica fueron emparejadas con las medidas de calidad ejecutadas en Lleida. Aunque las muestras eran físicamente diferentes, las parejas estaban formadas por piezas del mismo cultivo, recolectadas el mismo día y con idénticos períodos de frigoconservación y de “shelf life”. Las medidas realizadas en Lleida incluyeron técnicas clásicas y otras más novedosas. Las técnicas físico-químicas aplicadas fueron la acidez, la penetromía, los sólidos solubles, el índice de almidón y la colorimetría.

La utilización de un cromatógrafo permitió realizar medidas sobre la producción de volátiles que la fruta genera durante su maduración. Dicho instrumento permitió realizar mediciones sobre la identidad y cantidad de 14 volátiles aromáticos además de etileno y CO₂. En total, de cada medida de calidad se obtuvieron 21 indicadores. La tabla 6.9 (que se encuentra en la página 164) muestra, además de otros datos, los 14 aromas que fueron identificados y cuantificados cromatográficamente para las 75 medidas realizadas en Lleida.

6.3.3 Clasificación no supervisada

Con el fin de determinar como afectan los factores mencionados al proceso de maduración, en primer lugar se realizó un análisis PCA con las 75 medidas. Para comparar la resolución de la nariz electrónica frente a las técnicas establecidas se realizaron las proyecciones partiendo de las señales de los sensores y de los indicadores de calidad.

La figura 6.4 muestra la proyección sobre las dos primeras componentes principales de las 75 medidas definidas por los 21 parámetros de calidad. En dicha gráfica cada medida es codificada a partir del color y de un número. Así, las medidas realizadas el día 1 de “shelf life” están rotuladas en color verde, las de 4 días de color azul y las de

siete en color rojo. A su vez, el primer dígito de cada etiqueta representa la recolecta y el segundo el grupo al que pertenece la fruta medida.

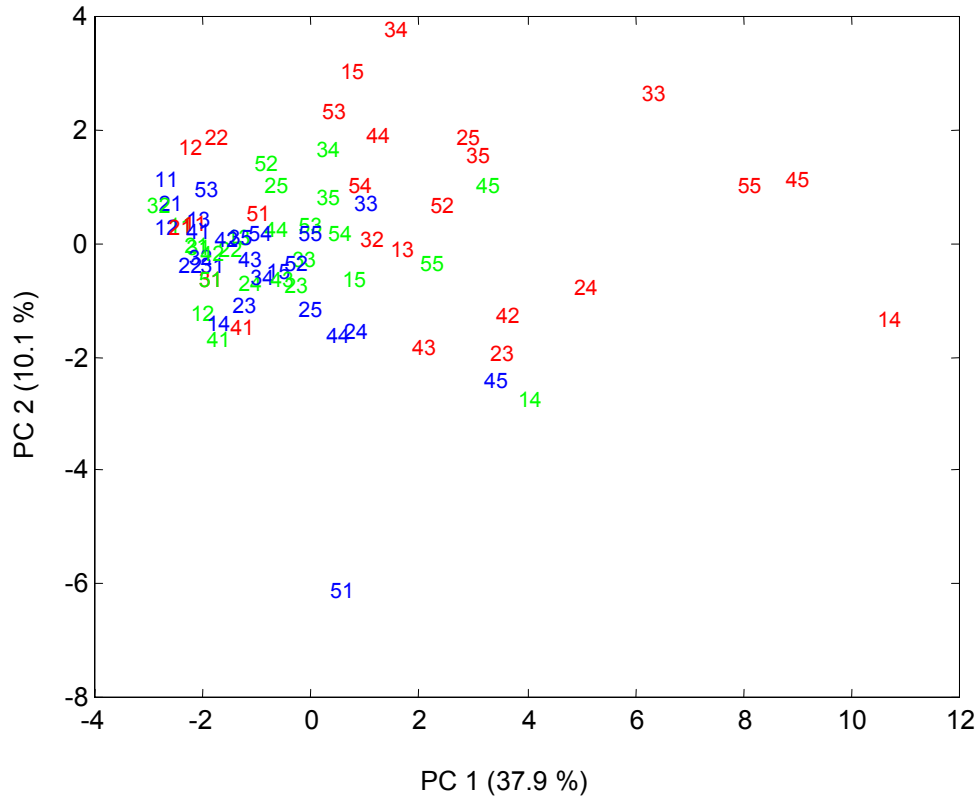


Figura 6.4: Diagrama PCA de las 75 medidas descritas por los 21 indicadores de calidad

Se puede observar que en la proyección obtenida con los indicadores de calidad no se puede definir ninguna agrupación clara. Lo único que se observa es la tendencia a que las medidas de uno o cuatro días de “shelf life” se mantengan a la izquierda y las de siete días se prolonguen hacia la derecha. Proyecciones con componentes principales de menor varianza dieron siempre resultados peores.

La proyección PCA de las medidas realizadas con la nariz electrónica se muestra en la figura 6.5. Se probaron diferentes combinaciones de parámetros y en dicha figura se muestra el mejor resultado considerando el parámetro “separación”. Como las medidas de “shelf life” se realizaron tras uno, cuatro y siete idas desde la salida de la cámara de concentración (o recolecta), se consideró que la agrupación natural debía seguir ese

criterio al definir dicho parámetro. La mejor resolución se consiguió con el parámetro ΔG para los sensores taguchi 800, 822, 882 y 825.

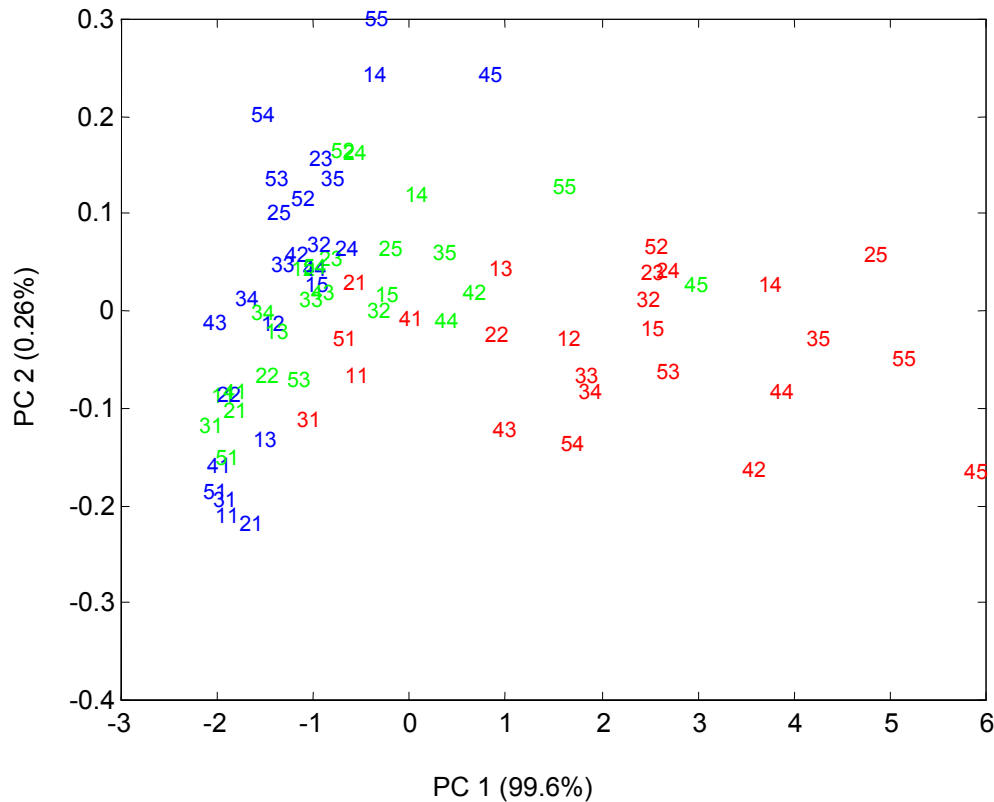


Figura 6.5: Proyección PCA de las medidas realizadas con la nariz electrónica

En la gráfica se puede observar que la mayoría de las medidas del séptimo día de shelf-life están bastante separadas de las medidas del primer y cuarto día. Solamente cinco medidas de fruta con siete días de “shelf life” están mezcladas con las medidas realizadas en días anteriores. Estos “outliers” pertenecen al primer grupo de cada recolecta, precisamente las muestras de fruta que no pasaron por el frigorífico. Un estudio realizado sobre esta variedad de peras llega a la conclusión de que para acelerar el proceso de maduración tras la recolecta es necesario un periodo de frigoconservación. Como la nariz electrónica sensa volátiles que están relacionados con dicho proceso estos “outliers” confirman las conclusiones de dicho estudio, ya que el sistema de olfato electrónico detecta que las peras que no han pasado por frigoconservación están en un estadio de maduración muy atrasado tras siete días de “shelf life”.

Aparte de estos “outliers” sólo dos muestras del día cuatro quedan proyectadas dentro de la agrupación del día 7 y una medida más está justo en la frontera. Las dos primeras pertenecen al quinto grupo (cuatro semanas de frigoconservación) de la penúltima y última recolecta, por lógica las muestras con un estado de maduración más avanzado tras cuatro días de shelf life por haber estado más tiempo en frigoconservación y haber sido recolectadas en las últimas fechas planeadas. Solo la medida de la frontera puede considerarse un verdadero “outlier” ya que pertenece al grupo dos y a la cosecha 4.

Buscando más tendencias en la gráfica se puede concluir que el factor que domina el proceso de maduración son los días de shelf life. Una mirada dentro del cluster “rojo” permite ver que medidas de un mismo grupo están más cerca entre sí que medidas de una misma cosecha, un comportamiento que permite deducir que el segundo factor en importancia es el tiempo de frigoconservación, mientras que el que afecta menos es la fecha de la recolecta. Estas conclusiones están en consonancia con las derivadas del estudio clásico con las muestras de Lleida.

Para eliminar toda identificación subjetiva de las agrupaciones en el diagrama PCA se aplicaron algoritmos de clasificación no supervisada Fuzzy Art. La tabla 6.5 muestra las clases a las que cada medida fue asignada de forma espontánea, sin aprendizaje previo.

Clase	Medidas
1	H1G1D1, H1G1D4, H1G1D7, H1G2D1, H1G2D4, H1G3D1, H1G3D4, H1G4D1, H1G4D4, H1G5D1, H1G5D4, H2G1D1, H2G1D4, H2G1D7, H2G2D1, H2G2D4, H2G3D1, H2G3D4, H2G4D1, H2G4D4, H2G5D1, H2G5D4, H3G1D1, H3G1D4, H3G1D7, H3G2D1, H3G2D4, H3G3D1, H3G3D4, H3G4D1, H3G4D4, H3G5D1, H3G5D4, H4G1D1, H4G1D4, H4G1D7, H4G2D1, H4G3D1, H4G3D4, H4G4D1, H4G4D4, H4G5D1, H5G1D1, H5G1D4, H5G1D7, H5G2D1, H5G2D4, H5G3D1, H5G3D4, H5G4D1, H5G4D4, H5G5D1
2	H1G2D7, H1G3D7, H1G4D7, H1G5D7, H2G2D7, H2G3D7, H2G4D7, H3G2D7, H3G3D7, H3G4D7, H4G2D4, H4G3D7, H4G5D4, H5G2D7, H5G3D7, H5G4D7, H5G5D4
3	H2G5D7, H3G5D7, H4G2D7, H4G4D7, H4G5D7, H5G5D7

Tabla 6.5: Clasificación Fuzzy Art con las señales de la nariz electrónica

Lo único que se controló fue el parámetro de vigilancia para que en cada prueba se clasificaran las medidas en tres categorías diferentes (teóricamente una para cada día de shelf life). En dicha tabla, cada medida está descrita por una etiqueta cuyo primer número (de H1 a H5) identifica la cosecha, el segundo el grupo (de G1 a G5) y el tercero el día (D1, D4 o D7). Los resultados confirman el comportamiento observado en el PCA, ya que las cinco medidas del grupo 1 realizadas el séptimo día de shelf life se clasifican en la categoría 1. Solo tres “outliers” están clasificados en la clase 2, los mismos que aparecieron en el análisis PCA. Al igual que en ese análisis, dos de estas tres medidas pertenecen al grupo que estuvo más semanas en frigoconservación, confirmando la tendencia de aceleración del proceso de maduración cuando las piezas están un tiempo prolongado en la cámara. La medida H4G2D4 es la única que puede ser considerada un verdadero “outlier” sin explicación biológica para su clasificación.

