

Tesis Doctoral

*Aplicación de las técnicas de Espectroscopia Vis/NIR
y de imágenes de retrodifusión de luz láser a la
evaluación del estado de madurez de melocotón,
manzana y cereza*

Para optar al título de Doctora por la Universidad de Zaragoza, presentada por la
Licenciada en Ciencias Químicas Victoria Lafuente Rosales

Directores

Dr. Ángel Ignacio Negueruela

Dr. Jesús Val Falcón

Instituciones

Estación Experimental de Aula Dei (CSIC)

Facultad de Veterinaria (Universidad de Zaragoza)

AGRADECIMIENTOS

Quiero agradecer a todas las personas que han hecho posible esta Tesis.

En primer lugar, agradecer a mis dos directores, Jesús Val e Ignacio Negueruela por su continuo apoyo.

A Rosa Oria, Coordinadora del grupo Consolidado de Investigación reconocido por el Gobierno de Aragón: Alimentos de Origen Vegetal (T41), por permitirme iniciar esta Tesis.

A Luis Javier Herrera, Doctor del Departamento de Arquitectura y Tecnología de los Computadores, de la Universidad de Granada por aparecer en el momento adecuado y haberme motivado a continuar.

A Juan García Olmo por introducirme en el mundo de la quimiometría y The Unscrambler, de la Universidad de Córdoba.

A Melchor Maestro, por dejarme usar el equipo Vis/NIR del Instituto Pirenaico de Ecología (CSIC).

A Concha Urzola, compañera de despacho, por su ayuda en el análisis de imágenes.

A mis Chicas Val, M^a Ángeles Gracia, Azahara Díaz, Marta Pérez, Leticia Pérez y Sara del Río por todo el trabajo, apoyo y ánimo que me han dado.

A los miembros del departamento de Óptica de la Universidad de Granada, en especial a M^a del Mar Pérez por su colaboración en los trabajos realizados y a Juancho y Razban por su ayuda en temas de tratamiento de imágenes.

A la gente del grupo de Alimentos de Origen Vegetal de la Facultad de Veterinaria Ana, Guille, Diego, Tere.

A Mike por su ayuda con el inglés.

A Elena Lahoz y Gabriel Sangüesa, compañeros de trabajo por el ánimo y la comprensión prestada.

A Tania Mochales por apoyarme y mostrarme su experiencia.

A las Chicas Igme (Cristina, Carlota, Paloma y Tania) por compartir penas y alegrías en la etapa de doctorandas.

A mi madre por el apoyo incondicional en todos los momentos de esta carrera ya hayan sido buenos o malos.

A Elías y Martina por comprenderme y saber esperar.

Este trabajo ha sido financiado con los proyectos de investigación:

Nuevas estrategias para evitar el desarrollo de calciopatías en cultivos frutales modelo (AGL2009-08501).

- Subproyecto 2 del “Plan Específico de Teruel. Mejora de la calidad del melocotón de Calanda” (PET2007-09-C05-02).
- 2 Convenios de colaboración entre el Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA) y el Consejo Superior de Investigaciones Científicas (CSIC), en el proyecto “mejora de la calidad del melocotón de Calanda: material vegetal, técnicas culturales, conservación y comercialización” en el ámbito del FITE 2013 y 2104.

RESUMEN

Aragón cultiva el 30% de la superficie total de frutales de España. Unas 50.000 ha dibujan el paisaje frutícola de Aragón. La demanda de calidad por parte de consumidores, agricultores, centrales hortofrutícolas es cada vez más exigente. Tradicionalmente, el control de calidad, se realiza con técnicas que destruyen el fruto y otras subjetivas, que incluyen la estimación visual del aspecto del fruto. Sin embargo, estas técnicas ralentizan el proceso de selección en Central no son suficientemente representativas de la población muestreada, además de constituir un gasto elevado por la fruta destruida.

Estos factores han impulsado el desarrollo de técnicas más rápidas y menos costosas, que puedan aplicarse a la mayoría de los frutos procesados y que reflejen el estado interno y externo de calidad del fruto. Las llamadas técnicas ópticas, se basan en medir la interacción de la luz con la materia, en este caso la fruta y permiten determinar en una sola medida varios parámetros físico-químicos del fruto, como firmeza y contenido en sólidos solubles totales. Entre estas técnicas cabe destacar la Espectroscopia de Infrarrojo Cercano y el Análisis de Imágenes Multiespectrales, que hemos utilizado en este trabajo.

Para poder relacionar los datos obtenidos por estas técnicas con los índices de madurez del fruto, es necesario construir modelos de calibración. Los métodos elegidos de regresión lineal (PLS) y no lineal (LS-SVM) proporcionan buenos resultados ya que permiten predecir el estado de madurez de la fruta a partir de la señal óptica obtenida de cada fruta.

Las técnicas ópticas proporcionan un elevado número de variables (tantas como longitudes de onda registradas de los espectros), entre las que existe una gran colinealidad e información irrelevante. En este trabajo, nos hemos planteado realizar una selección de variables con el fin de elaborar modelos de calibración más sencillos que puedan ser aplicados en las líneas de selección de las centrales hortofrutícolas.

Palabras clave: fruta, Vis/NIR, imagen multiespectral, PLS, LS-SVM, selección de variables.

ABSTRACT

More than 30% of the surface devoted to fruit production in Spain corresponds to Aragon. Around 50,000 ha constitute the landscape of fruit tree production of Aragón. Fruits and vegetables of high quality properties are increasingly required by consumers, and consequently by producers and by the rest of agro food chain. Traditionally, tests of quality are performed by techniques which imply whether the destruction of the sample (fruit) or other subjective non-quantitative methods as the estimation of external visual appearance. In general, these techniques slow down the selection processes; do not provide enough representatively of the whole sample population and also imply the waste of considerable amounts of the facilities used to monitor strictly quality.

Among others, all these facts led to the development of new techniques which include very short times of measurement and less expensive equipment in order to facilitate their inclusion into the selection lines at industrial level. The so called optical techniques are based on the measurement of the interaction between the light and the living sample, the fruit in our case, and from a single measurement several physical-chemical parameters of the fruit, as total soluble solids and firmness, can be estimated or calculated.

Near Infrared Spectroscopy and Multispectral Images Analysis are the optical techniques used in this work.

Statistical models are created to relate optical information with reference values obtained by conventional destructive methods. Linear regression (PLS) and nonlinear regression (LS-SVM) models were selected among others as we demonstrate that provide very good results to predict the stage of maturity by using optical signals obtained from each fruit. However, optical techniques produced a large number of variables, many of them with high collinearity and also huge amounts of irrelevant information. One of the achievements of this work has been the development of a variable selection procedure, which allows the development of simpler calibration models.

Keywords: fruit, Vis/NIR, Multispectral image, PLS, LS-SVM, variable selection

Índice de contenido

1.-	<i>Introducción</i>	1
1.1.-	LA FRUTICULTURA EN ESPAÑA	3
1.2.-	CALIDAD DE LA FRUTA	5
1.2.1.-	Factores que condicionan la calidad de la fruta	6
1.3.-	MADURACIÓN	9
1.4.-	EVALUACIÓN DE LA CALIDAD.....	12
1.4.1.-	Métodos clásicos que implican la destrucción del fruto	12
1.4.2.-	Métodos no destructivos	16
1.5.-	TRATAMIENTO DE DATOS.....	27
1.5.1.-	Métodos de regresión.....	29
1.5.2.-	Selección de variables.....	32
1.6.-	VENTAJAS E INCONVENIENTES DE LOS MÉTODOS NO DESTRUCTIVOS ...	35
2.-	<i>Objetivos</i>	37
2.1.-	OBJETIVOS GENERALES	39
2.2.-	OBJETIVOS ESPECÍFICOS	39
3.-	<i>Materiales y Métodos</i>	41
3.1.-	MATERIAL VEGETAL	43
3.2.-	INSTRUMENTACIÓN ÓPTICA.....	45
3.2.1.-	Quality Spec Pro 2006.....	45
3.2.2.-	MultiSpec Instrument AG Tec5.....	46
3.2.3.-	Cámara monocromática BASLER 310f	47
3.3.-	SOFTWARE	48
3.4.-	MÉTODOS Y ANÁLISIS DE REFERENCIA	49
3.5.-	TRATAMIENTO DE DATOS.....	51
3.5.1.-	Análisis de Mínimos Cuadrados Parciales (Partial Least Square analysis. PLS)	52
3.5.2.-	Máquina de Vectores de Soporte de Mínimos Cuadrados (Least Squares Support Vector Machine. LS-SVM).....	53
3.5.3.-	Selección de Variables hacia atrás mediante Información Mutua	54
4.-	<i>Manzana</i>	57
4.1.-	INTRODUCCIÓN.....	59
4.2.-	MATERIAL Y MÉTODOS	61
4.2.1.-	Material Vegetal	61

4.2.2.-	Parámetros de calidad del fruto convencionales (destructivos)	62
4.2.3.-	Métodos ópticos para estimar parámetros de calidad	62
4.2.4.-	Métodos estadísticos	66
4.3.-	DETERMINACIÓN DE FIRMEZA Y SST MEDIANTE ANÁLISIS DEL HISTOGRAMA DE LAS IMÁGENES MULTIESPECTRALES OBTENIDAS MEDIANTE CINCO LÁSERES (670, 785, 850, 905 y 980 nm).....	67
4.3.1.-	Resultados y discusión	67
4.3.2.-	Conclusión	78
4.4.-	DETERMINACIÓN DE FIRMEZA Y DE SST MEDIANTE EL ANÁLISIS DE LOS DIÁMETROS DE FERET DE LAS IMÁGENES MULTIESPECTRALES OBTENIDAS MEDIANTE TRES LÁSERES (670, 785 y 980 nm).....	79
4.4.1.-	Resultados y discusión	79
4.4.2.-	Conclusiones.....	82
4.5.-	CONCLUSIONES DEL CAPÍTULO	83
5.-	<i>Melocotón</i>	85
5.1.-	INTRODUCCIÓN	87
5.2.-	MATERIAL Y MÉTODOS	90
5.2.1.-	Material vegetal	90
5.2.2.-	Magness-Taylor test	90
5.2.3.-	Espectroscopia Vis/NIR	91
5.2.4.-	Medida de impacto acústico	92
5.2.5.-	Tratamiento estadístico	92
5.3.-	RESULTADOS Y DISCUSIÓN.....	93
5.3.1.-	Predicción de la firmeza MT mediante medida acústica	93
5.3.2.-	Predicción de la firmeza MT mediante espectroscopia Vis/NIR	94
5.3.3.-	Predicción de la firmeza MT mediante la combinación de señales acústicas y espectros Vis/NIR.	97
5.3.4.-	Selección de variables	99
5.4.-	CONCLUSIONES	101
6.-	<i>Cereza</i>	103
6.1.-	INTRODUCCIÓN	105
6.2.-	MATERIAL Y MÉTODOS	107
6.2.1.-	Material vegetal	107
6.2.2.-	Espectroscopia Vis/NIR	108
6.2.3.-	Firmeza	108
6.2.4.-	Contenido en Sólidos Solubles	109
6.2.5.-	Tratamiento estadístico	109

6.3.- RESULTADOS Y DISCUSIÓN	110
6.3.1.- Modelos de calibración.....	112
6.3.2.- Validación.....	113
6.3.3.- Selección de variables.....	114
6.4.- CONCLUSIONES.....	116
7.- <i>Discusión general</i>	119
8.- <i>Conclusiones finales</i>	127
9.- <i>Bibliografía</i>	133

Índice de Figuras

Figura 1. Frutas de mayor producción en España 2013. Fuente: MAGRAMA 2013	
Figura 2. Consumo de fruta per cápita en los años 2011 y 2012. Fuente: MAGRAMA, 2012.	4
Figura 3. Curva sigmoidea que representa el crecimiento de la fruta de pepita .	11
Figura 4. Curva doble sigmoidea que representa el crecimiento de la fruta de hueso.....	11
Figura 5. Fenómenos producidos por la luz reflejada.....	18
Figura 6. Espectro electromagnético en el rango de la luz visible y zonas adyacentes.	18
Figura 7. Equipo Vis/NIR QualitySpec Pro 2600.....	46
Figura 8. Equipo MultiSpec Instrument AG Tec5 y compartimento de medida (a la derecha)	47
Figura 9 a) Equipo de obtención y tratamiento de imágenes multiespectrales. b) Detalle de la disposición de los láseres, de la cámara y del posicionador de muestras.....	48
Figura 10. Refractómetro digital ATAGO PR-101.....	49
Figura 11. Penetrómetro Fruit Pressure Tester FT32	49
Figura 12. Texturómetro TA.XT.....	50
Figura 13. Durómetro digital Agrostá 100 Field.....	50
Figura 14. Medidor de firmeza por impulso acústico AFS.	51
Figura 15. Esquema del funcionamiento del equipo utilizado para la captura de imagen de la retrodifusión de la luz.....	63
Figura 16. a) Imagen de luz retrodifundida (670 nm), b) Histograma de píxeles de la imagen.....	64
Figura 17. Imagen capturada de la retrodifusión láser, subdividida en segmentos con diferentes niveles de gris	65
Figura 18. Evolución de los valores de SST (°Brix), firmeza (kgf) y calibre (cm) durante la campaña primera (a) y segunda (b). En la primera campaña el valor 0 en el eje de abscisas corresponde a la fecha de 01/05/2009 y el valor 8 a la fecha 15/09/2009. En la segunda campaña, el valor 0 en el eje de abscisas corresponde a la fecha de 15/08/2010 y el valor 3 a la fecha de 30/09/2010.....	68
Figura 19. Imágenes de luz retrodifundida de los 5 láseres utilizados (de izquierda a derecha: 670, 785, 850, 905, 980 nm) en tres frutos con diferente grado de madurez (de arriba abajo: verde, maduro y sobremaduro)	70
Figura 20. Histogramas de las imágenes de retrodifusión de cada uno de los 5 láseres en tres momentos de recogida de la primera campaña: 15 de mayo (verde), 19 de julio (maduro) y 15 de septiembre (sobremaduro).....	72
Figura 21. Análisis de componentes principales para el láser 2 (785 nm) en la primera campaña. A la izquierda están los frutos inmaduros (Mayo-Julio). A la derecha, frutos maduros.....	76
Figura 22. Análisis de componentes principales para el láser 2 (785 nm) para las dos campañas. Primer y cuarto cuadrante: frutos maduros 1ª campaña. Tercer	

cuadrante: frutos inmaduros 1ª campaña. Segundo cuadrante: frutos maduros 2ª campaña	77
Figura 23. Imágenes multiespectrales monocromáticas obtenidas por iluminación con láseres de diodo a tres longitudes de onda: 670, 785 y 980 nm.....	80
Figura 24. Estadísticos de predicción para firmeza y sólidos solubles.....	81
Figura 25. Dispositivo de adquisición del espectro	91
Figura 26. Espectros de absorción de melocotones en distintos estados de madurez.....	94
Figura 27. Coeficientes de regresión para firmeza en el grupo de calibración 1..	97
Figura 28. Ejemplo de selección de variable para ambos métodos de regresión, usando todas las variables (661) y usando sólo 10 variables.....	100
Figura 29. Espectros Vis/NIR de absorción para prunus avium 'Chelan' con el equipo QualitySpec (450-2500 nm).....	111
Figura 30. Estadísticos de predicción para firmeza y sólidos solubles para ambos métodos de regresión, PLS y LS-SVM	114
Figura 31. Ejemplo de selección de variables para SSC y firmeza.....	115
Figura 32. Selección de longitudes de onda más representativas en la determinación de firmeza (a) y SSC (b)	116

Índice de Tablas

Tabla 1. Índices de madurez para algunas frutas. (Fuente: Reid, 2007)	13
Tabla 2. Materiales comúnmente empleados en equipos NIRS (adaptado de Workman, 2004)	23
Tabla 3. Muestras utilizadas en los experimentos.....	44
Tabla 4. Características técnicas de los instrumentos de detección óptica utilizados en los estudios	45
Tabla 5. Equipos utilizados en los análisis de referencia	49
Tabla 6. Valores estadísticos para los parámetros de SST y firmeza en la primera y segunda campañas y en la combinación de ambas.	69
Tabla 7. Estadísticos de calibración usando método de regresión PLS para los parámetros SST y firmeza en la primera campaña	74
Tabla 8. Estadísticos de calibración usando método de regresión PLS para los parámetros SST y firmeza en las dos campañas unidas.....	75
Tabla 9. Valores estadísticos del conjunto de muestras para los parámetros de SST y firmeza obtenidos por métodos destructivos de análisis.....	79
Tabla 10. Estadísticos de calibración usando método de regresión PLS para los parámetros SST y firmeza	80
Tabla 11. Número de muestras (n), media (X), desviación estándar (DE), rango y coeficiente de variación (%CV) para la firmeza en los tres grupos de calibración y validación	93
Tabla 12. Estadísticos de calibración para firmeza MT usando método de regresión lineal (PLS) y no lineal (LS-SVM) para Vis/NIR	95
Tabla 13. Estadísticos de predicción para firmeza MT usando método de regresión lineal (PLS) y no lineal (LS-SVM) para Vis/NIR	96
Tabla 14. Estadísticos de calibración para firmeza usando método de regresión lineal (PLS) y no lineal (LS-SVM) usando Vis/NIR+AWETA	97
Tabla 15. Estadísticos de predicción para firmeza usando método de regresión lineal (PLS) y no lineal (LS-SVM) usando Vis/NIR+AWETA	98
Tabla 16. Estadísticos de predicción para firmeza MT usando método de regresión lineal (PLS) y no lineal (LS-SVM) usando Vis/NIR+AWETA con tres variables	100
Tabla 17. Número de muestras (n), intervalo, media, desviación estándar (DE) y coeficiente de variación (%CV) para los dos parámetros de estudio para los tres grupos de calibración (n=211) y validación (n=70)	110
Tabla 18. Estadísticos de calibración de tres grupos de calibración, para dos métodos de regresión PLS y LS-SVM, para la firmeza en el intervalo de 450-700 nm y para SSC 700-2420 nm y 700-1060 nm	113

ÍNDICE DE ABREVIATURAS

- AFS: Acoustic Firmness Sensor (Sensor de medida acústica)
- ANN: Artificial Neural Networks (Redes Neuronales Artificiales)
- CCD: Charge Coupled Device (Dispositivo de Carga Acoplada)
- CV: Coeficiente de Variación
- DE: Desviación Estándar
- DT: Detrenching
- FDP: Probability Density Function (Función de densidad de Probabilidad)
- FFT: Fast Fourier Transform (Transformada Rápida de Fourier)
- GA: Genetics Algorithm
- ICA: Independent Component Analysis (Análisis de Componentes Independientes)
- iPLS: Interval Partial Least Squares
- LS-SVM: Least Squares Support Vector Machine (Máquina de Vectores de Soporte de Mínimos Cuadrados)
- LV: Variables Latentes
- MI: Mutual Information (Información Mutua)
- MLR: Multiple Lineal Regression (Regresión Lineal Múltiple)
- MSC: Multiplicative Scatter Correction
- MT: Magness Taylor
- NIRS: Near Infrared Spectroscopy (Espectroscopia de Infrarrojo Cercano)
- OSC: Orthogonal Signal Correction
- PCA: Principal Components Analysis (Análisis de Componentes Principales, ACP)
- PCR: Principal Components Regression (Regresión de Componentes Principales)
- PLS: Partial Least Squares (Regresión de Mínimos Cuadrados Parciales)
- r^2 : coeficiente de validación cruzada
- R_{cv}^2 : Coeficiente de Determinación de Validación Cruzada
- RMSEP: Error Medio de Predicción
- RBF: Radial Basis Function (Función de Base Radial)
- ROI: Region of Interest (Regiones de interés en la imagen)
- RPD: Residual Predictive Deviation (Desviación Predictiva Residual)
- SCARS: Stability Competitive Adaptive Reweighted Sampling
- SECV: Error estándar de Validación Cruzada
- SEP: Error Estándar de Predicción
- SNV: Standard Normal Variate

SPA: Successive Projections Algorithm

SST: Sólidos Solubles Totales

UVE: Uninformative Variable Elimination

1.- Introducción

1.1.- LA FRUTICULTURA EN ESPAÑA

Según estadísticas de la FAO en 2013, los principales países europeos productores de fruta son Italia, España y Turquía. En la agricultura española, la producción frutícola tiene un peso muy importante, tanto desde el punto de vista del consumo interno como de la exportación. Actualmente, las frutas que tienen mayor importancia para el mercado en fresco son: naranja, manzana, pera, melocotón, albaricoque y cereza (Figura 1). Cabe señalar que España presenta una gran diversidad de producciones, que se cultivan en climatologías diferentes y con unos períodos de recolección amplios. Las zonas de mayor importancia son Murcia y Andalucía (para las producciones muy tempranas); y Cataluña y Aragón (para las producciones intermedias y tardías). España es el segundo productor europeo con 1.060.000 t (en 2013), siendo las principales comunidades autónomas productoras: Cataluña (34%), Aragón (28%) y Murcia (14%). La fruticultura es una actividad fundamental en la agricultura española. Las exportaciones de fruta durante el año 2014 han supuesto 6.675 M€. La actividad frutícola se extiende sobre 1.009.959 hectáreas, distribuidas por el Sur, Levante, Valle del Ebro y Canarias.

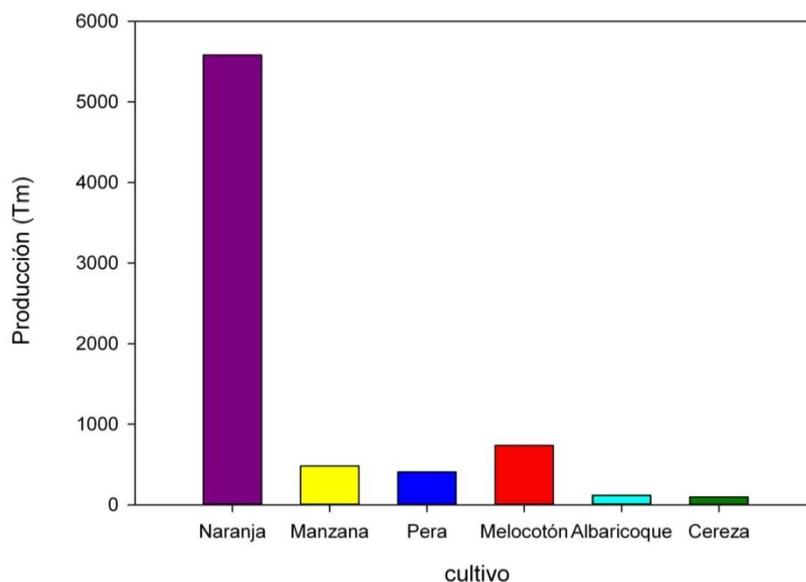


Figura 1. Frutas de mayor producción en España 2013. Fuente: MAGRAMA 2013

En el año 2014 el consumo de fruta per cápita fue de 103,4 kg/persona/año frente a los 101,46 kg/persona/año del 2011 (Figura 2), observándose una ligera tendencia al alza en los últimos años. La producción y consumo de fruta no sólo tiene repercusión económica sino también sobre la salud de la población.

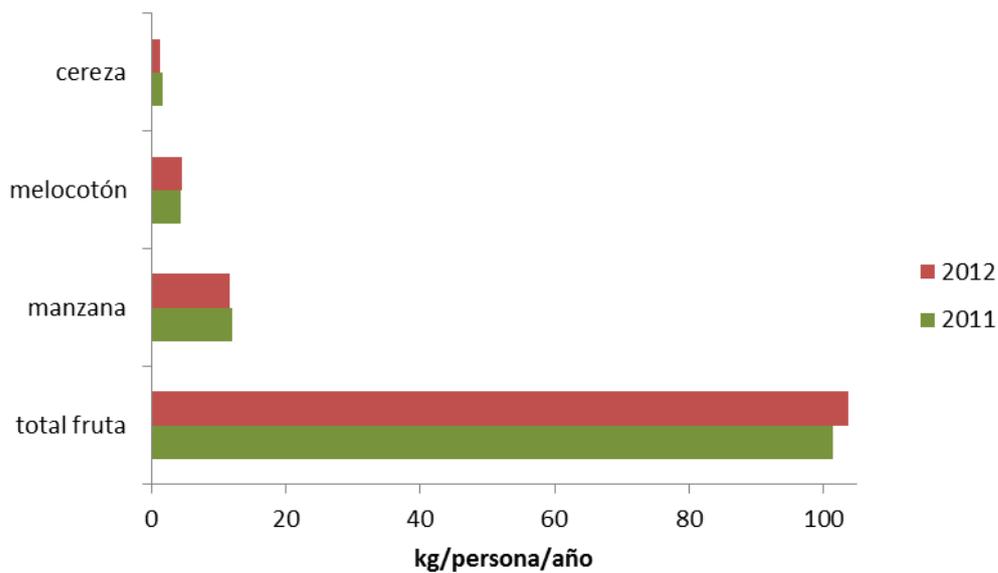


Figura 2. Consumo de fruta per cápita en los años 2011 y 2012. Fuente: MAGRAMA, 2012.

La importancia del consumo de frutas en la alimentación humana es esencial. Son componentes básicos de una dieta saludable, y un consumo diario suficiente puede contribuir a la prevención de importantes enfermedades, como las cardiovasculares y algunos tipos de cáncer.

La tendencia mundial se dirige a un mayor consumo de frutas y hortalizas, motivado fundamentalmente por una creciente preocupación por una dieta más equilibrada, con menor proporción de carbohidratos, grasas y aceites y con una mayor presencia de fibra, vitaminas y minerales. Esto se fundamenta, en parte, por las menores necesidades calóricas de la vida moderna, caracterizadas por un mayor confort y sedentarismo. El otro factor que determina esta tendencia es la

creciente concienciación de la importancia de la dieta en la salud y longevidad (Prior and Cao, 2000).

1.2.- CALIDAD DE LA FRUTA

Genéricamente se habla de consumidor, sin tener en cuenta la enorme variedad de tipos de sujetos, de sus gustos y preferencias. Los perfiles de consumo son específicos para cada país e incluso para una región, variando incluso con el sexo, edad, nivel educativo y nivel socioeconómico de las muestras de población sujetas a estudio.

La palabra “calidad” proviene del latín *qualitas*, que significa atributo, propiedad o naturaleza básica de un objeto. Sin embargo, en la actualidad y en sentido abstracto su significado es “grado de excelencia o superioridad” (Kader, 1985). Aceptando esta definición, se puede decir que un producto es de mejor calidad que otro cuándo es superior en uno o varios atributos que son valorados objetiva o subjetivamente. La calidad de un alimento es un término complejo puesto que no se puede determinar por una sola propiedad, sino que viene dada por la combinación de sus propiedades físicas, químicas y sensoriales. Todos los agentes que participan en la cadena alimentaria (agricultores, centrales hortofrutícolas, agentes comerciales) desean que el producto reúna unas características determinadas, pero es el consumidor quien tiene la última palabra en la decisión de compra de un producto u otro.

En los últimos tiempos se ha observado un aumento en los requisitos de calidad tanto externa como interna, aunque los aspectos externos (presentación, uniformidad, apariencia, madurez y frescura) son los principales factores que influyen en la decisión de la compra.

La calidad de la fruta se puede clasificar de distinta forma en función de los atributos considerados:

- Apariencia externa: aquí se incluyen las propiedades que hacen referencia al tamaño, color y forma
- Flavor: es la sensación química detectada por el gusto y por el olfato.

- **Textura:** Es la propiedad de los alimentos apreciada por los sentidos del tacto, la vista y el oído; se manifiesta cuando el alimento sufre una deformación.
- **Valor nutricional:** contenido en vitaminas, fibras, minerales y compuestos fenólicos

1.2.1.- Factores que condicionan la calidad de la fruta

Controlar la calidad de la fruta en el árbol es fundamental para asegurar una buena postcosecha. La calidad de los productos agrícolas comienza realmente con la elección adecuada de la planta, labores culturales y otras prácticas precosecha (poda, abonado, riego y protección de cultivo) y los factores agroambientales del cultivo (Moccia et al, 2006).

- **Condiciones climáticas**

La insolación y la temperatura afectan a varios atributos de la calidad del fruto (Kader, 2007). Las altas temperaturas asociadas a una radiación solar intensa en el periodo precosecha ocasionan una variedad de alteraciones en el color y pardeamientos, así como en las propiedades organolépticas debido a cambios en el contenido en sólidos solubles y acidez titulable (Romojaro et al, 2006).

En manzanas, fresas y peras, se ha puesto de manifiesto una relación indirecta entre la temperatura y la firmeza, manteniéndose o aumentando cuando el nivel térmico no es elevado (Sams, 1999). En el cultivo del melocotonero, las características de calidad están asociadas con la disponibilidad de luz, entre otros factores (Corelli-Grappadelli et al, 1991). Así, por ejemplo, la poda de verano contribuye a mejorar la iluminación de la planta (Myers, 1993).

La maduración también se ve afectada por la temperatura, inhibiéndola o acelerándola, e incrementando la desecación por pérdida acelerada de agua,

originando alteraciones tanto en el exterior como en el interior del fruto (Romojaro et al., 2006).

- **Nutrición mineral**

La nutrición mineral de la planta afecta directamente a la composición del fruto. El equilibrio entre nutrientes afecta al crecimiento y estado fisiológico del fruto pudiendo originar desórdenes fisiológicos tanto por deficiencias como por excesos (Soares-Gomes et al, 2005). Nitrógeno, fósforo, potasio y calcio son los nutrientes que participan de forma más activa en numerosos procesos metabólicos (Romojaro et al, 2006).

El nitrógeno es el nutriente que tiene mayor influencia en la calidad de la fruta. El aporte nitrogenado incrementa la concentración de SST del fruto, posiblemente al mejorar el suministro de carbono al mismo y acelerar la degradación del almidón (Xia et al, 2009) pero es contraproducente cuándo se aplica en exceso. El exceso de nitrógeno influye negativamente sobre la calidad y el crecimiento de los frutos, afecta la intensidad del color, la concentración de azúcares, firmeza, SST y el período de almacenamiento, entre otros parámetros (Neilsen et al, 1999; Raese et al, 2007).

Niveles adecuados de fósforo y potasio repercuten en el calibre, el color, la firmeza y contenido en sólidos solubles del fruto (Fallahi et al, 1997)

El calcio está involucrado en gran cantidad de rutas metabólicas y morfológicas de las plantas. Participa en numerosos procesos de desarrollo y mantenimiento de la estructura de la pared celular, por su capacidad de establecer enlaces iónicos con los grupos carboxilato de las pectinas, provocando fisiopatías relacionadas con la textura (Poovaiah et al., 1988). Como fisiopatías más representativas pueden mencionarse la vitrescencia en melón, “blossom end rot” en tomate y pimiento, “bitter pit” en manzana y “black Herat” en apio.

Ciertas fisiopatías por deficiencia de calcio pueden reducirse por una adecuada irrigación y manejo adecuado de la fertilización (Val et al, 2008; Blanco et al, 2010; Val and Fernández, 2011)

- **Riego**

El riego en los cultivos juega un papel fundamental en el rendimiento y en la calidad de las partes comestibles de la planta. Un adecuado suministro hídrico implica que los aportes, deben estar en concordancia con las necesidades del cultivo de manera que se produzca un apropiado desarrollo vegetativo y reproductivo. Sin embargo, en la búsqueda por mejorar la eficiencia del uso del agua, en los últimos años, ha habido un creciente interés en la aplicación de técnicas de riego conducentes a reducir la cantidad de agua aplicada por debajo del máximo utilizado por el cultivo, induciendo un estrés moderado con el mínimo impacto en los rendimientos y calidad del fruto. Al respecto, se señala que en los frutos del manzano obtenidos bajo déficit hídrico controlado son más firmes que los de plantas regadas al 100% de su demanda de transpiración debido a que al ser de menor tamaño, presentan mayor densidad celular (Ebel et al, 1993; Volz et al., 2003).

En general, un estrés hídrico moderado disminuye el rendimiento del cultivo pero favorece ciertos atributos de calidad de la fruta (Girona et al, 2005).

- **Poda y aclareo**

Estas prácticas culturales están relacionadas con el tamaño de los frutos. El aclareo aumenta el tamaño de los frutos pero disminuye el rendimiento total, por lo que tiene que alcanzarse un equilibrio (Crisosto and Mitchell, 2007). La carga de cosecha también afecta a la incidencia de alteraciones fisiológicas. El aclareo de flores y frutos se practica con fines comerciales para maximizar el valor de los cultivos mediante la optimización del tamaño comercial de los frutos, la mejora de los rendimientos, del color del fruto, de su forma y calidad, así como

para promover una mejor floración y para mantener la estructura del árbol (Byer, 2003).

- **Posición de la fruta**

Además de la carga de cosecha, la calidad de los frutos se ve afectada por la posición que éstos ocupan en el árbol. En manzanas, nectarinas y ciruelas, se han registrado grandes diferencias en el contenido de SST, acidez y tamaño entre los frutos de la parte exterior y los de la parte zona interior de la copa del árbol. La aplicación de sistemas de poda que permitan la penetración de la luz solar a las áreas centrales y bajas de la copa de los árboles son recomendables para reducir el número de frutas sombreadas (Flores, 2009)

1.3.- MADURACIÓN

El control de calidad comienza en el campo en el momento de la recolección. El estado de madurez en el momento de la recolección tiene gran influencia en la calidad del fruto. Cosechas demasiado tempranas proporcionan frutos de pequeño tamaño, color inadecuado, firmeza y acidez elevadas y baja concentración de sólidos solubles. Cosechas demasiado tardías ocasionan una pérdida importante de firmeza y acidez y sensibilizan al fruto a desórdenes fisiológicos. En el caso del melocotón, si se cosecha antes de que su estado fisiológico sea el adecuado, no podrá completar su evolución climatérica durante su conservación y su calidad será deficiente, el fruto perderá firmeza y no tendrá ni aroma ni sabor (Lill et al, 1989). Sin embargo, si se recolecta tardíamente, su vida útil se acortará y será más susceptible a los daños mecánicos y a desarrollar problemas de harinosidad. Para la manzana, cuando la recolección se realiza antes de tiempo, se desarrolla bitter pit y la coloración de la piel cambia hasta adquirir un tono amarillento. Por el contrario, una recolección tardía produce una descomposición interna del corazón (Murga y Palazón, 1984). Perring and Pearson, 1987, encontraron que la menor incidencia de bitter pit se producía cuando los frutos se recolectaban después de la iniciación del climaterio.

Los frutos son seres vivos que, mientras están en el árbol, respiran y se desarrollan gracias a los nutrientes de la planta madre. Una vez que se recolectan deben sobrevivir con sus propias reservas. Los frutos pueden ser climatéricos (manzanas, peras, melocotones y nectarinas), en los que la maduración puede completarse después de ser recolectados y no climatéricos (cítricos, uvas, fresas y cerezas) que deben madurar y adquirir la calidad organoléptica de consumo en la planta, puesto que su evolución, una vez separados de la planta, es muy lenta y no maduran adecuadamente.

En consecuencia, es necesario distinguir entre *madurez de recolección*, que sería el estado óptimo para recolectar, y la *madurez de consumo*, que sería cuando el fruto ha adquirido todas las características organolépticas propias que le hacen adecuado para el consumo (Kader, 2007).

El hecho de escindir el fruto del árbol implica el corte de suministro de agua con minerales y de azúcares y otros compuestos orgánicos e inorgánicos procedentes de la fotosíntesis. Por tanto, es necesario conocer el mejor momento para recolectar. Se considera el inicio de la formación del fruto, el momento de la antesis, cuando el polen penetra en el pistilo de la flor y a partir de ese momento se desencadenan una serie de procesos que regularán el crecimiento del fruto. Durante todo el período de crecimiento, el fruto aumentará de tamaño y de peso, y adquirirá la forma y el color característicos de la especie y la variedad. El crecimiento termina cuando deja de aumentar de peso o de tamaño. El crecimiento sigue en general una curva sigmoidea (Figura 3), que en el caso de la fruta de hueso es una doble sigmoidea (Figura 4), ya que tiene una parada durante el endurecimiento del hueso. Durante el crecimiento, además del aumento de peso, se produce un cambio en la estructura de los tejidos y en los componentes bioquímicos del fruto. Se producen cambios: en los pigmentos de la piel, cambiando el color del fruto; en los azúcares y ácidos que varían su sabor; en la estructura de las paredes celulares, que dan consistencia a la pulpa, y en la presencia de fenoles, que confieren astringencia a un fruto inmaduro, y de otras moléculas minoritarias como hormonas, vitaminas, antioxidantes, moléculas precursoras del aroma.

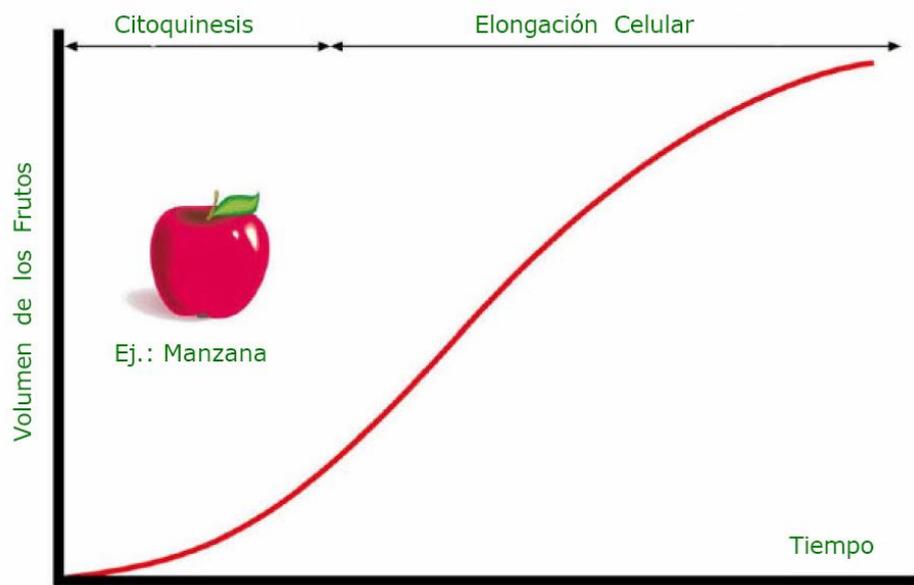


Figura 3. Curva sigmoidea que representa el crecimiento de la fruta de pepita

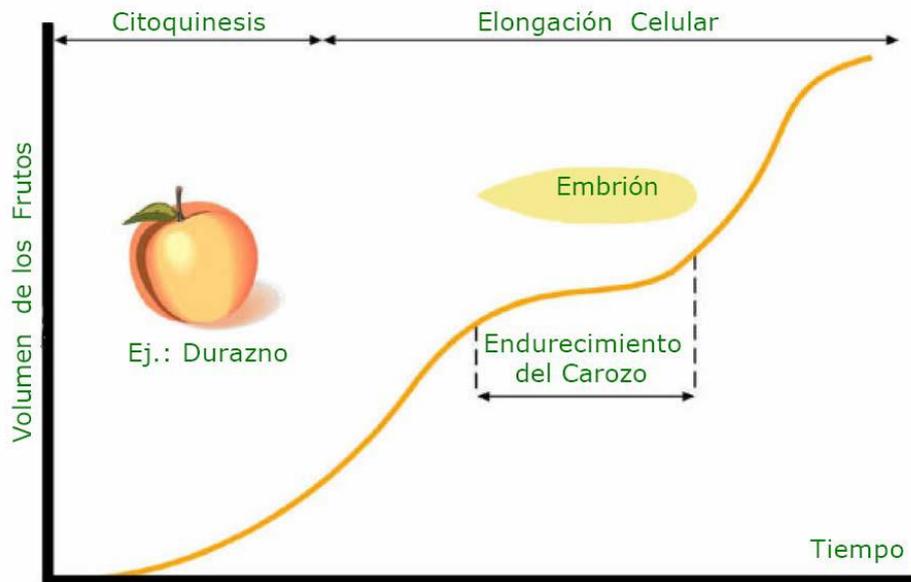


Figura 4. Curva doble sigmoidea que representa el crecimiento de la fruta de hueso

Tras la fase de crecimiento, comienza la de maduración, en la que el fruto sufre una serie de cambios bioquímicos que acaban en la senescencia (Wills, 1984; Lilien-Kipnis y Lavee, 1971; Chalmers and Van den Ende, 1975; Sedgley and Griffin, 1988)

Una vez que se ha alcanzado el momento óptimo de maduración se procede a la recolección del fruto.