Clase	Medidas
1	H1G1D1, H1G1D4, H1G1D7, H1G2D1, H1G2D4, H1G2D7, H1G3D1, H1G3D4, H1G3D7, H1G4D1, H1G4D4, H1G4D7, H1G5D1, H1G5D4, H1G5D7, H2G1D1, H2G1D4, H2G1D7, H2G2D1, H2G2D4, H2G2D7, H2G3D1, H2G3D4, H2G3D7, H2G4D1, H2G4D4, H2G5D1, H2G5D4, H2G5D7, H3G1D1, H3G1D4, H3G2D1, H3G2D7, H3G3D4, H3G4D1, H3G4D4, H3G5D1, H3G5D4, H5G2D4, H5G3D7
2	H2G4D7 H3G1D7 H3G2D4 H3G3D1 H3G3D7 H3G4D7 H3G5D7 H4G1D1 H4G1D4 H4G1D7 H4G2D1 H4G2D4 H4G2D7 H4G3D1 H4G3D4 H4G3D7 H4G4D1 H4G4D4 H4G4D7 H4G5D1 H4G5D4 H5G1D4 H5G1D7 H5G2D1 H5G3D1 H5G3D4 H5G4D1 H5G4D4 H5G4D7 H5G5D1 H5G5D4
3	H4G5D7 H5G1D1 H5G2D7 H5G5D7

Tabla 6.6: Clasificación Fuzzy Art utilizando indicadores de calidad

La tabla 6.6 muestra la misma clasificación utilizando los 21 indicadores de calidad, en donde se puede observar que no existe una agrupación clara. Lo único que se puede decir es que en la categoría 1 tiene más medidas del día 1 y que la categoría 3 contiene, básicamente, medidas del día 7.

Una vez clara la tendencia y los factores que influyen la velocidad de maduración, se intentó ver si era posible mejorar la clasificación utilizando un algoritmo supervisado.

Para ello, a cada una de las 75 medidas se le asignó un estado de maduración: verde o maduro. El criterio que se siguió para esta asignación ideal fue la de considerar verdes todas las muestras medidas tras uno o cuatro días de shelf life más las cinco medidas del grupo 1 tras siete días de shelf life. El resto de medidas del día 7 fueron consideradas maduras. Utilizando una red neuronal Fuzzy Artmap, con la misma combinación de parámetros de entrada utilizados en el PCA y la red Fuzzy Art, se aplicó un “leave-one-out” que permitió alcanzar un acierto del 94.6% (71 de 75 medidas). Las medidas H4G2D4, H4G4D4 y H5G5D5 fueron erróneamente clasificadas como maduras y la H1G3D7 como verde. Como se recordará, las medidas H4G2D4 y H5G5D5 fueron los mismos “outliers” encontrados en la red Fuzzy Art.

6.3.4 Correlación entre indicadores de calidad y señales de la matriz de sensores

Al igual que con el segundo estudio sobre manzanas, uno de los objetivos de esta experiencia con peras era comparar de forma objetiva las medidas con nariz electrónica frente a las realizadas con las técnicas de control de calidad. A diferencia del estudio anterior, en este caso un total de 21 indicadores fueron extraídos de cada medida.

Es importante recordar que en esta experiencia las muestras medidas con la nariz electrónica eran diferentes de las que se sacrificaron en Lleida para obtener indicadores de calidad. A pesar de ser recolectadas en la misma partida, en la misma fecha, pasar un intervalo de frigoconservación idéntico y esperar un mismo número de días de “shelf life”, la variabilidad inherente a la fruta es una interferencia difícil de obviar en este tipo de comparaciones.

Aunque inicialmente se consideró la posibilidad de utilizar redes neuronales para realizar predicciones, al final se descartaron debido a que para funcionar bien requieren de un nutrido número de medidas de entreno y el proceso de aprendizaje es computacionalmente muy costoso. Una vez más se creyó conveniente utilizar modelos PLS, mucho más rápidos de calcular y que ofrecían la posibilidad de ser acoplados a algoritmos genéticos para la selección de las variables de entrada.

Como el número de sensores era muy elevado, y de cada sensor se pueden extraer múltiples parámetros (estáticos y dinámicos) se planteó la necesidad de utilizar algún algoritmo de selección de variables. Se escogió un algoritmo genético por su facilidad de acople con el algoritmo PLS. Además de los parámetros de los sensores, se incluyeron variables con información de peso, superficie y humedad relativa de cada medida. La tabla 6.7 muestra los parámetros con los que se invocó al algoritmo genético. En la correlación de cada uno de los 21 indicadores dicho algoritmo intentó encontrar una combinación de parámetros que optimizase el poder de predicción del equipo.

Parámetro	Valor
Tamaño de la población	128
Número máximo de generaciones	250
Probabilidad de mutación	0.005
Criterio de convergencia	100%
Crossover	Doble
Nº máximo variables latentes	10
Orden de la validación cruzada	3
Iteraciones	1

Tabla 6.7: Parámetros del algoritmo genético

Como dicho algoritmo es computacionalmente eficiente, se pudo aplicar una correlación cruzada de orden 1. Para cada uno de los 21 indicadores se crearon 75 modelos PLS para realizar las predicciones. En cada iteración se autoescalaron los datos de las medidas de entreno y la medida restante de evaluación se centró y escaló con la media y varianza del conjunto de medidas de aprendizaje.

Parámetro	Ssq	Coef. Corr.	Lv
Firmeza	16.22	0.89	9
Contenido en sólidos solubles	102.12	0.32	7
Acidez tritratable	65.17	0.48	6
Índice de almidón	42.50	0.68	6
Saturación $(*a^2+*b^2)^{1/2}$	28.51	0.80	7

Tabla 6.8: Resultados en las predicciones físico-químicas

La tabla 6.8 muestra el coeficiente de correlación, el error cuadrático medio (Ssq) y el número de variables latentes óptimas (lv) en la predicción de indicadores clásicos. En dicha tabla se puede observar que la firmeza y la saturación del color son los indicadores mejor predichos. La figura 6.6 muestra una representación gráfica de dichas predicciones. Hay que recordar que los valores representados están centrados y escalados, por lo que los valores numéricos en la gráfica no representan los valores reales obtenidos al aplicar cada una de las técnicas.

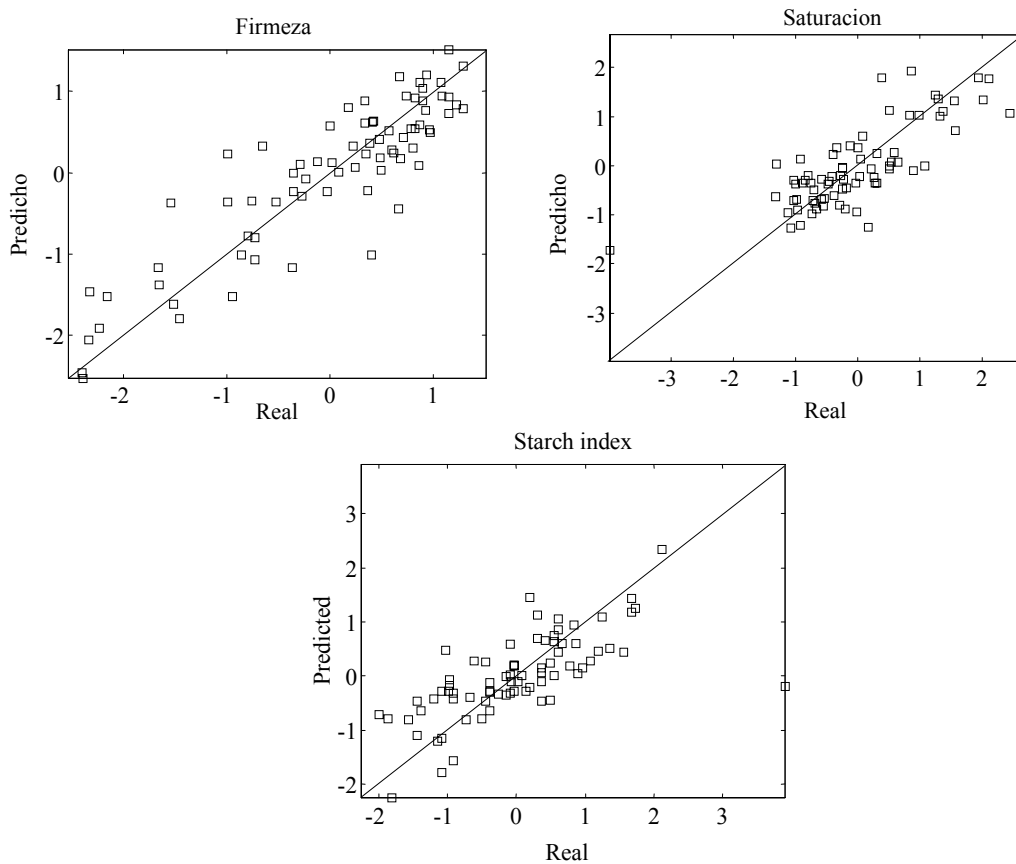


Figura 6.6: Gráficas de las mejores predicciones de parámetros físico-químicos

Las mejores predicciones para componentes volátiles se obtuvieron para el acetato de butilo y para el acetato de propilo. La figura 6.7 muestra las gráficas de ambas predicciones y la tabla 6.9 el coeficiente de correlación, el error cuadrático medio y el número óptimo de variables latentes para la predicción de cada indicador.

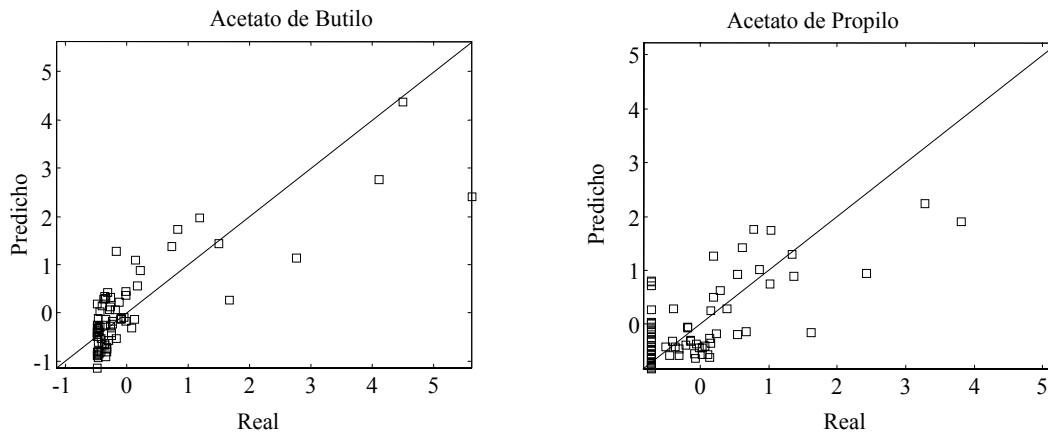


Figura 6.7: Representación gráfica de las mejores predicciones de volátiles

Aroma	Ssq	Corr. Coef.	Lv
Acetato de Etilo	74.90	0.33	4
Etanol	63.91	0.48	5
<u>Acetato de propilo</u>	43.68	0.70	2
Etileno	58.62	0.59	2
1-Propanol	84.58	0.30	4
<u>Acetato de butilo</u>	29.97	0.82	9
1-Butanol	51.59	0.62	1
CO ₂	52.70	0.61	10
Acetato de hexilo	60.06	0.63	1
hexanoato de etilo	99.46	0.2	4
Acetato 2-Metilpropilo	90.38	0.35	4
Acetato 2-metilbutilo	55.75	0.56	10
Acetato de pentilo	75.66	0.56	1
2-metilo-1-butanol	140.52	0.34	3
Butirato de butilo	60.79	0.59	7
Hexanol	85.18	0.54	1

Tabla 6.9: Resultados en la predicción de volátiles

6.3.5 Conclusiones

Las peras parecen ser especialmente adecuadas para las medidas con nariz electrónica, ya que generan una señal suficientemente clara y además muestran una tendencia muy

acusada durante su proceso de maduración. Es por ese motivo que en la segunda temporada la clasificación según estados de maduración funciona tan bien.

Las correlaciones entre indicadores y señales de la matriz de sensores no es tan buena, debido, fundamentalmente, a que las muestras medidas con la nariz electrónica no eran físicamente las mismas que las utilizadas para extraer indicadores de calidad. Además, a pesar de que se intentó que las condiciones de frigoconservación fueran idénticas, al estar en cámaras diferentes seguro que existieron pequeñas diferencias en temperatura y humedad entre Lleida y Tarragona.

REFERENCIAS

[6.1] Brezmes J.; Llobet E.; Vilanova X.; Saiz G.; Correig X., *Fruit ripeness monitoring using an electronic nose*, *Sensors and Actuators B-Chemical*, (2000), vol 69, num 3, 223-229

[6.2] Brezmes J; Llobet E; Vilanova X ; Saiz G; Correig X., *Evaluation of the electronic nose as a novel instrument to assess fruit ripeness*, Eurodeur-airodeur, Paris, Junio 2001

[6.3] Brezmes J; Llobet E; Vilanova; López M.L.; Recasens I, X ; Saiz G; Correig X., *Design and evaluation of an electronic nose to assess fruit ripeness*, IEEE Sensors (enviado)

CAPITULO 7: **MEDIDAS CON MELOCOTONES Y** **NECTARINAS**

7.1 INTRODUCCIÓN

La tercera variedad de fruta estudiada con el primer prototipo fue el melocotón. Como los resultados fueron muy alentadores [7.1,2], en la segunda temporada se realizó un nuevo estudio con una variedad que presentaba dificultades al determinar su fecha óptima de recolección. A diferencia del resto de variedades, las nectarinas sólo fueron estudiadas en la segunda temporada con el prototipo de flujo continuo [7.3,4].

Una vez más, en el primer año se quiso estudiar la viabilidad de utilizar una nariz electrónica mientras que con el prototipo perfeccionado se buscaron objetivos con un claro interés comercial debido a los buenos resultados obtenidos con el prototipo inicial.

7.2 ESTUDIO DE VIABILIDAD

7.2.1 Objetivos

Como en todos los ejemplos anteriores, el principal objetivo de la primera temporada era comprobar si la nariz electrónica detectaba el proceso de maduración de

melocotones. Para comprobarlo se realizaron múltiples medidas y se plantearon dos objetivos concretos: La clasificación en tres categorías de maduración de las piezas medidas con la nariz y la determinación de los días transcurridos desde la recolección a cada medida.

Una vez más hay que recordar que la clasificación ideal del estado de maduración la realizó un operario de forma subjetiva, aunque apoyado en técnicas destructivas como la penetromía. Sin embargo, la determinación de los días de “shelf life” de las piezas eran datos objetivos que, como veremos, la nariz fue capaz de estimar con bastante precisión.

7.2.2 Metodología experimental

Para el estudio inicial se escogieron melocotones de la variedad “blanca” recolectados en la zona de Cambrils, en la provincia de Tarragona. Se realizaron tres recolectas los días 14 y 28 de julio y el 18 de agosto. En cada una de ellas se seleccionaron 60 melocotones que se dividieron en 3 grupos de 20 muestras cada uno.

Cada grupo de 20 piezas se subdividió en un subgrupo de medida y otro de reserva para poder aplicar técnicas destructivas en paralelo a las mediciones con nariz electrónica. Como ya se aplicó en las otras variedades, las muestras de cada grupo se emparejaron intentando maximizar la semejanza en peso, tamaño y estado de maduración. Cada vez que una muestra era sacrificada para obtener parámetros de calidad era substituida por su pareja del grupo de reserva.

Una vez recolectados, los melocotones se almacenaron a temperatura ambiente (20 °C) y se iban midiendo a lo largo de su “shelf life”, desde el día de la recolecta hasta que se consideraban podridos. En cada jornada se realizaron 5 medidas con la nariz electrónica, de manera que dos de los tres grupos de cada recolecta fueron medidos dos veces mientras que el tercero sólo una. Entre medidas se dejaba una hora de reposo para que los sensores recuperasen su estado inicial. Al final de cada jornada se escogía al azar un

melocotón de cada uno de los tres grupos y se practicaba en él la penetromía, la medición de la acidez y el contenido en sólidos solubles.

Grupo	Número de medidas
1	10
2	10
3	9
4	9
5	8
6	9
7	9
8	8
9	8

Tabla 7.1: Medidas con nariz electrónica realizadas para cada grupo

Como se realizaron tres recolectas de 60 melocotones cada una, un total de 180 melocotones fueron utilizados en las mediciones, repartidos en 9 grupos diferentes que inicialmente tenían 20 muestras. El número total de mediciones con la nariz electrónica fue de 84. Tras una inspección inicial se descartaron 4 de ellas al presentar comportamientos anormales en los transitorios de los sensores, por lo que al final fueron 80 medidas las que se utilizaron en el estudio. La tabla 7.1 lista el número de medidas realizadas sobre cada uno de los 9 grupos de melocotones. Se realizaron mediciones a diario salvo limitaciones en los recursos, y en cada jornada se practicaron las técnicas destructivas.

7.2.3 Evolución de la señal

Las primeras mediciones con melocotones mostraron la debilidad de la señal que se obtenía, por lo que optó por realizar varias pruebas hasta encontrar la manera de garantizar una señal libre de interferencias para poder realizar el seguimiento de la maduración con garantías.

La figura 7.1 muestra la respuesta de los sensores ante sucesivas inyecciones en la cámara de medición de la atmósfera de la cámara que contiene la fruta tras una hora de

concentración. En ella se puede observar como tras cada inyección (hay 4) los sensores responden con un transitorio de conductancia. Por lo tanto, se consideró adecuado realizar cuatro inyecciones en cada medida para asegurar una señal adecuada para los posteriores análisis con algoritmos de reconocimiento de patrones.

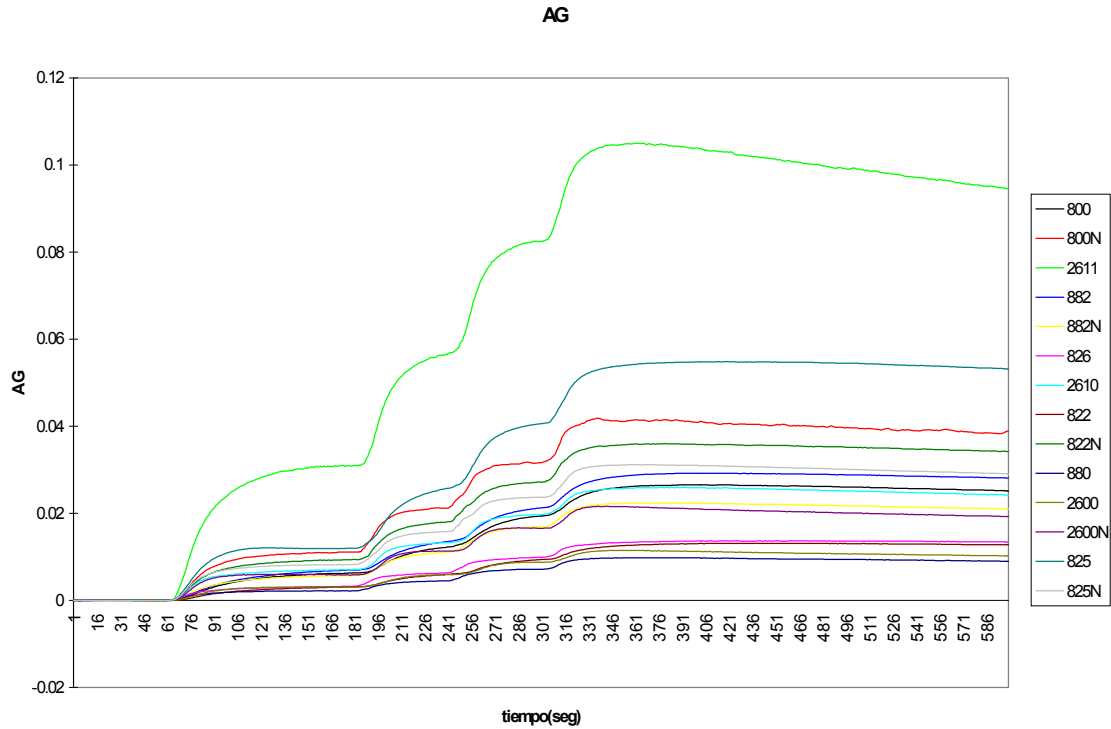


Figura 7.1: Evolución de la señal de los sensores tras sucesivas inyecciones de volátiles generados por melocotones

Una vez realizadas todas las medidas con cuatro inyecciones, el estudio de la evolución a lo largo del proceso de maduración permitió constatar que el parámetro que mejor seguía dicho proceso era el incremento de conductancia tras la tercera inyección. El utilizar como parámetro lo que ocurre tras sucesivas inyecciones tiene varias ventajas. Además de proporcionar una señal de mayor nivel (y por lo tanto con menos interferencias) [7.5] hay un estudio sobre el uso de inyecciones sucesivas en el que se concluye que los resultados son más repetitivos. La no utilización de la última inyección se debe a que las evoluciones que se observan no son tan claras, quizás por la saturación de los sensores ante tanta repetición del proceso de inyección de volátiles.