Recolectar demasiado pronto o demasiado tarde puede provocar problemas (Recasens et al, 1989). Si se recolecta demasiado pronto es posible que los frutos no hayan acabado de desarrollarse, lo que implica, menor peso, menor calibre, y unas características organolépticas típicas de un fruto inmaduro. Si se recolecta demasiado tarde, los frutos son más sensibles a los golpes y a las fisiopatías.

La cosecha se puede hacer de forma manual o mecánica. A pesar de numerosas investigaciones y desarrollo de equipos mecánicos de recolección, la recolección manual sigue prevaleciendo sobre la mecánica (Thompson, 2007), siendo la ventaja más importante el hecho de que se puede seleccionar con mayor precisión la madurez de la fruta.

1.4.- EVALUACIÓN DE LA CALIDAD

1.4.1.- Métodos clásicos que implican la destrucción del fruto

Para ayudar a estimar el momento óptimo de madurez se tienen en cuenta los índices de madurez (Tabla 1) que se toman como indicadores de la fecha óptima de recolección.

Algunos parámetros que se usan como índices de madurez:

Según Westwood (1982), el índice de madurez es el indicio o señal que permite calcular el grado de madurez que ha alcanzado un fruto durante el proceso de desarrollo del cultivo y, según Kader (1985), los índices de madurez comprenden parámetros para identificar el momento propicio para efectuar la recolección.

Tabla 1. Índices de madurez para algunas frutas. (Fuente: Reid, 2007)

Índices de madurez	Frutas
Acidez	Granada roja, cítricos, papaya, melón, kiwi
Astringencia (contenido en taninos)	Caqui
Color externo	Todas las frutas
Concentración interna de etileno	Manzana, pera
Contenido de azúcar	Manzana, pera, frutas de hueso, uvas
Contenido de almidón	Manzana, pera
Contenido de aceite	Aguacate
Desarrollo de la capa de abscisión	Melón, manzana
Días transcurridos de floración a cosecha	Manzana, pera
Firmeza	Manzana, pera, fruta de hueso
Jugosidad	Cítricos
Peso específico	Cereza, sandía
Proporción azúcar/ácido	Melón, kiwi
Tamaño	Todas las frutas

En recolección, un fruto de calidad, dependiendo de la especie y variedad debe contar con una serie de características y valores mínimos de alguno de los siguientes parámetros:

- Físicas: calibre, color, firmeza y forma
- Químicas: contenidos en azúcares, ácidos y vitaminas
- Organolépticas: sabor, olor y textura

El *calibre* corresponde al contorno del fruto en su diámetro máximo. Es uno de los componentes más fácil de medir, aunque no es un carácter decisivo de la calidad organoléptica. No obstante, este parámetro tiene un impacto importante en la aceptación del consumidor, ya que en general se prefieren formas globulares sin protuberancias ni irregularidades (Badenes et al, 2006). Para su medida, se disponen de plantillas calibradores para cada variedad que ayudan a la recolección.

El *color* puede ser un índice de calidad y de madurez. Muchos frutos cambian el color de su piel cuando maduran. El color verde de la fruta se debe principalmente a la presencia de clorofilas que durante la maduración van degradándose y dejan visibles otros pigmentos como xantofilas (amarillos) o carotenos (rojos). El color de fondo puede ser un buen índice en variedades verdes, como las manzanas Golden, que muestran un viraje del verde al amarillo durante la maduración y en variedades que pierden el color verde y desarrollan un color rojo. Este parámetro tiene un impacto importante en la aceptación del consumidor y en su comercialización (Badenes et al, 2006).

El color de la fruta puede evaluarse con cartas de color de una manera subjetiva. También puede medirse de forma más precisa, en campo, con instrumentos como el colorímetro.

La *firmeza* se mide como la resistencia de la pulpa, una vez retirada la piel, a la penetración de un émbolo de diámetro conocido hasta una pequeña profundidad de la pulpa del fruto, este método se denomina Magness-Taylor. Esta medida puede hacerse con un penetrómetro manual o de sobremesa, la medida se expresa en newtons (N) o kilopondios (kp). También se puede medir con aparatos de mayor precisión como el texturómetro, en el que la velocidad con que se ejerce la penetración del émbolo es constante y no depende de la persona que realiza la medida. Esta medida está relacionada con la estructura de los tejidos y la estabilidad de las paredes celulares, lo que nos permite obtener datos sobre las propiedades texturales del fruto, relacionadas con la sensación producida cuando se muerde o mastica. La firmeza es uno de los parámetros más ligados con la maduración, de manera que cuando un fruto madura la firmeza va disminuyendo (Crisosto, 1994).

El *contenido en sólidos solubles* del fruto se determina mediante un refractómetro manual o de sobremesa. El valor que da un refractómetro es el porcentaje de sólidos solubles totales (p/p), que se expresa en porcentaje de sacarosa o grados Brix. Los azúcares van aumentando mientras el fruto está en el

árbol. La medida del contenido de azúcares individuales requiere de técnicas cromatográficas.

La *acidez* valorable es otro de los componentes principales del sabor y del aroma de los frutos (Colaric, 2005) que junto con los azúcares son los principales responsables del sabor. Los ácidos orgánicos presentes en las células del fruto son muy variados y se encuentran en muy bajas proporciones. En el caso de manzanas, peras y melocotones, el más abundante es el ácido málico. Los ácidos van disminuyendo durante la maduración ya que pueden ser sustratos para reacciones ligadas a la respiración. La medida de la acidez se lleva cabo mediante una valoración ácido-base neutralizando con NaOH.

Todos estos índices van variando durante la maduración, unos aumentan y otros decrecen. Normalmente hay que llegar a un compromiso entre todos ellos. En general, la firmeza debe encontrarse entre unos valores límites establecidos, los sólidos solubles deben de tener un valor mínimo y la acidez un valor máximo. La síntesis de compuestos aromáticos junto con la pérdida de firmeza de la pulpa, es una de las características más notorias durante el proceso de maduración de cualquier fruto (Crisosto, 1994; Crisosto et al, 1995)

El *aroma* característico de cada fruta está formado por un grupo de compuestos aromáticos altamente diversificado, algunos muy reactivos, y casi todos presentes en el fruto en concentraciones extremadamente bajas. Se empezó a investigar sobre técnicas de valoración de calidad basadas en la medición de estos compuestos aromáticos, como la Cromatografía de Gases. La ventaja de estas técnicas es su buena correlación con las características organolépticas de la fruta (López et al, 2000). Pero las dificultades que ordinariamente se encuentran en el análisis cualitativo y cuantitativo de los compuestos aromáticos, además de la complejidad instrumental que tiene su determinación, cuantificación e identificación de estos compuestos, no las hacen viables para un uso rutinario de control de calidad. En los últimos años se están desarrollando las llamadas “narices electrónicas”, capaces de reconocer aromas

en diversos productos agrícolas. Sin embargo, estos sensores no son todavía suficientemente específicos y salen del escenario de esta tesis.

1.4.2.- Métodos no destructivos

En los últimos años, los laboratorios de desarrollo y algunas centrales hortofrutícolas han introducido en las líneas de manipulación nuevas tecnologías, técnicas avanzadas no destructivas que presentan múltiples ventajas: no destrucción del producto, detección en el lineal de selección, control de todas las unidades, fiabilidad de las determinaciones, automatización de los procedimientos.

Entre las técnicas no destructivas se encuentran:

- Técnicas de impulso acústico: evalúan la firmeza mediante resonancia acústica.
- De compresión: consisten en medir la deformación que sufre el fruto cuando se le aplica una fuerza.
- Ópticas: Espectroscopia Infrarrojo Cercano y Análisis de Imágenes Multiespectrales.

Actualmente las técnicas más versátiles para la evaluación no destructiva de la calidad y clasificación de productos agrícolas son las ópticas, basadas en las propiedades ópticas de los productos (Chen, 1996). Se trata de métodos cuyo fundamento se basa en medir la interacción entre la materia y la energía en forma de luz.

Basándonos en sus propiedades ópticas, un objeto puede ser traslúcido, semitraslúcido, turbio y opaco. Los productos agrícolas son turbios y transmiten a través de ellos la luz (Mireei, 2010). Los fenómenos de reflectancia, absorción y transmitancia pueden ocurrir cuando la radiación electromagnética interacciona con un tejido biológico turbio (Qin, 2007). La luz reflejada es un fenómeno complejo y se manifiesta de tres formas: reflectancia especular, reflectancia difusa externa y retrodifusión de la luz (light backscattering) (Figura 5)

1.4.2.1.- Interacción luz-materia.

Cuando la luz incide sobre una muestra, se producen varios fenómenos:

- *Reflexión*
 - Reflexión especular: la radiación incidente cambia su dirección al llegar a la superficie y es reflejada sin opción a penetrar en el interior de la muestra. Esta señal contiene información de escasa utilidad para estimar parámetros de calidad de fruto.
 - Reflexión difusa: Porción significativa de la luz que tras incidir sobre la muestra e interactuar con las partículas de la superficie, es devuelta al primer medio. Contiene información espectroscópica de propiedades químicas de la muestra.
- *Refracción*: Desviación de la luz provocada por la variación de los índices de refracción de las partículas de diferente composición con las que interactúa.
- *Absorción*: parte de la radiación que penetra en la muestra es absorbida en función del tipo de enlaces que presenta la muestra.
- *Transmitancia*: otra parte de la radiación no absorbida, emerge por el otro lado de la muestra.
- *Retrodifusión de la luz (light backscattering)*: La difusión se produce tras múltiples refracciones debido a cambio de fases en el interior del material, llegando a volver al primer medio. Mientras que la absorción se debe a algunos constituyentes químicos de la fruta (como son los pigmentos: clorofila, carotenos y antocianos), la difusión está influenciada por la densidad, composición, células y estructuras intracelulares del tejido, y en consecuencia puede ser útil para medir las características mecánicas del mismo.

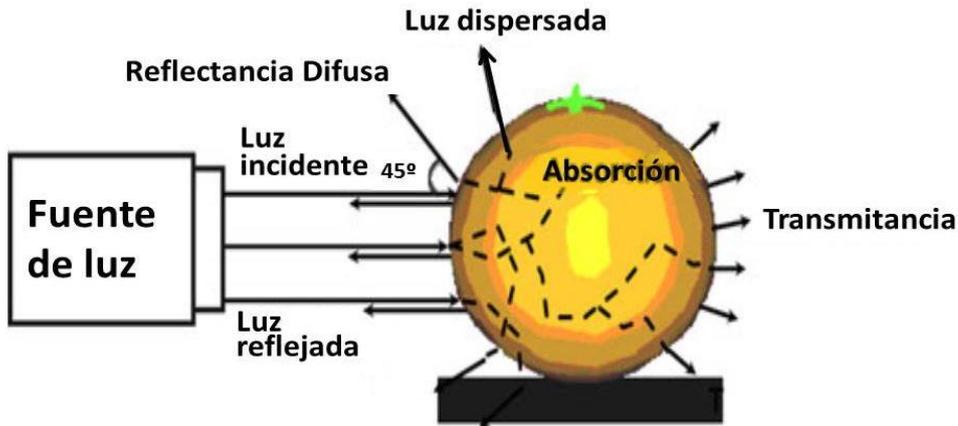


Figura 5. Fenómenos producidos por la luz reflejada

1.4.2.2.- Espectroscopia Vis/NIR

La espectroscopia de reflectancia Visible/Infrarrojo cercano (Vis/NIRS) presenta un enorme potencial para la caracterización de una gran variedad de productos de forma instantánea. En esta técnica, la región del visible (400 nm a 700 nm) nos da información acerca de material colorante de la muestra. Los pigmentos son los responsables de la coloración debido a la presencia de bandas de absorción en el rango visible. En el infrarrojo, en cambio, se utiliza la región de longitudes de onda entre los 780 y los 2500 nanómetros del espectro electromagnético (Shenk et al, 2001) (Figura 6).

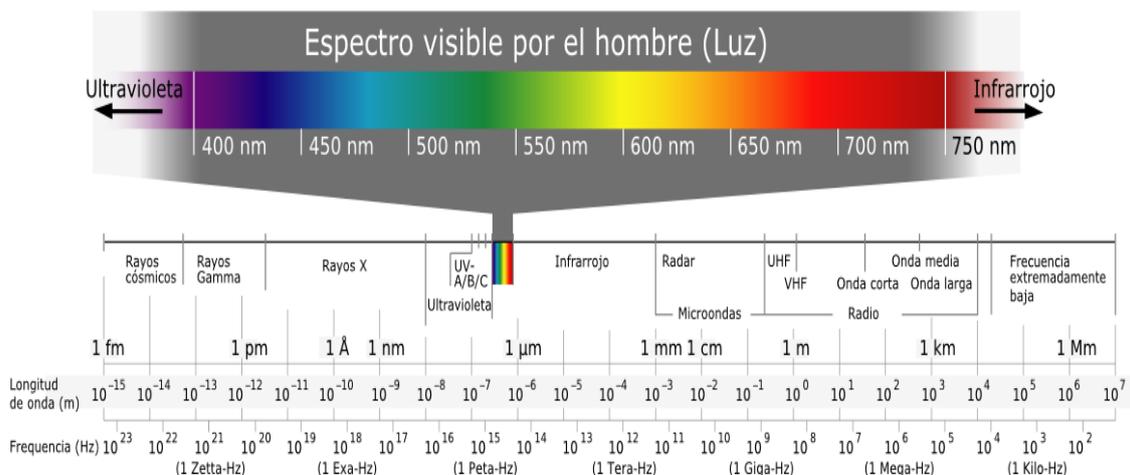


Figura 6. Espectro electromagnético en el rango de la luz visible y zonas adyacentes.

El descubrimiento de la radiación infrarroja fue realizado por William Herschel en 1800 cuando estudiaba el espectro de la luz solar. Pero no fue hasta 1881 cuando Abney y Festing documentaron los primeros espectros de líquidos orgánicos en el intervalo espectral 700 a 1200 nm (Abney y Festing, 1881). En la década de los 50 del pasado siglo, como consecuencia de los primeros espectrofotómetros comerciales NIR, se desarrollaron las primeras aplicaciones analíticas en esta región del infrarrojo cercano (Kaye, 1954). Sin embargo, el impulso definitivo de la tecnología y su aceptación como método analítico por la comunidad científica se produce en la década de los 60, gracias a los trabajos desarrollados por Karl Norris, quien comienza a experimentar las posibilidades de utilización de la espectroscopía NIR en el campo del análisis agroalimentario (Norris, 1964; Ben-Gera, 1968). A partir de los años 80, y sobre todo en los 90, esta tecnología experimentó un mayor desarrollo y sus aplicaciones se han extendido a nuevos campos, debido al trabajo de investigación y desarrollo llevado a cabo por grupos de investigación de distintos países.

La tecnología NIRS consiste en la incidencia de un haz de luz sobre la muestra y, en función de la naturaleza de los enlaces presentes en sus moléculas, fundamentalmente de aquellos tipo: -CH, -NH, -OH y -CO, interaccionará con ellos que absorberán una determinada cantidad de radiación NIR (Burns and Ciurczak, 1992)

La absorción de la energía por la materia obedece a la ley de Lambert-Beer, que establece que la absorbancia a cualquier longitud de onda es proporcional al número o concentración de las moléculas absorbentes presentes en el camino recorrido por la radiación.

$$A = \log \frac{1}{R} = \varepsilon \times \ell \times c$$

donde **A** es la absorbancia aparente y **R** la reflectancia relativa ($R=R_{\text{muestra}}/R_{\text{referencia}}$), ε coeficiente de extinción molar, **c** la concentración y **l** es el espesor de la muestra.

Esto determina que, para un material de naturaleza química heterogénea, el espectro obtenido en la región del infrarrojo cercano sea la combinación de bandas de absorción parciales superpuestas o muy cercanas, que suelen confundirse en una línea suavizada, en la que se encuentran picos, valles y curvaturas en forma de hombro.

La interpretación espectral de las bandas de absorción de los espectros NIR, presenta una gran complejidad, en comparación con los espectros de absorción del infrarrojo medio donde los picos o bandas obtenidas pueden ser interpretados directamente mediante su comparación con librerías espectrales de compuestos puros.

Además, tanto las características físicas del instrumento (compartimento de la muestra, geometría óptica del instrumento, etc) como las de la muestra (tamaño, forma, distribución las partículas, índices de refracción, etc) influyen en el espectro NIR de una determinada muestra. Como consecuencia, es imposible obtener un espectro NIR sin que se produzca dispersión de la radiación incidente (Burns and Ciurczak, 1992; Bertrand, 1993; Bertrand and Dufour, 2000). Esto hace prácticamente imposible elaborar librerías de espectros para cada instrumento.

- **Instrumentación**

Los primeros espectrofotómetros comerciales NIR aparecieron en la década de los 50. El esquema básico de estos instrumentos era el mismo que el utilizado en los espectrofotómetros UV – Visible: fuente de radiación, sistema de selección de longitudes de onda, compartimento para la muestra y detector. A continuación se describen con más detalle las diferentes partes integrantes de un instrumento NIRS.

Fuente de radiación

Actualmente las fuentes de radiación utilizadas en espectroscopia NIR se pueden clasificar en dos tipos, en función del tipo de radiación emitida:

- Espectro completo: la fuente más usada sigue siendo la lámpara halógena de wolframio con ventana de cuarzo. Se trata de una fuente de elevada intensidad de emisión y que cubre ampliamente la zona NIR del espectro electromagnético, proporcionando una radiación continua en el intervalo 320 – 2500 nm.
- De bandas discretas: se encuentran los diodos de emisores de luz o LEDs (*Light emitting diodes*). Se trata de dispositivos semiconductores que emiten radiación NIR en una banda estrecha de longitudes de onda. Entre los semiconductores más utilizados para NIRS, se encuentran los de Ga-As, que emiten en la región entre 900 y 970 nm y los de In-Ga-As que permiten llegar a longitudes de onda más largas (1600 nm) (Romero, 2001).

Sistema de selección de longitudes de onda

Son dispositivos que permiten la selección de una banda estrecha de longitudes de onda del espectro que proporciona la lámpara. Se clasifican en dos tipos: dispersivos y no dispersivos.

Dentro de los *sistemas dispersivos* se encuentran los monocromadores. En general, estos están formados por una rendija de entrada, que proporciona un haz estrecho de luz, un colimador, que hace paralela la radiación, el elemento dispersante de la radiación y un segundo colimador para enfocar la radiación sobre la rendija de salida. Lo que caracteriza a cada monocromador es el elemento dispersante. El más utilizado actualmente en espectroscopia NIR es la red de difracción. Este dispositivo provoca que cada longitud de onda de la radiación policromática incidente se difracte en diferentes ángulos en cada una de las estrías marcadas en la superficie de la red, provocando interferencias constructivas y destructivas. El barrido de longitudes de onda discretas sobre la

muestra se realiza de manera que en cada momento una longitud de onda seleccionada sea enfocada sobre la rendija de salida del monocromador.

El conjunto de *sistemas no dispersivos* es más amplio. Se dispone de instrumentos con filtros convencionales, filtros optoacústicos e instrumentos que ofrecen directamente la transformación de Fourier, con características bastante diferenciadas entre sí.

En los sistemas de filtros convencionales la selección de longitudes de onda se realiza interponiendo un elemento entre la fuente de radiación policromática y la muestra, que permite el paso de determinadas longitudes de onda. Los filtros más sencillos son los de absorción (Schenk and Diffe, citados por Pou, 2002), en los que las longitudes de onda transmitidas dependen del material con que estén hechos. Otro tipo de filtros utilizados son los de interferencia en los que las longitudes de onda transmitidas dependen del índice de refracción del material y del espesor óptico del filtro.

Los filtros optoacústicos están basados en la interacción del material del filtro con las ondas sonoras. Cuando el haz de luz policromática incide en el cristal, al aplicar una señal acústica, además del haz original se generan dos haces de luz monocromática polarizada, de los cuales solo se utiliza uno con fines analíticos.

Estos sistemas son rápidos, con una alta reproducibilidad de la longitud de onda y robustos, lo que los hace especialmente útiles en el desarrollo de instrumentación de planta industrial. A pesar de ello, su uso no se ha extendido debido al coste económico.

Los instrumentos NIR de transformada de Fourier (FT-NIR) se basan en la utilización de interferómetros como sistema de selección de longitudes de onda. El interferómetro más usado es el de prismas, que consiste en dos prismas de cuarzo de distinto tamaño que se encuentran enfrentados. Uno de ellos es móvil y el otro fijo.

Detectores

Los detectores utilizados inicialmente en espectroscopia NIR son detectores de fotones (fotoconductores, fotodiodos y fotomultiplicadores) y

detectores térmicos, cada vez menos utilizados. Actualmente, los detectores más comunes en los equipos NIRS son líneas (arrays) de diodos fabricados con los materiales que se indican en la Tabla 2.

Tabla 2. Materiales comúnmente empleados en equipos NIRS (adaptado de Workman, 2004)

Materiales	Intervalo espectral de detección (nm)
InAs (Arseniuro de Indio)	1700-5700
InGaAs (Arseniuro de Indio y Galio)	900-1700
InSb (Antimoniuro de Indio)	1800-6800
MCT (Teluro de Cadmio y Mercurio)	1000-1700
PbS (Sulfuro de Plomo)	1100-3000
PbSe (Seleniuro de Plomo)	1700-5500
PbTe (Teluro de Plomo)	1500-4500
Silicio	300-1100

Un aspecto importante es la disposición de los sistemas detectores. Para medidas de transmitancia, es suficiente situar el detector en línea con la muestra y el haz incidente. En las medidas de reflectancia, se quiere captar la radiación difusa reflejada por la muestra y colocando de modo que la luz que recibe de la muestra tenga una dirección que forme un ángulo con la radiación incidente, 45° normalmente (CIE, 2004).

Como se ha comentado, los instrumentos Vis-NIRS permiten registrar espectros de muestras sólidas y líquidas. La variabilidad en la presentación de la muestra conlleva que exista una gran variedad de accesorios para presentar la muestra al instrumento. En función de la naturaleza y tipo de producto puede ser más apropiado un modo de análisis u otro.

1.4.2.3.- Retrodifusión de la luz (light backscattering)

La reflexión especular se produce cuando el ángulo de la luz incidente sobre la superficie del objeto es igual al ángulo en el que es reflejado, mientras

que en la reflexión difusa externa la salida es a cualquier otro ángulo. Esta reflectancia lleva información de la superficie del objeto como es el color (Mireei, 2010). Solamente entre 4-5% de la luz incidente es reflejada de ambas formas. El resto de luz pasa hacia el interior del objeto (Birth, 1976)

Gran parte de la luz que penetra es reflejada por componentes biológicos internos y es dispersada hacia la superficie del objeto. Los fotones retrodifundidos llevan información acerca de la morfología y estructura de los tejidos. Lu (2004) manifestó que la luz dispersa es un fenómeno físico que depende del tamaño de célula y de las propiedades de la matriz del objeto. Los principales elementos de difusión en fruta y vegetales son las 'interfaces' en las paredes de la células puesto que producen cambios abruptos en el índice de refracción (McGlone et al., 1997), pero las partículas en suspensión, como los gránulos de almidón, cloroplastos y mitocondrias pueden también inducir difusión por difracción en las superficies de la partículas donde el índice de reflexión es diferente al de los alrededores. La difusión depende del tamaño, de la forma y de la microestructura de las partículas.

Numerosos trabajos han descrito el fenómeno físico de la retrodifusión de la luz para predecir índices de calidad en fruta y hortaliza, como los de Qing et al, 2007b; Peng and Lu, 2006; Tu et al, 1995 en tomate y manzana usando un láser de diodo a 670nm, o McGlone et al, 1997 determinando la firmeza en kiwi mediante un láser de 864 nm.

- **Instrumentación**

Una imagen es una representación visual de un objeto que emite luz o de un objeto iluminado por una fuente radiante. Las que se perciben en las actividades cotidianas provienen normalmente de la luz reflejada por los objetos. La naturaleza básica de una imagen está caracterizada por tres componentes: la distribución espectral de la luz incidente que procede de la fuente que ilumina la escena contemplada; la cantidad de luz reflejada por los objetos de la escena y el dispositivo formador de imagen: ojo, cámara, etc.

En el proceso de formación de una imagen digital intervienen los siguientes elementos: la fuente radiante, el objeto y el sistema de formación de la imagen que consiste, básicamente, en un sistema óptico, un sensor y un digitalizador. Se debe indicar que los sensores actuales permiten obtener imágenes en el visible y en el infrarrojo.

Fuente de luz

La fuente de luz debe tener la capacidad de producir un haz de luz en continuo. La composición espectral de la radiación, el tamaño del haz de luz y su ángulo de incidencia en la muestra son los principales factores para seleccionar el sistema de iluminación.

Los principales diámetros del haz de luz utilizados por la mayoría de los investigadores (manzana, Peng and Lu, 2005; melocotón, Lu and Peng, 2006; kiwi, Baranyai and Zude, 2009) son 1.6, 1.5 y 1 mm.

Como el objetivo es sólo capturar los fotones retrodispersos, el ángulo del haz incidente debe ser seleccionado de modo que la radiación especular no sea registrada por el detector. Un ángulo incidente entre 7 y 25° con respecto al eje vertical es el más recomendado (Lu, 2004; Qin and Lu 2007a, b).

Sensor óptico

Es un dispositivo capaz de capturar una escena y digitalizarla (puede ser una cámara fotográfica digital). Las cámaras más populares están basadas en sensores CCD (dispositivo de carga acoplada), que consisten en una serie de sensores (píxeles) compuestos cada uno de ellos por una fotocélula y un capacitor (Peterson, 2001). Las cámaras CCD pueden ser monocromáticas o capaces de registrar el color. Una cámara CCD de color es menos sensible que una monocromática. El sensor se coloca en la normal de la superficie iluminada. Cuando el haz de luz iluminante es pequeño y casi perpendicular a la superficie del objeto la imagen adquirida es prácticamente simétrica respecto al punto central de incidencia. Esto hace que el procesado de las imágenes sea más fácil.

Retrodifusión de la luz

En las técnicas de *backscattering*, en función del tipo de fuente de luz y del detector, el sistema se puede dividir en dos principales técnicas:

- Imágenes de retrodifusión de luz monocromática o láser
- Imágenes de retrodifusión de banda ancha que incluye: imágenes multiespectrales de retrodifusión e imágenes hiperespectrales de retrodifusión. (Lu and Peng, 2006; Qing et al, 2007a, 2008)

Todas estas técnicas tienen los mismos principios teóricos, pero hay diferencias en cuanto a la adquisición y tratamiento de las imágenes:

1. En una primera etapa del proceso, se adquiere la imagen digital utilizando sensores ópticos con capacidad para producir información numérica de la señal producida.
2. En segundo lugar se realiza el preprocesamiento de la imagen donde se aplican filtros y algoritmos mediante un programa adecuado.
3. A continuación, y con el mismo programa, se realiza un procesado de la imagen, que consiste en la división del fotograma en las distintas partes que la componen, de las que se obtienen regiones de interés en la imagen (ROI). Posteriormente, se segmenta el ROI clasificando cada pixel como perteneciente a una clase determinada en función de su nivel de gris, siguiendo algunos de los diversos métodos descritos en la bibliografía (González and Woods, 2002; Russ, 2006). La salida del proceso de segmentación es una imagen que contiene la frontera de la región o los puntos que forman la región. La representación por la frontera es apropiada cuando el objetivo se centra en las características de la forma externa como esquinas o concavidades y convexidades. La representación por regiones es apropiada cuando la atención se centra en propiedades internas, como la textura del fruto.
4. El paso siguiente es la parametrización que se dedica a extraer rasgos que producen las regiones de interés o rasgos que son básicos para diferenciar una clase de objetos de otra.

5. Por último, se realiza el reconocimiento y la interpretación: el reconocimiento es el proceso que asigna una etiqueta a un objeto basada en la información que proporcionan los descriptores (clasificación). La interpretación lleva a asignar significado al conjunto de objetos reconocidos.

1.5.- TRATAMIENTO DE DATOS

El término 'Quimiometría', desde hace 30 años aproximadamente, quiere resumir el concepto que engloba la medida en química. La Quimiometría trata, específicamente, de todos aquellos procesos que transforman señales analíticas y datos más o menos complejos en información. La Quimiometría empezó utilizando métodos ya existentes de origen matemático y/o estadístico y otros procedentes del campo de la lógica formal para conseguir sus fines de reducción de datos, métodos que se han ido ampliando con el paso del tiempo. Por todo ello, la Quimiometría tiene usos interdisciplinares (Massart et al, 1997; Wold and Sjöström, 1998).

La gran superposición de las bandas de absorción de los diferentes constituyentes presentes en la muestra, la gran cantidad de datos registrados y la elevada información redundante (colinealidad) provocan una escasa selectividad de los datos con los que se trabaja. Todo esto implica que sea necesario recurrir a métodos estadísticos de análisis multivariante. El análisis multivariante podría definirse como el conjunto de métodos, matemáticos, estadísticos o gráficos, empleados en el análisis de datos siempre y cuando se consideren varias variables de forma simultánea (Martens and Naes, 1989; Martens and Martens, 2001; Naes et al, 2002). En el caso de las tecnologías ópticas, el análisis multivariante se utiliza para estimar cualquier propiedad de la muestra a partir de múltiples variables espectrales o de análisis digital de imagen.

En estos datos suelen aparecer contribuciones o efectos no deseados, conocidos como ruido, que son debidos bien al proceso de registro de la señal, bien a la naturaleza misma de la muestra o bien al ruido ambiental.

Con el fin de extraer la información química más relevante de cada muestra se hace necesario recurrir a pretratamientos de la señal espectral, ya que permiten separar la información meramente química de las variaciones de origen físico.

Los pretratamientos más tradicionalmente utilizados se suelen dividir en cuatro grupos (Gaitán-Jurado, 2009):

- Reducción de ruido
- Corrección de la línea base: buscan minimizar o eliminar los efectos multiplicativos y aditivos. Los métodos más utilizados para ello son:
 - *Derivadas*; método Norris (Norris and Williams, 1984) y el de Savitzky-Golay (SG) (Savitzky and Golay, 1964).
 - *Multiplicative Scatter Correction (MSC)* (Geladi et al, 1985)
 - *Standard Normal Variate (SNV) y Detrending (DT)* (Barnes et al, 1989)
 - *Orthogonal Signal Correction (OSC)* (Wold et al, 1998)
- Mejora de resolución espectral: buscan eliminar el solape entre bandas espectrales. Destacan las derivadas y las sustracciones espectrales (Ozaki et al, 2007)
- Métodos de centrado y normalización: los primeros buscan realizar un ajuste al conjunto de datos para reposicionar el centroide de los mismos en el origen de coordenadas y la normalización trata de igualar los valores de las diferentes variables de cada muestra (Downey et al, 1992)

Después de estos métodos de pretratamiento, para completar el estudio de los datos se realiza un Análisis de Componentes Principales (PCA: Principal Components Analysis), que es uno de los métodos más empleados como etapa inicial de cualquier análisis multivariante para estudiar la población de muestras con la que se va a trabajar (Martens and Martens, 2001). La finalidad más importante del PCA es intentar reducir la dimensionalidad de la matriz X de datos. Para ello, busca las direcciones ortogonales que expliquen la máxima

variabilidad de las muestras y las utiliza como nuevos ejes de coordenadas. Estos nuevos ejes se llaman Componentes Principales (CPs) (Jackson, 1991). El primer CP es el que explica la máxima variabilidad. El segundo se escoge de tal forma que sea perpendicular al primero y que explique la máxima variabilidad una vez eliminada la explicada por el primer CP y así sucesivamente. Para poder definir matemáticamente estos nuevos ejes se usan las cargas (*loadings*), que son las coordenadas de los nuevos ejes en la antigua base y las coordenadas de las muestras en estos nuevos ejes se denominan puntuaciones (*scores*).

El PCA también se utiliza para detectar muestras anómalas (*outliers*), obtener información sobre las variables más importantes y detectar posibles agrupamientos entre las muestras.

1.5.1.- Métodos de regresión

Tras los procesos matemáticos de pretratamiento de la señal espectral, los siguientes pasos están orientados a desarrollar una ecuación de calibración capaz de predecir parámetros de otras muestras con características similares a las incluidas en el colectivo de aprendizaje.

Después, se procede a calcular una ecuación de calibración capaz de predecir parámetros de otras muestras con características similares a las incluidas en el colectivo de calibración (Shenk and Westerhaus, 1995 y 1996; Williams and Sobering, 1992).

Un buen método de calibración debe satisfacer estos requerimientos:

- Los modelos matemáticos deben realizar una buena estimación
- Resolver el problema de colinealidad, que aparece cuando se trabaja con información espectral NIRS, la que provoca inestabilidad en las predicciones (Naes, 1992; Pérez-Marín et al, 2007).

Una etapa importante en la obtención de la ecuación de calibración es la selección de muestras para el colectivo de calibración. Cuanto mayor sea el

número de muestras en este colectivo mayor será la capacidad de predicción de la ecuación resultante, aunque un número excesivo puede producir sobreajuste e inutilizar los modelos de predicción. Además del número, es importante que la variabilidad físico-química y óptica cubra el rango de variación de la población de la que se toman las muestras. Los datos de referencia de las muestras obtenidos en el laboratorio deben ser lo más precisos posibles, ya que los resultados obtenidos mediante un método secundario como las técnicas ópticas tendrán, como mucho, una precisión similar al del método de análisis convencional usado como referencia.

La selección de muestras para constituir el colectivo de calibración puede ser realizada a través del uso de herramientas matemáticas diseñadas para la estructuración de la población y la elección de muestras representativas del colectivo.

Para el desarrollo de una calibración se han descrito distintos métodos de regresión siendo los más utilizados:

1.5.1.1.- Métodos lineales

- Regresión Lineal Múltiple (RLM) (Shenk and Westerhaus, 1995, 1996; Puigdomenech, 1998) lleva a cabo una regresión a partir de un pequeño número de variables seleccionadas de entre todo el conjunto inicial.
- Regresión por Componentes Principales (RCP), combinación lineal de los datos espectrales obtenidos mediante PCA y empleo de estas nuevas variables en el desarrollo de la ecuación de regresión.
- Regresión por Mínimos Cuadrados Parciales (PLS)

El método de regresión por mínimos cuadrados parciales se introdujo hace unos ochenta años por Wold (Wold, 1982) y actualmente se utiliza en Espectroscopia, para extraer información de espectros complejos. El

PLS es especialmente útil para el análisis de datos multivariantes con un gran número de variables independientes. Varios trabajos describen el algoritmo básico PLS y las propiedades de los factores PLS como Geladi and Kowalski, 1986; Naes et al, 1986. El propósito del análisis PLS es encontrar una relación matemática entre un conjunto de variables independientes, la matriz X, y conjunto de variables dependientes, la matriz Y. El modelo resultante tiene la forma:

$$Y=XB+E$$

Donde B es la matriz de los coeficientes de regresión obtenida del análisis PLS y E es una matriz de residuales

En el caso de que exista colinealidad o redundancia entre las variables, la matriz B se usa para reducir dichas variables o sintetizarlas. Estas variables latentes son calculadas como combinaciones lineales de las variables de independientes.

A partir de este momento, PLS se puede convertir en una herramienta de predicción en la que se utiliza el modelo construido en la fase de entrenamiento junto con nuevas variables de entrada para predecir las variables dependientes.

1.5.1.2.- Métodos no lineales

- Least squares support vector machine (LS-SVM)

Es un método de regresión muy utilizado en los últimos años para predecir parámetros relacionados con la madurez de la fruta y otras propiedades físicas y químicas. Tiene la capacidad de trabajar con las relaciones lineales y no lineales que se establecen entre las variables de estudio y resolver estos problemas de forma rápida (Suykens and Vandewalle, 1999). Estudios previos han demostrado el potencial de este método de regresión no lineal para determinaciones cuantitativas en

productos agrarios (Sun et al., 2009, Shao et al., 2011 and Chauchard et al., 2004).

- Artificial Neural Networks (ANN)

- ANN es otro método de regresión no lineal. Se puede definir como un sistema iterativo de cálculo que intenta reproducir, de forma simple y sencilla, el sistema de conexiones que existe entre las neuronas del cerebro humano.
- Un proceso llevado a cabo con ANN consta de dos etapas: la de aprendizaje (learning) y la de respuesta. Durante el aprendizaje, la red neuronal 'aprende' a partir de los ejemplos que se le presentan, adaptando los pesos de las conexiones en respuesta a las señales que le llegan de la capa de entrada y, opcionalmente, de la respuesta deseada. La etapa de respuesta se refiere a cómo la red procesa globalmente las señales que llegan a su capa de entrada y proporciona la respuesta en la capa de salida.
- Hay diferentes tipos de redes neuronales y se utilizan unas u otras dependiendo del tipo de problema que se desee resolver. (Pérez-Marín et al, 2007; Taghadomi-Saberi et al, 2013).

1.5.2.- Selección de variables

En las técnicas ópticas, especialmente en Espectroscopia de Infrarrojo Cercano, se trabaja con un gran número de variables. Es por eso, que para la obtención de los modelos de calibración se utilizan métodos de regresión multivariante. El más utilizado es el método Partial Least Squares (PLS) que tiene la habilidad de reducir problemas de colinealidad y solapamiento de bandas (Xiaobo et al, 2012). Pero éste no evita la influencia de las variables que aportan información irrelevante en la construcción del modelo.

Por tanto, la selección de variables es muy importante en los métodos multivariantes, ya que la eliminación de las variables que aportan ruido y poca

información, así como las redundantes, ayudará a la obtención de modelos de calibración más precisos y sencillos y desarrollo de sistemas de medida más baratos (Andersen y Bro, 2010)

Los métodos de selección de variables están basados en la elección de un pequeño número de variables, para así obtener un conjunto con el que el error sea mínimo en la obtención del modelo (Xiabo et al, 2012).

La selección de variables comenzó teniendo en cuenta sólo el espectro medido y viendo cuales eran las más importantes mediante los coeficientes de regresión. En los últimos años se han desarrollado estrategias estadísticas y heurísticas diversas para esta elección. Todo método de selección de variables sigue un proceso de dos pasos: primero es detectar un conjunto óptimo de variables relevantes, y luego desarrollar un algoritmo de calibración que permita la optimización del modelo.

Existen diferentes estrategias para la selección de variables:

- Univariante: selecciona las variables que tienen la correlación mayor con la variable respuesta. Sin embargo esta estrategia no considera el fenómeno de la colinealidad.
- Selección multivariante: las variables de entrada son eliminadas progresivamente de un modelo.
 - Selección hacia adelante
 - Selección hacia atrás
 - Selección paso a paso (mezcla de las anteriores)

Estas estrategias son genéricas, pudiéndose encontrar algoritmos híbridos entre estas diferentes estrategias, y no son únicos, existiendo estrategias específicas distintas de estas.