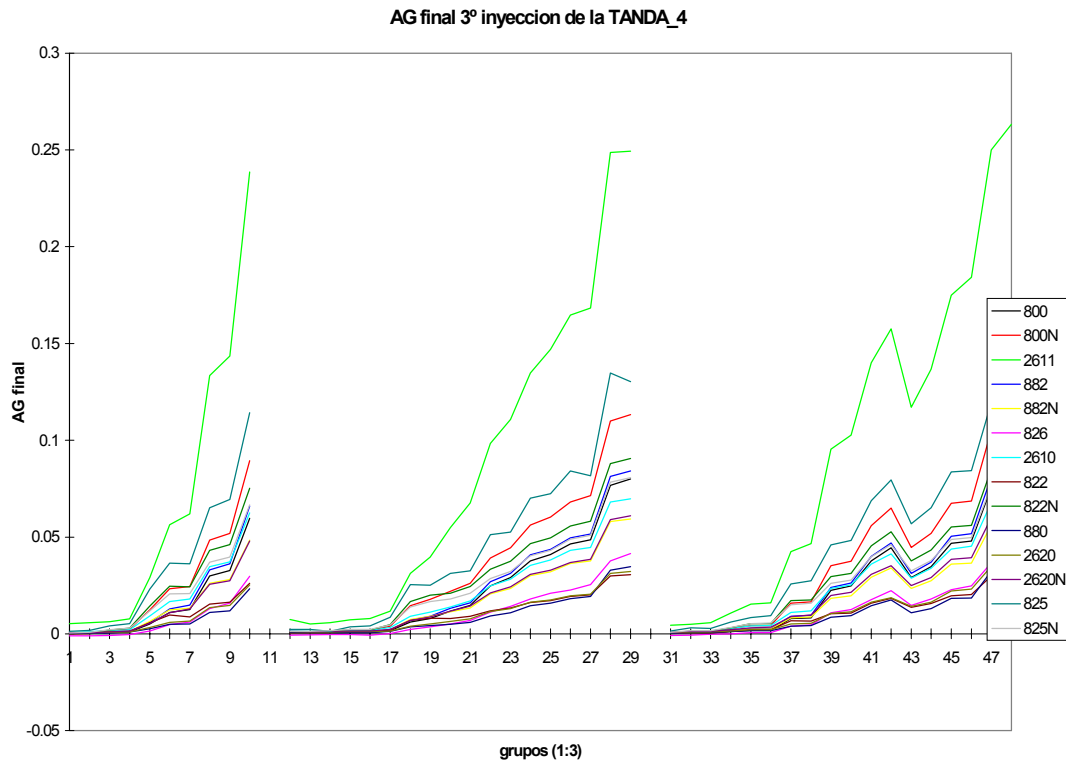


Figura 7.2: Evolución de la señal para los tres primeros grupos de melocotones

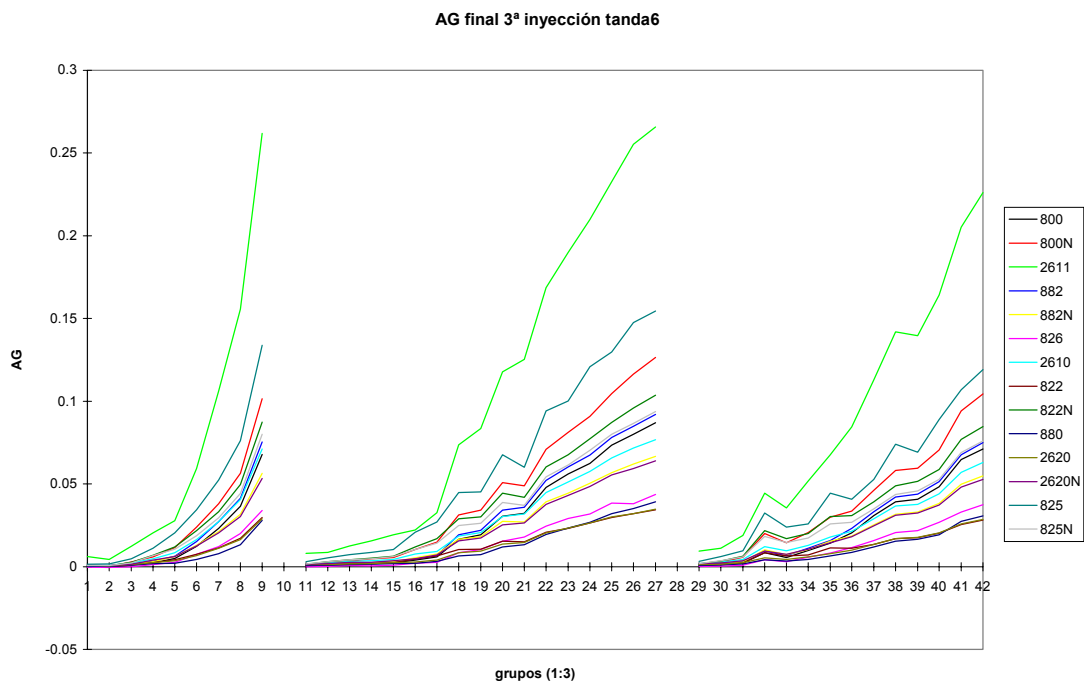


Figura 7.3: Evolución de la señal para los tres últimos grupos de melocotones

Una vez garantizada una señal libre de interferencias y con cierta repetitividad la característica más importante a encontrar es la existencia de una evolución en la señal que permita seguir el proceso de maduración. La figura 7.2 muestra la evolución que se observa con el parámetro ΔG de la tercera inyección (a partir de ahora $\Delta G3$). Para demostrar que el incremento de señal no es debido a la deriva de los sensores, las figuras 7.2 y 7.3 muestran las evoluciones de 6 de los 9 grupos, los tres de la primera recolecta y los tres de la tercera. En ellas se puede observar como los valores del primer día de las medidas de los grupos 7,8 y 9 son muy inferiores a los de los últimos días de las medidas de los primeros tres grupos y parecidos a los de los primeros días de los mismos, prueba clara de que la señal aumenta debido al incremento de volátiles producidos al madurar la fruta y no a derivas temporales.

7.2.4 Resultados

Para comprobar la resolución del primer prototipo se aplicaron análisis PCA con las medidas realizadas en cada una de las recolectas. Estos diagramas se realizaron utilizando el parámetro $\Delta G3$. Se exploraron proyecciones con varias componentes principales aunque las tendencias de maduración ya se encontraron en la primera componente principal.

La figura 7.4 muestra el diagrama PCA para las medidas con los grupos de la primera recolecta (grupos 1,2 y 3). Cada medida esta representada por un círculo de diferente color según el grupo (rojo para el primero, azul para el segundo, añil para el tercero) y un número que indica el número de orden de la medida. En ella se puede observar como las primeras medidas están concentradas a la izquierda del diagrama y posteriormente se van desplazando al centro de la proyección. Este comportamiento se repite para las tres recolectas, una clara indicación de que la nariz tiene suficiente resolución para seguir el proceso de maduración de los melocotones.

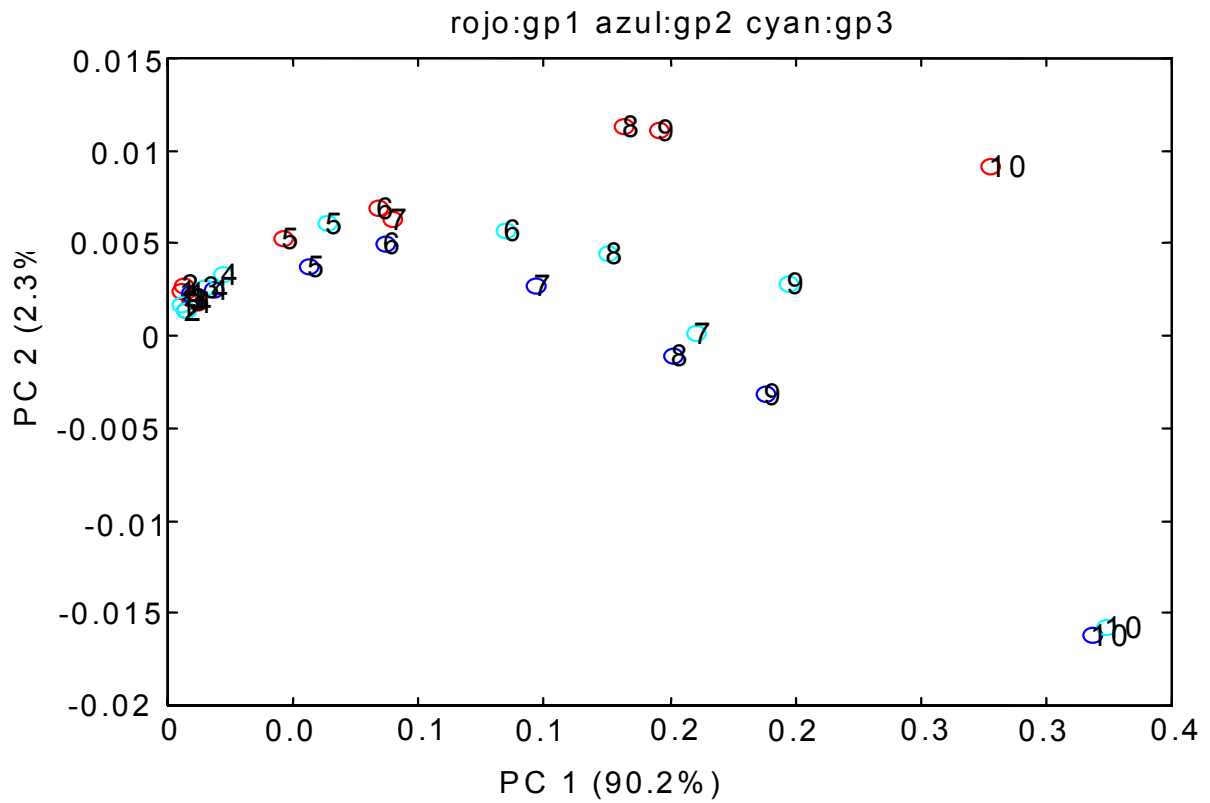


Figura 7.4: Análisis PCA para los tres grupos de la primera recolecta de melocotones

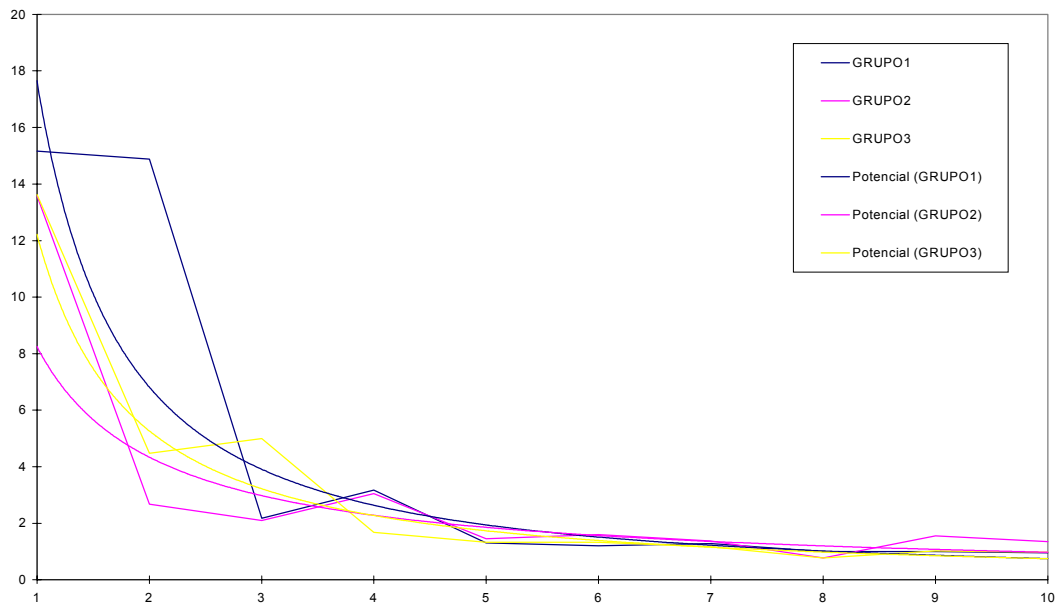


Figura 7.5: Penetración a lo largo del tiempo para los tres primeros grupos. Datos reales e interpolados

Al igual que con peras y manzanas, se practicaron análisis destructivos para orientar al operario que debía clasificar las piezas en verdes, maduras y sobremaduras. La figura 7.5 muestra los resultados de firmeza para los grupos 1, 2 y 3. En ella se muestran rectas que unen los valores obtenidos y tres curvas que interpolan los resultados para intentar extraer la tendencia general y filtrar las interferencias introducidas por la propia variabilidad de la fruta. Comportamientos parecidos se obtuvieron para el resto de grupos.

El resto de parámetros destructivos no presentaron tendencias claras que permitiesen ayudar a determinar las tres categorías de maduración que el operario debía establecer, por lo que se basó, fundamentalmente, en la firmeza de las piezas. La tabla 7.2 muestra los valores que arbitrariamente se consideraron frontera entre las tres categorías de madurez.

	Verdes	Maduros	Sobremaduros
Melocotones	$p > 4 \text{ kg}$	$1.5 \text{ kg} < p < 4 \text{ kg}$	$p < 1.5 \text{ kg}$

Tabla 7.2: Valores frontera de penetromía para las tres categoría de maduración

Para realizar las clasificaciones se utilizó una red neuronal con estructura similar a la descrita en capítulos anteriores. Los sensores se escogieron, fundamentalmente, atendiendo al criterio de sensibilidad aunque también se intentó incluir aquellos cuyo comportamiento se desmarcaba de la tendencia general, por si aportaban información útil. La tabla 7.3 lista los sensores escogidos para probar las diferentes combinaciones.