Muchos son los métodos de selección de variables desarrollados: Successive Projections Algorithm (SPA), Uninformative Variable Elimination (UVE), Genetics Algorithm (GA), Interval Partial Least Square (iPLS), Competitive Adaptive

Rewighted Sampling (CARS), siendo sin embargo la Información Mutua (MI) un criterio muy utilizado para el diseño de algoritmos de selección de variables. Esta medida, proveniente de la Teoría de la Información de Shannon, es una medida de correlación no lineal entre grupos de variables, que pretende medir el nivel de información que tiene una variable (o grupo de variables) respecto de otra variable (o grupo).

1.6.- VENTAJAS E INCONVENIENTES DE LOS MÉTODOS NO DESTRUCTIVOS

Las principales ventajas de estas técnicas ópticas son:

- No son destructivas ni invasivas.
- Permiten, en muestras vivas, seguir la evolución de los parámetros a lo largo del tiempo.
- La preparación de la muestra es nula
- La obtención de los espectros y las imágenes es rápida.
- El análisis de cada muestra presenta un bajo coste.
- Aumenta la capacidad analítica del laboratorio debido a la rapidez con la que se obtienen los resultados.
- Permite el análisis de varios parámetros de forma simultánea.
- Es una técnica idónea para procesos de control en planta debido a la ausencia de partes móviles en el aparato. Además hay una tendencia hacia la miniaturización y compactación de los instrumentos.

Como todas las técnicas también tienen inconvenientes:

- El precio de los equipos es elevado.
- La complejidad de la señal NIR y de las imágenes multiespectrales obliga a aplicar técnicas estadísticas complejas para la obtención de los modelos.
- Es necesario un modelo de calibración distinto para cada parámetro que se necesite medir.
- El calibrado es difícil debido al gran número de muestras que es necesario medir para que el modelo de calibrado sea fiable

2.- *Objetivos*

2.1.- OBJETIVOS GENERALES

El objetivo general de este trabajo de investigación es evaluar el potencial de dos técnicas, Espectroscopia de Infrarrojo Cercano y Análisis de Imágenes Multiespectrales, para su aplicación como herramientas de estimación de los parámetros más importantes de calidad del fruto, con el fin de determinar el punto óptimo de recolección por métodos no destructivos de muestra.

2.2.- OBJETIVOS ESPECÍFICOS

1. Desarrollo y evaluación de modelos de predicción NIRS para determinar parámetros de calidad en melocotón.
2. Desarrollo y evaluación de modelos de predicción NIRS para determinar parámetros de calidad en cereza.
3. Desarrollo y evaluación de nuevos modelos de predicción obtenidos por Análisis de Imágenes Multiespectrales para determinar parámetros de calidad en manzana.
4. Optimización y validación de los modelos de predicción calculados por métodos de regresión lineales y no lineales
5. Implementación y validación de métodos de selección de variables para simplificar y mejorar la capacidad predictiva de los modelos de calibración

Para alcanzar estos objetivos se han redactado tres capítulos, en los que se estudia cada una de las especies frutales por separado.

3.- *Materiales y Métodos*

Los experimentos desarrollados en este trabajo, y que conforman el cuerpo principal de la Tesis Doctoral, se describen en detalle en los capítulos 4, 5 y 6. Cada capítulo contiene un apartado dedicado específicamente a los materiales y métodos utilizados en esa parte. No obstante, con el fin de facilitar la comprensión de la totalidad del trabajo y simplificar la exposición del mismo, se ha incluido en este capítulo una descripción de: todos los colectivos muestrales analizados, la instrumentación utilizada, los métodos y análisis de laboratorio realizados, los métodos estadísticos y los softwares empleados.

3.1.- MATERIAL VEGETAL

En esta tesis se ha trabajado con tres frutas distintas: dos de hueso (cereza y melocotón) y una de pepita (manzana). Estas muestras proceden de localidades diferentes.

La cereza, procedía de una explotación comercial propiedad de Orchard S. L. situada en la localidad de la Almunia de Doña Godina (Zaragoza). La variedad con la que se ha trabajado ha sido Chelan sobre patrón híbrido.

Chelan fue lanzada a comienzos de los años noventa. Es la variedad de maduración más temprana del programa de mejora de la Washington State University que se recolecta de 10 a 12 días antes que Bing o Van. Chelan pertenece al grupo III de compatibilidad de polen y florece 2 a 4 días antes que Bing. Los frutos de Chelan tienen excelente firmeza y un tamaño moderado, con un promedio de diámetro de 24 a 26 milímetros (8 a 10 g). Para alcanzar buen tamaño de fruta, se recomienda poda agresiva en la dormancia invernal dada la alta productividad y precocidad de esta variedad que, además, muestra buena resistencia al rajado por lluvia.

La manzana fue cultivada en una parcela de la finca experimental de la Estación Experimental de Aula Dei (EEAD-CSIC) en Zaragoza. La variedad utilizada ha sido Smoothie Golden Delicious, se caracteriza por ser una variedad con el epicarpio fino y brillante, su color varía desde el verde al amarillo, según el

estado de maduración. Su carne es crujiente, jugosa, dulce y aromática. Los árboles de esta finca han sido injertados sobre Pajam1.

El melocotón provenía de una finca experimental propiedad de AFRUCAS (Asociación de Fruticultores de la Comarca de Caspe), situada en el paraje denominado “Mas de la Punta” del término municipal de Caspe. La variedad utilizada ha sido Calrico sobre GF677.

La variedad Calrico es uno de los clones más interesantes de la segunda prospección dentro de la variedad población Amarillos tardíos del Bajo Aragón, cuya colección se encuentra ubicada en Alcañiz (Teruel). Procedentes de esta colección, en 1999, el CTA de Aragón seleccionó y protegió los clones Jesca, Calante y Evaisa, actualmente autorizados por la DO Melocotón de Calanda, y en 2008 solicitó la protección de cuatro nuevos clones: Calejos (procedente de la primera prospección), Calemil (AL-19), Calprebor (AL-1) y Calrico (AL-4). Esta colección tiene un gran interés genético ya que comprende gran parte de la variabilidad de la población (Alonso y Espada, 2011).

La toma de muestras se realizó durante dos campañas, 2010 y 2011 para melocotón y 2011-2012 para cereza y manzana.

En la Tabla 3 se describen las distintas muestras utilizadas en los diferentes estudios que se han llevado a cabo.

Tabla 3. Muestras utilizadas en los experimentos

	Primera campaña			Segunda campaña		
	Variedad	Total muestras	Período de recolección	Variedad	Total muestras	Período de recolección
Melocotón	<i>‘Calrico’</i>	150	Quincenal	<i>‘Calrico’</i>	110	Semanal
Manzana	<i>‘Smoother Golden Delicious’</i>	215	Quincenal	<i>‘Smoother Golden Delicious’</i>	150	Quincenal
Cereza	<i>‘Chelan’</i>	150	Semanal	<i>‘Chelan’</i>	132	Cada 2 días

3.2.- INSTRUMENTACIÓN ÓPTICA

La información técnica sobre los equipos detectores de las señales ópticas (espectrofotómetros NIR y cámara digital en retroiluminación laser) utilizados en la presente Tesis se detalla en la Tabla 4

Tabla 4. Características técnicas de los instrumentos de detección óptica utilizados en los estudios

	QualitySpec Pro 2600	Multispec instrument AG Tec5	Cámara monocromática BASLER 310f
Intervalo óptico (nm)	350-2500	350-1160	400-1000
Detector	InGaAs	InGaAs	CCD
Resolución DI (nm)	1	1	No disponible
Modo medida	Reflexión no especular	Reflexión no especular	Reflexión no especular

3.2.1.- Quality Spec Pro 2006

El equipo QualitySpec Pro 2600 (Analytical Spectral Devices, INC. Colorado, USA) es un espectrofotómetro de barrido rápido con una velocidad de escaneo de 10 espectros/segundo (Figura 7). La fuente de iluminación es una lámpara de Wolframio (12V/45W). El sistema de detección está formado por un monocromador con doble detector de In-Ga-As.

Antes de cada serie de medidas, se realiza una calibración del equipo, midiendo un blanco suministrado por la empresa.



Figura 7. Equipo Vis/NIR QualitySpec Pro 2600

3.2.2.- MultiSpec Instrument AG Tec5

El espectrofotómetro Vis/NIR Multispec (Tec5 AG, Frankfurt A.M., Alemania), consta de un espectrómetro SC – NMS I (Zeiss, Jena, Alemania), una fuente de luz halógena (12 V y 20 Watt) y tanto la luz de iluminación como la de medida se transportan mediante fibra óptica. Para hacer las mediciones del espectro en la oscuridad, el fruto se coloca en un compartimento en el que se han situado las sondas de iluminación y recepción de luz y el fruto (Figura 8).

Previo a la obtención de los espectros, se realiza una medida con una muestra blanca para obtener la referencia de calibración con la que compararemos las absorbancias difusas relativas a cada fruto.

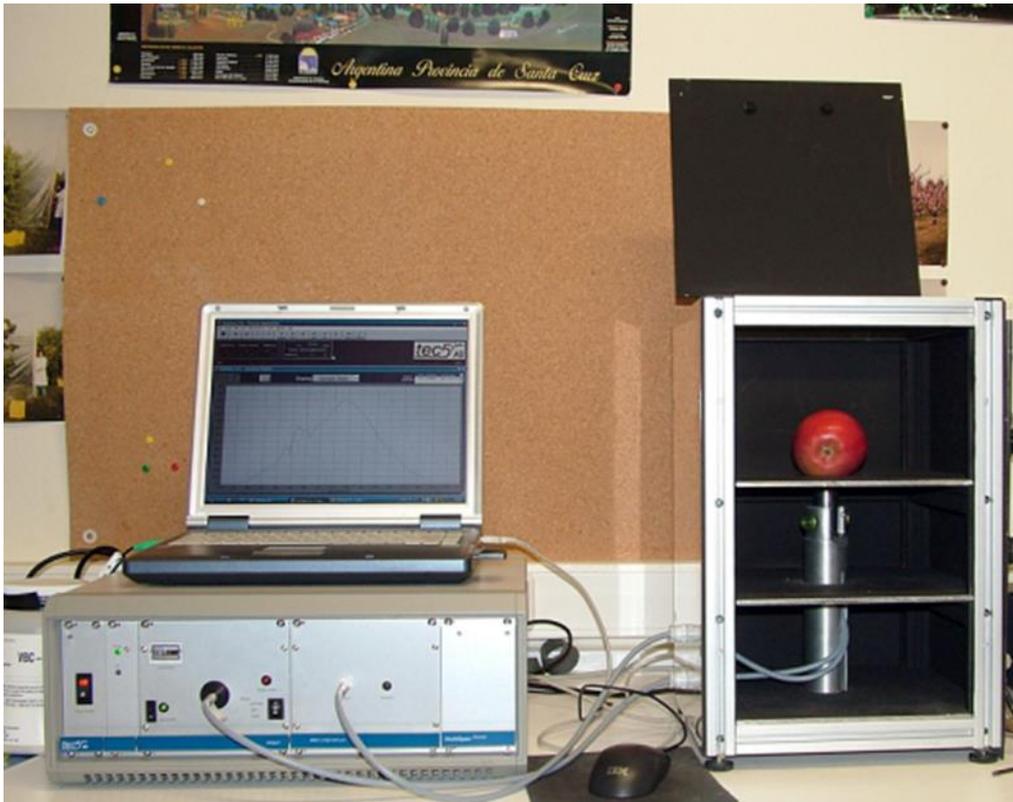


Figura 8. Equipo MultiSpec Instrument AG Tec5 y compartimento de medida (a la derecha)

3.2.3.- Cámara monocromática BASLER 310f

Las imágenes multispectrales se obtuvieron mediante una cámara monocromática CCD (Charge Coupled Device: Dispositivo de Carga Acoplado), BASLER A310f (Germany) sensible al Vis-NIR (Fig. 9.a). Las muestras se disponían en un dispositivo que permitía obtener un área de retrodifusión de hasta 25 mm de diámetro. El ángulo de incidencia del haz laser fue de 15° con la normal y la distancia de la cámara al fruto fue de 14 cm en la normal. El tamaño de las imágenes era de 782 x 582 píxeles. Las imágenes fueron tomadas en una habitación sin luz, a temperatura ambiente de 20°C .

Como iluminación, (Fig. 9.b) se utilizaron cinco láseres de diodo colimados (Power Technology), con longitudes de onda de 670, 785, 850, 905 y 980 nm y con 3 mW de potencia de salida los cuatro primeros y 28.3 mW, el último, para compensar la baja sensibilidad de la cámara en esta zona.

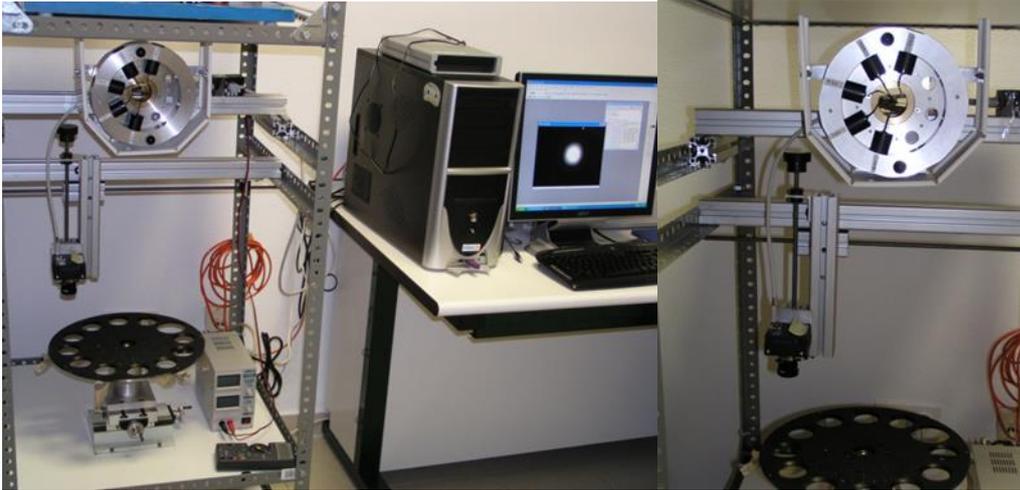


Figura 9 a) Equipo de obtención y tratamiento de imágenes multispectrales. b) Detalle de la disposición de los láseres, de la cámara y del posicionador de muestras.

3.3.- SOFTWARE

Los programas utilizados para el análisis digital de imágenes han sido Matrox 8.0 (Matrox Electronics Systems Ltd. Quebec, Canada) e ImageJ 1.46r (Open Code, Software). Los softwares utilizados para el tratamiento quimiométrico de los datos han sido Matlab (Mathworks, Inc. USA) y The Unscrambler X (CAMO Software AS, Norway).

Matlab es un programa de cálculo técnico que permite hacer operaciones con vectores y matrices. Tiene su propio lenguaje de programación y consta de un código básico y varias librerías especializadas que sirven para realizar programas de cálculo específicos de las necesidades del trabajo.

The Unscrambler X es un programa quimiométrico entre cuyas características se encuentra el cálculo de: distancias espectrales, regresiones globales, análisis discriminantes, PCA, PLS, etc.

3.4.- MÉTODOS Y ANÁLISIS DE REFERENCIA

En la Tabla 5 se describen los equipos utilizados para realizar los análisis de referencia en las muestras

Tabla 5. Equipos utilizados en los análisis de referencia

SÓLIDOS SOLUBLES TOTALES

Refractómetro digital



Figura 10. Refractómetro digital ATAGO PR-101

La medida del contenido en sólidos solubles totales se ha realizado con un refractómetro digital ATAGO PR-101 (Figura 10) que permite la determinación de los sólidos solubles presentes en el zumo. Estos sólidos solubles son en su mayoría azúcares. La unidad de medida del refractómetro es el grado Brix, que representa el % (p/p) de sacarosa disuelta en agua.

FIRMEZA

La firmeza Magness-Taylor, es el método destructivo clásico utilizado para medir la firmeza de la pulpa. Se han utilizado tres equipos.

Penetrómetro manual



Figura 11. Penetrómetro Fruit Pressure Tester FT32

Para medir la firmeza en puntos opuestos de la superficie de los frutos hemos utilizado un penetrómetro Fruit Pressure Tester FT32 (Istituto per la Valorizzazioni dei Prodotti Agricoli, Milano, Italy) (Figura 11). La forma de efectuar la medida ha sido: eliminar parte de la epidermis del fruto mediante un pelador y, tras colocar la parte pelada de la fruta bajo el penetrómetro, introducir el puntal de 8 mm de diámetro hasta la muesca indicada en el mismo. El aparato mide la firmeza en kgf que convertimos a Newtons (N) como unidad de fuerza

Texturómetro TA.XT plus (Stable Micro Systems, Godlming, UK).

El texturómetro está equipado con una sonda de una fuerza máxima de 5 kgf, tiene una resolución de 0.0001 kgf y una precisión de 0.0025% (Figura 12). La medida se realizó en lados opuestos de la fruta, tras retirar la piel. La sonda tenía 11,3 mm de diámetro y su velocidad de avance fue de 2 mm/s. Los datos obtenidos por el equipo se recogen en un PC.



Figura 12. Texturómetro TA.XT

Compresión de émbolo (método no destructivo)



Figura 13. Durómetro digital Agrost 100 Field

Las medidas se realizan con un durómetro digital no destructivo AGROSTA 100Field (Agro Technology, Forges Les Eaux, France). El equipo no penetra los frutos. Su medición se basa en cuantificar la fuerza necesaria para retraer un émbolo una cierta longitud al aplicarlo de forma manual contra la superficie de la fruta. Los resultados se expresan en unidades Durofel. Las medidas se hicieron en cada una de los dos extremos de un del fruto.

Impulso acústico (método no destructivo)



Figura 14. Medidor de firmeza por impulso acústico AFS.

El equipo AFS (Acoustic Firmness Sensor (Figura 14) evalúa la firmeza mediante resonancia acústica y un índice de firmeza propio, el IF. Este índice se obtiene a partir del análisis de Fourier de una señal acústica, tras atravesar el fruto analizado. El equipo mide la masa de la fruta y el espectro de frecuencias resonante. El índice se expresa:

$$IF=f_R^2 \times m^{2/3}$$

En este índice se relacionan, la masa, m (en kg) y la frecuencia de la máxima amplitud, f_R (en hercios)

3.5.- TRATAMIENTO DE DATOS

La metodología en el tratamiento de datos ha sido la misma en todos los estudios

Antes de calcular las ecuaciones de calibración se aplicaron diferentes pretratamientos en las variables ópticas, con el objetivo de mejorar los resultados de los modelos de calibración: Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), derivadas de primer y segundo orden y normalización (media cero, desviación estándar uno).

En los tres capítulos se realizó un Análisis de Componentes Principales (Principal Components Analysis, PCA) para la detección de posibles muestras anómalas (outliers), buscando reducir en lo posible la dimensionalidad de la matriz X y detectar posibles agrupamientos o pautas de comportamiento entre

los distintos grupos de muestras. Para la detección de anómalos se empleó el estadístico *leverage*, basado en el cálculo de distancias en espacios n-dimensionales, y que muestra la influencia de cada muestra en el modelo de regresión construido.

La obtención de las ecuaciones de calibración se realizó usando dos métodos de análisis multivariante: PLS y LS-SVM, en el caso de los capítulos de melocotón y cereza, mientras que para el de manzana tan sólo se utilizó el método de regresión PLS.

3.5.1.- Análisis de Mínimos Cuadrados Parciales (Partial Least Square analysis. PLS)

Este método es utilizado para la construcción de modelos predictivos de variables (y) cuando los factores (x) son numerosos y éstos son altamente colineales. La matriz X es sustituida por una matriz de variables latentes (LVs), las cuales son a su vez combinaciones lineales de los vectores x que maximizan la covarianza de la matriz Y, ajustando por mínimos cuadrados tanto las variables latentes como los coeficientes de regresión, con un coeficiente de determinación (r^2) elevado. La principal característica de esta composición, en nuestro trabajo, es que busca la máxima correlación entre los espectros (variables x) y la propiedad a determinar (variables y).

Esta técnica se ha utilizado en un número considerable de estudios relacionados con la predicción de índices de calidad en frutas y hortalizas (Nicolai et al., 2007; Sánchez et al, 2012; Maniwarra et al., 2014; Martínez-Valdivieso et al, 2014; etc.).

La validación cruzada de L iteraciones (L=20) se utilizó para calcular el número óptimo de variables latentes y para evitar el sobreajuste en el desarrollo de los modelos de calibración (Xiaobo et al, 2007; Zhang et al, 2013).

3.5.2.- Máquina de Vectores de Soporte de Mínimos Cuadrados (Least Squares Support Vector Machine. LS-SVM)

La técnica de Máquina de Vectores de Soporte de Mínimos Cuadrados (LS-SVM) es un método de regresión que tiene la capacidad de hacer frente a la modelización multivariante lineal y no lineal y la resolución de estos problemas de manera relativamente rápida (Suykens and Vandewalle, 1999). Estudios previos han demostrado el potencial de los métodos de regresión no lineales en productos agrícolas (Sun et al., 2009, Shao et al., 2011 y Chauchard et al. 2004). En el método LS-SVM se utiliza un kernel con función de base radial (RBF), cuyos hiperparámetros γ y σ se optimizan mediante la búsqueda en rejilla (Grid-search) y la validación cruzada (LS-SVMLab, Suykens et al, 2002).

Una vez obtenidos los modelos de calibración, mediante los dos métodos de regresión, éstos se prueban con un set de validación. Los mejores modelos de calibración fueron seleccionados basándose en el mayor coeficiente de determinación de validación cruzada (R_{cv}^2) y el menor valor de error estándar de validación cruzada (SEVC) (Williams, 2001).

Otro estadístico utilizado es la Desviación Predictiva Residual (Residual Predictive Deviation, RPD) calculado como el cociente entre la desviación estándar (DE) y el SECV de los valores del grupo de calibración (Williams, 2001). Este estadístico se calcula para facilitar la comparación de los resultados obtenidos para distintos grupos de calibración con diferentes medias (Williams, 2001). Williams (2001) señala que valores de RPD comprendidos entre 3 y 5 indican predicciones eficientes.

Respecto a la validación, los resultados de los diferentes grupos de predicción se evaluaron comparando el error medio de predicción (RMSEP), el coeficiente de validación externa (r^2), el bias y el error estándar de predicción (SEP).

3.5.3.- Selección de Variables hacia atrás mediante Información Mutua

La información mutua es una medida de correlación no lineal entre dos variables proveniente de la Teoría de la Información de Shannon (Cover, 1990). Para dos conjuntos de variables continuas, X e Y, se puede calcular como:

$$I(X,Y) = \iint \mu_{XY}(x,y) \log \frac{\mu_{XY}(x,y)}{\mu_X(x)\mu_Y(y)} dx dy$$

donde $\mu_{XY}(x,y)$ es la función de densidad de probabilidad (FDP) de X e Y, y $\mu_X(x)$ y $\mu_Y(y)$ son las funciones de densidad marginal de los conjuntos de variables X e Y. Su ventaja en comparación con otras medidas de correlación, es su capacidad para identificar relaciones no lineales entre los conjuntos de variables involucradas.

En la literatura reciente se pueden encontrar numerosos intentos para diseñar algoritmos de selección de los factores más relevantes (longitudes de onda y/u otros factores) para la predicción de propiedades químicas, muchos de los cuales están basados en la técnica de Información Mutua (Rossi et al. 2006, Pent et al 2005, Estévez et al. 2009). La selección de variables pretende identificar las variables redundantes e irrelevantes para proceder a su eliminación. La identificación de las variables relevantes es esencial en problemas espectrométricos, puesto que longitudes de onda cercanas pueden proporcionar información muy similar (colinealidad). La reducción de las variables necesarias para predecir cualquier magnitud es esencial en la disminución de los costes experimentales y de evaluación, pero también para aumentar la capacidad de generalización de los modelos, esto es, la capacidad de predicción de datos no vistos (Guyon and Elisseeff, 2003).

El método propuesto en este trabajo, es un enfoque publicado por primera vez por Koller and Sahami, 1996 y adaptado para variables continuas por Herrera et al., 2006, y que se basa en el concepto de “marco de Markov”: Dado un conjunto de variables X y una variable de salida Y, se dice que un conjunto de

variables M_i en X es un marco de Markov para una variable x_i en X con respecto a Y , si $I(\{M_i \cup x_i\}, Y) = I(M_i, Y)$, esto es, si M_i tiene en sí mismo toda la información que tiene x_i sobre Y . Un marco de Markov es por tanto un grupo de variables que subsume la información mutua contenida en una cierta variable; en la práctica, y para nuestros propósitos: con respecto a la variable objetivo. El algoritmo consiste en un método de selección de variables hacia atrás que comienza con el conjunto completo de variables, e iterativamente descarta aquellas variables para las que se detecta que se hay un marco de Markov en el conjunto de variables restantes X_G , esto es, aquellas para las que la información con respecto a Y está ya contenida en el conjunto X_G de variables que quedan.

El algoritmo establece los siguientes pasos:

- 1.- Calcular la MI entre todos los pares de variables de entrada $I(x_i, x_j)$
- 2.- Comenzando con el conjunto completo de variables de entrada $X_G = X$, iterar:
 - a) Para cada variable x_i , sea el candidato de *marco de Markov* M_i el conjunto p de variables en X_G para los que $I(x_i, x_j)$ es mayor.
 - b) computa para cada x_i

$$Loss_i = I(\{M_i \cup x_i\}, Y) - I(M_i, Y)$$
 - c) escoge la x_i para la que $Loss_i$ es menor, y elimina x_i de X_G .
- 3.- Continúa en el paso 2 hasta que no queden variables

De esta manera, se obtiene un ranking de relevancia de variables (en orden inverso). Nótese que, de esta manera, las variables que con menor influencia con respecto a la variable de salida (variables irrelevantes) se descartarán en las iteraciones iniciales del algoritmo. Igualmente, las variables redundantes se podrían eliminar en las primeras iteraciones. Las variables relevantes con menor redundancia serían las últimas en ser “seleccionadas” (eliminadas) por el algoritmo. Una discusión más extensa sobre la eficiencia, funcionamiento y otros aspectos del algoritmo se puede consultar en Herrera et al. (2006).

El estimador de la información mutua utilizado en este trabajo es el de los k-vecinos más cercanos de Kraskov (Kraskov et al. 2004), que demostraba que tenía un comportamiento más robusto con respecto a otras alternativas cuando se trabaja con conjuntos de variables. El valor de p utilizado para el algoritmo será $p = 1$, tal y como recomendaban Herrera et al. (2006).

4.- *Manzana*

Parte de este trabajo ha sido publicado en la revista SPIE vol 8785 con el título 'Determination of quality parameters in Apple *Smoothie Golden Delicious* using backscattering laser imaging, adjunto en el Anexo I.

4.1.- INTRODUCCIÓN

La industria hortofrutícola requiere sistemas efectivos de control y calidad como respuesta a las crecientes demandas del consumidor. Habitualmente, en los laboratorios de control de calidad y en algunas instalaciones industriales, se utilizan métodos físicoquímicos de análisis que comportan la destrucción del fruto. Además, en ciertos casos se hace uso de estimaciones subjetivas como métodos no destructivos empleando escalas hedónicas. En manzana, los parámetros más utilizados para determinar la calidad son la firmeza de la pulpa y su contenido en Sólidos Solubles Totales (SST). La firmeza es uno de las variables que mejor se relaciona con la madurez del fruto, ya que es un valioso parámetro que permite determinar propiedades de la textura de muchos alimentos sólidos, como la manzana. La firmeza puede definirse como la resistencia a la fuerza de penetración que ofrece la pulpa de un fruto, fuerza que es necesaria para romper los tejidos del mismo (Mohsenin, 1986). La concentración de SST es otro de los parámetros que se utilizan para determinar la madurez del fruto, ya que según avanza la maduración los sólidos solubles aumentan. A lo largo del desarrollo del fruto, se acumulan fotoasimilados en forma de almidón que más tarde se transforman en otros glúcidos más simples durante el proceso de maduración como la fructosa, glucosa o sacarosa. Una vez recolectada la fruta, su firmeza sufre un declive durante el proceso de almacenamiento mientras que los SST se mantienen (Peng, Y and Lu, R., 2006).

En los últimos años se están evaluando e implementando técnicas no destructivas para determinar la madurez del fruto y estimar su momento óptimo de recolección. Tanto las medidas de firmeza como las de SST son procedimientos destructivos, como se ha dicho anteriormente, por lo que sólo se

pueden aplicar a un limitado número de frutos. Por ello, estas técnicas no son de gran utilidad en las líneas de clasificación de las centrales frutícolas.

Un conjunto de técnicas, basadas en las propiedades ópticas de las frutas, pueden considerarse adecuadas para la evaluación no destructiva de la calidad de los productos agrícolas (Chen, 1996). Estas tecnologías parecen prometedoras, pero su relación con los métodos destructivos estándar de la manzana todavía no está bien establecida (Shmulevich et al., 2003; Lu, 2004). En este trabajo se ha evaluado una de estas técnicas no destructivas, el análisis de las imágenes capturadas de la retrodifusión de luz de láseres con longitudes de onda en las zonas del espectro del Visible y del Infrarrojo cercano (NIR). Las longitudes de onda del Visible son absorbidas por los diferentes pigmentos que componen la manzana como clorofila, carotenos y antocianos (Knee, 1972; Merzlyak et al, 1999; Zude-Sasse, et al 2000). Además, numerosos estudios consideran la banda de absorción de la clorofila, en el visible, y la del agua, en el NIR, como las más relacionadas con el estado de madurez y con la firmeza (Peng and Lu, 2007, McGlone et al 1997). Por ello, en este trabajo se utilizaron láseres con longitudes de onda de 905 y 980 nm, dentro de una de las bandas de absorción del agua y azúcares, otro láser con una longitud de onda de 670 nm emite dentro del espectro visible y se relaciona con la capacidad de absorción de luz de la clorofila y otros de 785 y 850 nm entre las dos bandas de absorción destacadas.

En este tipo de trabajos, tras la adquisición de las imágenes multiespectrales de las manzanas, se procede a un análisis digital de las mismas y se registran y cuantifican diferentes parámetros, que no son necesariamente idénticos, dependiendo de los autores consultados (Peng, et al 2006; Qing, et al, 2007a).

El objetivo que se pretende alcanzar con este trabajo es establecer un modelo estadístico que relacione las medidas de firmeza y de contenido en sólidos solubles con dos parámetros característicos obtenidos de las imágenes multiespectrales, ambos en función de distintos niveles de gris: el histograma de la imagen obtenida para cada láser, utilizado en el que llamamos “primer ensayo”, y los diámetros de Feret, correspondientes a las segmentaciones de las

imágenes, utilizado en el que llamamos “segundo ensayo”. De este modo, buscamos obtener un sistema no destructivo de clasificación individualizada de manzanas en función de su calidad, con referencia en los parámetros tradicionales de clasificación en la industria: firmeza y SST.

4.2.- MATERIAL Y MÉTODOS

4.2.1.- Material Vegetal

En este estudio hemos trabajado con la variedad de manzana ‘Smoothie Golden Delicious’, que es una mutación natural de “Golden Delicious”, obtenida en 1958 en EE.UU. Los frutos analizados se han obtenido de una plantación cultivada en una parcela experimental de Estación Experimental de Aula Dei (CSIC) en Zaragoza. La finca cuenta con riego por goteo y sus árboles están formados en eje central. El marco de plantación es de 3 metros entre árboles y 5 metros entre filas. La recogida de los frutos y las medidas experimentales del ensayo fueron realizadas durante las campañas 2009 y 2010.

En el primer año de estudio, se recolectaron 25 frutos cada 15 días desde el mes de Mayo hasta Octubre. Se cosecharon 225 manzanas en total.

En el segundo año, 2010, se realizaron tres muestreos durante los meses de Septiembre y Octubre con el fin de agrupar la variabilidad en la madurez de las manzanas en tres grados. Se recolectaron 50 muestras en cada fecha con una frecuencia quincenal, recogiendo un total de 150 manzanas.

Las manzanas fueron recolectadas por la mañana y se llevaron inmediatamente al laboratorio para realizar los análisis destructivos y no destructivos. En todos los frutos se obtuvieron imágenes multiespectrales de la zona de impacto del láser. Además, en esos mismos puntos, se analizaron los parámetros sólidos solubles y firmeza, mediante los tradicionales métodos destructivos.

4.2.2.- **Parámetros de calidad de fruto convencionales (destructivos)**

4.2.2.1.- **Firmeza**

Mediante el método de Magness-Taylor, se midió la firmeza como la máxima fuerza necesaria para perforar la pulpa una manzana una profundidad de 10 mm, después de quitar la piel de esa zona, en un equipo TA.XT plus (Stable Micro Systems, Godlming, UK) controlado por PC. El texturómetro estaba equipado con una sonda de una fuerza máxima de 5 kgf, con una resolución de 0.0001 kgf y una precisión de 0.0025%. La medida se realizó en los puntos donde se habían obtenido las imágenes, en caras opuestas de la manzana, tras retirar la piel. La sonda tenía 11 mm de diámetro y su velocidad de avance fue de 2 mm/s.

4.2.2.2.- **Sólidos Solubles Totales**

El análisis de sólidos solubles totales se llevó a cabo mediante la utilización de un refractómetro digital PR-101ATAGO CO. (Tokyo, Japón). Cada muestra se licua y se utilizan unas gotas para medir el promedio de los azúcares presentes en todo el fruto, en grados Brix, la medida se hizo por duplicado. El grado Brix se expresa como el tanto por ciento (%) (p/p) de sacarosa disuelta en agua

4.2.3.- **Métodos ópticos para estimar parámetros de calidad**

4.2.3.1.- **Adquisición y tratamiento de imágenes.**

La Figura 15 representa un esquema del dispositivo utilizado para la adquisición de imágenes. Las imágenes multiespectrales se obtuvieron mediante una cámara monocromática CCD (Charge Coupled Device: Dispositivo de Carga Acoplado), BASLER A310f (Germany) sensible al Vis-NIR (400 - 1000 nm). Las muestras se colocaron en un dispositivo que permite obtener un área de retrodifusión de hasta 25 mm de diámetro. El ángulo de incidencia del haz laser fue de 15° con la normal y la distancia de la cámara al fruto fue de 14 cm en la

normal. El tamaño de las imágenes es de 782 x 582 píxeles. Las imágenes fueron tomadas en una habitación sin luz, a temperatura ambiente de 20°C.

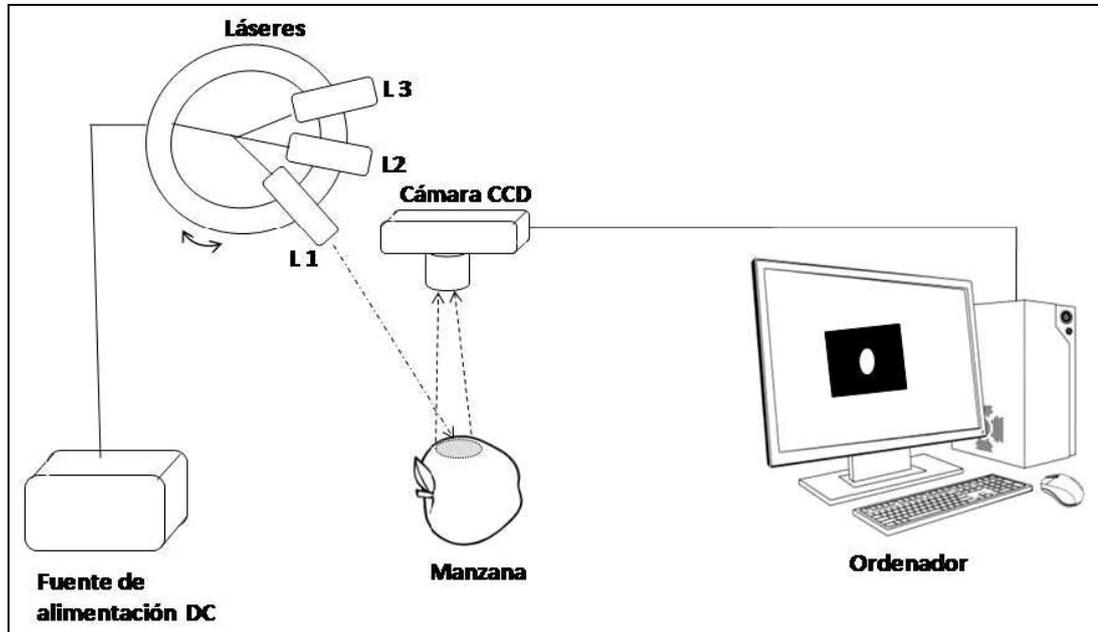


Figura 15. Esquema del funcionamiento del equipo utilizado para la captura de imagen de la retrodifusión de la luz.

Como fuente de iluminación, se utilizaron cinco láseres de diodo colimados (Power Technology Inc. Alexander, USA), con longitudes de onda de 670, 785, 850, 905 y 980 nm, con una potencia de salida de 3 mW, los 4 primeros y de 28.3 mW el último. La mayor potencia de salida del láser de 980 nm tiene como motivo compensar la pérdida de sensibilidad de la cámara en esta zona del espectro respecto a las otras longitudes de onda. Las cinco longitudes de onda fueron seleccionadas en base a otros estudios (Peng et al, 2006; Qing et al, 2007b). La retrodifusión de la luz que proviene de la superficie del fruto, fue registrada en imágenes monocromáticas de 8 bits (de 0 (negro) a 255 (blanco) niveles de gris). Cada una de las imágenes obtenidas fue analizada con dos programas de tratamiento de imágenes: ImageJ 1.46r (Open Code, Software) y Matrox 8.0 (Matrox Electronic Systems Ltd., Canada).

El primer parámetro que se calculó de la imagen fue el histograma de niveles de gris, que corresponde al número de píxeles por intervalos de un bit. (Figura 16). Para ello, se utilizó el Software ImageJ 1.46r. Con este programa se selecciona un área que incluye la zona de luz retrodifundida. A esta área se le da el nombre de ROI (Region Of Interest) en la bibliografía de tratamiento digital de imágenes (González et al, 2002).

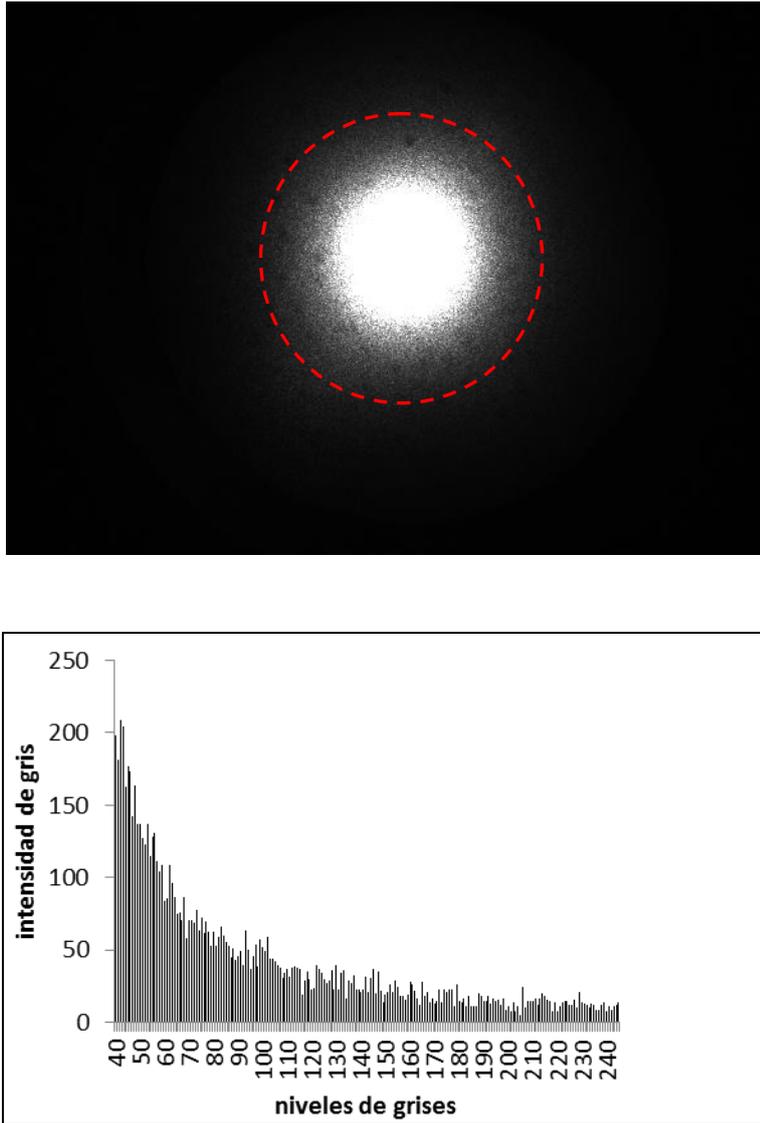


Figura 16. a) Imagen de luz retrodifundida (670 nm), **b)** Histograma de píxeles de la imagen

En este caso se consideraron los valores por encima de 245 como zona de saturación y los valores por debajo de 40 como ruido.

El programa Matrox 8 se utilizó para estudiar la evolución del diámetro medio de Feret de cada segmentación de la imagen, obtenido para distintos niveles de gris, como se puede ver en la Figura 17. Dado que las figuras correspondientes a las segmentaciones no son circulares, se calcula el diámetro medio de Feret como la media de los distintos diámetros de Feret de la imagen, que son los diámetros de las proyecciones de un objeto medidos a diferentes ángulos con la horizontal, en sentido anti horario entre 0º y 180º, realizando 36 medidas, una cada 5º. En este caso equivaldría al diámetro de la circunferencia con el mismo perímetro que la segmentación. En este ensayo se utilizaron tan sólo tres láseres para la obtención de los modelos de calibración, porque comprobamos que los datos de los láseres 850 y 905 nm no aportaban ningún resultado positivo al fin que nos proponíamos.

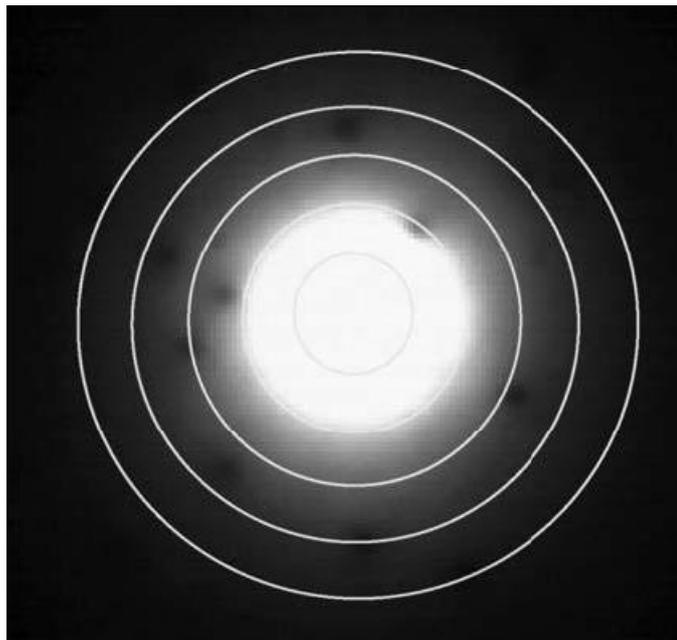


Figura 17. Imagen capturada de la retrodifusión láser, subdividida en segmentos con diferentes niveles de gris

Las segmentaciones de la imagen se realizaron disminuyendo los niveles de gris cada 10 unidades desde el nivel 245 hasta el 40.

4.2.4.- Métodos estadísticos

El análisis estadístico se realizó con el Software The Unscrambler X.1 (CAMO Software AS, Norway).

Tras aplicar diferentes pre-tratamientos en los datos obtenidos, se observó que ninguno de ellos mejoraba mejoraban los resultados finales por lo que se optó por utilizar los datos originales.

En primer lugar se realizó un análisis de componentes principales (Principal Components Analysis, PCA) para la detección de posibles muestras anómalas. En este caso, no se detectaron muestras anómalas.

Para obtener los modelos de calibración se utilizó el método de regresión lineal de Mínimos Parciales Cuadrados (Partial Least Square, PLS).

Para evitar el sobreajuste en el desarrollo de los modelos de calibración, se usó la validación cruzada en la evaluación de los modelos. El mejor modelo de calibración se seleccionó usando los siguientes parámetros estadísticos: R_{cv}^2 y SECV. El estadístico RPD también se tuvo en cuenta al evaluar el modelo de calibración. En el caso de la validación se utilizaron para seleccionar el mejor modelo, RMSEP, r^2 y el 'bias'.

4.3.- DETERMINACIÓN DE FIRMEZA Y SST MEDIANTE ANÁLISIS DEL HISTOGRAMA DE LAS IMÁGENES MULTIESPECTRALES OBTENIDAS MEDIANTE CINCO LÁSERES (670, 785, 850, 905 y 980 nm)

4.3.1.- Resultados y discusión

4.3.1.1.- Descripción de los datos

En la Figura 18, se representan los valores individuales de los parámetros físico-químicos obtenidos de las manzanas de cada grupo de muestras correspondientes a las diferentes fechas de muestreo para las dos campañas. Por motivos de escala, en las gráficas se han expresado los valores de firmeza en kgf.

En la primera campaña, se confirma tanto la disminución de la firmeza como el aumento de los SST y del calibre conforme avanza el estado de maduración del fruto, aunque, como puede observarse, no siguen tendencias lineales. La firmeza comienza en valores medios de aproximadamente 9 kgf (88,29 N) y termina con una media de 3 kgf (29,43 N), disminuyendo tres veces su valor. En cambio, los valores de los SST no llegaron a duplicarse comenzando en un valor medio de 8° Brix y alcanzando los 15° Brix.

En la segunda campaña se observan menores diferencias en los valores iniciales y finales de firmeza, SST y calibre ya que el período de muestreo es más corto. Los SST parten de 12,6°Brix evolucionando hasta 21,7°Brix. Respecto a la firmeza, los valores parten 56 N disminuyendo hasta los 16 N.

El calibre (medido en cm) aumenta en ambas campañas, siendo menos pronunciado el aumento en la segunda, por corresponder manzanas muestreadas en un intervalo más corto hasta alcanzar la madurez final.

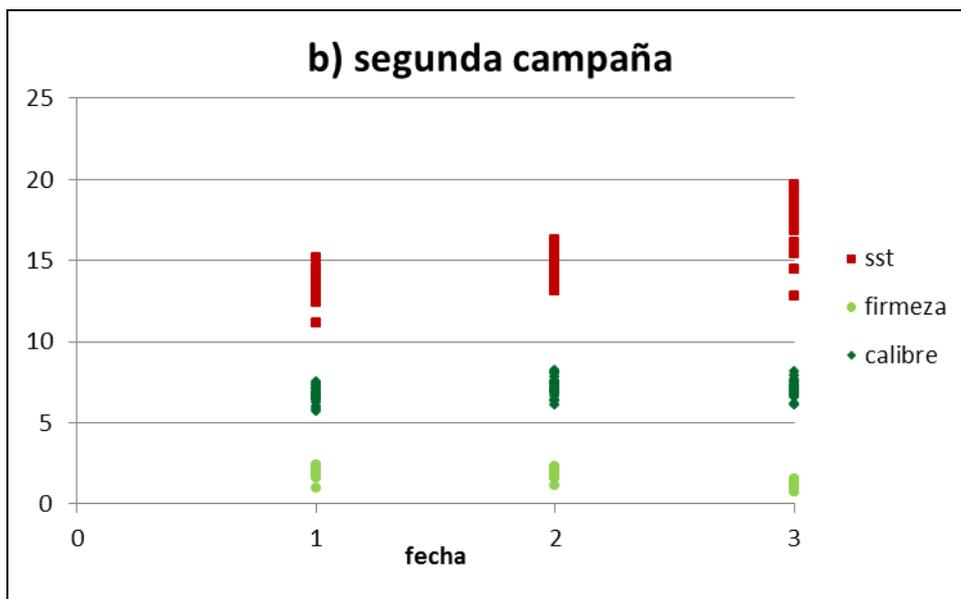
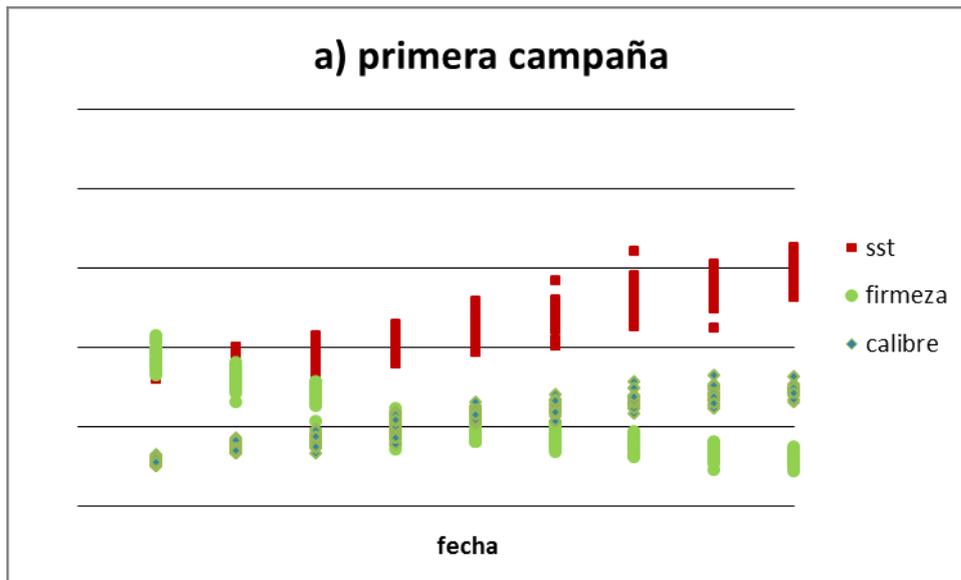


Figura 18. Evolución de los valores de SST ($^{\circ}$ Brix), firmeza (kgf) y calibre (cm) durante la campaña primera (a) y segunda (b). En la primera campaña el valor 0 en el eje de abcisas corresponde a la fecha de 01/05/2009 y el valor 8 a la fecha 15/09/2009. En la segunda campaña, el valor 0 en el eje de abcisas corresponde a la fecha de 15/08/2010 y el valor 3 a la fecha de 30/09/2010.

En la Tabla 6 se muestran los valores de la media, desviación estándar (DE), rango de valores y coeficiente de variación (%CV) para las dos campañas por separado y para el estudio combinado de ambas.

Tabla 6. Valores estadísticos para los parámetros de SST y firmeza en la primera y segunda campañas y en la combinación de ambas.

	SST (°Brix)			Firmeza (N)		
	2009	2010	2009-2010	2009	2010	2009-2010
n	214	150	364	214	150	364
Media	11.1	16.1	13.2	54.2	37.3	49.1
DE	2.1	1.9	3.2	21.9	7.6	20.5
Rango	8-16.3	12.6-21.7	8-21.7	21.5-107	16-55.8	16-107
%CV	18.9	11.8	24.2	40.4	20.4	41.7

Se obtiene una mayor variación en los parámetros en el estudio combinado de ambas campañas debido al aumento de la variabilidad en las muestras.

4.3.1.2.- Imágenes

En la Figura 19 se muestran imágenes de luz retrodifundida de los 5 láseres utilizados. El láser que presenta un mayor campo de dispersión de la luz es el que corresponde a una longitud de onda de 785 nm y el de menor campo es el láser de 670nm, lo que podría explicarse por la mayor absorción de luz en esta longitud de onda por parte de la clorofila. Conforme avanza el estado de madurez aumenta claramente la dispersión de luz en todos los casos. En las imágenes que presentan mayor dispersión de luz se pueden observar unas manchas circulares negras alrededor del punto central, que corresponden a las

lenticelas de la piel de la manzana, tema que no se trata en este trabajo, pero que se indudable interés para futuros estudios.

Debemos señalar que algunas de las fotografías no se realizaron correctamente, lo que da distinto número de imágenes para cada láser.



19/05/2009



01/07/2009



15/09/2009

Figura 19. Imágenes de luz retrodifundida de los 5 láseres utilizados (de izquierda a derecha: 670, 785, 850, 905, 980 nm) en tres frutos con diferente grado de madurez (de arriba abajo: verde, maduro y sobremaduro)

4.3.1.3.- Histogramas

De cada una de las imágenes se obtuvo el histograma que clasificaba los píxeles de la imagen en función de cada cada nivel de gris, mediante el programa ImageJ, con un ROI del mismo tamaño para todas las imágenes. En la Figura 20 se muestra el histograma para cada láser en distintos momentos de madurez del fruto: verde, maduro y sobremaduro. Los láseres 1 y 2 proporcionan mayores diferencias en los histogramas según la fecha de recolección, mientras que para el resto de láseres los histogramas apenas varían en las tres fechas de muestreo. En el caso del láser 1, se observa que para la primera fecha de muestreo (15 de

Mayo) se produce menor retrodifusión debido al alto contenido de clorofila en la manzana que absorbe esta radiación. El número de píxeles en cada nivel de gris es más alto a medida que madura la manzana concomitante a la disminución del contenido en clorofila. Para el láser 2 ocurre lo mismo, conforme aumenta la maduración, se obtiene una mayor retrodifusión. La variación de la retrodifusión en esta longitud de onda se atribuye a cambios en la estructura de los tejidos, debido a la degradación de la pectina y la variación del contenido de agua (Martínez-Vega et al, 2012).

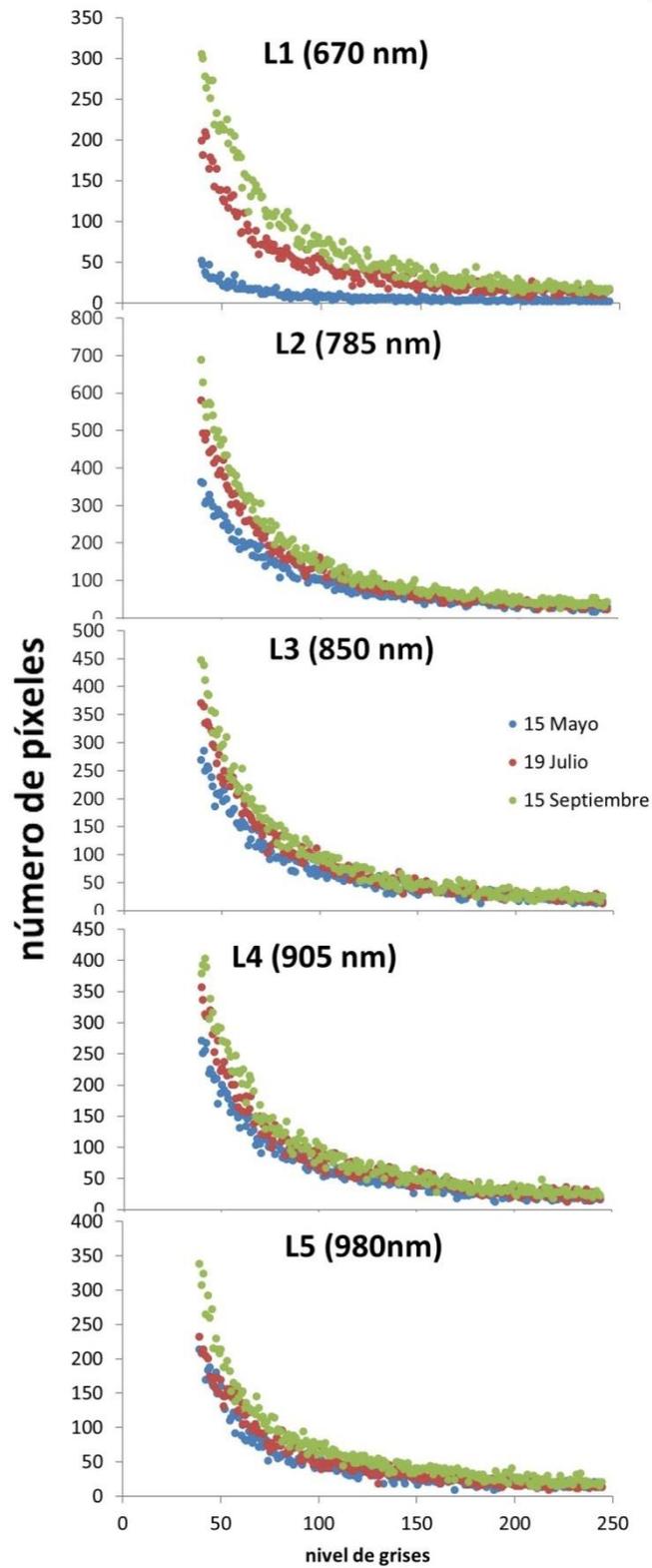


Figura 20. Histogramas de las imágenes de retrodifusión de cada uno de los 5 láseres en tres momentos de recogida de la primera campaña: 15 de mayo (verde), 19 de julio (maduro) y 15 de septiembre (sobremaduro)

4.3.1.4.- Calibración firmeza y SST

- **Campaña 2009**

En la Tabla 7 se muestran los estadísticos de calibración obtenidos para SST y firmeza, mediante la aplicación del método de regresión PLS a los datos de las muestras recolectadas en la primera campaña con los cinco láseres. En el caso de los sólidos solubles se observa que los valores más elevados de R_{cv}^2 son 0.67 y 0.69, correspondientes a los láseres 4 y 2 respectivamente. Los errores de calibración son del orden de 1.2 °Brix para ambos láseres. Para la firmeza, se obtuvieron valores aún más altos, R_{cv}^2 , entre 0.7 y 0.81. El hecho de que los mejores resultados se hayan obtenidos con los láseres de longitudes de onda 785 y 905 nm puede deberse a que en la zona del espectro 751-915 nm parece que hay menor absorción de los componentes químicos de la fruta, por tanto es posible que la dispersión de la luz por la estructura física de los tejidos de la fruta proporcione más información (Qing et al, 2007b). El peor resultado fue para el láser de 670 nm, lo que resulta extraño por la estrecha relación entre esta longitud de onda que correspondiente a la absorción de la clorofila con la firmeza (Tu et al, 1995; Moons et al, 1997; Zude et al, 2002). Los valores de SECV fueron altos excepto el de 9.5 N que corresponde al láser 2. Los resultados obtenidos para ambos parámetros son muy similares a los descritos por otros autores (Qing et al, 2007a) (SST, $R_{cv}^2= 0.64$ y firmeza, $R_{cv}^2=0.76$).

Tabla 7. Estadísticos de calibración usando método de regresión PLS para los parámetros SST y firmeza en la primera campaña

Parámetro	láser	n	PLS	SECV	R_{cv}^2
SST (°Brix)	L1	213	1	1.14	0.53
	L2	215	1	1.2	0.69
	L3	213	1	1.38	0.59
	L4	215	1	1.24	0.67
	L5	213	2	1.32	0.65
Firmeza (N)	L1	213	1	16.5	0.36
	L2	215	2	9.5	0.81
	L3	213	1	11.8	0.71
	L4	215	1	11.4	0.73
	L5	213	2	10.6	0.77

n: número de muestras

PLS: número de componentes principales con las que se ha obtenido el modelo de calibración

SECV: error estándar de validación cruzada

R_{cv}^2 : coeficiente de determinación para la validación cruzada

Obtenidos los modelos de calibración para sólidos solubles y firmeza se realizó una validación externa de los mejores modelos para ambos parámetros. Los modelos validados fueron los obtenidos con los láseres de 785 y 905 nm para ambos parámetros. Los resultados de la validación externa tanto para SST como para firmeza fueron muy pobres. Esto puede ser debido a que, a pesar de que existe una gran variabilidad en los parámetros destructivos, al aplicar los modelos obtenidos para la primera campaña, como predictor en la segunda, gran parte de los valores de los parámetros de esta campaña queden fuera del rango de los de la primera. Esto puede deberse a que en la segunda campaña se recolectaron muestras durante el mes de Septiembre cuando las manzanas estaban en un estado de madurez más avanzado. Por tanto, los modelos obtenidos en este trabajo en el primer año, no pueden aplicarse directamente para predecir los valores de SST y firmeza de la segunda campaña.

- **Estudio combinado de las campañas 2009 y 2010**

Como consecuencia de no haber podido realizar la validación de los modelos de calibración obtenidos con los datos de la primera campaña, nos planteamos obtener modelos de calibración para SST y firmeza uniando los resultados de ambas campañas por si se pudieran genera modelos con una posible mayor generalidad de aplicación. Los estadísticos de calibración obtenidos se muestran en la Tabla 8.

Tabla 8. Estadísticos de calibración usando método de regresión PLS para los parámetros SST y firmeza en las dos campañas unidas

Parámetro	láser	n	PLS	SECV	R_{cv}^2
SST (°Brix)	L1	360	1	2.72	0.31
	L2	364	3	2.13	0.51
	L3	363	2	2.69	0.21
	L4	363	2	2.27	0.5
	L5	362	1	3.24	0.00
Firmeza (N)	L1	360	1	18.0	0.22
	L2	364	3	14.4	0.51
	L3	363	2	16.5	0.36
	L4	363	2	15.2	0.42
	L5	362	2	16.8	0.32

n: número de muestras

PLS: número de componentes principales con las que se ha obtenido el modelo de calibración

SECV: error estándar de validación cruzada

R_{cv}^2 : coeficiente de determinación para la validación cruzada

Se observa que los valores de R_{cv}^2 para todos los láseres son muy bajos cuando se consideran los datos de ambas campañas para los dos parámetros destructivos, no pudiendo obtener buenos modelos de calibración. El láser que proporcionó mejores resultados fue de nuevo el láser 2 (785 nm). Esto indica que

los modelos de predicción de una campaña no son aplicables a la otra, siendo necesario ampliar estos estudios.

4.3.1.5.- Análisis de componentes principales

Al realizar el análisis de componentes principales para detectar posibles muestras anómalas en la primera campaña, se observó, que el láser 2 (785 nm), producía una separación de las muestras en dos grupos (Figura 21) en función de los valores (scores) de las muestras en el primer Componente Principal (PC1). A la izquierda (valores negativos) son las que corresponden a las primeras cuatro fechas de recogida (de mediados de Mayo a mediados de Julio) y a la derecha (valores positivos) las muestras recogidas a partir del 20 de Julio. Las muestras situadas a la derecha tienen valores de sólidos solubles por encima de 10^9 Brix y de firmeza por debajo de 70 N, por tanto presentan valores aceptables para la recolección. Sin embargo, estas fechas no se corresponden con las previstas de cosecha para esta variedad ya que en esas fechas no se habían alcanzado los tamaños recomendados para la comercialización.

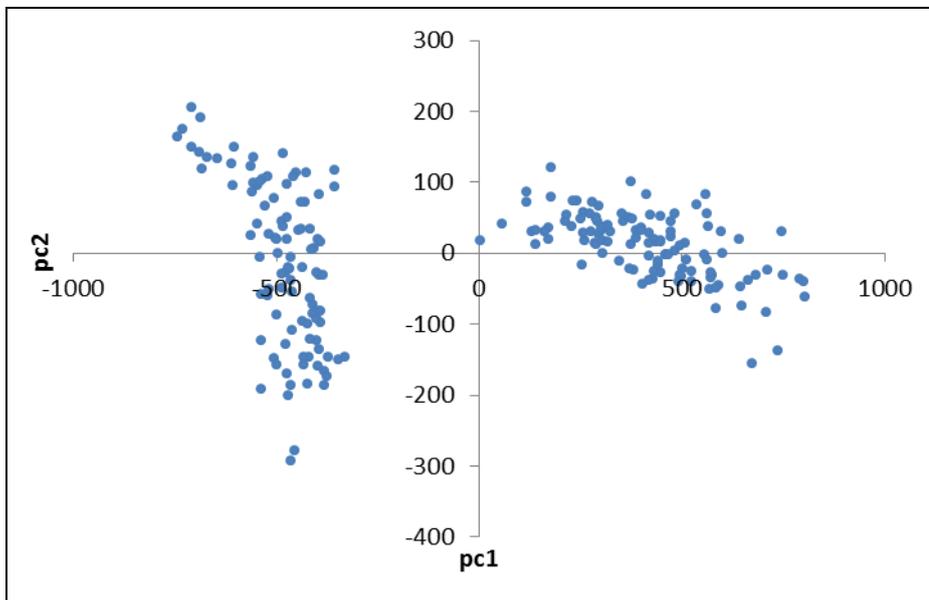


Figura 21. Análisis de componentes principales para el láser 2 (785 nm) en la primera campaña. A la izquierda están los frutos inmaduros (Mayo-Julio). A la derecha, frutos maduros.

En la Figura 22 se presenta el análisis de componentes principales para las manzanas de las dos campañas juntas. Se observan tres grupos, los dos de color azul, pertenecientes a la primera campaña, con la separación entre muestras maduras e inmaduras y el grupo de color rojo que corresponde a la segunda campaña. La aparición de este tercer grupo puede explicarse porque, como se ha comentado, la segunda campaña presenta muestras con grados de maduración diferentes a los de la primera campaña. En el resto de láseres no se ha observado esta separación.

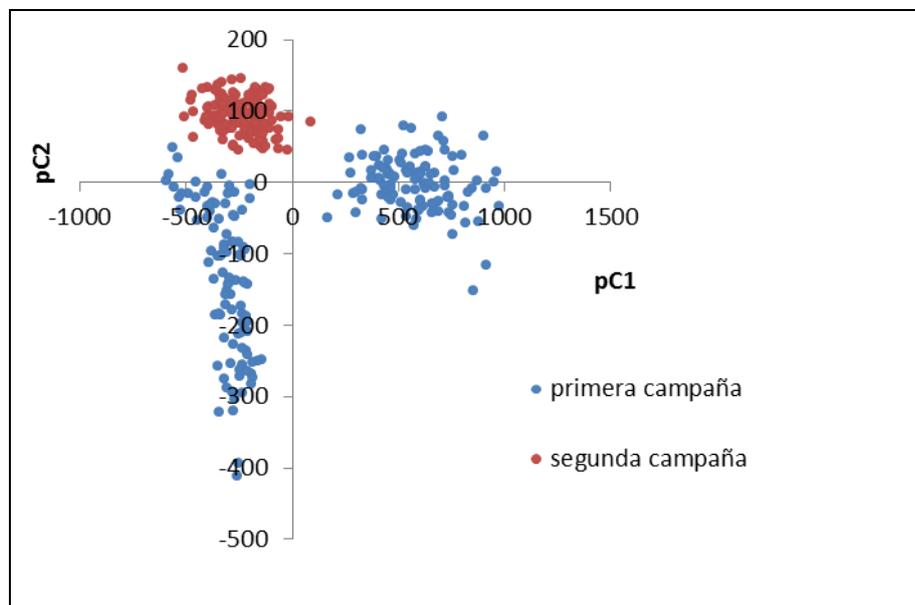


Figura 22. Análisis de componentes principales para el láser 2 (785 nm) para las dos campañas. Primer y cuarto cuadrante: frutos maduros 1ª campaña. Tercer cuadrante: frutos inmaduros 1ª campaña. Segundo cuadrante: frutos maduros 2ª campaña

Por tanto, se deduce que el histograma de las imágenes de retrodifusión producidas por el láser de 785 nm puede producir una separación de los frutos en función de su grado de madurez. A partir de esta idea, y aumentando el número de muestras y de las fechas de recolección, se podría obtener un modelo mediante PCA, a partir de los histogramas obtenidos de las imágenes de las manzanas iluminadas por esta fuente láser para poder clasificar otras muestras.

4.3.2.- Conclusión

En este estudio, basado en la manzana *Smoothie Golden Delicious*, se ha pretendido desarrollar modelos de calibración, mediante técnicas PLS, para la determinación de parámetros físico-químicos útiles para determinar la madurez del fruto, mediante su correlación con el histograma de las imágenes de retrodifusión obtenidas con varios láseres.

Se han obtenido, para los láseres de longitudes de onda 785 y 905 nm, buenos modelos de calibración con las manzanas de la primera campaña, donde se encontró una gran variabilidad en los valores de los parámetros de firmeza y sólidos solubles totales.

La validación de estos modelos con los frutos de la segunda campaña no ha sido posible ya que los valores de los parámetros físico-químicos de las muestras usadas para la validación se encontraban fuera del rango del modelo.

La combinación de los datos de ambas campañas para obtener nuevos modelos de calibración no ha dado lugar a buenos resultados, por lo que se propone ampliar el muestreo a nuevas campañas con el objetivo de optimizar resultados.

Por último, la aplicación del análisis de componentes principales (PCA) a los histogramas de las imágenes de retrodifusión del láser 2 (785 nm) permite una separación de los frutos según su estadio de madurez. Para el resto de láseres no se produce una separación visible entre momentos de madurez.

Consideramos que sería bueno seguir trabajando en la creación nuevos modelos de predicción con otros parámetros característicos de las imágenes multiespectrales.

4.4.- DETERMINACIÓN DE FIRMEZA Y DE SST MEDIANTE EL ANÁLISIS DE LOS DIÁMETROS DE FERET DE LAS IMÁGENES MULTIESPECTRALES OBTENIDAS MEDIANTE TRES LÁSERES (670, 785 y 980 nm).

4.4.1.- Resultados y discusión

En la tabla 9 aparecen los valores de media, desviación estándar, rango y coeficiente de variación para los grupos de calibración y validación para los dos parámetros de estudio, sólidos solubles y firmeza, para las muestras de la primera campaña. Se puede observar que las medias de ambos parámetros están incluidas en el rango seleccionado, por tanto se confirma el acierto al seleccionar grupos de calibración y validación ya que entre ambos conjuntos se hay una adecuada agrupación de valores.

Tabla 9. Valores estadísticos del conjunto de muestras para los parámetros de SST y firmeza obtenidos por métodos destructivos de análisis.

	SST(°Brix)		Firmeza(N)	
	Calibración	Validación	Calibración	Validación
n	139	69	139	69
Media	11.1	11.3	55.1	50
DE	2.2	2.2	22.8	17.8
Rango	8.2-16.3	8.2-15.5	21.6-105.5	27.4-91.8
%CV	19.4	19.1	41.46	35.59

Mediante el método de regresión PLS se construyeron los modelos de calibración a partir de las medidas destructivas de azúcares y firmeza y los diámetros medios de Feret a distintos niveles de gris, de las imágenes multiespectrales de los tres láseres empleados (Figura 23). En la Tabla 10 aparecen los estadísticos de calibración para ambos parámetros. Para el contenido de sólidos solubles totales se obtuvo un valor de $R_{cv}^2=0.93$ y para firmeza de 0.87 mucho más elevados que los obtenidos por otros autores usando metodología similar como Qing et al, 2007b con valores de $R_{cv}^2=0.76$ y 0.80 para sólidos solubles y firmeza respectivamente y Lu, 2004 con coeficiente de regresión de validación cruzada para firmeza de 0.77. Para ambos parámetros

el número de variables latentes del modelo fueron muy bajas, 1 y 2 para SSC y firmeza respectivamente, y los valores de RPD fueron muy elevados para ambos índices de madurez, alcanzando el valor más alto en el caso de la firmeza (Tabla 10).

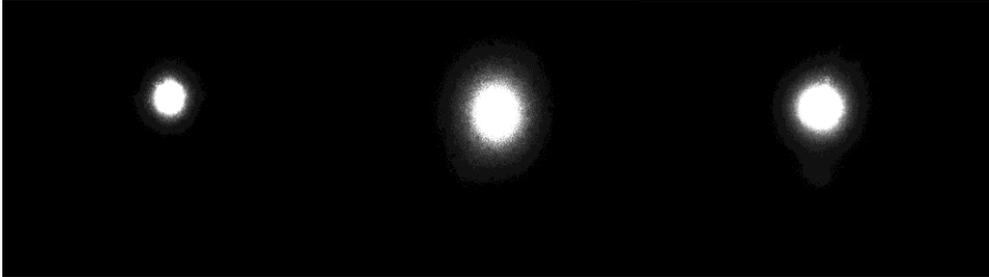


Figura 23. Imágenes multiespectrales monocromáticas obtenidas por iluminación con láseres de diodo a tres longitudes de onda: 670, 785 y 980 nm.

Tabla 10. Estadísticos de calibración usando método de regresión PLS para los parámetros SST y firmeza

	PLS	
	SSC(°Brix)	Firmeza(kg)
N	131	131
Lv	1	2
SECV	0.78	8.5
R_{cv}²	0.93	0.87
RPD	2.17	2.65

Se llevó a cabo una validación externa con 69 muestras. En la figura 24 se muestran las representaciones gráficas de ambos parámetros con los valores medidos frente a los predichos. En la figura también aparecen los estadísticos de validación.

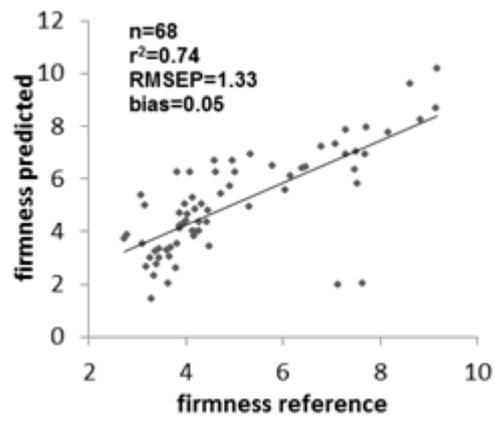
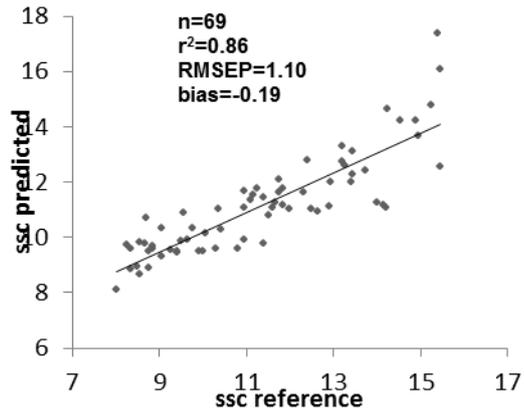


Figura 24. Estadísticos de predicción para firmeza y sólidos solubles

4.4.2.- Conclusiones

Este estudio ha demostrado la posibilidad de aplicar una técnica no destructiva como es el tratamiento de imágenes multiespectrales para determinar dos parámetros de calidad en la variedad de manzana 'Smoothie Golden Delicious' a partir de los diámetros de Feret de los distintos niveles de gris en las imágenes.

Estos resultados son prometedores por lo que se hace necesario ampliar el número de campañas, buscando la posibilidad de diseñar un aparato que estableciera los parámetros de calidad en campo.

Otra cuestión a tener en cuenta, sería profundizar en la aplicación de otros emisores láser de distintas longitudes de ondas que pudieran presentar mejoras en el desarrollo de los instrumentos portátiles de aplicación en campo para evaluar el estado de madurez del fruto del manzano en el árbol.

También sería conveniente ampliar las variedades susceptibles de estudio.

4.5.- CONCLUSIONES DEL CAPÍTULO

- En este capítulo se ha comprobado la posibilidad de aplicar las técnicas no destructivas de análisis de imágenes de retrodifusión de luz láser para predecir los parámetros de calidad SST y firmeza en manzana de la variedad '*Smoothie Golden Delicious*'.
- La utilización del láser de 785 nm da los mejores resultados al relacionar el histograma de las imágenes multiespectrales con los SST y la firmeza MT, en la primera campaña.
- Sin embargo, la extensión de los modelos de predicción de Sólidos Solubles Totales y Firmeza en manzana obtenidos en la primera campaña mediante el uso del histograma de las imágenes obtenidas con los 5 láseres no es posible en la campaña posterior.
- La utilización del Análisis de Componentes Principales utilizando el histograma de las imágenes, proporciona una buena clasificación en distintos estados de madurez.
- La utilización de los diámetros de Feret obtenidos de las imágenes de los láseres 670, 785 y 980 nm, a la predicción de SST y firmeza proporcionó buenos resultados en el modelo de predicción de la primera campaña.
- Consideramos que hay que proseguir el trabajo en sucesivas campañas y ampliar a otros parámetros de calidad el estudio de su relación con las imágenes multiespectrales para construir modelos de predicción más completos, que permitan su aplicación a instrumentos utilizables.

5.- *Melocotón*

Este capítulo es una adaptación del artículo publicado en 'Journal of the Science of Food and Agriculture' con el título 'Firmness prediction in *Prunus persica* Calrico peaches by visible/short-wave near infrared spectroscopy and acoustic measurements using optimised linear and non-linear chemometric models', adjunto en el Anexo II.

5.1.- INTRODUCCIÓN

El melocotón al ser recolectado queda separado de su fuente de nutrientes, pero sus tejidos siguen respirando y son fisiológicamente activos. La maduración de las frutas está ligada a complejas modificaciones físicas y químicas de sus propiedades, como son el ablandamiento, el aumento de la concentración de azúcares solubles y de aromas y, además, de cambios de coloración. Todos estos procesos tienen gran importancia porque influyen en el comportamiento de las frutas durante su almacenamiento, transporte y comercialización, afectando también a su valor nutritivo y características organolépticas (Kader, 2008). Además, alguno de estos procesos metabólicos está influenciado por factores externos como las condiciones meteorológicas del año de campaña y factores genéticos como el cultivar y el patrón sobre el que está injertado. Definir la fecha de recolección no solamente es útil para obtener una mejor calidad del producto sino que además contribuye a incrementar la producción y a minimizar los costes de las labores culturales (Thompson, 2007).

Uno de los parámetros que se correlaciona mejor con la madurez de la fruta y que, en definitiva, es de gran utilidad para determinar el momento óptimo de recolección, es la firmeza. Ésta se puede definir como la resistencia a la fuerza de penetración que ofrece la pulpa del fruto necesaria para romper sus tejidos (Mohsenin, 1986). Tradicionalmente la firmeza se ha determinado por el método destructivo 'Magness-Taylor' (MT). El test de MT consiste en un dinamómetro acoplado a un vástago cilíndrico que se introduce 8 mm en la pulpa de la fruta tras haber retirado una porción de piel. Este test es barato y

rápido, no obstante, los valores son muy variables y producen y presentan una alta incertidumbre (por encima de 30%; Barreiro y Ruiz Altisent, 1996) dependiendo del operador que ejerce la fuerza y del punto del fruto donde se toman las medidas. Adicionalmente, este test requiere la destrucción del fruto, lo que limita su utilidad.

En los últimos años se han desarrollado métodos no destructivos para la estimación de la firmeza de la fruta (más prácticos y reproducibles) con el objetivo de sustituir a los destructivos. Uno de estos métodos consiste en el uso de señales acústicas provocadas por impactos mecánicos sobre el fruto. La frecuencia resonante que emite un cuerpo depende directamente de su geometría, su masa y del módulo de elasticidad del material del que se compone (Studman, 2001). En los ensayos acústicos la muestra se excita mediante un impacto de baja intensidad y se produce una respuesta vibratoria en un rango audible (20-20000 hercios). La respuesta se registra con un micrófono y la señal en el tiempo es sometida a la Transformada Rápida de Fourier (FTT) para obtener la correspondiente señal en frecuencias, a partir de la cual se obtiene un índice de firmeza acústico. En la literatura se han realizado numerosos estudios sobre la utilización de métodos acústicos para la determinación de la firmeza en diversos tipos de fruta y hortalizas. No obstante, los resultados obtenidos muestran una alta variabilidad en la correlación entre ambas magnitudes dependiendo del tipo y variedad del producto analizado (variedad '*Delicious*' de manzana $r^2=0.25$, variedad de manzana '*Golden Delicious*' $r^2=0.72$, Abbott et al, 1995; aguacate $r^2=0.4$, Galili et al, 1998).