Sensores
800N
825
825N
2610
800
882N
822N

Tabla 7.3: Sensores escogidos para la red neuronal

El entrenamiento se hizo mediante una validación cruzada por grupos (de orden 9). De esta forma, se realizaron 9 iteraciones en cada una de las cuales se evaluó con las medidas de uno de los grupos mientras el aprendizaje se hizo con los ocho restantes. Una vez más, esta aproximación redujo a 9 veces el costoso proceso (computacionalmente hablando) de entrenar una red neuronal de las 80 que serían necesarias con un “leave-one-out”.

Variables	n° sensor	Sensores	n° inputs	Neuronas ocultas	Aciertos/ medidas	% ok
Go,G3,p,s	4	800N, 2610, 825, 825N	10	6	67/80	83.75
Go,G3,p,s	7	800, 800N, 882N, 2610, 822N, 825, 825N	16	8	66/80	82.5
$\Delta G3,p,s$	4	800N, 2610, 825, 825N	6	6	66/80	82.5
$\Delta G3,p,s$	7	800, 800N, 882N, 2610, 822N, 825, 825N	9	6	67/80	83.75
$\Delta G3,p,s$	7	800, 800N, 882N, 2610, 822N, 825, 825N	9	2,4,6,8,10	66,63,68,69,69 /80	86.25
$\Delta G3,p$	7	800, 800N, 882N, 2610, 822N, 825, 825N	8	6	60/80	75
$\Delta G3,s$	7	800, 800N, 882N, 2610, 822N, 825, 825N	8	6	65/80	81.25
$\Delta G3$	7	800, 800N, 882N, 2610, 822N, 825, 825N	7	6	57/80	71.25
$\Delta G3,p,s$	2	825, 822N	4	2,4,6	70,73,72/80	<u>91.25</u>
$\Delta G3,p,s$	3	825, 822N, 800	5	2,4,6	69,70,70/80	87.5

Tabla 7.4: Mejores resultados obtenidos en la clasificación por madurez de los melocotones

La tabla 7.4 muestra los mejores resultados obtenidos. De entre los mismos destaca la comparación, para una misma combinación de parámetros de sensores, de los resultados obtenidos con la presencia o ausencia de las señales de normalización. En esa comparación se puede observar perfectamente como estas señales ayudan claramente a mejorar los resultados. Por otro lado, en la tabla se refleja el hecho de que los mejores resultados son aquellos en los que se utiliza el incremento de conductancia (ΔG) de cada sensor, en detrimento de la combinación conductancia inicial y conductancia final (G_i , G_f), que obtiene peores resultados.

Variables	n° sensores	Sensores	n° inputs	Neuronas ocultas	Aciertos/medidas	% ok
$\Delta G3,p,s$	2	822N, 825	3	2,4,6	58,56,59/80	73.75
$\Delta G3,p,s$	2	822N, 825	3	2,4,6	65,70,66/80	87.5
AG3	2	822N, 825	2	2,4,6	60,55,57/80	75
$\Delta G3,p,s$	2	800, 800N	4	2,4,6	73,70,70/80	91.25
$\Delta G3,p,s$	2	800, 882N	4	2,4,6	72,72,72/80	90
$\Delta G3,p,s$	2	800, 822N	4	2,4,6	71,71,69/80	88.75
$\Delta G3,p,s$	2	800, 2610	4	2,4,6	73,69,72/80	91.25
$\Delta G3,p,s$	2	800, 825	4	2,4,6	73,71,71/80	91.25
$\Delta G3,p,s$	2	800, 825N	4	2,4,6	66,67,68/80	85
$\Delta G3,p,s$	2	800N, 882N	4	2,4,6	72,68,70/80	90
$\Delta G3,p,s$	2	800N, 822N	4	2,4,6	69,70,73/80	91.25
$\Delta G3,p,s$	2	800N, 2610	4	2,4,6	71,70,70/80	88.75
$\Delta G3,p,s$	2	800N, 825	4	2,4,6	69,69,70/80	87.5
$\Delta G3,p,s$	2	800N, 825N	4	2,4,6	70,69,70/80	87.5
$\Delta G3,p,s$	2	882N, 822N	4	2,4,6	71,69,71/80	88.75
$\Delta G3,p,s$	2	882N, 2610	4	2,4,6	74,68,69/80	92.5
$\Delta G3,p,s$	2	882N, 825	4	2,4,6	71,72,71/80	90
$\Delta G3,p,s$	2	882N, 825N	4	2,4,6	68,68,68/80	85
$\Delta G3,p,s$	2	822N, 2610	4	2,4,6	73,72,70/80	91.25
$\Delta G3,p,s$	2	822N, 825	4	2,4,6	70,73,72/80	91.25
$\Delta G3,p,s$	2	822N, 825N	4	2,4,6	70,68,70/80	87.5
$\Delta G3,p,s$	2	2610, 825	4	2,4,6	72,74,70/80	92.5
$\Delta G3,p,s$	2	2610, 825N	4	2,4,6	71,70,70/80	88.75
$\Delta G3,p,s$	2	825, 825N	4	2,4,6	70,68,69/80	87.5

Tabla 7.5: Resultados de clasificación con dos sensores

Teniendo en cuenta que los mejores resultados se producen cuando sólo se utilizan dos sensores, se realizó una búsqueda más exhaustiva de la combinación óptima de sensores probando todas las combinaciones posibles con dos sensores. Los mejores resultados quedan reflejados en la tabla 7.5, en la que se observa que se llega a un 92.5% de acuerdo entre la clasificación hecha por el operario y la realizada por la red (recordemos que, en la fase de aprendizaje, la red aprende a clasificar según el criterio del operario). Hay que destacar que en la mejor combinación, ninguno de los seis errores es entre

categorías extremas, ya que ninguna medida verde es clasificada como sobremadura ni a la inversa.

Un último objetivo con el primer prototipo fue el intentar que la nariz determinase el número de días de “shelf life” de las muestras medidas. Una vez más se utilizó una red neuronal que en este caso solo tenía una neurona en la capa de salida. El valor numérico de la salida debía proporcionar el número de días que habían transcurrido desde la recolecta hasta la medida con la nariz.

Bajo un entrenamiento “leave-one-group-out” se utilizaron todos los sensores de la tabla 7.3 más peso y superficie. La salida de la red fue escalada para que los valores analógicos designasen el número de días con valores comprendidos entre 0 y 1. Los resultados fueron suficientemente buenos como para no buscar combinaciones mejores, ya que estimaron con un error inferior a un día los intervalos de shelf life de las 80 medidas evaluadas.

7.2.5 Conclusiones

Los resultados obtenidos son tan buenos como los obtenidos para las peras, por lo que es posible afirmar que el primer prototipo es capaz de seguir sin problemas el proceso de maduración de melocotones blancos. Por otro lado, la determinación del número de días de shelf life es una clara prueba de que con mediciones objetivas la nariz electrónica también es capaz de proporcionar buenos resultados.

La única dificultad radica en la obtención de una señal lo suficientemente fuerte como para que las tendencias no queden enmascaradas por otro tipo de interferencias. Aunque la solución aportada (realizar varias inyecciones) permite obtener buenos resultados, el segundo prototipo se diseñará con el objetivo de aumentar la señal automatizando el proceso de muestreo.

7.3 MEDIDAS EN FLUJO CONTINUO

7.3.1 Objetivos

Con el primer prototipo se demostró que con los melocotones también es posible seguir el proceso de maduración a temperatura ambiente. De todas formas, para obtener una buena señal (sobre todo durante los primeros días, cuando la fruta todavía no ha madurado) es necesario realizar varias inyecciones. El segundo prototipo elimina esta necesidad además de automatizar todo el proceso.

Por ese motivo el objetivo de las experiencias con melocotones y nectarinas realizadas en la segunda temporada dejó de ser el simple seguimiento del proceso de maduración para buscar problemas reales que la nariz pudiese solucionar. Por ese motivo, se fijó como objetivo principal el determinar el momento óptimo de recolección para variedades en las que era difícil tomar dicha decisión por métodos convencionales.

Como objetivo secundario, al igual que en todas las experiencias realizadas en la segunda temporada, se realizó un estudio de correlación entre los parámetros de calidad de fruta y las señales de la matriz de sensores del sistema de olfato electrónico para ver si el prototipo era capaz de extraer datos de forma no destructiva que pudiesen servir para extrapolar información que sólo puede ser obtenida por técnicas convencionales que, o bien son destructivas o su aplicación dista mucho de ser práctica.

7.3.2 Metodología experimental

Para este estudio se utilizaron melocotones “Royal Glory” (*Prunus persica* L, Batsch; cv Royal Glory) y nectarinas “Big Top” (*Prunus persica* L, Batsch; cv Big Top). La fincas en las que se cosecharon estaban en Tarragona. Se realizaron nueve recolectas, desde el 10 de junio hasta el 7 de Julio. La separación entre diferentes recolectas fue variando para maximizar la cantidad de información en torno a la teórica fecha óptima de recolecta .

En cada recolecta se recogieron 60 piezas de cada variedad y se dividieron en cuatro grupos de quince piezas. Un primer grupo de cada variedad se asignó a medidas de etileno. Un segundo grupo se destinó a analizar componentes aromáticos; el tercero se utilizó para aplicar medidas físico-químicas y el último se mantuvo en condiciones de “shelf life” para seguir su maduración con la nariz electrónica. Todos los grupos fueron medidos con la nariz electrónica el mismo día de la recolecta. Tras las medidas, los tres primeros se enviaron a Lleida para extraer de ellos todos los parámetros de calidad planeados. La tabla 7.6 resume la planificación experimental que se aplicó a cada recolecta. Para la mayoría de los parámetros se obtuvo una pareja por cada recolecta, lo que significa que se disponía de nueve puntos de correlación para la gran mayoría de parámetros de calidad.

Grupo	Número de muestras	Medidas con la nariz electrónica	Medidas de etileno	Medidas de aromas	Medidas físico-químicas
1	15	Día de cosecha	Desde el día 2 al 7		
2	15	Día de cosecha		Día 3	
3	15	Día de cosecha			Día 4
4	15	Desde el día 1 al 7			

Tabla 7.6: Medidas realizadas tras cada recolecta (válido para las dos variedades)

Las medidas aromáticas se realizaron el tercer día tras la recolecta y se emparejaron con las medidas de nariz electrónica realizadas con las mismas muestras el día de la recolecta. Las medidas físico-químicas se realizaron cuatro días después de cada recolecta y fueron emparejadas con las realizadas con la nariz electrónica para las mismas muestras (grupo tres) el día de la recolecta.

Como las medidas de etileno con el grupo 1 se realizaron diariamente durante toda la semana siguiente a la recolecta, se pudieron obtener más puntos de correlación al emparejar dichas medidas con las de “shelf life” realizadas en Tarragona. Hay que destacar que estas correlaciones se realizaron entre medidas realizadas el mismo día pero con muestras diferentes. Además, aunque en la Universidad de Lleida estas

medidas se realizaron diariamente, los recursos en Tarragona no permitieron un seguimiento diario de la maduración en shelf life del grupo 4 de cada recolecta.

7.3.3 Clasificación no supervisada

El objetivo primordial de este experimento era ver si con la nariz electrónica se capta algún comportamiento que permita determinar el momento óptimo de recolecta de las variedades de melocotón y nectarinas estudiadas. Para ello es necesario que la nariz electrónica sea capaz de detectar diferencias entre las medidas realizadas el día de cada recolecta. Como para cada variedad hubo un total de nueve recolectas y en cada una de ellas se midieron cuatro grupos diferentes de fruta, se dispone de 36 medidas para realizar agrupaciones no supervisadas. El resultado ideal sería poder agrupar las medidas por cosecha, formando nueve categorías en cada una de las cuales estuviesen las cuatro medidas realizadas el día de cada recolecta.

En primer lugar se probaron diferentes proyecciones PCA utilizando el parámetro incremento de conductancia (ΔG) de todos los sensores. Se utilizó este parámetro para minimizar cualquier influencia debida a derivas o condiciones iniciales de los sensores, ya que esta clasificación está íntimamente ligada al paso del tiempo. Al igual que para el resto de los estudios realizados, los sensores FIS-SB ya no se utilizaron por presentar derivas temporales. Estas derivas fueron encontradas con las medidas de calibración realizadas durante las medidas con manzanas “pinklady”. El mismo estudio sirvió para corroborar que el resto de sensores no presentaban derivas apreciables en un período aproximado de un mes.

La figura 7.6 muestra, a la izquierda, el análisis PCA con las 36 medidas realizadas con la nariz electrónica el mismo día de recolecta para las nectarinas Big Top. La proyección se realiza sobre las dos primeras componentes principales. En dicha gráfica se puede observar como las medidas que pertenecen a las cosechas 1, 2 y 3 se agrupan perfectamente en tres clusters diferentes, con poca varianza entre las medidas de una

misma agrupación. Las medidas realizadas en cosechas posteriores no se agrupan debido a la gran varianza existente entre medidas de una misma recolecta.

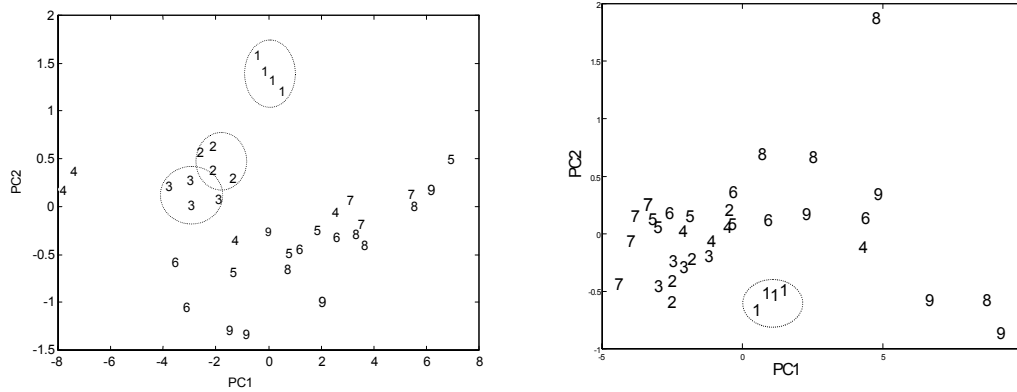


Figura 7.6: Diagrama PCA de las medidas hechas con el prototipo para nectarinas (izquierda) y melocotones (derecha)

La parte derecha de la figura 7.6 muestra los resultados obtenidos con las 36 medidas realizadas con melocotones Royal Glory. En este caso las agrupaciones no están tan claras. La única agrupación clara es la de la primera cosecha. A partir de ese momento las medidas se dispersan por todo el gráfico, por lo que a partir de la segunda recolecta se aprecia una variabilidad creciente en sucesivas recolectas.

Como ya se ha comentado, tras las mediciones con nariz electrónica la mayoría de las piezas medidas fueron enviadas a Lleida para un completo análisis de calidad. Tras ese análisis se concluyó que las nectarinas Big Top empezaron su proceso de maduración entre la tercera y cuarta cosecha, mientras que en el caso de los melocotones el proceso empezó entre la primera y segunda cosecha.

Curiosamente, la distribución de las medidas en los diagramas PCA fueron la clave para determinar las cosechas óptimas utilizando la nariz electrónica. Como el proceso de maduración no despierta al mismo momento en todas las piezas de fruta, es lógico observar un aumento de la variabilidad cuando las primeras piezas empiezan a madurar. Por el contrario, mientras dicho proceso no ha comenzado y todas las piezas son verdes las medidas deben ser mucho más repetitivas. Si nos fijamos en las proyecciones

podemos comprobar como el aumento de la variabilidad empieza precisamente en el momento en el que empieza la maduración según los indicadores de calidad, tanto en el caso de los melocotones como en el de las nectarinas. Por lo tanto, se puede concluir que la nariz es capaz de detectar el comienzo de dicho proceso que es justo cuando deben ser recolectadas las piezas si se quiere que lleguen en buen estado al consumidor.

Clase	Cosecha
1	1,1,1,1,2
2	2,2,2,3
3	3,3,3
4	4,4,5,7,8,9
5	4,5,5,5,6,9
6	6,6
7	4,6,7,7,7,8
8	8,8
9	9,9

Tabla 7.7: Clasificación Fuzzy Art para nectarinas Big Top

Clase	Cosecha
1	1,1,3,3,3,6
2	1,1,2
3	2,2,2,3
4	4,4,4,5,5,5,6,7,7
5	4,6,6,8
6	7,7
7	8,8
8	8,9,9,9
9	9

Tabla 7.8: Clasificación Fuzzy Art para melocotones Royal Glory

Para corroborar este comportamiento con otros métodos se realizó una clasificación no supervisada utilizando una red Fuzzy Art. La tabla 7.7 muestra los mejores resultados obtenidos para las muestras Big Top. El algoritmo crea nueve categorías (supuestamente para las nueve recolectas) y en la clasificación se puede ver el aumento de la varianza a partir de la cuarta cosecha. Así, mientras las cosechas 1, 2 y 3 están bastante agrupadas el resto de medidas están esparcidas por el resto de categorías. La tabla 7.8 muestra las clasificaciones obtenidas para los melocotones. En este caso prácticamente no se puede

observar ninguna tendencia, algo que ya se intuía debido a que en su diagrama PCA este comportamiento no es tan claro.

7.3.4 Correlación entre indicadores de calidad y señales de la matriz de sensores

Como objetivo secundario se propuso correlar las señales de los sensores del prototipo con los valores obtenidos para los indicadores de calidad. La tabla 7.9 muestra las relaciones entre las medidas con nariz electrónica y las medidas de calidad que fueron correladas. Se detalla, por ejemplo, si las medidas fueron sobre las mismas piezas, si fueron el mismo día y el número de puntos que se pueden emparejar.

Tipo de medida	Día de la medida	Mismo día que medida con nariz?	Mismas muestras que medida con nariz?	Puntos de medida
Etileno	2-7	Sí	No	24
Aromáticas	3	No	Sí	9
Físico-químicas	4	No	Sí	9

Tabla 7.9: Características de las correlaciones entre indicadores de calidad y señales del sistema

Como se puede observar, salvo para las mediciones con etileno (en las que se pueden correlar 24 puntos) en el resto de los casos solo hay nueve parejas de datos con las que trabajar. La idea es ver si la nariz electrónica es capaz de predecir cada uno de los indicadores a partir de la señal registrada en la matriz de sensores del prototipo.

Como el número de puntos para correlar era muy reducido se optó por no utilizar redes neuronales y se utilizaron modelos PLS para realizar las predicciones. Para aprovechar al máximo las medidas se utilizó una aproximación “leave-one-out”, de manera que para cada indicador se ajustaron nueve modelos PLS con 8 medidas para evaluarlos con la medida restante. Como en todos los estudios anteriores, los datos de entrenamiento fueron autoescalados y los de evaluación centrados y escalados por la media y varianza del conjunto de aprendizaje.

Como se incluyeron todos los sensores de la matriz (de los que se extrajeron los parámetros típicos), humedad en el interior de la cámara de medida, valores de peso y superficie medios de las piezas, se aplicó un algoritmo genético para encontrar una buena combinación de parámetros con la que realizar predicciones ajustadas. El algoritmo genético utilizado está acoplado a las rutinas PLS y realiza validaciones cruzadas para encontrar la combinación que ofrece un error cuadrático medio más bajo. Los parámetros aplicados a la rutina genética son los mismos que se utilizaron con peras.

Parámetro	Ssq	Coef.Corr.	Lv
<u>Firmeza</u>	1.69	0.94	5
SSC	11	0.22	4
Acidez	7.1	0.68	2
Tono (cara más colorada)	17	0.03	2
<u>Saturación</u> (cara más colorada)	4.3	0.79	4
<u>Tono</u> (cara menos colorada)	0.3	0.99	7
Saturación (cara menos colorada)	2.57	0.63	4

Tabla 7.10: Predicción de parámetros físico-químicos para las nectarinas

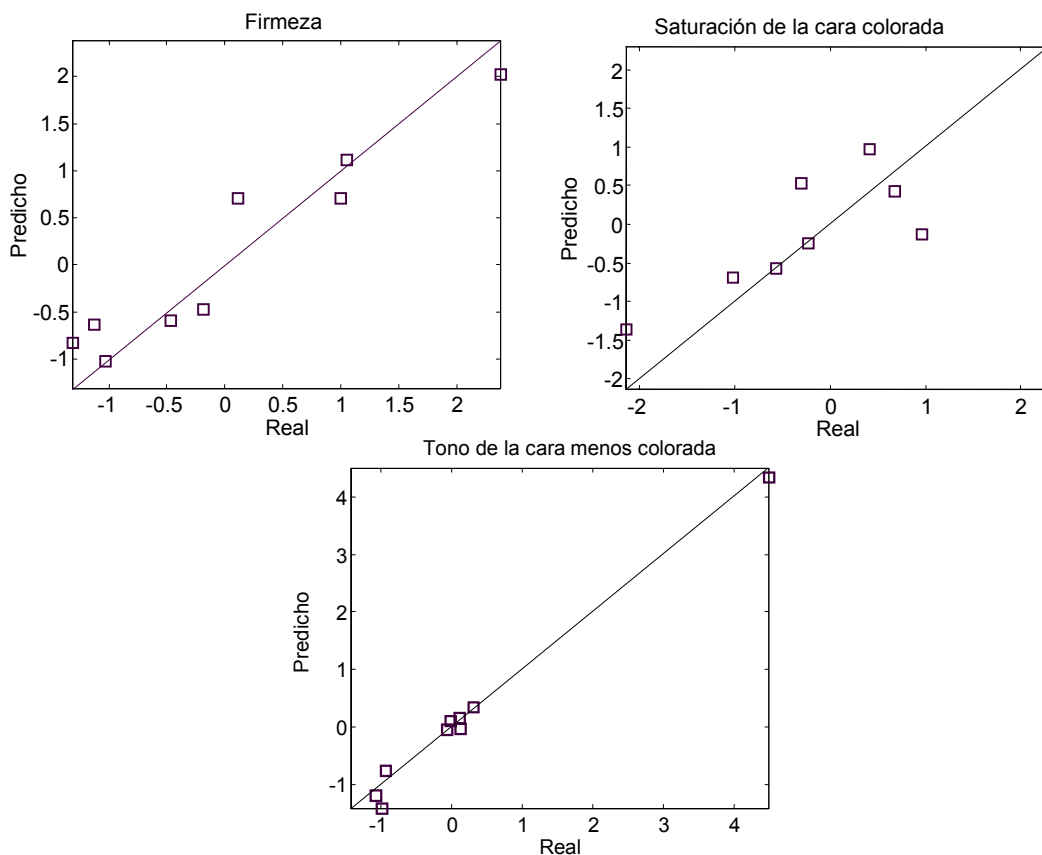


Figura 7.7: Mejores predicciones para parámetros físico-químicos para las nectarinas

Parámetro	Ssq	Coef. Corr.	Lv
Firmeza	3.5	0.94	6
SSC	4.88	0.95	5
Acidez	7.05	0.65	2
Tono (cara más colorada)	1.53	0.98	4
Saturación (cara más colorada)	14	-0.5	1
Tono (cara menos colorada)	14	0.68	4
Saturación (cara menos colorada)	15	0.39	2