Otra de las técnicas no destructivas emergentes usada para la predicción de la firmeza es la espectroscopia de Visible e Infrarrojo Cercano (Vis/NIR). Se han descrito diversos métodos basados en esta técnica para predecir valores de firmeza de varias frutas (Zude, 2006, en manzanas; Sánchez et al., 2011, en nectarinas; Shao et al., 2011, en melocotones; Clarck et al., 2003, en aguacate). En la técnica Vis/NIR, un haz de luz incide sobre la superficie de la fruta penetrando unos milímetros. Una parte de esta radiación es absorbida por la muestra y el resto es reflejado difusamente hacia unos detectores que convierten esta luz en un espectro característico de cada muestra, dependiendo

de su composición físico-química, ya que ciertos compuestos como azúcares, agua y pigmentos absorben parte de radiación en la zona del visible e infrarrojo.

Las señales obtenidas por estos métodos no destructivos se relacionan con el parámetro a estimar mediante el cálculo de ecuaciones de calibración. En el caso de la firmeza, se considera como parámetro de referencia el valor obtenido mediante el método MT, pues pese a su variabilidad, es el más usado en trabajos de campo y los textos legales relacionados con la calidad de la fruta, lo mencionan específicamente. Entre los métodos de regresión utilizados habitualmente, los más adecuados son los que hacen uso de relaciones lineales como 'Multiple Linear Regression' (MLR) (Clark et al, 2003), 'Principal Components Regression' (PCR) (Hernández-Gómez et al, 2006) y Partial Least Squares (PLS) (Peirs et al, 2003) y los métodos no lineales como: 'Least Squares Support Vector Machines (LS-SVM) (Liu et al, 2009) y 'Artificial Neuronal Networks' (ANN) (Kim et al, 2000).

En ese tipo de problemas de modelización de propiedades físico-químicas relacionadas con la madurez de frutas a partir de otros parámetros medidos de forma destructiva, tiene especial importancia el uso de métodos de selección de variables. Estos métodos tienen como objetivo determinar la contribución real de los factores considerados (longitudes de onda, frecuencias acústicas) en la estimación del parámetro objetivo. Estas técnicas facilitan el desarrollo de los modelos predictivos al eliminarse factores espurios que complican el modelo y pueden afectar al rendimiento. Además, pueden facilitar la implantación de estos métodos en la agroindustria. En el caso del NIRS, distintos autores han utilizado la selección de variables para mejorar sus resultados (Liu et al, 2009 y Huirong et al, 2012). Estas metodologías abarcan métodos como el Análisis de Componentes Independientes (ICA) (Liu et al 2009), Sucessive Projections Algorithm (SPA) (Xu et al, 2012) y Uninformative Variable Elimination (UVE) (Xiaobo et al, 2010) y tienen especial relevancia los métodos basados en la 'información mutua' (Mutual Information, MI) (Rossi, 2006; Pent, 2005). Estos métodos utilizan la MI como medida de correlación no lineal entre los factores que intervienen y la variable a estimar y, presentan la ventaja de que no

modifican el espacio original de características considerado, detectando de forma directa los factores a considerar.

El objetivo principal de este trabajo es desarrollar un método de predicción de la firmeza en la variedad de melocotón 'Calrico' mediante la utilización conjunta de información obtenida mediante métodos no destructivos, Vis/NIR y señal acústica (AWETA). Para ello, se ha aplicado un método de selección de características con dos objetivos: a) identificar la relevancia de ambas fuentes de información y b) seleccionar los factores más influyentes para la predicción, que permitan obtener modelos precisos y reproducibles.

5.2.- MATERIAL Y MÉTODOS

5.2.1.- Material vegetal

El estudio se realizó durante las campañas 2010 y 2011 sobre frutos de una finca experimental, situada en "Badilla del Corral de Dolader", próxima a la localidad de Caspe (Zaragoza). El ensayo se realizó con melocotón de la variedad 'Calrico' sobre un patrón híbrido. Los árboles se plantaron en 2000 en un marco de plantación de 5x3 metros. El riego fue de 25.000 L/ha/h). El total de muestras recolectadas fue de 260, 150 en el año 2010 y 110 en la segunda campaña. Los frutos una vez cosechados se transportaron al laboratorio para realizar los análisis. En primer lugar, se realizaron las medidas no destructivas (AWETA and NIRS) y después la medida de la firmeza Magness-Taylor en los mismos puntos que se tomaron las anteriores medidas.

5.2.2.- Magness-Taylor test

El equipo utilizado para medir la firmeza fue un penetrómetro Fruit Pressure Tester FT32 (Istituto per la Valorizzazioni dei Prodotti Agricoli, Milano, Italy) con una sonda de 8 mm. La medida se realizó en dos lados opuestos del fruto retirando previamente la piel de la zona sobre la que se introduce el

vástago de la sonda. La firmeza se midió en los mismos puntos en los que se obtuvo el espectro Vis/NIR.

5.2.3.- Espectroscopia Vis/NIR

Los espectros del melocotón intacto se obtuvieron mediante el equipo Vis/NIR Multispec (Tec5 AG, Frankfurt A.M., Alemania), que consta de un espectrómetro SC – NMS I (Zeiss, Jena, Alemania), una fuente de luz halógena (12 V y 20 Watt) y tanto la luz de iluminación como la luz de medida se transportan mediante fibra óptica. El fruto se coloca en un compartimento en el que se han situado las sondas de iluminación y recepción de luz. Se obtuvieron dos espectros de cada muestra. El intervalo de longitudes de onda utilizado ha sido de 400-1060 nm ($\Delta\lambda=1$ nm). La geometría de iluminación-medida es de 45°/0° (Figura 25)

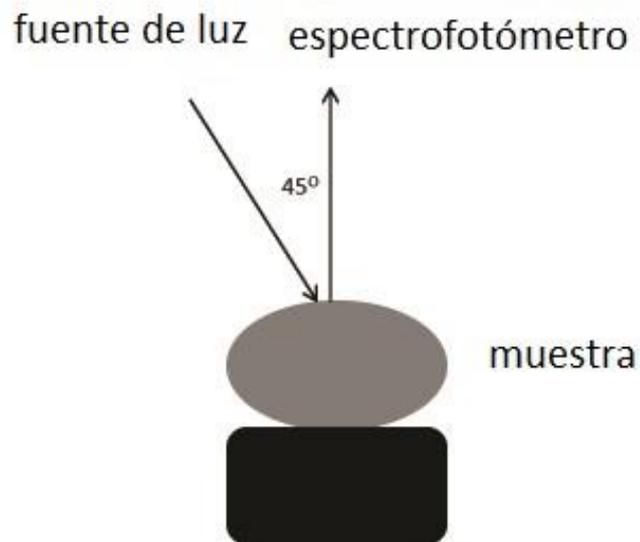


Figura 25. Dispositivo de adquisición del espectro

5.2.4.- Medida de impacto acústico

Las medidas de impulso acústico se realizaron con un equipo comercial (model AFS, AWETA, The Neterlands). El equipo utilizado mide la masa de la fruta y el espectro de las frecuencias resonantes generadas cuando la fruta es sometida a un impacto y proporciona un índice de firmeza (IF) (De Ketelaere et al, 2006; Valente et al, 2009)

$$IF=f_R^2 \times m^{2/3}$$

En este índice se relacionan, la masa, m (en kg) y la frecuencia de la máxima amplitud, f_R (en hercios)

Se realizaron tres medidas en la zona ecuatorial del fruto.

5.2.5.- Tratamiento estadístico

El análisis quimiométrico se llevó a cabo como se indica en el capítulo 3 Materiales y Métodos, utilizando dos métodos de regresión, PLS y LS-SVM, para la obtención de los modelos de calibración. El Software utilizado fue Matlab R2009a (Mathworks, Inc. USA).

La normalización, media cero y desviación estándar uno, fue el pretratamiento utilizado en cada una de las variables independientemente (Rossi et al, 2006; Suykens et al, 2002).

La separación de las muestras en diferentes grupos: calibración y validación, para obtener los modelos de calibración es uno de los pasos críticos. Se realizaron tres subdivisiones al azar, cada grupo estaba formado por muestras de calibración y validación en una ratio de 2:1 para calibración y validación respectivamente (Sánchez et al, 2011). Con el Análisis de Componentes Principales no se detectó ninguna muestra anómala.

Una vez obtenidos los modelos de calibración se realizó una selección de variables, usando el método de Mutual Information, como se describe en el Capítulo 3 de Materiales y Métodos.

Para determinar las mejores ecuaciones de predicción se utilizaron los diferentes estadísticos: R^2 , SEVC, RPD, RMSEP y bias, descritos en el Capítulo 3 de Materiales y Métodos.

5.3.- RESULTADOS Y DISCUSIÓN

En la Tabla 11 se detallan las muestras que componen los grupos de calibración y de validación, los valores de media, desviación estándar, rango y coeficiente de variación para las tres subdivisiones del conjunto inicial de muestras del parámetro firmeza MT. Se observa que los grupos de calibración presentan valores de media, rango y desviación estándar similares. El rango de los subgrupos de validación está comprendido en el de los de calibración lo que confirma que la selección de grupos se ha hecho de forma adecuada. Los valores de desviación estándar y coeficiente de variación indican gran variabilidad dentro de cada uno de los grupos de estudio.

Tabla 11. Número de muestras (n), media (X), desviación estándar (DE), rango y coeficiente de variación (%CV) para la firmeza en los tres grupos de calibración y validación

	Grupo 1		Grupo 2		Grupo 3	
	Calibración	Validación	Calibración	Validación	Calibración	Validación
N	194	66	194	66	194	66
X (kg)	3.9	3.6	3.7	3.5	3.8	3.7
DE(kg)	1.79	1.61	1.7	1.57	1.76	1.7
Rango(kg)	1-8.7	0.6-8.7	1.2-8.7	1-8.7	0.6-8.7	1.2-8.7
%CV	47.26	44.72	46.21	44.86	46.37	46.41

5.3.1.- Predicción de la firmeza MT mediante medida acústica

En primer lugar se determinó la correlación entre los datos de firmeza MT y los de señal acústica. Los coeficientes de regresión de los tres grupos de estudio se sitúan entre $r^2=0.55-0.52$. Se realizó una validación externa donde los valores de r_p^2 fluctuaron entre 0.58-0.45, los valores de RMSEP=1.04-1.16 y el bias=-0.03-0.02. En estudios similares con melocotón, los valores obtenidos en la

correlación entre ambos parámetros fueron algo más bajos $r^2=0.38$ (Golding et al., 2005) y $r^2=0.45$ (Herrero et al., 2009). Los resultados muestran, que pese a tener valores de r^2 más altos que los obtenidos por otros autores, la señal acústica se correlaciona mal con la firmeza MT en comparación con otros métodos no destructivos, como la espectroscopia Vis/NIR (Fu et al., 2008).

5.3.2.- Predicción de la firmeza MT mediante espectroscopia Vis/NIR

En la figura 26 se muestran los espectros de absorción Vis/NIR de melocotones en distintos estadios de madurez. La principal característica de estos espectros es el pico de absorción de la clorofila (680 nm). El contenido de la clorofila decrece a medida que la fruta madura, observándose cómo disminuye el máximo de absorción, hasta que la fruta alcanza un punto de maduración en el que desaparece la absorción a dicha longitud de onda. Los cambios en la zona del espectro por debajo de 700 nm se atribuyeron a modificaciones de color del fruto, debido a la absorción de pigmentos, desaparición de clorofila y síntesis de pigmentos coloreados, principalmente carotenoides y antocianos.

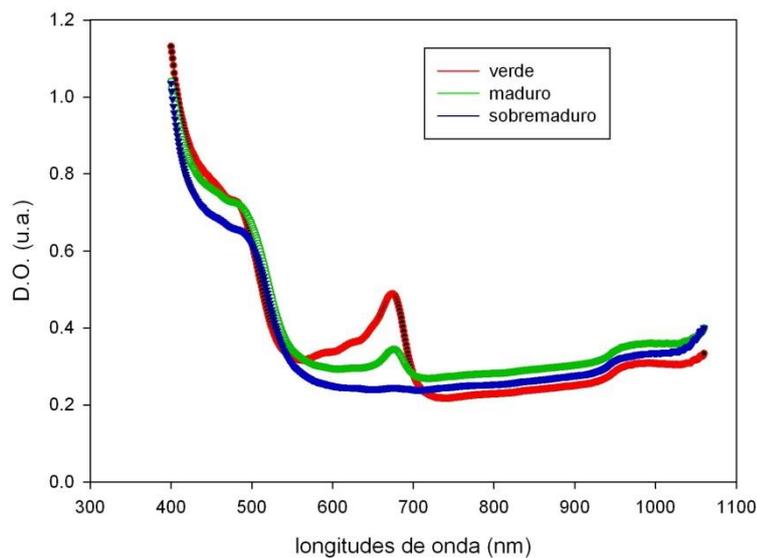


Figura 26. Espectros de absorción de melocotones en distintos estados de madurez.

Los modelos de calibración para la predicción de la firmeza MT a partir de los espectros Vis/NIR del melocotón fueron desarrollados tanto por regresión lineal (PLS) como no lineal (LS-SVM). En la Tabla 12 se muestran los parámetros obtenidos en el análisis estadístico por ambos métodos de regresión y para los tres grupos de calibración. Los valores de R_{cv}^2 , para PLS y LS-SVM, tienen prácticamente el mismo valor (0.7) obteniéndose una buena correlación entre espectros y firmeza. El modelo no lineal LS-SVM presenta valores de SECV más bajos y de RPD más altos para los tres grupos de calibración, lo que significa que con el método de regresión no lineal se obtiene una mejor correlación entre la firmeza MT y el espectro de absorción.

Tabla 12. Estadísticos de calibración para firmeza MT usando método de regresión lineal (PLS) y no lineal (LS-SVM para Vis/NIR)

PLS			
	Grupo 1	Grupo 2	Grupo 3
N	194	194	194
PLS	5	5	5
SECV(kg)	0.97	0.99	0.92
R_{cv}^2	0.69	0.71	0.74
RPD	1.84	1.82	1.91
LS-SVM			
N	194	194	194
PLS	5	5	5
SECV(kg)	0.89	0.87	0.91
R_{cv}^2	0.64	0.71	0.61
RPD	2.00	2.07	1.93

En la tabla 13 se muestran los estadísticos de las validaciones externas realizadas para los tres grupos de muestras. Los valores de r_p^2 son similares para ambos métodos de regresión, en cambio, los valores RMSEP y %CV son más bajos para el modelo PLS. De los resultados obtenidos se infiere la posibilidad de una buena predicción de la firmeza MT mediante el uso de espectroscopia NIR, lo que concuerda con los de otros autores en trabajos similares con melocotón (Fu et al. 2008; $R_{cv}^2=0.74$ y RMSEP=5.42N) que también usaron modelos de regresión

lineal y no lineal y obtuvieron resultados similares con ambos métodos. Costa et al, 2002 utilizaron regresiones lineales obteniendo $R_c^2=0.83$ y $SEC=1.70 \text{ kg/cm}^2$.

Tabla 13. Estadísticos de predicción para firmeza MT usando método de regresión lineal (PLS) y no lineal (LS-SVM) para Vis/NIR

PLS			
	Grupo 1	Grupo 2	Grupo 3
N	66	66	66
RMSEP(kg)	0.85	0.84	0.71
r_p^2	0.71	0.71	0.72
Bias	0.04	-0.01	0.07
%CV	23.94	24.21	19.29
LS-SVM			
N	66	66	66
RMSEP(kg)	0.9	0.92	0.86
r_p^2	0.68	0.74	0.74
Bias	-0.08	-0.07	0.12
%CV	26.09	27.54	23.56

Los coeficientes de regresión extraídos del modelo PLS para uno de los conjuntos de calibración se muestran en la Figura 27. Las longitudes de onda con más relevancia en el espectro de absorción del melocotón se identifican por los valores más altos de los coeficientes de regresión, en valor absoluto. En concreto, se observa que las bandas de 650-690 nm y de 470-550 nm tienen gran importancia, los valores más altos de coeficiente de regresión se encuentran en la banda de 600-700 nm, en conformidad con los trabajos de Zude-Sasse et al, 2006; Zude-Sasse et al, 2002, Ruiz-Altisent et al, 2006 y Shao et al 2011, que consideran esta zona del espectro visible como la que está más directamente relacionada con la madurez de los frutos. Estas bandas de absorción se relacionan con los pigmentos que proporcionan coloración a la fruta.

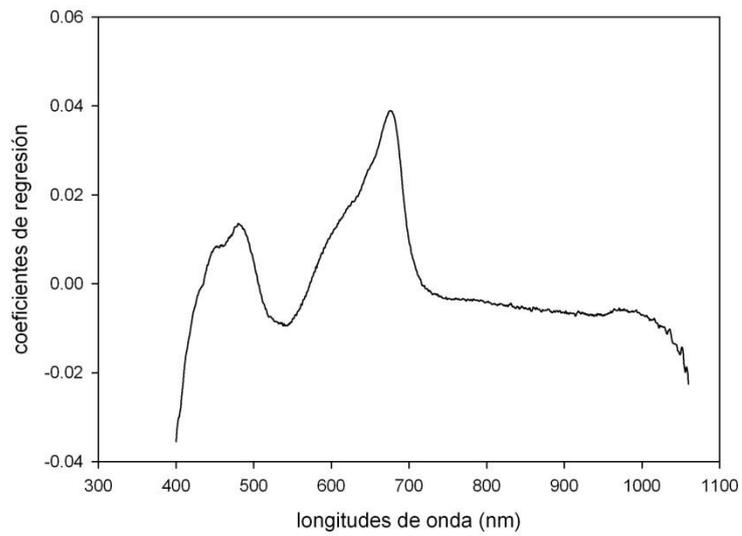


Figura 27. Coeficientes de regresión para firmeza en el grupo de calibración 1

5.3.3.- Predicción de la firmeza MT mediante la combinación de señales acústicas y espectros Vis/NIR.

Tabla 14. Estadísticos de calibración para firmeza usando método de regresión lineal (PLS) y no lineal (LS-SVM) usando Vis/NIR+AWETA

PLS			
	Grupo 1	Grupo 2	Grupo 3
N	194	194	194
PLS	5	5	5
SECV	0.91	0.92	0.9
R_{cv}²	0.74	0.74	0.74
RPD	1.97	1.96	1.95
LS-SVM			
N	194	194	194
PLS	5	5	5
SECV	0.83	0.83	0.84
R_{cv}²	0.86	0.84	0.79
RPD	2.16	2.17	2.09

El índice de señal acústica se unió al espectro de cada una de las muestras, obteniéndose una matriz de entrada con una variable más (194x661).

Las ecuaciones de calibración se volvieron a obtener con los métodos de regresión propuestos (Tabla 14) siguiendo la misma metodología que en los casos anteriores. Se obtuvo que los valores de R_{cv}^2 fueron más altos por ambos métodos (LS-SVM, 0.8-0.9, y PLS 0.7). Los valores de SECV fueron más bajos para PLS y LS-SVM. Los valores de RPD estuvieron en los tres casos por encima del valor 2. En este caso también se llevaron a cabo validaciones externas (Tabla 15). Los RMSEP, los coeficientes de regresión y los coeficientes de variación fueron mejores.

Tabla 15. Estadísticos de predicción para firmeza usando método de regresión lineal (PLS) y no lineal (LS-SVM) usando Vis/NIR+AWETA

PLS			
	Grupo 1	Grupo 2	Grupo 3
N	66	66	66
RMSEP	0.79	0.81	0.81
r_p^2	0.75	0.73	0.77
Bias	-0.005	-0.05	0.01
%CV	22.25	23.34	22.01
LS-SVM			
N	66	66	66
RMSEP	0.86	0.84	0.79
r_p^2	0.71	0.71	0.78
Bias	-0.1	-0.16	0.05
%CV	24.93	25.07	21.18

Los resultados obtenidos demuestran la capacidad de un modelo de calibración muy efectivo para la predicción de parámetros fisicoquímicos de frutas y hortalizas. Se combinan los datos de medidas acústicas con espectros Vis/NIR. Estrategias similares llevaron a cabo Valente et al, 2009 con mango ($R_{cv}^2= 0.61$, RMSECV= 3.20 Nmm) y Zude et al, 2006 con manzana ($R_{cv}^2=0.75-0.8$, SECV= 10.32-11.28 N/cm²).

5.3.4.- Selección de variables

Para reducir el número de variables (longitudes de onda y respuesta acústica) y analizar la posibilidad de mejorar la capacidad de generalización de los modelos, se aplicó el método de selección de variables 'backwards' basado en la Mutual Information (MI) (Herrera et al, 2006). Numerosos investigadores han utilizado métodos de selección para la detección de las variables más relevantes con las que se obtienen modelos más sencillos y con mejores rendimientos. Así Qing et al. (2007) hicieron uso del coeficiente de correlación y Xu et al (2012) aplicaron el método denominado Successive Projection Algorithm (SPA). En la Figura 28 se muestra el rendimiento de los modelos obtenidos mediante los dos métodos de regresión respecto al número de variables utilizadas. Se observa un aumento muy acusado de r^2_p respecto al uso de dos a tres variables para la obtención de las ecuaciones. A partir de la tercera variable el incremento del número de factores no implica un aumento del rendimiento inmediato. El valor del coeficiente de regresión obtenido mediante el uso de tres variables es del orden de 0.77 y 0.73 para LS-SVM y PLS respectivamente. Este comportamiento se observa en los tres grupos de calibración. Los modelos LS-SVM respondieron mejor a la creación de modelos con menor número de variables, como se muestra en otros trabajos similares (Liu et al, 2008, 2009; Li et al, 2013). Esto puede deberse a que en los métodos no lineales se considera la información tanto la lineal como la no lineal, mientras que los modelos lineales (PLS) sólo trabajan con la relación lineal entre los datos espectrales y los físico-químicos. En la tabla 16 se muestran los resultados de validación del modelo construido con tres factores. En ambos casos para LS-SVM y PLS, los valores de r^2_p y RMSEP son mejores que para los modelos construidos con todas las variables y los %CV son similares. Las principales ventajas de la selección de variables son reducir la colinealidad, la redundancia y el ruido que supone el empleo del espectro completo y la obtención de un modelo con sólo tres variables que proporciona mejores resultados.

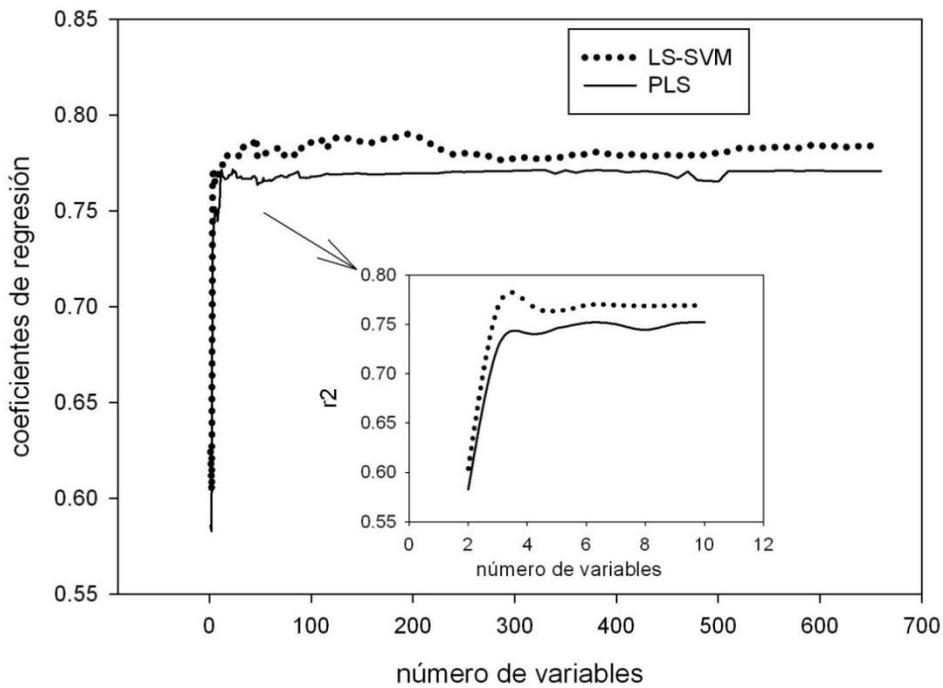


Figura 28. Ejemplo de selección de variable para ambos métodos de regresión, usando todas las variables (661) y usando sólo 10 variables

Tabla 16. Estadísticos de predicción para firmeza MT usando método de regresión lineal (PLS) y no lineal (LS-SVM) usando Vis/NIR+AWETA con tres variables

PLS			
	Grupo 1	Grupo 2	Grupo 3
N	66	66	66
RMSEP	0.77	0.9	0.88
r_p^2	0.76	0.66	0.73
Bias	-0.001	-0.13	-0.005
%CV	21.69	26.55	24.04
LS-SVM			
N	66	66	66
RMSEP	0.76	0.84	0.81
r_p^2	0.77	0.71	0.77
Bias	0.018	-0.13	-0.019
%CV	21.29	24.78	22.13

De las tres variables seleccionadas dos de ellas corresponden a longitudes de onda del espectro Vis/NIR, siendo la tercera el índice de firmeza

obtenido de la señal acústica. Las dos longitudes de onda seleccionadas en los tres casos son similares (679 y 696; 675 y 697; 674 y 693). Esto corrobora la importancia de la zona de absorción de la clorofila para la determinación de la firmeza MT (Figura 33). Resultados similares fueron obtenidos para manzanas (Moons et al, 1997, Tu et al, 1995 and Zude et al, 2002). De esta forma se deduce la posibilidad de poder estimar el estado de maduración con una instrumentación mucho más sencilla en la que sólo se requieren dos longitudes de onda en la zona del visible y un equipo de medida acústica, aportando rapidez al proceso de medida, además de trabajar con modelos mucho más sencillos.

5.4.- CONCLUSIONES

Este trabajo presenta el desarrollo de un método no destructivo, de los cada vez más demandados por las centrales hortofrutícolas y productores, para la firmeza MT en la variedad de melocotón 'Calrico'. La variedad de melocotón tratada es de especial interés en la zona del Bajo Aragón por la intención de introducirla en la denominación de origen 'Calanda'. El método propuesto está basado en la combinación de información no destructiva heterogénea, en concreto una medida de impulso acústico, y datos espectrales Vis/NIR.

Inicialmente, los resultados han demostrado que la conjunción de ambos métodos no destructivos, AWETA y espectroscopia Vis/NIR mejora la fiabilidad de la predicción de firmeza MT, respecto al uso de una única fuente de información. Con esta nueva propuesta, los valores de r^2_p y RMSEP mejoraron en todos los casos. Además, se ha observado que el espectro Vis/NIR proporciona mayor fiabilidad en la estimación de la firmeza MT que la medida AWETA.

Finalmente, el método de selección de variables basado en la Información Mutua utilizado ha conseguido seleccionar tres parámetros óptimos para realizar la calibración de los modelos, provenientes de ambas fuentes de información: la medida acústica y dos longitudes de onda correspondientes a la zona de absorción de la clorofila en el visible. Con la utilización de modelos LS-SVM con tres variables, se observó un mantenimiento (e incluso ligero aumento) del

rendimiento de los modelos de calibración (también en comparación con los modelos PLS). Además, se consiguió una simplificación considerable de los modelos de predicción previos. La ecuación final solo hace uso de 3 de las 661 variables de entrada disponibles, lo que permite proponer una metodología mucho más sencilla y menos costosa.

Hay que resaltar en último lugar que esta metodología propuesta sería potencialmente aplicable a cualquier fruta y hortaliza, lo que sin duda, podría constituir una futura línea de investigación.

6.- Cereza

Este capítulo es una adaptación del artículo enviado a 'Food Sciences and Technology International' con el título 'SSC and Firmness prediction in *Prunus avium* Chelan by Vis/NIR spectroscopy', adjunto en el Anexo III.

6.1.- INTRODUCCIÓN

El cerezo (*Prunus avium*) es una de las especies frutales de hueso más importante. En el año 2013 la producción de cereza en España fue de 96.946 toneladas. Una de las principales comunidades productoras de cereza es Aragón, alcanzando al año unas 20.562 toneladas. Las cerezas son frutos no climatéricos, no maduran una vez recogidas del árbol, por lo que deben recolectarse cuando alcanzan las características físico-químicas y organolépticas deseadas y se dispone de unos 7 días para su venta y consumo. A medida que la cereza madura en el árbol, aumenta el contenido de sólidos solubles mientras que la acidez y firmeza disminuyen. Los índices más utilizados para definir el momento de recolección de la cereza son: el contenido en sólidos solubles, la firmeza y el color de la epidermis. El aumento del contenido en sólidos solubles, debido a la degradación del almidón, se ha determinado tradicionalmente de forma destructiva mediante refractometría. La percepción de la firmeza por compresión manual sigue siendo el método más utilizado por productores y consumidores (Mitchman et al, 1998). El color es considerado el factor de calidad externo que más influye en el consumidor (Francis, 1995), cambia de verde a rojo brillante para pasar a colores rojos púrpuras. Durante la maduración, en el cambio de color que sufre la cereza se produce por acumulación de antocianinas (Webster and Looney, 1996) y carotenoides y la degradación de la clorofila. Para la determinación del color, algunos investigadores han utilizado cartas de color como guía para determinar el momento óptimo de cosecha (Timm et al, 1993), mientras que otros lo han hecho mediante las coordenadas CIELAB (Gao et al, 1995).

La empresa hortofrutícola requiere de sistemas efectivos y no destructivos de control de calidad. Desde los años 90, para estudiar la madurez de las frutas, se están desarrollando técnicas ópticas como la espectroscopia de infrarrojo cercano (NIRS), que ofrece una serie de ventajas frente a las técnicas tradicionales: evita la destrucción del fruto, proporciona una medida in situ más rápida, reduce el coste y puede determinar varios parámetros de calidad en una sola medida. Esta técnica consiste en un haz de luz que incide sobre la muestra; una parte sufre reflexión especular y es la responsable del brillo, otra parte, después de penetrar ligeramente en el fruto es absorbida de forma selectiva por los pigmentos, y el resto de la luz es reflejada por la muestra en modo difuso produciendo los espectros Vis/NIR de los que se obtienen la información deseada. Se han llevado a cabo numerosos estudios utilizando esta técnica por ejemplo: en manzana (Fan et al, 2009; Pissard et al, 2013), pera (Li et al, 2013), tomate (Tiwari et al, 2013), aguacate (Clark et al, 2003), naranja (Cayuela and Weiland, 2010) y cereza (Lu, 2001)

La tendencia actual se dirige a la búsqueda de métodos no destructivos de control de madurez relacionando la información espectral Vis-NIR de las muestras con valores de referencia obtenidos de forma destructiva. De esta forma se crean modelos estadísticos para predecir el parámetro de calidad deseado. Para la construcción de estos modelos se utilizan métodos de calibración multivariante lineales como, Multiple Linear Regression (MLR, Jha et al, 2014); Principal Components Regression (PCR, Hernández-Gómez et al, 2006), Partial Least Squares (PLS, Liu and Ying, 2005). En muchos casos, esta relación no es estrictamente lineal, por tanto hay que utilizar métodos no lineales, como Least Squares Support Vector Machines (LS-SVM, Chauchard et al, 2004) y Artificial Neural Networks (ANN) (Shao et al, 2008; Pérez-Marín et al, 2007).

En general, los métodos de regresión multivariante hacen uso de todas las variables del espectro para la obtención de los modelos de calibración. Estos métodos tienen la capacidad parcial de reducir el impacto de problemas como la colinealidad, superposiciones de banda e interacciones. Sin embargo, la influencia de variables colineales o que no contienen información relevante puede perjudicar la obtención de dichos modelos (Xiaobo et al, 2010). Algunas

variables aportan información útil y otras sólo ruido. Por tanto, es necesaria la selección previa de un número reducido de variables que permitan obtener una calibración mejor y más sencilla. Alguno de los métodos de selección de variables más utilizados en espectroscopia Vis/NIR son: Sucessive projections algorithm (SPA) (Costa-Pereira et al, 2008); Genetic algorithm (GA) (Wu et al, 2014); Uninformative variables elimination (UVE) (Wu et al, 2012); Stability competitive adaptive reweighted sampling (SCARS) (Zhang et al, 2013); Monte Carlo uninformative variables elimination (MC-UVE) (Xu et al, 2014); Mutual Information (MI) (Rossi et al, 2006; Herrera et al, 2015).

El objetivo principal de este trabajo ha sido investigar la relación entre los parámetros de calidad firmeza y contenido en sólidos solubles, con los espectros de reflexión Vis/NIR en *Prunus avium* cv 'Chelan'. Los modelos de calibración se han obtenido mediante los métodos de regresión PLS y LS-SVM. Además, se propone una selección de variables con el objetivo de detectar las longitudes de onda más influyentes en los parámetros físico-químicos y la obtención de calibraciones más estables y sencillas.

6.2.- MATERIAL Y MÉTODOS

6.2.1.- Material vegetal

Se recolectaron 282 muestras de *Prunus avium* 'Chelan' durante las campañas 2011 y 2012, cultivadas en una explotación comercial propiedad de Orchard, S.L. Se encuentra situada en el término de la Almunia de Doña Godina (Zaragoza, Spain). Los árboles se plantaron en el año 2000 en un marco de plantación de 5x3 m. El riego se aplicó por goteo con un caudal de 25.000L/ha/h. Las muestras fueron recogidas durante la cosecha semanalmente durante la primera campaña y cada dos días durante un período de dos semanas de cosecha, para la segunda campaña (47 muestras por día de recolección). Inmediatamente se transportaron al laboratorio donde se registraron los espectros Vis/NIR en dos puntos opuestos del fruto. A continuación se les

determinó firmeza y contenido en sólidos solubles en los mismos puntos de medida de la fruta.

6.2.2.- Espectroscopia Vis/NIR

El equipo utilizado fue un equipo modular de reflectancia QualitySpec Pro 2600 (Analytical Spectral Devices, INC. Colorado, USA). Los módulos utilizados fueron: espectrofotómetro de barrido rápido con un intervalo de medida de 350-2500nm ($\Delta\lambda = 1\text{nm}$), lámpara de iluminación de Wolframio (12V/45W) y sistema de detección formado por un monocromador con doble detector de InGaAs. La velocidad de escaneo es de 10 espectros/segundo. El ángulo formado entre la entrada y salida de luz es de 25° . El equipo mide los espectros de reflexión en unidades de densidad óptica ($D.O = \log[1/\text{Reflectancia}]$).

Se realizó la medida de dos espectros en la zona ecuatorial por fruto. Para el análisis quimiométrico se utilizó la media de ambos espectros. Se eliminó la parte inicial y final de los espectros para eliminar ruido, con lo que el intervalo de trabajo del equipo QualitySpec 2600 fue 400-2420 nm.

6.2.3.- Firmeza

La firmeza se determinó mediante un durómetro digital no destructivo AGROSTA 100Field (Agro Technology, Forges Les Eaux, France). El equipo no penetra los frutos. Su medición se basa en cuantificar la fuerza necesaria para retraer un émbolo una cierta longitud al aplicarlo de forma manual contra la superficie de la fruta dando una idea de firmeza a través de la deformación del mismo. Los resultados se muestran en unidades DI_{10} (Durofel Indice), estando ajustado el equipo con una prueba de 0.10 cm^2 . Las medidas se hicieron en cada una de las dos caras de la cereza en las que se midieron los espectros Vis/NIR.

6.2.4.- Contenido en Sólidos Solubles

El análisis de sólidos solubles totales se llevó a cabo mediante el uso de un refractómetro digital (Modelo ATAGO PR-101 Co., Tokyo, Japan). La medida se realizó sobre una gota de zumo de cereza. El análisis se hizo por duplicado en cada fruto.

6.2.5.- Tratamiento estadístico

El cálculo de las ecuaciones de calibración se realizó usando dos métodos de análisis multivariante: PLS y LS-SVM. El Software fue Matlab R2009a (The MathWorks, Natick, United States) con su propia programación.

El pretratamiento aplicado a cada una de las variables del espectro fue una normalización usando media cero y desviación estándar uno (Rossi et al, 2006; Suykens et al, 2002).

A continuación, para la detección de anomalos espectrales se empleó el estadístico *leverage*, calculado mediante un PCA. El resultado mostró que no había muestras anómalas. Este análisis de PCA se realizó mediante The Unscrambler X.1 (CAMO Software AS, Norway). Estos análisis se encuentran descritos en el Capítulo 3 de Materiales y Métodos.

La selección de muestras de calibración y validación del conjunto total es un punto crítico en la construcción del modelo de calibración. Para realizar la calibración, se ha dividido el total de las muestras en dos grupos; uno está formado por 2/3 de las muestras, seleccionadas aleatoriamente por el Software, y el tercio restante se ha destinado a muestras para la validación (Flores et al, 2009; Paz et al, 2009). Este proceso se ha repetido tres veces.

Los mejores modelos de calibración fueron seleccionados en base al valor más alto del coeficiente de determinación de validación cruzada (R_{cv}^2) y al menor valor de error estándar de validación cruzada (SEVC y RPD (Residual Predictive Deviation)) (Williams, 2001). Este mismo autor, señala que valores de RPD comprendidos entre 3 y 5 indican predicciones Vis/NIR eficientes.

Respecto a la validación, los resultados de los diferentes grupos de predicción se evaluaron comparando el error medio de predicción (RMSEP), el coeficiente de validación externa (r^2), bias y el error estándar de predicción (SEP).

Tras obtener los mejores modelos de calibración con todas las longitudes de onda, se realizó un proceso de selección de variables para simplificar los modelos. Para ello, se aplicó el método de Información Mutua, descrito en el Capítulo 3, Materiales y Métodos.

6.3.- RESULTADOS Y DISCUSIÓN

En la tabla 17 se muestran los valores medios, desviación estándar, rango y coeficiente de variación para las tres subdivisiones de muestras utilizadas en la calibración y validación. El rango para el conjunto de validación se encuentra dentro del rango del conjunto de calibración, esto indica que la selección de los grupos se ha realizado de forma adecuada por el software.

Tabla 17. Número de muestras (n), intervalo, media, desviación estándar (DE) y coeficiente de variación (%CV) para los dos parámetros de estudio para los tres grupos de calibración (n=211) y validación (n=70)

Parámetro	Grupo	Set	n	intervalo	media	DE	%CV
Durofel	1	Calibración	211	18-58	41.1	10.0	24.4
		Validación	70	19-56	42.4	9.8	23.3
	2	Calibración	211	18-58	42.1	9.9	23.6
		Validación	70	21-55	41.3	10.1	24.6
	3	Calibración	211	21-58	41.2	9.9	24.1
		Validación	70	18-57	42.4	10.1	24
SST (°Brix)	1	Calibración	211	9.5-30	17.7	4.8	27.1
		Validación	70	8.7-27.3	16.2	4.3	26.5
	2	Calibración	211	8.7-30	17.2	4.7	27.3
		Validación	70	9.8-28.1	17.5	4.7	26.9
	3	Calibration	211	8-7-28.1	17.4	4.7	27.0
		Validation	70	9.5-30	17.1	4.6	26.9

En la figura 29 se muestran ejemplos de los espectros de absorción de la cereza obtenidos con el equipo Vis/NIR utilizado. En la región del visible, se observa el pico de absorción de la clorofila ($\lambda \approx 675$ nm) que disminuye al madurar la cereza. En la región de infrarrojo (700-2000) aparecen varias bandas de absorción: 980, 1470 y 1940 nm, relacionadas con el alto contenido de agua que presenta la fruta (Osborne et al, 1993; Lammertyn et al, 2000; Nicolai et al, 2007). Además se observa una banda en torno a 1200 nm característica de la absorción de carbohidratos (Osborne et al, 1993). Espectros con bandas similares para cereza fueron obtenidos por Lu, 2001.

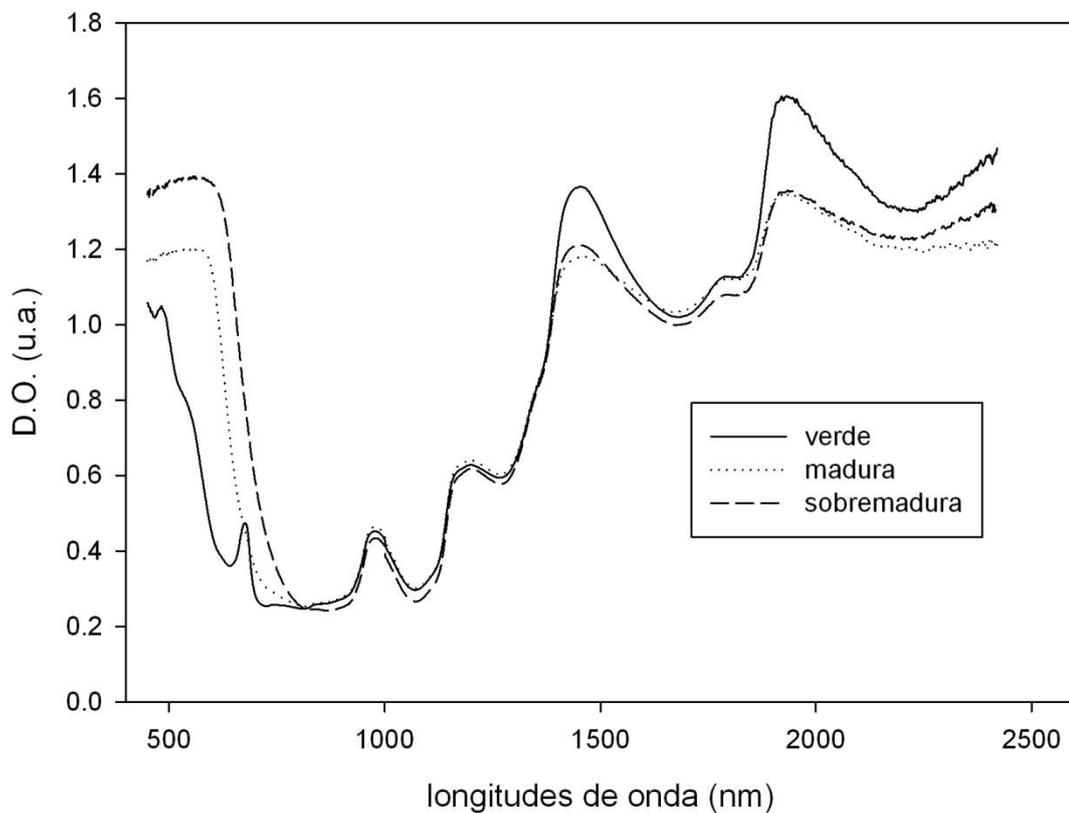


Figura 29. Espectros Vis/NIR de absorción para *prunus avium* 'Chelan' con el equipo QualitySpec (450-2500 nm).

6.3.1.- Modelos de calibración

En las Tabla 18 se muestran los estadísticos de calibración para los tres grupos de muestras seleccionados para la determinación de firmeza y SST respectivamente, usando para ello el intervalo de 450-700 nm para determinar firmeza y dos intervalos para SST: 700-2420 y 700-1060 nm.

Respecto al método de regresión, para el caso de la firmeza se obtuvieron resultados similares para ambos métodos de regresión PLS y LSSVM ($R_{cv}^2 = 0.82$ vs 0.76 , $SECV = 4.78$ vs 4.73 , respectivamente), a diferencia de lo obtenido por Li et al (2013); Liu et al (2008, 2009). Los valores de RPD obtenidos con el método LS-SVM están entre 2.1-2.2. En SST se obtienen estadísticos muy similares para ambos métodos de regresión LS-SVM y PLS respectivamente ($R_{cv}^2 = 0.96$ vs 0.97 , $SECV = 0.82$ y 0.83°Brix). Se obtienen mejores resultados para el intervalo de 700-1060 nm que para el de 700-2420 nm con ambos métodos como puede observarse en la Tabla 18. Los valores de RPD son mucho mayores para el intervalo de 700-1060 nm, estando entre 5.5-5.8. Se puede observar que la firmeza se relaciona con la banda del espectro visible y los SST con la banda del infrarrojo cercano del espectro, resultados muy acordes con los de otros estudios en los que se obtienen conclusiones similares (Carlini et al, 2000; McGlone et al, 2002; Martínez Vega et al, 2013).

Tabla 18. Estadísticos de calibración de tres grupos de calibración, para dos métodos de regresión PLS y LS-SVM, para la firmeza en el intervalo de 450-700 nm y para SST 700-2420 nm y 700-1060 nm

		PLS			LS-SVM		
		Grupo 1	Grupo 2	Grupo 3	Grupo 1	Grupo 2	Grupo 3
Firmeza (450-700 nm)	N	211	211	211	211	211	211
	LVs	5	5	5			
	SECV	5.86	6.12	5.67	4.51	4.74	4.69
	R_{cv}²	0.66	0.65	0.69	0.76	0.76	0.74
	RPD	1.71	1.62	1.75	2.22	2.09	2.11
SST (700-2420 nm)	N	211	211	211	211	211	211
	LVs	17	13	15			
	SECV(°Brix)	1.32	1.28	1.3	1.25	1.22	1.21
	R_{cv}²	0.98	0.96	0.97	0.92	0.92	0.92
	RPD	3.64	3.67	3.61	3.84	3.85	3.88
SST (700-1060 nm)	N	211	211	211	211	211	211
	LVs	11	11	11			
	SECV(°Brix)	0.82	0.83	0.83	0.83	0.86	0.82
	R_{cv}²	0.97	0.97	0.98	0.96	0.96	0.96
	RPD	5.85	5.66	5.66	5.78	5.46	5.73

6.3.2.- Validación

Para llevar a cabo la validación de precisión de los modelos de calibración se utilizaron los estadísticos RMSEP y r_p^2 . En la Figura 30, se muestran estos estadísticos para uno de los grupos de predicción en los intervalos de longitudes de onda en los que se han obtenido mejores resultados aplicando las ecuaciones de calibración, para firmeza (450-700nm) y sólidos solubles (700-1060nm).

Los estadísticos SECV y RMSEP fueron similares a los obtenidos en la calibración, lo que indica que el uso de estos modelos de calibración es aplicable para la predicción de firmeza y SST.

En ambos parámetros, los estadísticos de predicción fueron similares con independencia del método de regresión, siendo ligeramente mejores los obtenidos por método lineal (PLS) para el contenido el sólidos solubles totales (PLS, RMSEP=4.61°Brix, $r_p^2=0.79$; LSSVM, RMSEP=5.03°Brix, $r_p^2=0.75$).

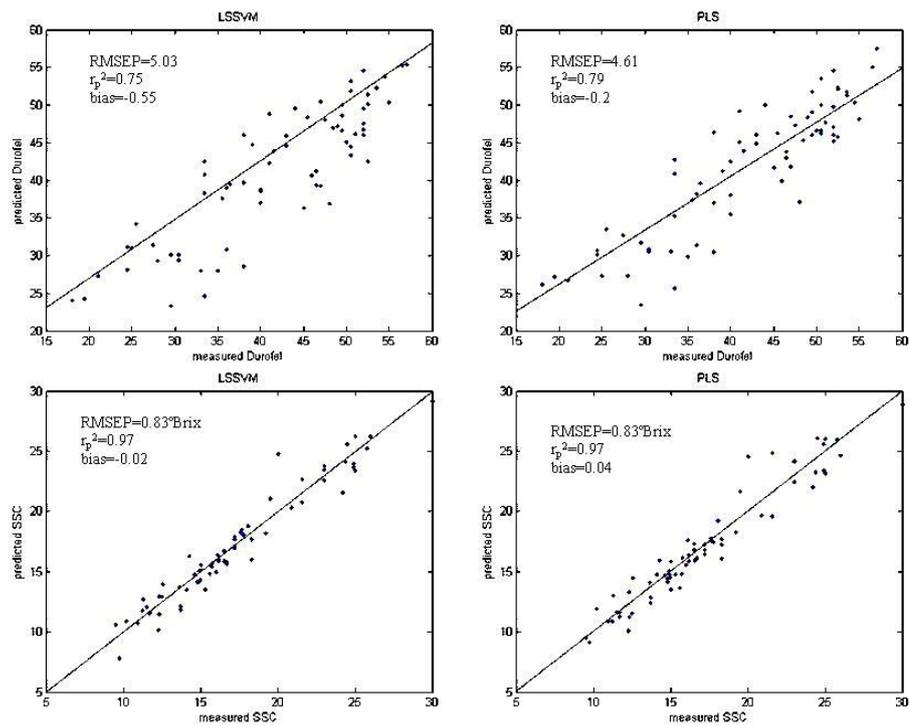


Figura 30. Estadísticos de predicción para firmeza y sólidos solubles para ambos métodos de regresión, PLS y LS-SVM

6.3.3.- Selección de variables

En la Figura 31 se muestran para SST y firmeza Durofel, los rendimientos de los modelos respecto al número de variables utilizadas. En ambos casos, se estabilizan los rendimientos a partir del uso de cuarenta longitudes de onda aproximadamente. Los valores de R_{cv}^2 obtenidos son 0.74 y 0.97 para firmeza y SST respectivamente, semejantes a los obtenidos con los modelos que utilizan todas las variables. Es de señalar que los resultados de la validación para la firmeza Durofel, en algunos de los subgrupos, son incluso más ajustados que en la calibración, lo que demostraría que esta selección de variables proporciona buenos modelos de calibración.

En la Figura 32 aparecen las longitudes de onda que se correlacionan mejor con la firmeza y contenido en sólidos solubles. En el caso de la firmeza, las

primeras variables seleccionadas corresponden a la banda de absorción de la clorofila, confirmando de nuevo que la firmeza está relacionada con el contenido en clorofila (Lafuente et al, 2014; Zude et al, 2011; Zude et al, 2002). Para SST, las longitudes de onda destacadas corresponden a la zona de espectro entre 700-820 nm que es uno de los rangos del espectro menos sujeto a variación en el espectro de absorbancia (Qing et al, 2007b). Además, determinadas longitudes de onda como la de 980 nm, se ha descrito que tiene relación con el contenido de agua en la cereza (Osborne et al, 1993; Lammertyn et al, 2000; Nicolai et al, 2007).

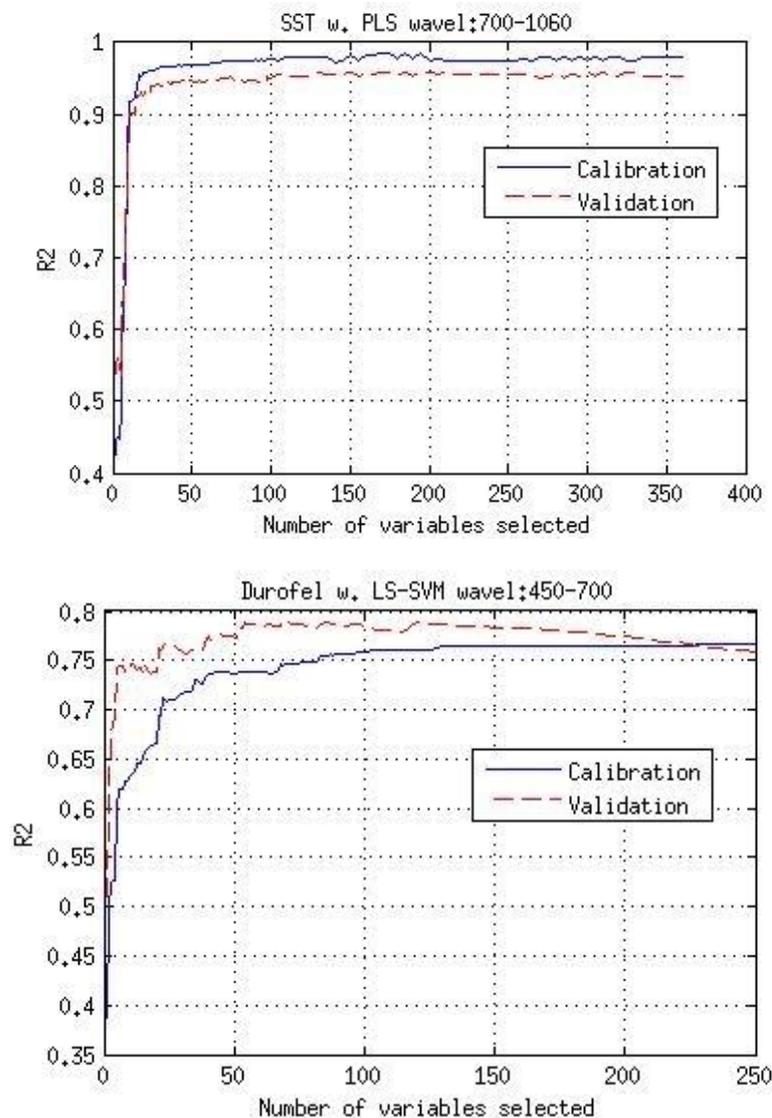


Figura 31. Ejemplo de selección de variables para SST y firmeza

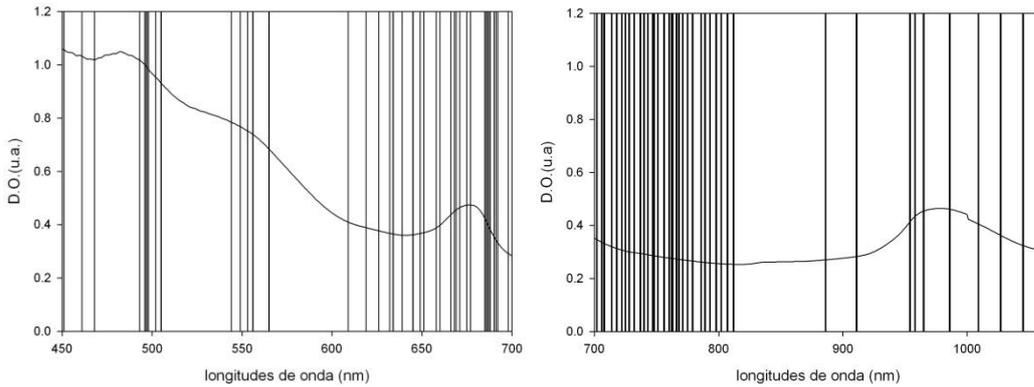


Figura 32. Selección de longitudes de onda más representativas en la determinación de firmeza (a) y SSC (b)

6.4.- CONCLUSIONES

En este trabajo se presenta el desarrollo de modelos de predicción para determinar SST y firmeza Durofel en *Prunus Avium cv. Chelan* mediante espectroscopia Vis/NIR, usando métodos de regresión lineal (PLS) y no lineal (LS-SVM). El desarrollo de esta metodología no invasiva es de gran interés para el sector frutícola en general y en particular para la cereza que es una de las especies frutales con mayor potencial de exportación.

La selección de intervalos del espectro para la determinación de firmeza (450-700 nm) y SST (700-1060 nm) mejoran la bondad de la predicción. Así, los valores de r_p^2 y RMSEP mejoran respecto al uso del espectro completo.

El uso del método de regresión no lineal LS-SVM ha mejorado los estadísticos de los modelos de predicción para de la firmeza (PLS, $r_p^2=0.71$, RMSEP=5; LS-SVM, $r_p^2=0.76$, RMSEP=4.5). En el caso de los SST, los resultados fueron similares por ambos métodos de regresión (PLS, $r_p^2=0.97$, RMSEP=0.83^oBrix; LS-SVM, $r_p^2=0.97$, RMSEP=0.86^oBrix).

Finalmente, la selección de variables por el método MI, ha permitido reducir el número de longitudes de onda de medida de unas 300 a 40 aproximadamente, obteniéndose en la calibración, resultados similares a los obtenidos usando todas las longitudes de onda registradas en el espectro. Esto

permite el desarrollo de equipos mucho más simples para la determinación del índice de madurez de la cereza.

Es necesario seguir trabajando en el desarrollo de los modelos de calibración basados en la metodología NIRS, aumentando el número de campañas y extendiéndolo a otras variedades de cereza.

7.- Discusión general

El trabajo descrito en la presente Tesis Doctoral ha pretendido profundizar en el conocimiento de tecnologías no destructivas: ópticas (Espectroscopia de Infrarrojo Cercano y Análisis de Imágenes Multiespectrales), de impulso acústico (AFS) y de deformación (Durofel) para la determinación de los parámetros de la calidad, entendiendo como tales, el contenido en sólidos solubles totales (SST) y la firmeza de la fruta, con el objetivo de predecir el momento óptimo de cosecha.

Las técnicas de impulso acústico y de deformación se utilizan desde hace años en la determinación del punto óptimo de cosecha en el campo.

La espectroscopia de Infrarrojo Cercano presenta una buena capacidad de predicción para SST (Martínez-Vega et al, 2013; Lu, 2001), firmeza (Moons et al, 1999; Fu et al, 2007), pH (Li et al, 2013; Shao et al, 2011), acidez (Peirs et al, 2000; Cen et al, 2007), obteniéndose buenas correlaciones entre espectros Vis/NIR y estos índices de calidad.

Sin embargo, de las técnicas de Análisis de Imágenes Multiespectrales apenas se han descrito estudios, en los que se correlacionen parámetros de las imágenes (diámetros, histograma, anchuras, parámetros de centrado) con los índices de calidad. Los trabajos desarrollados presentan buenos resultados en lo que respecta a la predicción de la madurez en frutas (Peng et al, 2007; Hashim, et al, 2013; Lorente et al, 2013).

Para la mejora de las ecuaciones de calibración, se aborda la aplicación de diferentes métodos de regresión lineal (Partial Least Squares, PLS) y no lineal (Least Squares Support Vector Machine, LSSVM), tomando como referencia estudios previos (Liu et al, 2005, 2009; Chauchard et al, 2004). Varios autores, defienden que el método de regresión no lineal mejora la capacidad predictiva de los modelos de calibración. En cambio, en el presente estudio se pone de manifiesto que el método de regresión LSSVM no siempre mejora los estadísticos, dependiendo del tipo de fruta con el que se trabaje.

Además, se han utilizado métodos de selección de variables, con el objetivo de simplificar los cálculos y la complejidad de los equipos destacando las variables más influyentes en los parámetros de calidad de la fruta. En el presente

trabajo se demuestra que es posible la mejora de resultados al reducir el número de variables utilizando un método de selección, como el de Mutual Information (MI). En distintos trabajos se utilizan métodos de selección de variables aplicados a técnicas ópticas (Costa-Pereira et al, 2008; Wu et al, 2014; Wu et al, 2012; Zhang et al, 2012; Xu et al, 2014; Herrera et al, 2006, 2015).

Para llevar cabo estos estudios, hemos utilizado las siguientes variedades de fruta durante dos campañas:

- Cereza: *prunus avium cv. Chelan*.
- Melocotón: *prunus persica cv. Calrico*.
- Manzana: *malus domestica cv. Smoothee Golden Delicious*.

Referente al estudio de la manzana, (Capítulo 4) se decidió trabajar con la técnica de Análisis de Imágenes Multiespectrales, con la utilización de 5 láseres distintos como fuente de iluminación, que ha sido desarrollada especialmente en manzanas de diferentes variedades: Gala, Fuji, Golden Delicious, Elstar y Pinova. Como novedad, se emplea esta técnica en la variedad 'Smoothee Golden Delicious' para la predicción de SST y firmeza.

En varias publicaciones, (Lu, 2004; Qing et al, 2007a; Romano et al, 2008), se relacionan parámetros de la imagen multiespectral con índices de calidad. En la presente tesis, se evaluaron dos parámetros de las imágenes multiespectrales: el histograma y los diámetros de Feret

La revisión exhaustiva de la literatura especializada ha revelado que apenas se han publicado trabajos en los que se haya estudiado la correlación entre el histograma de la imagen multiespectral con el estado madurez del fruto. En este trabajo hemos utilizado cinco láseres de partida. Los resultados obtenidos han sido adecuados para el contenido en sólidos solubles ($R_{cv}^2=0.69$) y firmeza ($R_{cv}^2=0.81$), especialmente cuando se trabaja con el que emite a 785 nm. Resultados similares se obtuvieron en el trabajo de Qing et al (2007a), usando el histograma con un láser de la misma longitud de onda. La aplicación de esta técnica para predecir la madurez en manzanas de campañas posteriores no

resultó viable. Sin embargo, se realizó una clasificación mediante un análisis de componentes principales en función del estado de madurez. Con esta clasificación se pudieron diferenciar tres grupos de muestras en distintos estadios de maduración y en dos campañas sucesivas. El hecho de no haber predicho los parámetros físico-químicos de la campaña posterior, se debe a que las muestras de la primera campaña fueron muy diferentes, en cuanto al estado de madurez, a las de la segunda campaña. Por tanto, proponemos seguir trabajando en la obtención de modelos de predicción con campañas sucesivas, para obtener un conjunto de muestras variables.

Respecto a los diámetros de Feret se observa una gran variación según la madurez del fruto. Gracias a esta variabilidad, se utiliza una combinación de tres láseres (670, 785, 980 nm) para poder predecir el contenido en sólidos solubles y firmeza. Los resultados obtenidos son satisfactorios, $r_p^2=0.86$ y $RMSEP=1.10^\circ\text{Brix}$ para SSC y $r_p^2=0.74$ y $RMSEP=13.3$ N para firmeza. En trabajos con metodología similar se han obtenido peores resultados, Lu, 2004a ($r_p^2=0.59$ y $SEP=0.78^\circ\text{Brix}$ para SSC y $r_p^2=0.76$ y $SEP=5.8$ N para firmeza); Lu, 2004b ($r_p^2=0.58$ y $SEP=6.18$ N).

En el capítulo 5 se abordó la variedad de melocotón '*Calrico*', candidata a ser incluida en la DO Melocotón de Calanda. Se estudió el desarrollo de las técnicas no destructivas, espectroscopia Vis/NIR e impulso acústico, para desarrollar modelos de calibración que nos permitan predecir la firmeza en el punto óptimo de recolección. Apenas se disponen de referencias respecto al uso de métodos quimiométricos en esta fruta (Fu et al, 2008; Shao et al, 2011; Fang et al, 2013; Liu et al, 2014).

En este ensayo se utilizaron un equipo Vis/NIR, con intervalo de trabajo de 400-1060 nm y un equipo de impulso acústico. Se desarrollaron ecuaciones de calibración con cada una de las técnicas por separado para predecir la firmeza. Pero la combinación de ambas metodologías no destructivas para la determinación de este índice, proporcionó mejores resultados ($r_p^2=0.78$ y

RMSEP=0.79 kgf). Metodologías similares se han desarrollado con éxito en otros estudios (Ruiz-Altisent et al, 2006; Valente et al, 2009).

El uso de un método de selección de variables contribuyó a simplificar la obtención de resultados. En este caso, se seleccionaron dos variables del equipo Vis/NIR y el índice correspondiente a la medida acústica. Las ecuaciones de calibración construidas con las tres variables seleccionadas proporcionaron estadísticos similares a los obtenidos con todas las variables ($r_p^2=0.77$ y RMSEP=0.76 kgf), consiguiendo reducir la matriz de variables de 661 a tan sólo 3. De esta manera, se simplifican notablemente el diseño de los equipos y los cálculos. En trabajos de otros autores, la selección de variables también proporcionó una mejora de los resultados (Qing et al, 2007b; Xu et al, 2012).

También se ha realizado una comparación entre métodos de regresión lineal y no lineal. En el caso del melocotón, los resultados se mejoraban mínimamente con el empleo del método de regresión no lineal (LS-SVM: $R_{cv}^2=0.86$, SECV=0.83 kgf; PLS: $R_{cv}^2=0.74$, SECV=0.91 kgf). En algunos trabajos de comparación de métodos de regresión consiguen mejorar los resultados con el método no lineal (Chauchard et al, 2004; Liu et al, 2008, 2010)

En el capítulo 6 se ha planteado el estudio de la cereza por el gran volumen de negocio que supone la exportación de esta especie para nuestra provincia. La cereza es una fruta que es necesario recolectarla en el momento exacto de madurez por lo que el desarrollo de la tecnología Vis/NIR, es crucial para determinar su punto óptimo de recolección. Además, este estudio contribuye a incrementar los escasos conocimientos disponibles acerca de la aplicación de la espectroscopia Vis/NIR en cereza (Carlini et al, 2000; Lu, 2001).

En este estudio, a pesar de haber utilizado un equipo Vis/NIR con un intervalo de longitudes de onda 350-2500 nm, se ha demostrado que el intervalo del espectro de 450 a 1060 nm basta para la determinación de la calidad de la cereza (SST y firmeza). Más concretamente, la zona del visible (450-700 nm) para la firmeza y el intervalo del infrarrojo cercano (700-1060 nm) para contenido en

sólidos solubles. Otros autores también relacionan estas bandas del espectro con los parámetros físico-químicos (Carlini et al, 2000; McGlone et al, 2002; Martínez Vega et al, 2013; Lafuente et al, 2014).

También se han evaluado métodos de regresión, lineal y no lineal, para verificar si existen diferencias en la capacidad predictiva de los modelos de calibración. En el presente estudio, no se han obtenido ventajas apreciables entre la aplicación de un método de regresión u otro.

En este caso, también se ha utilizado el método de selección de variables Información Mutua, sin mejorar tampoco la capacidad predictiva de las ecuaciones. Sin embargo, la aplicación de esta metodología ha permitido reducir el número de variables relevantes lo que a su vez facilita la construcción de equipos de medida más sencillos y sin que necesariamente deban incluir en su construcción redes de difracción que es siempre uno de los componentes más costosos debido a la perfección necesaria para construir ópticas de calidad analítica. En este caso, se pasó de usar las 2150 longitudes de onda del espectro total que proporciona el equipo a solamente 40 obteniendo resultados similares.

Las variables seleccionadas corresponden para el caso de la firmeza con la banda del visible, en la que absorben clorofila y antocianos mientras que la banda que mejor refleja el contenido en sólidos solubles es la del infrarrojo cercano (700-1000 nm) que corresponde a una banda de absorción de agua. Estos resultados están de acuerdo con los de Martínez-Vega et al, 2013 y Zude et al, 2006.

En esta tesis hemos usado técnicas clásicas destructivas y otras, más modernas no destructivas, comprobando la eficacia de las no destructivas para determinar el momento óptimo de cosecha en melocotón cv. *'Calrico'*, manzana cv. *'Smoother Golden Delicious'* y cereza cv. *'Chelan'*.

El uso de la metodología de Espectroscopia de Infrarrojo Cercano permite la predicción de los parámetros físico-químicos usados en melocotón y cereza.

La determinación de índices de calidad de la manzana mediante la técnica de Análisis de Imágenes Multiespectrales también ha resultado viable.

El uso de métodos de regresión lineal, PLS, y no lineales, LS-SVM, ha permitido mejorar la capacidad predictiva de las ecuaciones de calibración.

En el caso del melocotón, el método no lineal produce ligeros mejores resultados. Sin embargo, en cereza, no se aprecian diferencias entre los resultados obtenidos por ambos métodos.

En los capítulos dedicados a la cereza y al melocotón, la selección de variables del espectro Vis/NIR destaca la relación directa que existe entre la firmeza de la fruta con la longitud de la absorción de la clorofila (670 nm), y el contenido en azúcares con la banda del infrarrojo cercano. Además, la selección de variables nos permite trabajar con un número reducido de longitudes de onda, lo que facilita el diseño de equipos y reduce el tiempo de medida

8.- Conclusiones finales

Siguiendo la línea de los objetivos expuestos en el capítulo 2 y, a partir de los resultados obtenidos en los diferentes capítulos de esta tesis, podemos establecer las conclusiones que exponemos a continuación. Estas conclusiones tienen como característica común que hemos sustituido los métodos tradicionales destructivos de parámetros de calidad de la fruta por otros no destructivos que, además, hemos tratado de simplificar al máximo los cálculos de las ecuaciones de calibración utilizables en posibles aplicaciones industriales posteriores.

Estas conclusiones son:

MANZANA cv 'Smoothie Golden Delicious'

1. El análisis de imágenes multiespectrales de retrodifusión de láseres en manzana, permite predecir la firmeza y el contenido de sólidos solubles con una buena precisión, mediante modelos construidos a partir de parámetros de la imagen digital como: histogramas y diámetros de Feret.
2. La aplicación del modelo de predicción del momento de cosecha, que daba buenos resultados con la primera campaña, no ha resultado adecuado para determinar el estado de madurez de las manzanas en la segunda campaña. Será necesario hacer medidas en más campañas para mejorar los resultados.
3. Los histogramas de las imágenes de retrodifusión, producidos por el láser 785 nm, han permitido una clasificación de las manzanas, en ambas campañas, según el grado de madurez, mediante la aplicación de un Análisis de Componentes Principales.

MELOCOTÓN cv ‘Calrico’

4. Los modelos de predicción Vis/NIR (450-1060 nm) desarrollados para la determinación no destructiva de la firmeza en melocotón, presentan una buena capacidad predictiva.
5. La utilización conjunta del método Vis/NIR y el de impulso acústico (AWETA) mejora la capacidad predictiva de la firmeza frente a la de cada uno de los métodos por separado.
6. La aplicación de técnicas de selección de variables en el espectro Vis/NIR ha simplificado el modelo de predicción obtenido (de 661 a 3 variables) con resultados similares.
7. En el caso del melocotón, la aplicación de los métodos de regresión no lineales (LS-SVM) no presenta ventajas sobre los lineales (PLS), lo que permite utilizar estos cálculos, más simples, en los modelos.

CEREZA cv ‘Chelan’

8. Una buena determinación del parámetro de calidad “contenido en sólidos solubles totales” en cerezas cv ‘Chelan’ ha podido ser realizada con modelos de predicción que utilizan sólo el intervalo de longitudes de onda 700-1060 nm, en el infrarrojo cercano, banda espectral relacionada con la absorción de azúcares y agua.
9. Los modelos de predicción obtenidos con el intervalo de longitudes de onda 450-700 nm, relacionado con la absorción de clorofila, ha permitido una buena determinación del parámetro de calidad “firmeza”.
10. En el caso de la cereza, el uso de métodos de regresión lineales (PLS) y no lineales (LS-SVM) no ha presentado diferencias respecto a la capacidad predictiva de los diferentes modelos utilizados, lo que, de nuevo, permite utilizar modelos más simples.

11. Una optimización de los modelos de predicción de ambos parámetros, se ha conseguido mediante técnicas de selección de variables, reduciendo el número de las mismas de unas 300 a, aproximadamente, 40 variables en cada parámetro.

9.- Bibliografía

- Abbott, J.A. Massie, D.R. Upchurch, B.L., Hruschka, W.R. 1995. Nondestructive sonic firmness measurement of apples. *Transactions of the ASAE* 38:1461-1466.
- Alonso, J.M. y Espada, J.L. 2011. Melocotón de Calanda, cómo conseguir una oferta de elevada calidad mediante una producción sostenible. *Vida Rural*, 15 Febrero.
- Andersen, C.M., Bro, R., 2010. Variable selection in regression: a tutorial. *Journal of Chemometrics* 24, 728–737.
- Andersen, T. 2013. A sampling approach for predicting the eating quality of apples using visible–near infrared spectroscopy. *Journal of the Science of Food and Agriculture*, 93, 3710–3719.
- Badenes, M.L., Llácer, G., Crisosto, C.H, 2006. Mejora de la calidad de frutales de hueso. En: G. Llácer, M.J Díez, J.M. Carrillo y M.L. Badenes (eds), *Mejora genética de la calidad en plantas. Sociedad Española de Ciencias Hortícolas y Sociedad Española de Genética*, Valencia, 551-578.
- Baranyai, L and Zude, M. 2009. Analysis of laser light propagation in kiwifruit using backscattering and Monte Carlo simulation. *Computer and Electronics in Agriculture*, 69, 33–39.
- Barnes, R.J., Dhanoa, M.S. y Lister, S.J. 1989. Standard normal variate transformation_and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, 43, 772-777.
- Barreiro, P y Ruiz-Altisent, M. 1996. Propiedades mecánicas y calidad de frutos , definiciones y medidas instrumentales. *Fruticultura Profesional* nº 77 Marzo/Abril.
- Ben-Gera, I. and Norris, K. H. 1968. Direct spectrophotometric determination of fat and moisture in meat products. *Journal of Food Science*. 33, 64-67.
- Bertrand, D. 1993. Data pre-treatment and original analysis in spectroscopy. En *advanced comet chemometrics schools*. Libramont Belgium April 26-28
- Bertrand, D and Dufour, E. 2000. *La spectroscopie infrarouge et ses applications analytiques*. Editions TEC and DOC, Paris, France.
- Birth, G. S. 1976. How light interacts with foods. In J.J. Gaffney. (Ed.), *Quality detection in foods*. p6-11. St. Joseph: ASAE

- Blanco, A., Fernández, V., Val, J. 2010. Improving the performance of calcium-containing spray formulations to limit the incidence of bitter pit in apple (*Malus x domestica*). *Scientia Horticulturae*, 127, 23-28.
- Burns, D. A. and Ciurczak, E. W. 1992. Handbook of near infrared analysis. Burns D. A. and Ciurczak, E. W. (eds.). Marcel Dekker, New York, NY, USA.
- Byer, R. E. 2003. Flower and fruit thinning and vegetative: fruit balance. In: Ferree, D. C. and Warrington, I.J. (eds), Apples: botany, production and uses CAB international, London, UK, 409-436.
- Camps, C., Christen, D. 2009. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. *LWT- Food Science and Technology* 42, 1125-1131.
- Carlini, P., Massantini, R., Mencarelli, F. 2000. Vis-NIR measurement of soluble solids in cherry and apricot by PLS regression and wavelength selection. *Journal Agricultural Food Chemistry*, 48, pp. 5236-5242.
- Cayuela, J. A. Weiland, C. 2010. Intact orange quality prediction with two portable NIR spectrometers. *Postharvest Biology and Technology*. 58, 113-120
- Cen, H., Bao, Y., He, Y., Sun, D.W. 2007. Visible and near infrared spectroscopy for rapid detection of citric and tartaric acids in orange juice. *Journal of Food Engineering*,
- Chalmers, D. J and Van den Ende B. 1975. A reappraisal of the growth strategy and development of peach fruit. *Australia Journal Plant Physiology*, 2, 623-634.
- Chauchard, F., Cogdill, R., Roussel, S., Roger, J.M. and Bellon-Maurel, V. 2004. Application of LS-LVM to non linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemometrics and Intelligent Laboratory Systems*, 71, 141-150. CIE (Commission Internationale de l'Eclairage). 2004. Colorimetry, 3rd edn. Vienna, Austria: CIE Central Bureau, Publ. 15:2004
- Chen, P. 1996. Quality evaluation technology for agricultural products. In: Proc. Int. Conf. on Agric. Machinery Engineering. Seoul. Korea.
- Cho, R. K., Lee, K. H., Iwamoto, M. 1992. Determination of firmness of intact apple fruits by near infrared reflectance spectroscopy. In: Taylor, R. A. (Ed.), Proc 6th Int. Diffuse Reflect. Spectr. Conf. Counc. Dor Near IR Spectr., Chambersburg, PA., pp. 71-76.

- CIE (Commission Internationale de l'Eclairage). 2004. Colorimetry, 3rd edn. Vienna, Austria: CIE Central Bureau, Publ. 15.
- Clark, C.J., McGlone, V.A., Requejo, C., White, A. and Woolf, A.B. 2003. Dry matter determination in Hass avocado by NIR spectroscopy. *Postharvest Biology and Technology*. 29, 300-307.
- Colaric, M., Veberic, R., Stampar, F., Hudina, M. 2005. Evaluation of peach and nectarine fruit quality and correlations between sensory and chemical attributes. *Journal of the Science of Food and Agriculture*. 85, (15), 2611-2616.
- Corelli-Grappadelli, L., Coston, D. C. 1991. Thinning pattern and light environment in peach tree canopies influences fruit quality. *HortScience*. 26, 1464-1466.
- Costa, M., Noferini, M., Fiori, G., Miserocchi, O., Bregoli, A.M. 2002. NIRS evaluation of peach and nectarine fruit quality in pre- and post- harvest conditions. *Proc. 5th IS on Peach*. Eds. R.S. Johnson & C.H. Chrisosto. *Acta Horticulturae*. 592, ISHS.
- Costa-Pereira, A. F., Coelho-Pontes, M. J., Gambarra-Neto, F. F., Bezerra-Santos, S. R., Harrop-Galvao, R. K., Ugulino-Araújo, M. C. 2008. NIR spectrometric determination of quality parameters in vegetable oils using iPLS and variable selection. *Food Research International*. 41, 341-348.
- Cover, T.M. and Thomas, J.A. 1991. *Elementary Information Theory*. Wiley, New York.
- Crisosto, C.H. 1994. Stone fruit maturity indices: a descriptive review. *Postharvest News Inform*, 5 (6), 65-68.
- Crisosto, C.H. Mitchell, F.G., Johnson, R.S. 1995. Factors in fresh market stone fruit quality. *Postharvest News and Information*. 6, (2), 17-21.
- Crisosto, C. H. and Mitchell, F. G. 2007. Factores precosecha que afectan la calidad de frutas y hortalizas. In *Tecnología Poscosecha de Productos Hortofrutícolas*. 3ª Edición.
- De Keteleare, B., Howarth, S., Creeze, L., Lammertyn, J., Viaene, K., Bulens, I., De Baerdemaeker, J. 2006. Postharvest firmness changes as measured by acoustic and low-mass impact devices: a comparison of techniques. *Postharvest and Biology Technology*. 41, 275-284.
- Downey, G., Robert, P., Bertrand, D. 1992. *Analytical Proceedings*. 29, 8.