Tabla 7.11: Predicción de parámetros físico-químicos para los melocotones

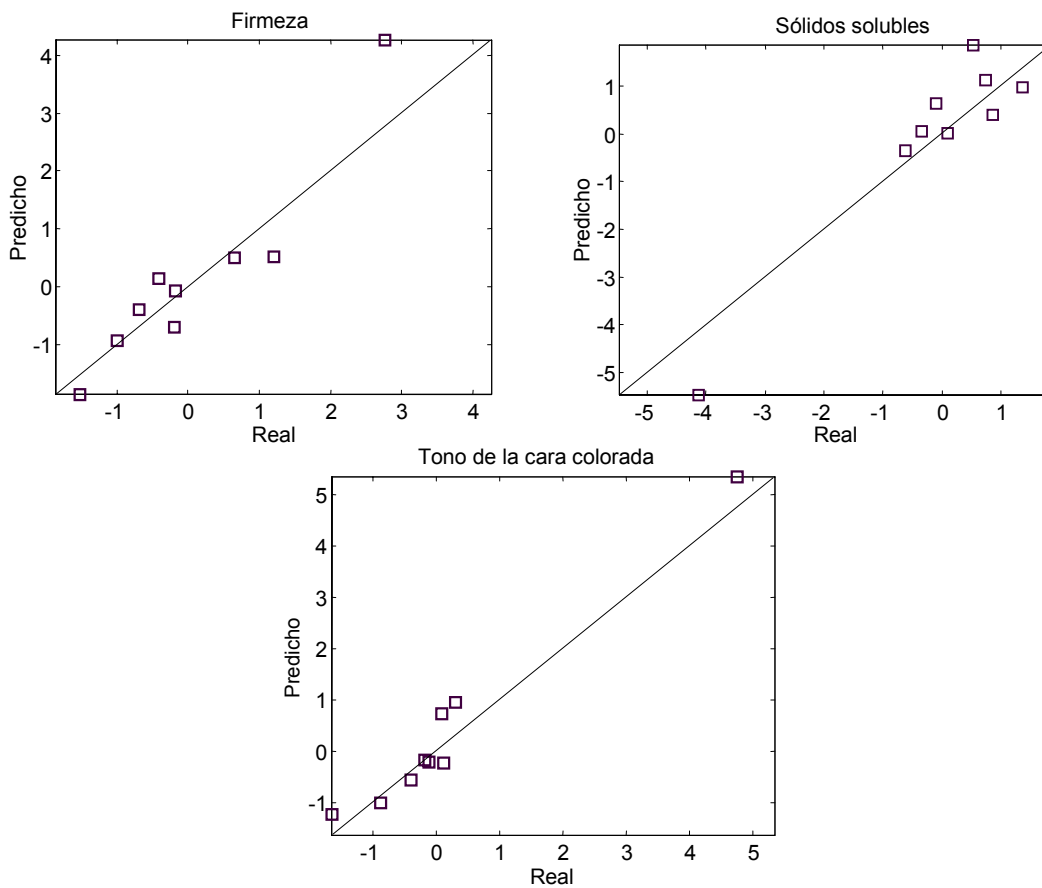


Figura 7.8: Mejores predicciones para parámetros físico-químicos para los melocotones

La tabla 7.10 muestra el coeficiente de correlación, el error cuadrático medio y el número de variables latentes del mejor modelo PLS encontrado por el algoritmo genético para la predicción de cada uno de los parámetros físico-químicos estudiados para las nectarinas Big Top. Se puede observar como las mejores predicciones son para la firmeza, el tono de la cara menos colorada, y, ya con menor acierto, la saturación de la cara más coloreada. La figura 7.7 muestra una representación gráfica para estos tres parámetros. Una vez más, hay que recordar que los valores numéricos han sido centrados y escalados, por lo que no tienen un significado físico directo. La tabla 7.11 resume los resultados obtenidos para los melocotones Royal Glory. De los datos se puede concluir que las buenas predicciones son para la firmeza, el tono de la cara más colorada y, en menor medida, a los contenidos en sólidos solubles. La figura 7.8 ilustra las predicciones realizadas para estos parámetros.

En lo que respecta a los componentes aromáticos, el etanol, acetato de hexilo, acetato z-3 de hexenilo y hexanal fueron los mejor predichos para las nectarinas Big Top. La figura 7.9 muestra las cuatro mejores predicciones y la tabla 7.12 muestra los resultados para todos los volátiles. En el caso de los melocotones también se realizó alguna predicción ajustada como el butirato de etilo, la γ -octalactona el 1-propanol y, en menor medida, el limonelo. La figura 7.10 muestra estas predicciones y la tabla 7.13 resume los resultados.

En el caso del etileno se tenían 24 puntos para correlar. Se procedió con la misma metodología que con los componentes aromáticos y se obtuvo una buena predicción para las nectarinas (con un coeficiente de correlación de 0.92). Sin embargo, con los melocotones Royal Glory los resultados fueron decepcionantes. La figura 7.11 muestra ambas predicciones.

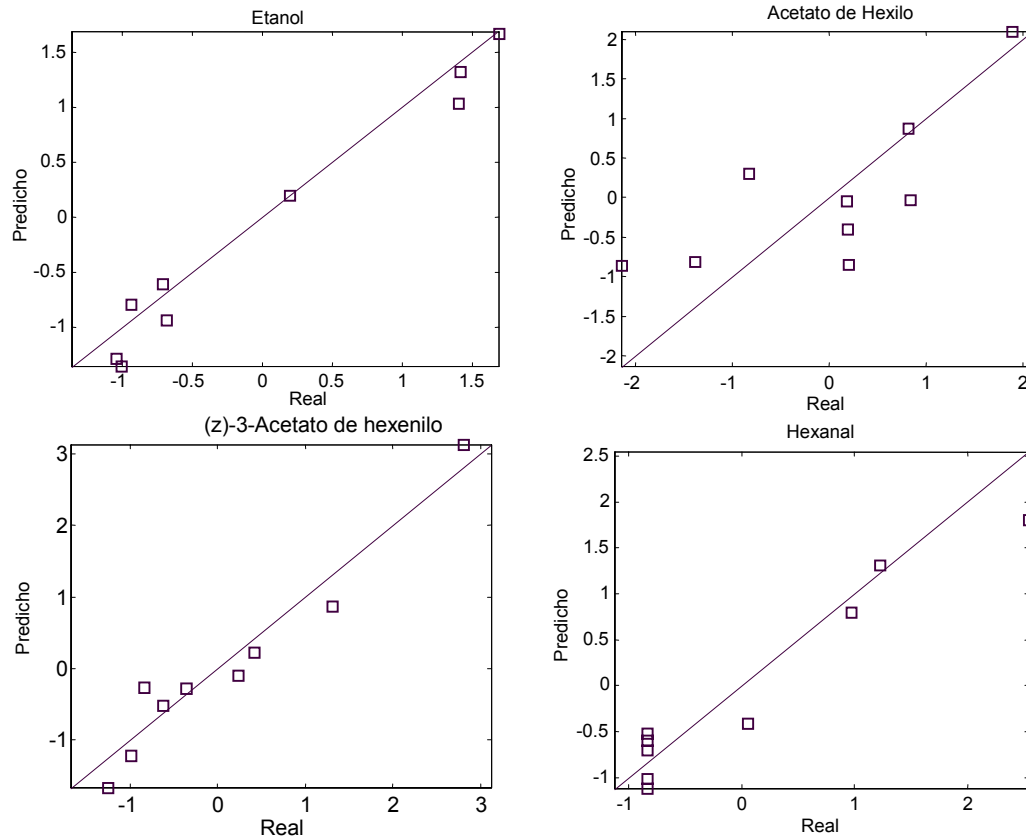


Figura 7.9: Mejores predicciones aromáticas para las nectarinas Big Top

Aroma	Ssq	Coef.Corr.	Lv
Acetato de Etilo	15.72	0.47	4
<u>Etanol</u>	0.42	0.98	8
Acetato de Propilo	21.04	0.43	9
Acetato 2-mepropilo	24.99	0.51	1
1-Propanol	22.89	-0.05	1
Acetato de Butilo	35.83	0.47	1
1-Butanol	8.24	0.63	7
Limonelo	8.14	0.67	8
<u>Acetato de hexilo</u>	3.21	0.88	4
<u>Acetato (Z)-3-hexenil</u>	1.15	0.96	6
(E)-2-hexenol	156.29	0.36	7
Linalool	9.76	0.60	1
Benzaldehido	121.09	0.28	3
γ -octalactona	17.75	0.06	2
δ -octalactona	7.90	0.55	2
γ -decalactona	11.46	0.14	1
<u>Hexanal</u>	1.097	0.96	6

Tabla 7.12: Resultados de predicción para los componentes aromáticos de las nectarinas Big Top

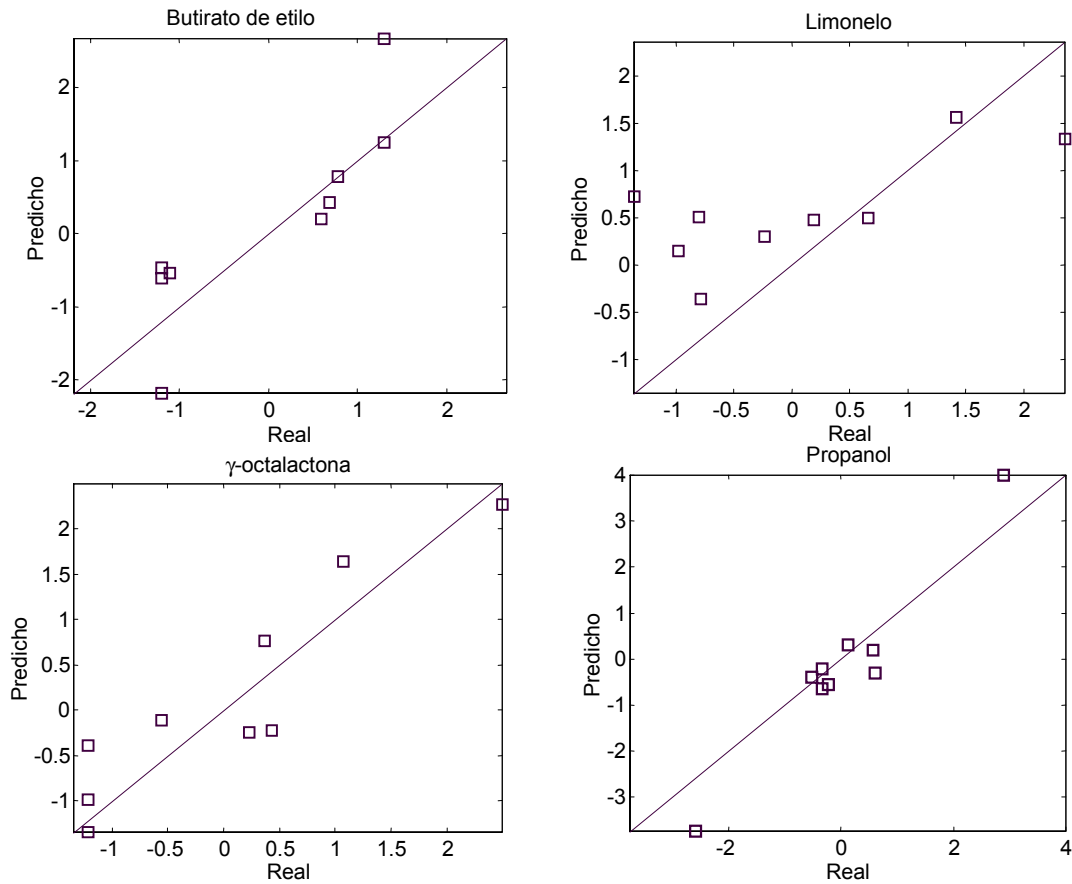


Figura 7.10: Mejores predicciones aromáticas para los melocotones Royal Glory

Aroma	Ssq	Corr. Coef.	Lv
Acetato de Etilo	12.99	0.56	4
Etanol	13.38	0.21	3
Acetato de Propilo	32.87	0.01	5
<u>1-Propanol</u>	3.84	0.87	6
<u>Butirato de etilo</u>	4.28	0.85	7
Acetato de Butilo	12.57	0.68	1
1-Butanol	9.9	0.40	2
<u>Limonelo</u>	8.98	0.72	8
Acetato de hexilo	16.5	0.03	3
Acetato (Z)-3-hexenil	16.19	0.43	3
(E)-2-hexenol	11.44	0.49	3
Linalool	12.33	0.35	2
Benzaldehido	32.40	0.55	4
Nerol	36.84	0.19	1
<u>γ-octalactona</u>	2.13	0.91	4
δ-octalactona	24.57	-0.04	3
γ-decalactona	48.38	0.01	1