- Ebel, R., Proebsting, E., Patterson, M. 1993. Regulated deficit irrigation may alter apple maturity, quality and storage life. *HortScience*. 28, 141-143.
- Estévez, P.A., Tesmer, A., Perez, C.A., Zurada, J.M. 2009. Normalized mutual information feature selection. *IEEE Trans. Neural Network*. 20, 189–201.
- Fallahi, E., Conway, W. S., Hickey, K. D., Sams, C. E. 1997. The role of calcium and nitrogen in postharvest quality and disease resistance of apples. *Hortsciencie*. 32, 831–835.
- Fan, G., Zha, J., Du, R. and Gao, L. 2009. Determination of soluble solids and firmness of apples by Vis/NIR transmittance. *Journal of Food Engineering*. 93, 416-420
- Fang, L., Li, H., Liu, Z., Xian, X. 2013. Online evaluation of yellow peach quality by visible and near infrared spectroscopy. *Journal of Food Science and Technology*. 5, volume 5, 606-612.
- Flores, K., Sánchez, M.T., Pérez-Marín, D., Guerrero, J.E., Garrido-Varo, A. 2009. Feasibility in NIRS instruments for predicting internal quality in intact tomato. *Journal of Food Engineering*. 91, 311-318.
- Flores, Katherine, 2009. Determinación no destructiva de parámetros de calidad de frutas y hortalizas mediante espectroscopia de reflectancia en el infrarrojo cercano. Tesis Doctoral. Universidad de Córdoba.
- Francis, F.J. 1995. Quality as influenced by colour. *Food Quality Prefer*. 6, 149-155
- Fu, X., Ying, X., Lu, H., Xu, H., Yu, H. 2008. FT-NIR diffuse reflectance spectroscopy for kiwifruit firmness detection. *Sens and Instrument. Food Quality*. 1, 29-35.
- Gaitán Jurado, 2009. Aportaciones de la Tecnología NIRS a la calidad de productos_agroalimentarios. Tesis Doctoral. Universidad de Córdoba.
- Galili, N., Shmulevich, I., Benichou, N. 1998. Acoustic testing of avocado for fruit ripeness evaluation. *Trans. ASAE* 41:399-407.
- Gao, L., Mazza, G., 1995. Characterization, quantization and distribution of anthocyanins and colorless phenolics in sweet cherries. *Journal of Agricultural and Food Chemistry*. 43, 343–346.
- Geladi, P and Kowalski, B. 1986. Partial least squares regression: a tutorial. *Analytica Chimica Acta*. 185, 1-17.

- Geladi, P., McDougall, D. y Martens, H. 1985. Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat. *Applied Spectroscopy*. 39, 491.
- Girona, J., Gelly, M., Mata, M., Arbone's A., Rufat, J., Marsal, J. 2005. Peach tree response to single and combined deficit irrigation regimes in deep soils. *Agricultural Water Management*. 72, 97-108.
- Golding, J. B. and Spohr, L. 2005. Nondestructive assessment of peach and nectarine firmness. *Acta Horticulturae*. (ISHS) 687, 261-270.
- González, R.C. and Woods, R.E. 2002. *Digital image processing*. Prentice Hall, USA.
- Guyon, I. and Elisseeff, A. 2003. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*. 3, 1157-1182.
- Hashim, N., Pflanz, M., Regen, C., Janius, Rahman, R., Osman, A., Shitan, M., Zude, M. 2013. An approach for monitoring the chilling injury appearance in bananas by means of backscattering imaging. *Journal of Food Engineering*. 116, issue 1, 28-36.
- Hernández-Gómez, A., He, Y., García-Pereira, A. 2006. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis-NIR spectroscopy techniques. *Journal of Food Engineering*. 77, 313-319.
- Herrera, L. J., Pomares, H., Rojas, I., Verleysen, M., Guillén, A. 2006. Effective input variable selection for function approximation. Springer Verlag Part 1 LNCS 4131. Pp. 41-60
- Herrera, L.J., Lafuente, V., Ghinea, R., Pérez, M.M., Negueruela, I., Pomares, H., Rojas, I., Guillén, A. 2015. Mutual information-based feature selection in spectrometric data for agricultura applications. *Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol I, IMECS 2015, March 18-20, 2015, Honk Kong*.
- Herrero, A., Diezma, B., Lleó, L., Valero, C., Ruiz-Altisent, M. 2009. Evaluación de técnicas acústicas para la determinación de firmeza en melocotón. *REF cong 2009-92. Agroingeniería*.
- Huirong, X., Bing, Q., Tong, S., Xiaping, F., Yibin, Y. 2012. Variable selection in visible and near infrared spectra: Application to on-line determination of sugar content in pears. *Journal of Food Engineering*. 109, 142-147.

- Jackson, J.E. 1991. *An User's Guide to Principal Components*, John Wiley and Sons. New York, USA. Kader, A. (Ed.). University of California, Division of Agriculture and Natural Resources, Oakland, California, USA. pp. 55-62
- Jha, S. N., Narsaiah, K., Jaiswal, R., Bhardwaj, M., Gupta, R., Kumar, R., Sharma, R. 2014. Nondestructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*. 124, 152-157
- Kader, A. A. 1985. Postharvest Biology and Technology: an overview. In: *Postharvest Technology of Horticultural Crops*. Kader, A. A; R. F. Kasmire; a F.G. Mitchel; M.S. Reid; N.F. Sommer y J.M. Thompson, 2, pp. 3-7
- Kader, A.A. 2007. *Biología y tecnología poscosecha: Un panorama*. En *Tecnología_Poscosecha de Productos Hortofrutícolas*. 3ª edición. Kader, A. (Ed.). University of California, Division of Agriculture and Natural Resources, Oakland, California, USA. pp. 43-54.
- Kader, A.A. 2008. Flavor quality of fruits and vegetables. *Journal of the Science of Food and Agriculture*. 88, 1863-1868.
- Kaye, W. 1954. Near-infrared spectroscopy: I. Spectral identification and analytical applications. *Spectrochimica Acta*. 6, 257.
- Kim, J. Mowat, A., Poole, P., Kasabov, N. 2000. Linear and non-linear pattern recognition models for classification of fruit from visible-near infrared spectra. *Chemometrics and Intelligent Laboratory Systems*. 51, 201–216.
- Knee, M. 1972. Anthocyanin, carotenoid, and chlorophyll changes in the peel of Cox's Orange Pippin apples during ripening on and off the tree. *Journal of Experimental Botany*. 23, 184–196.
- Koller, D. and Sahami, M. 1996. Toward Optimal Feature Selection. *Proc. Int. Conf. on Machine Learning*, 284-292.
- Kraskov, A., Stgbauer, H., Grassberger P. 2004. Estimating mutual information. *Phys Rev E*. 69: 066138.
- Lafuente, V., Herrera, L.J., Pérez, M.M., Val, J., Negueruela, I. 2014. Firmness prediction in *Prunus Persica* 'Calrico' peaches by visible/short wave near infrared spectroscopy and acoustic measurements using optimised linear and non-linear chemometric models. *Journal of the Science of Food and Agriculture*. 95, Issue 10, 2033-2040.

- Lammertyn, J., Peirs, A., De Baerdemaeker, J., Nicolai, B. 2000. Light penetration properties of NIR radiation in fruit respect to non-destructive quality assessment. *Postharvest Biology and Technology*. 1, 121-132.
- Li, J., Huang, W., Zhao, C and Zhang, B. 2013. A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy. *Journal of Food Engineering*. 116, Issue 2, 324-332.
- Lilien-Kipnis, H., Lavel j. 1971. Anatomical changes during the development of "Ventura" peach fruits. *Journal Horticultural Science*. 46, 103-110.
- Lill, R.E., O'Donoghue, E.M., King, G.A. 1989. Postharvest physiology of peaches and nectarines. *Hortic Rev*, 11, 413-452.
- Linberg, W., Persson, J. A., Wold, S. 1983. *Analytica Chemistry*. 55, 643.
- Liu, F., He, Y., Wang, L. 2008. Determination of effective wavelengths for discrimination of fruit vinegars using near infrared spectroscopy and multivariate analysis. *Analytica Chimica Acta*. 615, 10-17.
- Liu, F., Jiang, Y., He, Y. 2009. Variable selection in visible/near infrared spectra for linear and nonlinear calibrations: A case study to determine soluble solids content of beer. *Analytica Chimica Acta*. 635, 45-52
- Liu, Y., Sun, X., Zhou, J., Zhang, H., Yang, C. 2010. Linear and nonlinear multivariate regressions for determination sugar content of intact Gannan navel orange by Vis-NIR diffuse reflectance spectroscopy. *Mathematical and Computer Modelling*. 51, 1438-1443.
- Liu, W., Wang, Z., Qing, J., Li, H., Xiao, W. 2014 Classification and quantification analysis of peach kernel from different origins with near-infrared diffuse reflection spectroscopy. *Pharmacognosy Magazine*, 10, Issue 1, 441-114.
- Liu, Y and Ying, Y. 2005. Use of FT-NIR spectrometry in non-invasive measurements of internal quality of 'Fuji' apples. *Postharvest Biology and Technology*. 37, 65-71.
- Lleó, L. Barreiro, P., Ruiz-Altisent, M., Herrero, A. 2009. Multispectral images of peach related to firmness and maturity at harvest. *Journal of Food Engineering*. 93, 229-235.
- López, M.L., Lavilla, M.T., Recasens, I., Graell, J., Vendrell, M. 2000. 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, 311-324.

- Lorente, D., Zude, M., Regen, C., Palou, L., Gómez-Sánchez, J., Blanco, J. 2013. Early decay detection in citrus fruit using laser-light backscattering imaging. *Postharvest Biology and Technology*. 86, 424-430.
- Lu, R. 2001. Predicting firmness and sugar content of sweet cherries using near-infrared diffuse reflectance spectroscopy. *Transactions of the ASAE. American society of agricultural engineers* ISSN 0001-2351. Vol.44, pp. 1265-1271.
- Lu, R. 2004a. 2004. Prediction of apple fruit firmness by near infrared multispectral scattering. *Journal of Texture Studies*. 35, 263-276.
- Lu, R. 2004b. Multispectral imaging for predicting firmness and soluble solids content of apple fruit. *Postharvest Biology and Technology*. 31, 147-157.
- Lu, R. and Peng, Y. 2006. Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering*. 93(2), 161–171.
- Maniwara, P., Nakano, K., Boonyakiat, D., Ohashi, D., Hiroi, M., Tohyama, T. 2014. The use of visible and near infrared spectroscopy for evaluating passion fruit postharvest quality. *Journal of Food Engineering*. 143, 33-43
- Martens, H and Naes, T. 1989. *Multivariate Calibration*. John Wiley & Sons Inc., Chichester, UK
- Martens, H. and Martens, M. 2001. *Multivariate analysis of quality: an introduction*. John Wiley and Sons Inc., Chichester, UK.
- Martínez-Valdivieso, D., Font, R., Blanco-Díaz, M.T., Moreno-Rojas, J. M., Gómez, P. Alonso-Moraga, A., Del Río Celestino, M. 2014. Application of near-infrared reflectance spectroscopy for predicting carotenoid content in summer squash fruit. *Computers and Electronics in Agriculture*. 108, 71-79.
- Martínez-Vega, M., Sharifzadeh, S., Wulfsohn, D., Skov, T., Clemmensen, L., Toldam-Andersen, T. 2013. A sampling approach for predicting the eating quality of apples using visible–near infrared spectroscopy. *Journal of the Science of Food and Agriculture*. 93, 3710–3719.
- Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., De Jong, S., Lewi, P.J., Smeyers-Verdeke, J. 1997. *Handbook of Chemometrics and Qualimetrics (vol.2)*. Elsevier Science Pub., Amsterdam, The Netherlands.
- McGlone, V. A., Abe, H., Kawano, S. 1997. Kiwifruit firmness by near infrared light scattering. *Journal of Near Infrared Spectroscopy*. 5(1), 83-89.

- McGlone, V. A., Jordan, R. B., Martinsen, P. J. 2002. Vis/NIR estimation at harvest of pre- and post- storage quality indices for Royal Gala apple. *Postharvest Biology and Technology*. 25, 135-144.
- McGlone, V.A. and Kawano, S. 1998. Firmness, dry-matter and soluble solids assessment of postharvest kiwifruit. *Postarvest Biology and Technology*. 13, 131-141
- Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. 1999. Non destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*. 106(1), 135–141.
- Mireei, S. A. 2010. Nondestructive determination of effective parameters on maturity of mozafati and shahani date fruits by NIR spectroscopy technique. PhD thesis, Department of Mechanical Engineering of Agricultural Machinery, University of Tehran, Iran. In Persian.
- Mitcham, E.J., Clayton, M. and Biasi, W. V. 1998. Comparison of devices for measuring cherry fruit firmness. *Hort. Sci.*, 33: 723-727.
- Moccia, S., Chiesa, A., Oberti, A., y Tittonell, P.A. (2006). Yield and quality of sequentially grown cherry tomo and lettuce under long-term conventional, lowinput and organic soil management systems. *European Journal of Horticultural Science*. 71, 183-191.
- Mohsenin, N. N. 1986. *Physical properties of plant and animal materials: structure, physical, characteristics and mechanical properties*. New York: Gordon and Breach Science Publishers. 664 p.
- Moons, E. and Dardenne, P. 1999. Determination of internal apple quality by non-destructive visible and near infrared spectroscopy. In *Near Infrared Spectroscopy: Proceedings of the 9th International Conference*. Davies, A.M.C., Giangiacomo, R. (Eds.). NIR publications, Norwich UK. pp. 785-789
- Moons, E., Dardenne P., Dubois, A., Sindic, M. 1997. Non-destructive visible and NIR spectroscopy measurement for the determination of apple internal quality. *Acta Horticulturae*. 517, 441-448.
- Murga, J. y Palazón, I. 1984. *Manual de inspección de peras y manzanas*. S.O.I.V.R.E.- I.N.I.A. Madrid.
- Myers, S. C. 1993. Preharvest watersprout removal influences canopy light relations, fruit quality, and flower bud formation of 'Redskin' peach

- trees. *Journal of the American Society for Horticultural Science*. 118, 442–445.
- Naes, T., Irgens, C., Martens, H. 1986. *Applied Statistics*. 35, 195.
- Naes, T. 1992. Progress in multivariate calibration. In *Near Infrared Spectroscopy. Bridging the Gap between Data Analysis and NIR Applications*. Hildrum, K.I., Isaksson, T., Naes, T., Tandberg, A. (Eds.). Ellis Horwood Limited, Chichester, West Sussex, UK, pp. 51-60.
- Naes, T., Isaksson, T., Fearn, T., Davies, T. 2002. *A user-friendly guide to multivariate calibration and classification*. NIR Publications, Chichester, UK.
- Neilsen, G., E. Hogue, M. Meheriuk, M. 1999. Nitrogen fertilization and orchard-floor vegetation management affect growth, nutrition and fruit quality of gala apple. *Canadian Journal of Plant Science*. 79, 379- 385.
- Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K., Lammertyn, J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*. 46, 99-118.
- Norris, K. H. 1964. *Transaction of the American Society of Agricultural Engineers*. 7, 240
- Norris, K.H. and Williams, P.C. 1984. Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. Influence of particle size. *Cereal Chemistry*. 61, 158.
- Osborne, B. G., Feam, T., Hindle, P. H. 1993. Practical NIR spectroscopy with applications. In: Osborne, B. G., Feam, T., Hindle, P. H. (Eds.) *Food and Beverage Analysis*. Longman, Essex, pp. 11-35.
- Ozaki, Y., Morita, S. y Du, Y. 2007. *Near-Infrared Spectroscopy in Food Science and Technology*. Ozaki, Y., McClure, W. y Christy, A. (Eds.). John Wiley & Sons Inc., Hoboken, New Jersey, USA, p.47
- Paz, P., Sánchez, M.T., Pérez-Marín, D., Guerrero, J.E., Garrido-Varo, A. 2009. Evaluating NIR instruments for quantitative and qualitative assesment of intact Apple quality. *Journal Science Food and Agriculture*. 89, 781-790
- Peirs, A., Tirry, J., Verlinden, B., Darius, P., Nicolaï, B. 2003. Effect of biological variability on the robustness of NIR models for soluble solids content of apples. *Postharvest Biology and Technology*. 28, 269-280.

- Peng, Y. and Lu, R. 2005. Modeling multispectral scattering profiles for prediction of apple fruit firmness. *Transactions of ASAE*. 48 (1), 235–242.
- Peng, Y and Lu, R. 2006. Improving apple fruit firmness predictions by effective correction of multispectral scattering images. *Postharvest Biology and Technology*. 41, 266–274.
- Peng, Y. and Lu, R. 2007. Prediction of apple fruit firmness and soluble solids content using characteristics of multispectral scattering images. *Journal of Food Engineering*. 82, 142-152
- Pent, H., Long, F., Ding, C. 2005. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *Trans. Pattern Analysis and Machine Intelligent*. 27,1226-1238.
- Pérez-Marín, D., Garrido-Varo, A., Guerrero, J.E. 2007. Non-linear regression methods in NIRS quantitative analysis. *Talanta*. 72, 28-42.
- Perring, M. A. and Pearson, K. 1987. Suppression of bitter pit by calcium injected into the core cavities of harvested apples. *Journal of Horticultural Science*. 62, 303-304.
- Peterson, C. 2001. How it works: the charged- coupled device or CCD. *Journal of young investigators*.
- Pissard, A., Fernández-Pierna, J.A., Baeten, V., Sinnaeve, G., Lognay, G., Mouteau, A., Dupont, P., Rondia, A., Lateur, M. 2013. Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near infrared spectroscopy. *Journal of the Science of Food and Agriculture*. 93, 238-244.
- Poovaiah, B. W., Glenn, G.M., Reddy, A.S.N. 1988. Calcium and fruit softening: Physiology and biochemistry. *Horticultural Reviews*. 10, 107-152.
- Pou, N. 2002. Análisis de control de preparados farmacéuticos mediante espectroscopia en el infrarrojo próximo. Universidad autónoma de Barcelona. Tesis doctoral publicada en <http://www.tesisenred.net/TDX-0214103-224454>.
- Prior, R.L. and Cao, G.H. 2000. Antioxidant phytochemicals in fruits and vegetables: diet and health implications. *HortScience*. 35, (4), 588-592.
- Puigdoménech, A. 1998. Optimització del control de qualitat en productes agrícoles mitjançant espectrometria d'infraroig proper (NIR): aplicació als farratges de Catalunya. Transferència d'equacions de calibratge

- multivariant en xarxes instrumentals. Tesis Doctoral. Universitat de Barcelona
- Qin, J. 2007. Measurement of the optical properties of horticultural and food products by hyperspectral imaging. PhD thesis, Michigan State University, East Lansing, MI, USA.
- Qin, J., and Lu, R. 2007a. Measurement of the absorption and scattering properties of turbid liquid foods using hyperspectral imaging. *Applied Spectroscopy*, 61(4), 388–396.
- Qin, J., and Lu, R. 2007b. Monte Carlo simulation of light propagation in Apples. ASABE Annual International Meeting, paper no 073058. St. Joseph: ASABE.
- Qing, Z., Ji, B., Zude, M. 2007a. Predicting soluble solid content and firmness in apple fruit by means of laser light backscattering image analysis. *Journal of Food Engineerin.* 82, 58–67.
- Qing, Z., Ji, B., Zude, M. 2007b. Wavelength selection for predicting physicochemical properties of apple fruit based on near-infrared spectroscopy. *Journal of Food Quality.* 30, 511-526.
- Qing, Z., Ji, B., Zude, M. 2008. Non-destructive analysis of apple quality parameters by means of laser-induced light backscattering imaging. *Postharvest Biology and Technology.* 48, 215–222.
- Raese, J., Drake., S, Curry, E. 2007. Nitrogen fertilizer influences fruit quality, soil nutrients and cover crops, leaf color and nitrogen content, biennial bearing and cold hardiness of 'Golden Delicious'. *Journal of Plant Nutrition.* 30, 1585- 1604.
- Recasens, I., Roig, J., Graell, J. 1989. The effect of harvest date on 'Flor de invierno' pears in cold storage. *Acta Horticulturae.* 256, 213-222.
- Reid, M. S. 2007. Maduración e índices de madurez. In *Tecnología poscosecha de productos hortofrutícolas.* 3ª edición. Kader, A. (Ed.). University of California, Division of Agriculture and Natural Resources, Oakland, California, USA. pp. 63-72.
- Roger, J. M and Bellon-Maurel, V. 2000. Using genetic algorithms to select wavelengths in near-infrared spectra: application to sugar content prediction in cherries. *Applied Spectroscopy.* Volume 54, 9.
- Romano; G., Baranyai, L., Gottschalk, K., Zude, M. 2008. An Approach for Monitoring the Moisture Content Changes of Drying Banana Slices with

- Laser Light Backscattering Imaging. Food and Bioprocess Technology. 1, Issue 4, 410-414.
- Romero, M.A. 2001. Desarrollo de nuevas tecnologías analíticas en el control de calidad de la industria farmacéutica. Universidad autónoma de Barcelona. Tesis doctoral publicada en <http://www.tesisenred.net/TDX-0712102-102928>
- Romero, F., Martínez-Madrid, M.C., Pretel, M.T. 2006. Factores precosecha determinantes de la calidad y conservación en poscosecha de productos agrarios. V Simposio Ibérico VIII Nacional de Maduración y Post-Recolección, Orihuela Alicante, pp. 91-96.
- Rossi, F., Lendasse, A., François, D., Wertz, V., Verleysen, M. 2006. Mutual Information for the selection of relevant variables in spectrometric nonlinear modelling. Chemometrics Intelligent Laboratory Systems. 80, 215-226
- Ruiz-Altisent, M., Lleó, L., Riquelme, F. 2006. Instrumental quality assessment of peaches: Fusion of optical and mechanical parameters. Journal of Food Engineering. 74, 490-499.
- Russ, J.C. 2006. The Image Processing Handbook, Fifth Edition. CRC Press, Boca Raton, USA.
- Sams, C.E. 1999. Preharvest factors affecting postharvest texture. Postharvest Biology and Technology. 15, 249-254.
- Sánchez, M.T., De la Haba, M.J., Benítez-López, M., Fernández-Novales, J., Garrido-Varo, A., Pérez-Marín, D. 2012. Non-destructive characterization and quality control of intact strawberries based on NIR spectral data. Journal of Food Engineering. 110, 102-108
- Savitzky, A. and Golay, M.J.E. 1964. Analytical Chemistry. 36, 1627.
- Sedgley, M. and Griffin, A. R. 1989. Sexual reproduction of tree crops. Academic Press, London. 378 pp
- Shao, Y., Bao, Y., He, Y. 2011. Visible/Near-Infrared Spectra for Linear and Nonlinear Calibrations: A case to predict Soluble Solid Contents and pH Values in Peach. Food Bioprocess Technology. 4, 1376-1383
- Shao, Y., He, Y., Bao, Y., Mao, J. 2008. Near- Infrared spectroscopy for classification of oranges and prediction of the sugar content. International Journal of Food Properties. 12, 644-658.

- Shenk, J. S., Workman, J., Westerhaus, M. 2001. Application of NIR spectroscopy to agricultural products. In Handbook of Near Infrared Analysis. 2nd Edition. Burns, D.A. Ciurczac, E. W. (Eds.). Marcel Dekker, Nueva York. (USA). pp. 419-474.
- Shenk, J.S., Westerhaus, M.O. 1995. Routine Operation, Calibration, Development_and Network System Management Manual. NIRSystems, Inc., 12101 Tech Road, Silver Spring. MD 20904. USA.
- Shenk, J.S., Westerhaus, M.O. 1996. Calibration the ISI way. In Near Infrared_Spectroscopy: The Future Waves. Davies, A.M.C., Williams, P.C. (Eds.). NIR Publications, Chichester, UK. pp.198-202.
- Shmulevich, I., Galili, N., Howarth, M.S., 2003. Non-destructive dynamic testing of apples for firmness evaluation. Postharvest Biology and Technology. 29, 287–299.
- Soares-Gomes, A., Trugo, L.C., Botrel, N., Da Silva-Souza, L.F. 2005. Reduction of internal browning of pineapple fruit (*Amanas comusus* L.) by preharvest soil application of potassium. Postharvest Biology and Technology. 35, 201-207.
- Studman, C. J. 2001. Computers and electronics in postharvest technology: a review. Computers and Electronics in Agriculture. 30, 109-124.
- Sun T., Lin H., Xu H., Ying Y. 2009. Effect of fruit moving speed on predicting soluble solids content 'Cuiguan' pears (*Pomaceae pyrifolia* Nakai cv. Cuiguan) using PLS and LS-SVM regression. Postharvest Biology and Technology. 51, 86-90
- Sun, T., Lin, H., Xu, H., Ying, Y. 2009. Effect of fruit moving speed on predicting soluble solids content of 'Cuiguan' pears (*Pomaceae pyrifolia* Nakai cv. Cuiguan) using PLS and LS-SVM regression. Postharvest Biology and Technology. 51, 86-90.
- Suykens J.A.K, and Vandewalle J. 1999. Neural Process. Lett. 9, 293.
- Suykens J.A.K, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. 2002. Least Squares Support Vector Machines, World Scientific, Singapore.
- Taghadomi-Saberi, S., Omid, M., Emam-Djommeh, Z., Ahmadi, H. 2013. Evaluating the potential of artificial neural network and neuro-fuzzy techniques for estimating antioxidant activity and anthocyanin content of sweet cherry during ripening by using image processing. Journal of the Science of Food and Agriculture. 94, 95-101

- Thompson, J. F. 2007. Sistemas de cosecha. In *Tecnología Poscosecha de Productos Hortofrutícolas*. 3ª edición. Kader, A. (Eds). University of California, Division of Agriculture and Natural Resources, Oakland, California, USA. pp.73-76.
- Timm, E. J., G. K. Brown, P. R. Armstrong, and R. M. Beaudry. 1993. A portable instrument for measuring firmness of cherries and berries. ASAE Paper No. 93-6539. St. Joseph, Mich.: ASAE.
- Tiwari, G., Slaughter, D. C., Cantwell, M. 2013. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biology and Technology*. 86, 221-229.
- Travers, S., Bertelsen, B., Kucheryavskiy, S. 2013. Predicting apple (cv. Elshof) postharvest dry matter and soluble solids content with near infrared spectroscopy. *Journal of the Science Food and Agriculture*. 94 (5), 955-962.
- Tu, K., De Busscher, R., De Baerdemaeker, J., Schrevers, E. 1995. Using laser beam as light source to study tomato and apple quality non-destructively. *Food Processing Automation IV Conference*, Chicago, IL, 3-5.November. Pp. 528-536.
- Val, J and Fernández, V. 2011. In-season calcium-spray formulations improve calcium balance and fruit quality traits of peach. *Journal Plant Nutrition Soil Science*. 174, 465-472.
- Val, J., Monge, E., Risco, D., Blanco, A. 2008. Effect of preharvest calcium sprays on calcium concentrations in the skin and flesh of apples. *Journal of Plant Nutrition*, 31, issue 1, 1889-1905.
- Valente, M., Leardi, R., Self, G., Luciano, G., Pain, J.P. 2009. Multivariate calibration of mango firmness using vis/NIR spectroscopy and acoustic impulse method. *Journal of Food Engineering*. 94, 7-13.
- Vanderwalle, J. 1999. Least squares support vector machine classifiers. *Neural process. Lett.* 9, 293-300.
- Volz, R., Harker, F., Lang, S. 2003. Firmness decline in 'Gala' apple during fruit development. *Journal of the American Society for Horticultural Science*. 128, 797-802.
- Webster, A. D. and Looney, N.E. 1996. *The cherries: crop physiology. Production and uses*. CAB International. Wallingford, Oxon, Reino Unido.

- Westwood, M. N. 1982. Fruticultura de zonas tempranas. Madrid, España. Ediciones Mundi-Prensa. pp.251-267.
- Williams, P.C. 2001. Implementation of near-infrared technology. In: Williams, P.C.Norris, K.H. (Eds.), Near-infrared Technology in the Agricultural and Food Industries. AACC, Inc., St. Paul, MN,145–169.
- Williams, P. C. and Sobering, D. 1992. Objective prediction of functionality by near infrared technology. En: Making Light Work: Advances in Near Infrared Spectroscopy. Murray, I., Cowe, I. (Eds), VCH, Weinheim, Germany, pp. 217-222.
- Wills, R.H.H., Lee, T.H., McGlasson, W.B., Hall, E.G., Graham, D. 1984. Fisiología y manipulación de frutas y hortalizas postrecolección. Ed. Acribia. Zaragoza, España.
- Wold, H. 1966. Multivariate analysis, Academic Press, New York, p. 391.
- Wold, H. 1982. Soft Modeling. The basic design and more extensions. En: Systems under indirect observation. Jöreskog K. G. y Wold, H. Amsterdam.
- Wold, S. and Sjöström, M. 1998. Chemometrics, present and future success. Chemometrics and Intelligent Laboratory Systems, 44,3-14.
- Workman, J.J. 2004. Near-Infrared spectroscopy in Agriculture. Roberts, C. A., Workman, J.J. y Reeves, J. B. (Eds.). American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, Madison, Wisconsin, USA, p.11.
- Wu, D., Nie, P., He, Y., Bao, Y. 2012. Determination of calcium content in powdered milk using near and mid-infrared spectroscopy with variable selection and chemometrics. Food and ioprocess Technology. Volume 5, Issue 4, 1402-1410.
- Wu, Z., Wu, E., Wang, F., Long, J., Xu, X., Jiao, A., Jin, Z. 2014. Rapid determination of rocess
- Wu, Z., Wu, E., Wang, F., Long, J., Xu, X., Jiao, A., Jin, Z. 2014. Rapid determination of process variables of Chinese rice wine using FT-NIR spectroscopy and efficient wavelengths selection methods. Food Analitical Methods DOI 10.1007/s12161-014-0021-6.
- Xia, G., L. Cheng, A. Lakso., M. Goffinet. 2009. Effects of nitrogen supply on source-sink balance and fruit size of `gala´apple trees. Journal of the American Society for Horticultural Science. 134, 126-133.

- Xiaobo, Z., Jiewen, Z., Povey, M., Holmes, M., Hanpin, M. 2010. Variables selection methods in near-infrared spectroscopy. *Analytica Quimica Acta*. 667, 14-32.
- Xiaobo, Z., Jiewen, Z., Yanxiao, L. 2007. Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models. *Vibrational Spectroscopy* 44: 220-227.
- Xu, F., Huang, X., Dai, H., Chen, W., Ding, R., Teye, E. 2014. Nondestructive determination of bamboo shoots lignifications using FT-NIR with efficient variables selection algorithms. *Analytical Methods*. 6, 1090-1095.
- Xu, H., Qi, B., Sun, T., Fu, X., Ying, Y. 2012. Variable selection in visible and near-infrared spectra: Application to on-line determination of sugar content in pears. *Journal of Food Engineering*. 109, 142-147.
- Zhang, X., Li, W., Yin, B., Chen, W., Kelly, D. P., Wang, X., Zheng, K., Du, Y. 2013. Improvement of near infrared spectroscopy (NIRS) analysis of caffeine in roasted Arabica coffee by variable selection method of stability competitive adaptive reweighted sampling (SCARS). *Spectrochimica Acta parta A: Molecular and Biomolecular Spectroscopy*. 114, 350-356
- Zude, M., Herold, B., Roger, J.M., Bellon-Maurel, V., Landahl, S. 2006. Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. *Journal of Food Engineering*, 77, 254–260.
- Zude-Sasse, M., Herold, B., Geyer, M. 2000. Comparative study on maturity prediction in Elstar and Jonagold_apples. *Gartenbauwissenschaft*. 65 (6), 260–265.
- Zude-Sasse, M., Pflanz, M., Spinelli, Dosche, C., Torricelli, A. 2011. Non-destructive analysis of anthocyanins in cherries by means of Lambert-Beer and multivariate regression based on spectroscopy and scatter correction using time-resolved analysis. *Journal of Food Engineering*. 103, 68-75
- Zude-Sasse, M., Truppel, I., Herold, B. 2002. An approach to non-destructive apple fruit chlorophyll determination. *Postharvest Biology and Technology*. 25, 123-133

ANEXO I

‘Determination of quality parameters in Apple *Smoothie Golden Delicious* using backscattering laser imaging’

Victoria Lafuente, Jesús Val, Concha Urzola, A. Ignacio Negueruela

Determination of quality parameters in apple ‘Smoothie Golden Delicious’ using backscattering laser imaging

Victoria Lafuente^{1,2}, Jesús Val¹, Concha Urzola², Ignacio Negueruela²

¹Estación Experimental de Aula Dei (CSIC). Departamento de Nutrición Vegetal. Avda. Montañana, 1005 C.P. 50059, Zaragoza

²Departamento de Física Aplicada. Facultad de Veterinaria. Universidad de Zaragoza. Calle Miguel Servet 177, C.P. 50013, Zaragoza

mvlafuente@ipe.csic.es

ABSTRACT

The objective of this study was to evaluate ‘Smoothie Golden Delicious’ apples, using backscattering images as a non-destructive method to estimate some quality parameters during the ripening process of fruit.

In this experiment, 208 apples were collected from Aula Dei Experimental Station (CSIC) in Zaragoza (Northeastern, Spain). Backscattering images were taken from samples collected every 15 days for 7 months until harvest. Fruits soluble solids contents (SSC) and firmness were measured by traditional destructive methods. The light source was a solid state laser diode module emitting at three different spectral bands (670, 785 and 980 nm)

PLS calibration methods were used to create the statistical model to predict SSC and firmness. Good results were obtained when PLS calibration was applied to Feret diameter corresponding at different levels of gray. The R_{cv}^2 values were 0.87 for firmness and 0.93 for SSC, which are higher than other values previously reported in the literature for the same crop.

Keywords: Multispectral imaging, Backscattering, fruit

INTRODUCTION

Fruit quality has been evolving throughout time. At present, fruit production and commercialization extend to deliver criteria to which the consumer imposes (Vallejo, 1990). Setting the proper harvest date is not only useful in obtaining a better product, but also increases production. The main changes during ripening are color, flavour and texture. The quality indices most often used to determinate fruit quality are color, firmness and soluble solids content (SSC). Traditionally, firmness was determined by the destructive method Magness-Taylor and refractometry respectively. The main advantages of these techniques are that they present a good correlation with the organoleptic characteristics of the fruit (López et al, 2000), and their problem is they are destructive.

In recent years, non-destructive methods have been developed for estimating fruit parameters with the aim of replacing the destructive ones. The most practical and satisfactory techniques for the non destructive quality evaluation of agricultural products are the techniques obtained from optical spectrometry (Chen, 1996) such as Near Infrared Spectroscopy (NIR), Fast Fourier Transform NIR (FTNIR) and laser light backscattering image analysis. A lot of researchers have studied laser light backscattering which are obtained by a laser to predict quality indexes in fruit (Qing et al, 2007; Peng et al, 2006; Tu et al, 1995; McGlone et al 1997).

Scattering is a physical phenomenon that is more closely associated with the structure of fruit tissue (Lu, 2004). Light scattering is dependent on the density, cell structures and intra-cellular matrices of fruit, captured onto an image.

Therefore, light scattering is expected to be useful in quantifying certain mechanical and textural properties of fruit, such as firmness (Lu and Peng, 2006).

Multivariable regression methods ('Multiple Linear Regression' (MLR), 'Principal Components Regression' (PCR) y Partial Least Squares (PLS)) are used to correlate backscattering laser images to the physical and chemical properties of fruit.

The objective of this work was to get the soluble solid content and firmness prediction for '*Smoother Golden Delicious*' apple, using monochromatic backscattering imaging. For this purpose, a laser light source providing three wavelengths. Partial Least Square (PLS) regression was carried out and applied on a test-set to evaluate the performance of calibration models using average diameter of Feret corresponding at different levels of gray.

MATERIALS AND METHODS

Samples

The experiment was performed on the Apple cultivar '*Smoother Golden Delicious*', in an experimental orchard, in Zaragoza, Spain. Trees were planted in a 5x3 frame. Fertigation was performed by local drip. Samples were harvested every 15 days between May and October 2012. Backscattering images were taken and destructive methods were performed to determinate SSC and firmness.

Multispectral images

Multispectral images were obtained by a monochromatic camera CCD, (BASLER A310f (Germany) sensitive Vis-NIR (360-1000 nm)). The samples get ready in a device that allowed obtaining a backscattering area of 25 mm. The beam incident was 15° and the working distance from camera to sample was 14 cm. The image size was 782 x 582 pixels. The images were kept at room temperature 20°C. The light source was a solid state laser diode (Power Technology) module, emitting at 670, 785 and 980 nm with an output power of 3 mW and 28.3 mW for, respectively. The large output power of the laser 980 nm is to compensate the loss of camera sensity in this spectral zone. These wavelengths were selected based on other studies (Peng et al, 2006; Qing et al, 2007). Backscattering light from samples was kept by monochromatic images (Figure 1) where it was possible observe to the two zones. Each image was analyzed with Matrox 8.0 Software. The gray scale of light intensity is expressed as the integer, which ranged from dark (0) to white (255).

Texture

Firmness was measured as the resistance needed by the fruit pulp, to penetrate (by force) 10mm to the apple. The texture analyzer used (model TA.XT Plus Texture Analyzer, Texture Technologies (North American)) had a loading rate of 2mm/sec. Fruit skin was removed prior to taking values in the area measured. The texture analyzer was equipped with a load cell of 5 kgf.

Solid Soluble Content

After firmness measurements, juice was extracted and SSC was measured using a digital refractometer (Model ATAGO PR-101 Co., Tokyo, Japan).

Statistical treatment

Partial Least Square analysis (PLS)

PLS analysis is a regression method used often to construct prediction models of reference parameters taken from destructive methods. PLS allows the correlation between backscattering images with ripeness parameters. PLS has been used frequently in studies related to the prediction of quality indexes in various fruits and vegetable: avocados (Clark et al, 2003), nectarines (Sánchez et al, 2011) and asparagus (Pérez et al, 2002).

In this work, calibration models were developed to predict firmness and SST of intact apples. Cross-validation was used to avoid overfitting in the development of calibration models. Data management was performed using The Unscrambler X (CAMO) software.

Selection of samples for calibration and validation purposes is one of the critical steps when developing a calibration model. A random subdivision on calibration and validation data was performed, using a 2:1 ratio for calibration and validation respectively. A Principal Component Analysis (PCA) was performed to detect anomalous samples obtained from calibration models. The samples detected as anomalous were not included in the calibration groups.

Statistics used to select the best equations were: coefficient of determination for cross validation (R^2) and standard error of cross validation (SEVC)(Williams, 2001). Other statistics used were the coefficient of variation (%CV), defined as the percentage ratio of the SECV to the mean value of the reference data for the calibration set, and the RPD (Residual Predictive Deviation) calculated as the ratio of the standard deviation (SD) of the reference data to the SECV (Williams, 2001). %CV and RPD enable SECV to be standardized, facilitating in comparing the results obtained from sets of different means (Williams, 2001).

The best predictive models obtained were subsequently subjected to external validation. The effect of different settings on model performance was evaluated by comparing the root mean square error of prediction (RMSEP), coefficient of external validation (r^2), bias and standard error of prediction (SEP).

RESULTS AND DISCUSSIONS

PLS analysis constructed calibration models from destructive (SSC and firmness) and non-destructive measurements (backscattering images). Table 1 shows the number of samples for calibration, validation, mean values, standard deviation, range and coefficient of variation. The validation group range of validation is included in calibration, which demonstrates the group selection was made correctly.

Table 2 shows the statistical parameters obtained from the regression analysis. The R_{cv}^2 values obtained were 0.93 and 0.87 for SSC and firmness, respectively. The number of principal components found in our calibration model were greater than values found in previous studies (Zude et al. 2007, $R_{cv}^2=0.76, 0.80$; and Lu, 2003, $R_{cv}^2=0.77$). PLS results presented in this study were found to be more precise when compared to PLS values previously reported in other studies. Our large sample population size along with our diverse range of physical and chemical values assisted in obtaining more precise values.

Figure 2 shows the external statistical validation performed on 69 samples. r^2 values were $r^2=0.86-0.74$ for SSC and firmness. With these results it is deduced that it is possible to obtain a good predictions in apple for the first, because the time of sampling has been very wide. It is necessary to continue working to construct better calibration models.

CONCLUSIONS

This study has demonstrated the possibility to apply nondestructive technique for the analysis of backscattering images to destructive quality parameters in apple ‘*Smoothie Golden Delicious*’.

In the future, it is necessary to construct prediction models which are more precise.

It should be noted that this proposed methodology could be applied to any variety of apple.