Tabla 7.13: Resultados de predicción para los componentes aromáticos de los melocotones Royal Glory

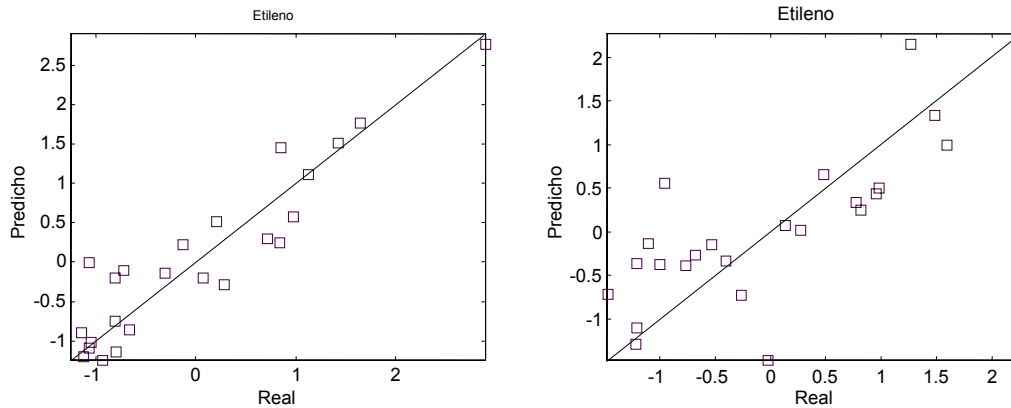


Figura 7.11: Predicciones de etileno para nectarina (izquierda) y melocotones (derecha)

7.3.5 Conclusiones

Con las medidas realizadas con el segundo prototipo se ha intentado demostrar que la nariz electrónica puede ser un instrumento útil en la determinación de la fecha óptima de cosecha de las nectarinas Big Top, variedad difícil de estudiar con técnicas tradicionales. Los resultados para los melocotones Royal Glory no son tan claros al estar la fecha óptima de recolecta entre la primera y segunda cosecha. Hubiese sido deseable que los estudios se hubiesen centrado más (es decir, haber empezado a realizar medidas varios días antes de la fecha óptima de maduración) para poder confirmar el comportamiento obtenido con las nectarinas.

En cuanto a la correlación, parece claro que la firmeza es un buen indicador de los volátiles generados por ambas variedades, ya que la nariz es capaz de predecir con bastante exactitud su valor sin necesidad de destruir las piezas de fruta. Por otro lado, la nariz ha sido capaz de predecir correctamente alguno de los componentes aromáticos de ambas variedades, algo importante ya que las propiedades organolépticas (y por tanto la calidad) de los melocotones y nectarinas están asociadas a la producción aromática de las muestras. Este dato sugiere que la utilización de una nariz electrónica para valorar las propiedades organolépticas de la fruta de forma no destructiva puede ser factible.

REFERENCIAS

- [7.1] Brezmes J.; Llobet E.; Vilanova X.; Saiz G.; Correig X., *Non-destructive fruit ripeness monitorization using Tin dioxide gas sensors and pattern recognition algorithms*, Proceedings of ISOEN, Tübingen (Germany), september 1999, ISBN 3-00-004819-7, 259-262
- [7.2] Brezmes J.; Llobet E.; Vilanova X.; Saiz G.; Correig X., *Fruit ripeness monitoring using an electronic nose*, *Sensors and Actuators B-Chemical*, (2000), vol 69, num 3, 223-229
- [7.3] Brezmes J; Llobet E; Vilanova X ; Saiz G; Correig X., *Evaluation of the electronic nose as a novel instrument to assess fruit ripeness*, Eurodeur-airodeur, Paris, Junio 2001
- [7.4] Brezmes J; Llobet E; Vilanova; López M.L.; Recasens I, X ; Saiz G; Correig X., *Design and evaluation of an electronic nose to assess fruit ripeness*, IEEE Sensors (enviado)
- [7.5] Llobet E., Vilanova X., Brezmes J. , Alcubilla R., Calderer J., Sueiras J.E., Correig, X., *Analisis of the conductance transient of thick-film tin dioxide gas sensors under successive gas injection steps*, Meas. Sci. Technol., (1997) Vol. 8, Iss 10, 1133-38

CAPITULO 8:

CONCLUSIONES

El objetivo primordial de este estudio ha sido valorar la posibilidad de utilizar una nariz electrónica para monitorizar el proceso de maduración de diferentes variedades de fruta. Para alcanzar dicha meta se planificaron dos etapas: un primer estudio de viabilidad y un segundo estudio mucho más completo, con objetivos concretos y con una valoración más objetiva de resultados, comparando la eficacia del sistema de olfato electrónico con técnicas de control de calidad de fruta

La primera etapa requería la construcción de un prototipo en el que las consideraciones operativas eran obviadas para centrarse en su objetivo: ver si la fruta emitía suficientes volátiles como para ser detectados por la matriz de sensores y ver si las variaciones en la producción aromática durante el proceso de maduración eran captadas por el instrumento.

El prototipo captó perfectamente las emisiones generadas por las peras, tanto en su intensidad como en su evolución, lo que generó buenas expectativas en cuanto a la posibilidad de utilizar una nariz electrónica para monitorizar el proceso de maduración de peras. Las pruebas realizadas con redes neuronales demostraron que el sistema de olfato electrónico podía clasificar las muestras tan bien como un operario especializado,

a pesar de entrenar con un grupo de fruta y realizar la evaluación en otro con piezas totalmente distintas.

Los melocotones presentaron problemas desde el punto de vista de nivel de señal, lo que se solucionó temporalmente realizando varias inyecciones en la cámara de medida. Sin embargo, una vez solucionado ese problema, las tendencias eran claras y así lo confirmaron las concordancias entre las clasificaciones hechas por el operario y las realizadas por las redes neuronales del prototipo. Además, la determinación del número de días de “shelf life” fue una tarea objetiva que la nariz realizó con gran exactitud .

En el caso de las manzanas, el prototipo captaba bien la presencia de volátiles pero la evolución de la señal no era tan clara. Por ese motivo la clasificación no tuvo tanto éxito como en el resto de fruta. De las tres variedades es en la que los resultados fueron peores, algo confirmado por estudios realizados por otros equipo de investigación.

En definitiva, las conclusiones obtenidas tras el primer año de trabajo fueron muy interesantes. En primer lugar, se confirmó que era viable la monitorización del proceso de maduración con una nariz electrónica, aunque también dejó bien claro que las conclusiones obtenidas para una variedad no tenían por que funcionar en otra. Otros resultados importantes fueron el descubrir la importancia del peso y superficie como señales de normalización de los datos y la necesidad de diseñar un proceso de muestreo que permitiese aprovechar mejor los volátiles generados por la fruta, con el propósito de generar una señal más fuerte en los sensores, sobre todo cuando las piezas estaban verdes.

El segundo prototipo se diseñó a partir de la experiencia obtenida con el primero. Se automatizó gracias al diseño de un circuito de flujo continuo, que a su vez perfeccionaba el método de muestreo al arrastrar todos los volátiles de la cámara de concentración a la de medida utilizando una bomba de aire y varias electroválvulas. Este sistema permitió obtener buenas señales en las tres variedades estudiadas el segundo año. Además, un diseño modular facilitó experimentar con gran variedad de sensores.

El estudio de la deriva de los sensores fue uno de los aspectos diferenciales de los estudios realizados en la segunda temporada. Se comprobó que los sensores Taguchi (serie 8) y FIS (serie SP) no presentaban derivas apreciables tras un mes de funcionamiento, mientras que los sensores FIS (serie SB) fueron apartados del estudio debido a su fuerte deriva temporal, característica que puede falsear los resultados al ser la maduración un proceso que está íntimamente ligado al paso del tiempo.

De todas formas, el equipo permitió obtener señales muy claras para las tres variedades de fruta estudiadas, confirmando que el sistema de flujo continuo muestrea mejor los volátiles de la fruta, además de ser fácilmente automatizable.

Otra de las novedades que introdujeron los estudios del segundo año fue la correlación entre indicadores de calidad y señales de la matriz de sensores. Los resultados demuestran que para cada variedad se pueden predecir algunos indicadores de forma no destructiva pero, salvo la firmeza, estos indicadores no son necesariamente los mismos en cada variedad.

Algunos de estos indicadores sirven, fundamentalmente, para determinar el nivel de maduración de las piezas medidas. Cuando la nariz es capaz de estimar su valor correctamente a partir de sus señales se está demostrando que es posible monitorizar el proceso de maduración de forma no destructiva. Aunque estos resultados puedan parecer sorprendentes, lo cierto es que los cambios fisiológicos que se producen en la fruta climatérica a lo largo de su maduración son fruto de procesos químicos que también liberan componentes volátiles a la atmósfera, por lo que existe una clara relación entre los parámetros físico-químicos como la firmeza y los volátiles generados por la fruta en concentración.

Otros indicadores, fundamentalmente los que componen el análisis del perfil aromático de la fruta, además de poder indicar el nivel de maduración, son excelentes parámetros para medir la calidad organoléptica de la fruta, ya que sus cualidades gustativas están íntimamente relacionadas con su aroma. Por lo tanto, la correcta predicción de algunos de ellos está indicando que un sistema de olfato electrónico podría constituir un instrumento no destructivo de medición de la calidad organoléptica de la fruta.

Los buenos resultados obtenidos en este trabajo describen la eficiencia de un sistema de laboratorio, equipo que trabaja en unas condiciones diferentes a las que debería trabajar un sistema comercial. Para que una nariz electrónica pueda ser utilizada como instrumento de medición en la industria frutícola es necesario estudiar y minimizar algunos problemas que son permisibles en un sistema de laboratorio, pero que no son controlables en el mundo industrial.

Uno de los aspectos que debe perfeccionarse es el proceso de calibración. La configuración actual del equipo requiere realizar calibraciones para cada variedad de fruta, por muy parecidas que sean las variedades entre sí. Además, las derivas a largo plazo de los sensores pueden invalidar las calibraciones entre dos temporadas consecutivas. Para solucionarlo es necesario perfeccionar la tecnología de los sensores semiconductores o diseñar algoritmos de procesamiento de señal que minimicen los efectos de las derivas a largo plazo en el funcionamiento del equipo.

En segundo lugar, las mediciones requieren un tiempo de concentración elevado, lo que hace poco práctico realizar un control de calidad o maduración sobre toda una producción. Por lo tanto, en su configuración actual, la monitorización se realizaría sobre una muestra significativa de la población, aproximación que, de todas formas, utilizan todas las técnicas de control de calidad que se aplican hoy en día.

Por último, las influencias de humedad y temperatura deben ser estudiadas en profundidad ya que sólo en condiciones de laboratorio pueden ser controladas satisfactoriamente. En una utilización comercial deben monitorizarse y compensar su efecto, algo que puede complicar el proceso de calibración para cada variedad de fruta.

En su estado actual, el prototipo de flujo continuo diseñado en este trabajo tendría aplicación en laboratorios con temperatura y humedad controlada y su uso comercial se debería limitar a variedades de fruta de gran valor económico que justificasen la pérdida de tiempo en el control no destructivo de su maduración o calidad. Siguiendo esta argumentación se hicieron algunas pruebas preliminares con sandías.

Desgraciadamente, su gruesa piel no permite obtener ninguna señal aromática que permita determinar si esta en su punto óptimo de recolección o consumo.

Una posible aplicación de una variante del prototipo diseñado sería el control de cámaras de conservación, en las que el enrarecimiento de su atmósfera puede ser síntoma de un proceso no deseado de maduración. De todas maneras, esta aplicación requiere ser estudiada ya que los sensores deberían trabajar en condiciones extremas de temperatura, humedad y de nivel de oxígeno si la atmósfera de la cámara está controlada.

La conclusión final del estudio es que, aunque es factible la utilización de un sistema de olfato electrónico para monitorizar el grado de maduración de la fruta, su interés comercial está supeditado a la solución de los problemas mencionados en este capítulo.