ACKNOWLEDGMENTS

This work has been partially supported by the AGL2009-08501/AGR project from National Program of Research Projects.

I would like to thank Michael Krawitzky, Azahara Díaz, M^a Ángeles Gracia for their help

REFERENCES

- [1]Chen, P. 1996. Quality evaluation technology for agricultural products. In: *Proc. Int. Conf. on Agric. Machinery Engineering*. Seoul. Korea
- [2]Clark, C.J., McGlone, V.A., Requejo, C., White, A., Woolf, A.B. 2003. Dry matter determination in ‘Hass’ avocado by NIR spectroscopy. *Postharvest Biology and Technology*. Volume 29, pp. 300-307. Cover, T.M. and Thomas, J.A. 1991. *Elements of Information Theory*, Wiley, New York.
- [3]López M.L., Lavilla M.T., Recasens I., Graell J., Vendrell M., 2000. Changes in aroma quality of ‘Golden Delicious’ apples after storage at different oxygen and carbondioxide concentrations. *Journal of the Science of Food and Agriculture*, 80, 311-324
- [4]Lu, R., 2004. Multispectral imaging to predicting firmness and soluble solids content of apple fruit. *Postharvest Biology and Technology* 31:147-157.
- [5]Lu, R., Peng, Y., 2006. Hyperspectral scattering for assessing peach fruit firmness. *Biosystem Engineering* 93 (2), 161–171.
- [6]McGlone, V. and Kawano, S. 1998. Firmness, dry matter and soluble solids assessment of postharvest kiwifruit by NIR spectroscopy. *Postharvest Biology and Technology* 13:131-141.
- [7]Peng, Y and Lu, R. 2007. Prediction of apple fruit firmness and solid soluble content using characteristics of multispectral scattering images. *Journal of Food Engineering* 82 142–152.

[8]Pérez-Marín, D., Sánchez, M.T., Cano, G. and Garrido, A. 2002. Prediction of texture in green asparagus by near infrared spectroscopy (NIRS). *Journal Food Quality* 25: 277-287.

[9]Qing, Z., Ji, B. and Zude, M. 2007. Wavelength selection for predicting physicochemical properties of apple fruit based on near-infrared spectroscopy. *Journal of food quality*. Volume 30, Issue 4, pp.511-526.

[10]Sánchez, M.T., De la Haba, M.J., Guerrero, J.E., Garrido-Varo, A. and Pérez-Marín, D. 2011. Testing of a local approach for the prediction of quality parameters in intact nectarines using a portable NIRS instrument. *Postharvest Biology and Technology* 60:130-135.

[11]Tu, K., De Busscher, R., De Baerdemaeker, J., Schrevens, E. 1995. Using laser beam as light source to study tomato and apple quality non-destructively. *Food Processing Automation IV Conference*, Chicago, IL, 3-5. November. Pp. 528-536.

[12]Williams, P.C., 2001. Implementation of near-infrared technology. In: Williams, P.C.,Norris, K.H. (Eds.), *Near-infrared Technology in the Agricultural and Food Industries*. AACC, Inc., St. Paul, MN, pp. 145–169.

FIGURES

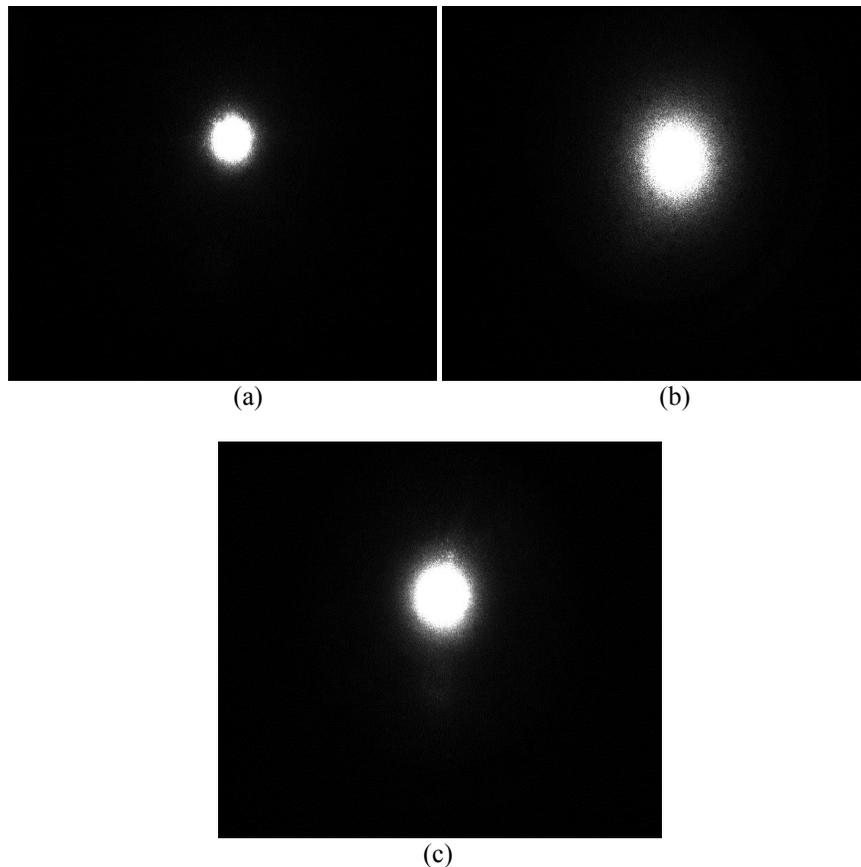


Figure 1. Multispectral images obtained with 670(a), 785(b) y 980(c) nm, respectively in ‘*Smoothee Golden Delicious*’ apple.

Table 1. Mean, standard deviation, range and coefficient of variation for calibration and validation groups for SSC and firmness in ‘*Smoother Golden Delicious*’ apple

	SSC		Firmness	
	calibration	Validation	calibration	validation
N	139	69	139	69
Mean	11.1	11.3	5.51	5
DS	2.2	2.2	2.28	1.78
Range	8.2-16.3	8.2-15.5	2.16-10.55	2.74-9.18
%CV	19.4	19.1	41.46	35.59

Table 2. Calibration statistics for SSC and firmness in ‘*Smoother Golden Delicious*’ apple

	PLS	
	SST	Firmness
N	131	131
PLS	1	2
SECV	0.78	0.85
R_{cv}²	0.93	0.87
RPD	2.17	2.65

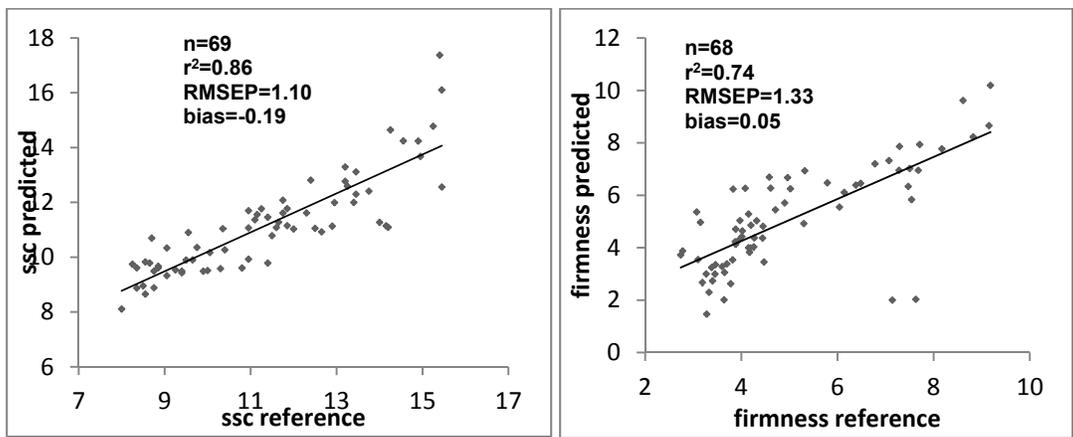


Figure 2. Prediction model for SSC and firmness in ‘*Smoother Golden Delicious*’ apple

ANEXO II

'Firmness prediction in *Prunus persica Calrico* peaches by visible/short-wave near infrared spectroscopy and acoustic measurements using optimised linear and non-linear chemometric models'

Victoria Lafuente, Luis Javier Herrera, Jesús Val, A. Ignacio Negueruela

Firmness prediction in *Prunus persica* 'Calrico' peaches by visible/short-wave near infrared spectroscopy and acoustic measurements using optimised linear and non-linear chemometric models

Victoria Lafuente,^{a*} Luis J Herrera,^b María del Mar Pérez,^c Jesús Val^d and Ignacio Negueruela^e

Abstract

BACKGROUND: In this work, near infrared spectroscopy (NIR) and an acoustic measure (AWETA) (two non-destructive methods) were applied in *Prunus persica* fruit 'Calrico' ($n = 260$) to predict Magness–Taylor (MT) firmness.

METHODS: Separate and combined use of these measures was evaluated and compared using partial least squares (PLS) and least squares support vector machine (LS-SVM) regression methods. Also, a mutual-information-based variable selection method, seeking to find the most significant variables to produce optimal accuracy of the regression models, was applied to a joint set of variables (NIR wavelengths and AWETA measure).

RESULTS: The newly proposed combined NIR-AWETA model gave good values of the determination coefficient (R^2) for PLS and LS-SVM methods (0.77 and 0.78, respectively), improving the reliability of MT firmness prediction in comparison with separate NIR and AWETA predictions. The three variables selected by the variable selection method (AWETA measure plus NIR wavelengths 675 and 697 nm) achieved R^2 values 0.76 and 0.77, PLS and LS-SVM.

CONCLUSION These results indicated that the proposed mutual-information-based variable selection algorithm was a powerful tool for the selection of the most relevant variables.

© 2014 Society of Chemical Industry

Keywords: NIR; acoustic system; firmness; peaches

INTRODUCTION

From the time peaches (*Prunus persica*) are harvested, the fruits are separated from the source of nutrients but the fruit tissues continue to breathe and are physiologically active. Ripening is linked to complex physical and chemical changes, such as softening, increased concentrations of soluble sugars, and changes in flavour and colour. These processes are important because they influence changes that occur during storage, transportation and marketing and, to some extent, affect the nutritional value and sensorial characteristics of peaches.¹ In addition, these metabolic processes are influenced by external factors such as seasonal and genetic factors, e.g. crop variety or situation. Harvest date not only influences product quality but also boosts production while minimising agricultural costs.²

Firmness is one of the parameters that correlates best with fruit ripeness and is ultimately very useful in determining the optimal harvest date.^{3,4} Firmness can be defined as the resistance by the fruit pulp⁵ to penetration force. Firmness has long been determined by the destructive Magness–Taylor method (MT). The MT test involves attaching a dynamometer to a cylindrical rod,

which is inserted 8 mm into the fruit pulp after part of the skin has been removed. Although this test is inexpensive and rapid, it is of limited usefulness because it destroys the fruit. Also, it renders uncertain values and is highly variable, depending on the force exerted by the operator and the location on the fruit from where the measurements are taken.

* Correspondence to: Victoria Lafuente, Consejo Superior de Investigaciones Científicas (CSIC), Nutrición Vegetal, Zaragoza, 50059, Spain E-mail: mvlafuente@ipe.csic.es

a Consejo Superior de Investigaciones Científicas (CSIC), Nutrición Vegetal, Zaragoza, Spain

b Departamento de Arquitectura y Tecnología de los computadores, Universidad de Granada, Granada, Spain

c Departamento de Óptica, Universidad de Granada, Granada, Spain

d Estación Experimental de Aula Dei, CSIC, Plant Nutrition, Avda, Montañana 1005, Zaragoza, Spain

e Física aplicada, Universidad de Zaragoza, Zaragoza, Spain

Since 2003, more practical and reproducible non-destructive methods have been developed for estimating fruit firmness with the aim of replacing destructive ones. One such method involves the use of acoustic signals caused by vibrations or mechanical impacts to the fruit. The resonant frequency that is emitted by an object depends directly on the geometry, mass and modulus of elasticity of the material being analysed.⁶ Acoustic tests performed on the sample are stimulated by a low-intensity impact, producing a vibratory response within an audible range (20–20 000 hertz). The response is recorded with a microphone and the signal in time is processed using fast Fourier transform to produce the corresponding signal in terms of the frequency, which produces an acoustic firmness index. Many studies are available on the use of acoustic methods in determining the firmness in various types of fruit and vegetables. However, the results show a high variability in the correlation between the two variables, depending on the cultivar and variety of the analyte (e.g. apple variety 'Delicious', $R^2 = 0.25$; apple variety 'Golden Delicious', $R^2 = 0.72$;⁷ avocado,⁸ $R^2 = 0.43$).

Another emerging non-destructive technique currently used for firmness prediction is near infrared spectroscopy (NIRS). Many researchers are developing methods based on this technique to predict firmness values for different fruits (e.g. apples,⁹ nectarines,¹⁰ peaches,¹¹ avocados¹²). In the NIRS technique, a light beam strikes the fruit, penetrating a few millimetres. Part of this radiation is absorbed by the sample and the remainder is reflected towards a detector, which converts this light into a spectral characteristic for each sample. Depending on their physical and chemical composition, different compounds such as sugars, water and pigments absorb part of the radiation in the visible and infrared region.

Signals recorded in these non-destructive methods are related to the desired estimation parameter from which calibration equations are formulated. The resulting firmness MT value is considered to be the reference. Despite the variability, it is an acceptable reference in fieldwork. Linear model connections are some of the more commonly used regression methods, such as multiple linear regression,¹³ principal components regression,¹⁴ and partial least squares (PLS).¹⁵ Nonlinear methods such as least squares support vector machines (LS-SVM)¹⁶ and artificial neural networks¹⁷ are used also.

The use of variable selection methods is of particular relevance in these types of physico-chemical property-modelling problems related to fruit ripeness. These methods aim to determine the actual contribution of the factors (wavelengths, acoustic frequencies) considered in estimating the objective parameter. These techniques enable the development of predictive models by eliminating spurious factors that complicate the model and that may affect output. In the case of NIRS, different authors have used the selection of variables to improve their results, such as $R = 0.97$, root mean square error calibration (RMSEC) = 0.2°Brix, given by Liu *et al.*,¹⁶ and the root mean square error of prediction (RMSEP) = 0.02 mg L⁻¹, given by Araújo *et al.*¹⁸ These approaches include methods such as independent component analysis,¹⁶ successive projections algorithm¹⁸ and uninformative variable elimination.¹⁹ In addition, methods based on mutual information are of particular relevance.^{20,21} These methods use mutual information as a measure of non-linear correlation among intervening factors and the variable to be estimated. They have the advantage of not altering the original features considered and directly detect the factors.

The main objective of this work was to investigate peach firmness prediction from the variety 'Calrico' using information gained through non-destructive methods, i.e. visible/NIR and acoustic signalling equipment (model AFS; AWETA Technologies, Nootdorp, the Netherlands). A mutual-information-based variable selection method was developed to identify the relevance of both sources of information and select the most prominent factors to make the prediction and build the most accurate and reproducible models.

MATERIAL AND METHODS

Samples

This study was performed during 2010 and 2011, in an experimental peach (*Prunus persica*) orchard, called AFRUCAS, located in the area of Badilla del Corral de Dolader, belonging to Caspe (Zaragoza, Spain). The experiment was performed on peach variety 'Calrico' over a hybrid pattern. This variety is of particular interest in the area of Bajo Aragón, which has applied for the certificate of origin 'Calanda'. The trees were planted in 2000, in a pattern of 5 × 3 m. Fertigation was applied by local drip (25 000 L ha⁻¹ h⁻¹). Samples (260 fruits) were harvested in 2010 (150), 50 samples were harvested weekly for 3 weeks and in 2011 (110), 22 peaches were tested weekly for 5 weeks. In this way, maturity ranged from hard green (unripe) to soft mature (over-ripe) stages. The samples were harvested early in the morning, transported at 20 °C to the laboratory for analysis (same afternoon). Non-destructive determinations (AWETA and NIRS) were made prior to destructive MT firmness determination.

Visible/near infrared spectroscopy

Spectra from intact peaches were measured with a reflectance modular equipment Multispec instrument (AG Tec5; AM Frankfurt, Germany) equipped with a spectrometer SC-NEM I (Zeiss, Jena, Germany) (range, 400–1060 nm, $\Delta\lambda = 1$ nm), a halogen light source (12 V and 20 W), and one fibre optic to conduct the light from the source on the sample and another fibre to carry the light reflected by the samples to the detector. The illumination-measurement geometry was set up at 45°/0° (Fig. 1)

The measured spectra were the average of 50 spectra (integration time 400 ms). Two spectral measurements were made per fruit at two fixed opposed positions on the equator, using the mean of the two spectra for the calibration processes.

Acoustic measurements

Acoustic firmness measurements were made by means of a commercial desk-top acoustic-firmness sensor (model AFS; AWETA).

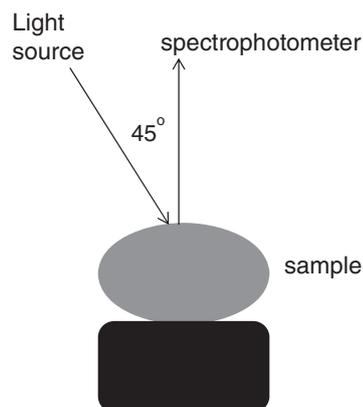


Figure 1. Set-up for the acquisition of reflectance spectra.

The sensor recorded the weight and resonant frequencies of the acoustic vibration generated by gently tapping the fruit on the equatorial area, from which an acoustic firmness index (FI) was provided by the AWETA equipment:^{22–24}

$$FI = f_R^2 \times m^{2/3}$$

In this index, the mass, m (in kilograms) and the resonant frequency of maximum amplitude, f_R (in hertz) are related. Acoustic measurements were made three times per fruit and the average value was calculated.

Magness–Taylor test

The Magness–Taylor firmness test was performed using a hand-held penetrometer Fruit Pressure Tester FT32 (Istituto per la Valorizzazione dei Prodotti Agricoli, Milano, Italy) with a probe 8 mm in diameter. The fruit skin was removed with a blade at two positions around the equator and firmness measured (kg). The firmness was tested in the same area where NIR reflectance spectra were measured.

Data analysis

Chemometric data analysis was performed using the Matlab R2009a software (The MathWorks, Natick, USA) with own programming. Different pre-treatment combinations were evaluated for calibration models such as multiplicative scatter correction, standard normal variate, and first- and second-order derivatives. Finally, the best results were found using normalisation with a mean of 0 and variance of 1 for the values for each wavelength and acoustic measurement. This type of normalisation is applied individually to each variable considered, whether it be spectral or otherwise; in this way it is independent of whether or not fusion is variable.^{20,25}

Selection of samples for calibration and validation purposes is one of the critical steps when developing a calibration model. To detect the spectra of outliers, a principal components analysis was performed. The samples detected as outliers were not included in the calibration groups. Three different random sub-divisions on calibration and validation data were performed, using a ratio 2:1 for calibration and validation, respectively.¹⁰

Partial least squares analysis

PLS is a regression method often used to construct prediction models for reference parameters established by destructive methods. A high number of variables which may have high collinearity are used. PLS, which allows the visible/NIR spectra to correlate with physico-chemical parameters of the fruit, has been used in a considerable number of studies related to the prediction of indices in fruits and vegetables.^{26–28}

In the present study, calibration models were developed to predict firmness in intact peaches. L -fold cross-validation ($L = 10$) was used to avoid over-fitting in the development of the calibration models.^{3,10,29,30}

Least squares support vector machines

Least squares support vector machines (LS-SVM) is a regression model that has been used in recent years to predict parameters related to fruit ripening and other chemical and physical properties. It has the ability to deal with linear and non-linear multivariate modelling and resolving these problems in a relatively quick way.³¹ Previous research has demonstrated the potential of this

non-linear regression model for several quantitative applications in agro-food products.^{11,32,33} In the development of the LS-SVM model, the radial basis function kernel was used. Grid-search and cross-validation (LS-SVMLab²⁵) were used to achieve the optimal combination of γ and σ^2 (δ) hyper-parameters of the model.

In both regression methods, the statistics used to select the most accurate models were: the coefficient of determination for cross-validation (R_{cv}^2) and standard error of cross-validation (SECV).³⁴ Other statistics used were the coefficient of variation (%CV), defined as the percentage ratio of the SECV to the mean value of the reference data for the calibration set, and the residual predictive deviation (RPD) calculated as the ratio of the standard deviation (SD) of the reference data to the SECV.³⁴ These latter two statistics facilitate the comparison of results found with sets of different means.³⁴

Predictive models formulated were subsequently subjected to internal validation. The effect of different settings on the performance of the model was evaluated by comparing root mean square error of prediction (RMSEP) determination coefficient in validation (R^2), bias and standard error of prediction (SEP).

Mutual information backwards variable selection

Mutual information is a non-linear correlation measurement from the information theory.³⁵ For two sets of continuous variables, X and Y , it can be calculated by:

$$I(X, Y) = \int \frac{\mu_{X,Y}(x, y) \log [\mu_{X,Y}(x, y)]}{\mu_X(x) \mu_Y(y)} dx dy$$

where $\mu(x, y)$ is the joint probability density function of X and Y , and $\mu_x(x)$ is the marginal density function of the set of variables X . Its advantage in comparison with other correlation criteria is that it is able to identify non-linear relations among the variables involved.

Several attempts have been reported in recent literature for designing algorithms to identify the most relevant factors (wavelengths and/or other factors) for predicting chemical properties, many of those being based on mutual information.^{20,21,36} Variable selection aims at identifying irrelevant and redundant variables for their rejection. Identification of redundant variables is critical in spectrometric problems, as nearby wavelengths provide usually similar information. The reduction in variables needed to predict any magnitude is essential to reduce experimental and evaluation costs, but also for increasing the so-called generalisation capability of the models, i.e. prediction capability on unseen data.³⁷

The method proposed in this work is an approach first put forward by Koller and Sahami³⁸ and adapted for continuous variables,³⁷ which is based on the Markov blanket concept. Given a set of input variables X and an output variable Y , a set of variables M_i in X is said to be a Markov blanket for a variable x_i in X with respect to Y , if $I(\{M_i \cup x_i\}, Y) = I(M_i, Y)$, provided that M_i itself has all the information that x_i has about Y . A Markov blanket is thus a group of variables that subsumes the mutual information content in a certain variable, in practice (and for our purposes) with respect to the objective variable. The algorithm consists of a backwards variable selection method which starts with the complete set of variables, and iteratively discards those which are detected to have a Markov blanket in the remaining set X_G of variables; that is, those whose information with respect to Y is already present in the remaining set X_G of variables.

Table 1. Number of samples (N), mean (X), standard deviation (SD), range and coefficient of variation (%CV) for firmness in the three calibration and validation groups

Parameter	Group 1		Group 2		Group 3	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
N	194	66	194	66	194	66
X	3.9	3.6	3.7	3.5	3.8	3.7
SD	1.79	1.61	1.7	1.57	1.76	1.7
Range	1–8.7	0.6–8.7	1.2–8.7	1–8.7	0.6–8.7	1.2–8.7
%CV	47.26	44.72	46.21	44.86	46.37	46.41

The algorithm states the following steps:

- 1 Calculate the mutual information between each pair of input variables $I(x_i, x_j)$.
- 2 Starting from the complete set of input variables $X_G = X$, iterate:
 - i For each variable x_i , let the candidate Markov blanket M_i be the set of p variables in X_G for which $I(x_i, x_j)$ is highest.
 - ii For each x_i compute $Loss_i = I(\{M_i \cup x_i\}, Y) - I(M_i, Y)$.
 - iii Choose the x_i for which $Loss_i$ is lowest and eliminate x_i from X_G .
- 3 Continue with step 2 until no variables remain.

In this way, the relevance of variables (in reverse order) is ranked. Under this operation, it should be noted that variables that have low influence with respect to the output variable (irrelevant variables) will be soon discarded, as the $Loss_i$ value should tend to 0. Similarly, redundant variables are iteratively discarded at earlier stages. Relevant variables with low redundancy are the last to be 'chosen'. Further discussion about efficiency, character and operation of the algorithm can be found in Herrera *et al.*³⁹

The mutual information estimator used in this work is the k -nearest neighbours one,⁴⁰ as it has been shown in the literature to have a more robust behaviour with respect to other alternatives when working with groups of variables. The p parameter of the algorithm (in step 2i of the algorithm) will take the value $p = 1$, as recommended.³⁹

RESULTS AND DISCUSSION

Data description

Table 1 shows the number of calibrated and validated samples. Mean values, standard deviation, range and coefficient of variation for three sub-divisions of the initial set of samples from the MT firmness parameter are also shown for all three groups. It is observed that all three calibration groups present similar mean values, range and standard deviation. The range of validation of the sub-groups is included in those of calibration, which demonstrates that group selection has been made correctly. The values of standard deviation and coefficient of variation indicate high variability in each group.

Magness–Taylor firmness prediction through acoustic measurement

The correlation between the data of MT firmness and acoustic signal was studied. Regression coefficients of the three study groups were between $R^2 = 0.52$ and $R^2 = 0.55$. Internal validation was performed in which R^2 values ranged from 0.45 to 0.58, the values of RMSEP = 1.04–1.16 and bias = –0.03 to 0.02. In similar studies

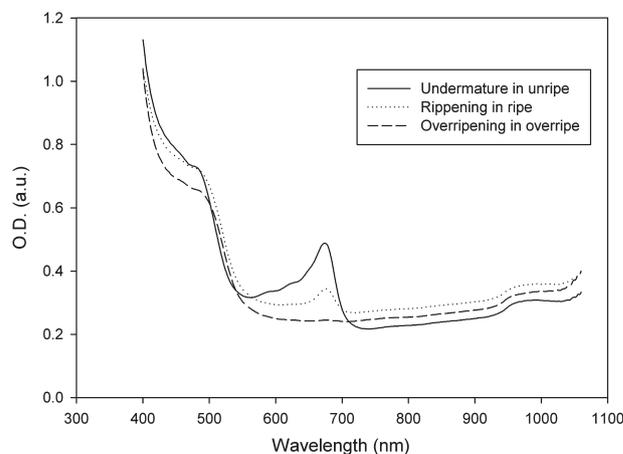


Figure 2. Reflected spectra of peaches in three ripening stages.

with peaches, values found for the correlation of both parameters were slightly lower ($R^2 = 0.38$)⁴¹ and $R^2 = 0.45$.⁴² The results show that, despite having higher R^2 values than previously reported, the acoustic signal had correlations with MT firmness which were relatively low in comparison with other non-destructive methods, such as visible/NIR spectroscopy.⁴³

Magness–Taylor firmness prediction through visible/near infrared spectroscopy

Figure 2 shows the average visible/NIR spectra of peaches at different stages of ripeness. The main component of these spectra is the peak chlorophyll absorption (680 nm). Chlorophyll content decreases as fruit ripens. The decline in the absorption peak was observed until the fruit reached a ripeness point at which the peak disappeared. Changes in the spectral region below 700 nm were attributed to changes in fruit colour, due to pigment absorption, disappearance of chlorophyll and synthesis of coloured pigments, mainly carotenoids and anthocyanins.

Calibration models for MT firmness prediction of peach based on visible/NIR spectra were developed with both linear regression (PLS) and non-linear regression (LS-SVM). Table 2 shows the statistical parameters determined from the analysis through both methods of regression and for the three calibration groups. The R_{cv}^2 values for PLS and LS-SVM had the same value (0.7), a good correlation between spectra and firmness. The non-linear model, LS-SVM, presented lower SECV values and higher RPD values for the three calibration groups, signifying that the non-linear regression method provided a better correlation between MT firmness and the absorption spectrum.

Table 3 shows the statistics of internal validations performed for the three sample groups. R^2 values are similar for both regression

Table 2. Calibration statistics for firmness using partial least squares (PLS) and least squares support vector machines (LS-SVM) regression methods for visible/NIR and visible/NIR + AWETA

Method and parameter	Visible/NIR			Visible/NIR + AWETA		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
Partial least squares						
<i>N</i>	194	194	194	194	194	194
No. of components	5	5	5	5	5	5
SECV (kg)	0.97	0.99	0.92	0.91	0.92	0.9
R_{cv}^2	0.69	0.71	0.74	0.74	0.74	0.74
RPD	1.84	1.82	1.91	1.97	1.96	1.95
Least squares support vector machines						
<i>N</i>	194	194	194	194	194	194
No. of components	5	5	5	5	5	5
SECV (kg)	0.89	0.87	0.91	0.83	0.83	0.84
R_{cv}^2	0.64	0.71	0.61	0.86	0.84	0.79
RPD	2.00	2.07	1.93	2.16	2.17	2.09

N, number of samples; SECV, standard error of cross-validation; R_{cv}^2 , coefficient of determination for cross-validation; RPD, residual predictive deviation.
NIR, near infrared spectroscopy; AWETA refers to the acoustic signalling equipment used in this study.

Table 3. Validation statistics using partial least squares (PLS) and least squares support vector machines (LS-SVM) regression methods for the prediction for firmness using visible/NIR, visible/NIR + AWETA, and visible/NIR + AWETA with three variables

Method and parameter	Visible/NIR			Visible/NIR + AWETA			Visible/NIR + AWETA (variable selection)		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
Partial least squares									
<i>N</i>	66	66	66	66	66	66	66	66	66
RMSEP (kg)	0.85	0.84	0.71	0.79	0.81	0.81	0.77	0.9	0.88
R^2	0.71	0.71	0.72	0.75	0.73	0.77	0.76	0.66	0.73
Bias	0.04	-0.01	0.07	-0.005	-0.05	0.01	-0.001	-0.13	-0.005
%CV	23.94	24.21	19.29	22.25	23.34	22.01	21.69	26.55	24.04
Least squares support vector machines									
<i>N</i>	66	66	66	66	66	66	66	66	66
RMSEP (kg)	0.9	0.92	0.86	0.86	0.84	0.79	0.76	0.84	0.81
R^2	0.68	0.74	0.74	0.71	0.71	0.78	0.77	0.71	0.77
Bias	-0.08	-0.07	0.12	-0.1	-0.16	0.05	0.018	-0.13	-0.019
%CV	26.09	27.54	23.56	24.93	25.07	21.18	21.29	24.78	22.13

N, number of samples; RMSEP, root mean square error of prediction; R^2 , coefficient of external validation; %CV, coefficient of variation.
NIR, near infrared spectroscopy; AWETA refers to the acoustic signalling equipment used in this study.

methods; however, RMSEP values and %CV are lower for the PLS model. The results imply the possibility of a good MT firmness prediction through the use of NIR spectroscopy, as the findings were consistent with those of other authors in similar works on peaches⁴⁴ ($R_{cv}^2 = 0.74$ and RMSEP = 5.42 N) using linear and non-linear regression models and reporting similar results with both methods²³ in which linear regression methods were used to formulate models, giving $R_c^2 = 0.83$ and SEC = 1.70 kg cm⁻² as a result.

Regression coefficients from the PLS model for one of the calibration sets are shown in Fig. 3. The more relevant wavelengths in the absorption spectrum of peach are identified by higher values, in absolute value, of the regression coefficients. Specifically, it is observed that the bands 470–550 nm and 650–690 nm are of great importance. Bands at 400 and 1060 nm were not considered because they contributed excessive noise and did not improve

model. The higher regression coefficient values in the range of 600–700 nm are consistent with previous studies,^{4,9,11,23} that consider the visible region to be directly related to fruit ripeness. These absorption bands are related to the pigments that provide the fruit with colour.

Magness–Taylor firmness prediction by combining the acoustic signal and visible/near infrared spectra

The acoustic index was combined with each of the spectral vector samples visible/NIR, giving extended input data (194 × 661). With the same objective, acoustic measurement was added to the visible/NIR spectrum. Similar strategies carried out²⁴ with mango ($R_{cv}^2 = 0.61$, RMSECV = 3.20 N mm) and⁹ with apple ($R_{cv}^2 = 0.80$ to 0.75, SECV = 10.32–11.28 N cm⁻²) gave favourable results with the combination of variables in all cases.

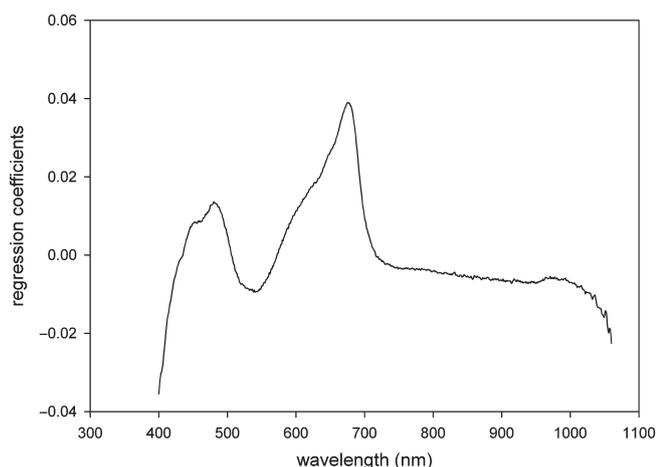


Figure 3. Regression coefficients for firmness for calibration group 1.

The results demonstrated the suitability of a new method involving the pooling of heterogeneous information gathered using several methods to construct more effective calibration models for predicting physico-chemical parameters of fruits and vegetables. Therefore, fusion of the firmness index and visible/NIR spectra are assumed to give more robust predictions for fruit-flesh firmness. Calibration equations were formulated from the proposed regression methods (Table 2) following the same methodology as in the previous cases. R_{cv}^2 values were found to be greater for both methods (LS-SVM, 0.8–0.9; and PLS, 0.7). SECV values were lower for PLS and LS-SVM. RPD values were above 2 in each of the three groups. Internal validation was also performed (Table 3). RMSEP and coefficients of variation were lower and regression coefficients were higher.

Selection of variables

To reduce the number of variables (wavelength and acoustic response) and analyse the possibility of improving the generalisation ability of the models constructed, we applied the variable selection method backwards; based on mutual information.³⁹ Many researchers use selection methods to detect the most relevant variables for which simpler models and higher outputs result,⁴⁵ making use of the correlation coefficient and with the method called the successive projection algorithm.¹⁸ Figure 4 shows the output of the models built through two methods of regression on the number of variables. There is an increase in R^2 with respect to the use of two to three variables. In the third variable, the increase in factors does not imply an immediate output increase.

The regression coefficient calculated by using three variables was 0.77 and 0.73 for LS-SVM and PLS, respectively. This pattern is observed in all three groups of calibration. LS-SVM models responded better to the creation of models with fewer variables, as shown in similar works.^{3,16,46} This may have occurred because, in the non-linear method, both linear and non-linear are taken, whereas linear models (PLS) work only with a linear relationship between spectral and physico-chemical data. Table 3 shows the results of validation of the model constructed with three factors. In both cases for LS-SVM and PLS the R^2 and RMSEP values were improved and %CV was similar. The main advantages of variable selection reduced collinearity, redundancy and noise. The use of the spectrum involved building a model with only three variables that produced better results.

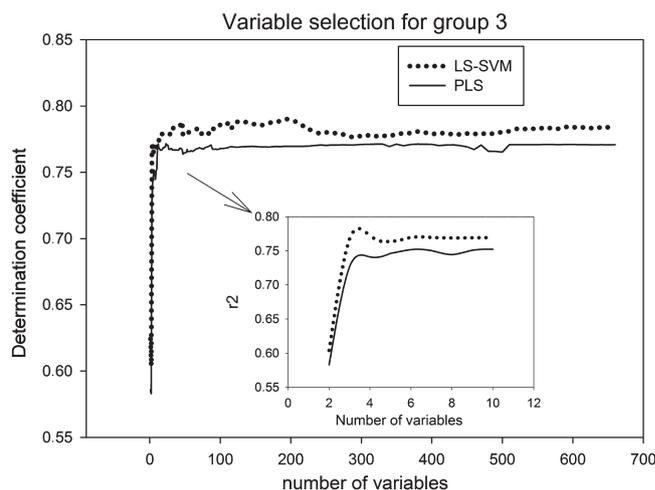


Figure 4. Example of variable selection for both regression models (LS-SVM, dotted line; and PLS, solid line), using all variables (661) and using only 10 variables.

From the three variables selected, two correspond to wavelengths of the visible/NIR spectrum and the third is the firmness index formulated from the acoustic signal. The two selected wavelengths in the three groups proved similar (679 and 696 nm, 675 and 697 nm, 674 and 693 nm). This confirms the importance of the chlorophyll absorption region to determine MT firmness (Fig. 4). Similar results have been reported for apples.^{23,47,48} In this way, to calculate the state of ripeness with much simpler instrumentation that requires only two simple optic devices in the visible region, with a filter and a detector for each wavelength selected. The diffraction grating and detector array are eliminated. AWETA equipment becomes necessary.

CONCLUSIONS

This work, using the peach variety ‘Calrico’, develops a non-destructive method, increasingly demanded by horticultural and agricultural production centres, for MT firmness. The peach variety treated is of particular interest in the area of Bajo Aragón in Spain, which seeks its certificate of origin ‘Calanda’. The proposed method is based on the pooling of non-destructive heterogeneous information, specifically a measure of acoustic pulse and visible/NIR spectral data.

Initially, the results show that the combination of two non-destructive methods, AWETA and visible/NIR spectroscopy, improve the reliability of MT firmness prediction with respect to a single source of information used alone. With this new approach, R^2 and RMSEP values improved in all cases. Furthermore, it was found that the visible/NIR spectrum provided greater reliability in estimating MT firmness than did the AWETA measure.

Finally, the variable selection method based on mutual information used succeeded in finding three optimal parameters for model calibration from both information sources: the acoustic measurement and two wavelengths and for each the corresponding visible chlorophyll absorption region. While using three variables LS-SVM maintenance was observed (and even slightly increased) in the output of the calibration models (in comparison with the PLS models). Furthermore a considerable simplification of the constructed prediction models was obtained. The equation formulated uses only three of the 661 available input variables, providing a simpler methodology and being more affordable in the field of agriculture.

Lastly, it should be noted that this proposed methodology could be applied to any species and variety of fruit and vegetable, as expected in the future.

ACKNOWLEDGEMENTS

This work was partially supported by the AGL2009-08501/AGR project from the National Program of Research Projects, and the GENIL-PYR-2014-12 project from the GENIL Program of the CEI BioTic, Granada, Spain. We are indebted to volunteers in this study for their outstanding commitment and cooperation.

REFERENCES

- Kader, AA, Flavor quality of fruits and vegetables. *J Sci Food Agric* **88**:1863–1868 (2008).
- Thompson JF, Sistemas de cosecha, in *Tecnología Postcosecha de Productos Hortofrutícolas*, 3rd edition, ed. by Kader A. University of California, Division of Agriculture and Natural Resources, Oakland, California, pp. 73–76 (2007).
- Li J, Huang W, Zhao C and Zhang B, A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy. *J Food Eng* **16**:324–332 (2013).
- Ruiz-Altisent M, Lleo L and Riquelme F, Instrumental quality assessment of peaches: Fusion of optical and mechanical parameters. *J Food Eng* **74**:490–499 (2006).
- Mohsenin NN, *Physical Properties of Plant and Animal Materials: Structure, Physical, Characteristics and Mechanical Properties*. Gordon and Breach Science Publishers, New York (1986).
- Studman CJ, Computers and electronics in postharvest technology: A review. *Comp Elec Agric* **30**:109–124 (2001).
- Abbott JA, Massie DR, Upchurch BL and Hruschka WR, Nondestructive sonic firmness measurement of apples. *Trans ASAE* **38**:1461–1466 (1995).
- Galili N, Shmulevich I and Benichou N, Acoustic testing of avocado for fruit ripeness evaluation. *Trans ASAE* **41**:399–407 (1998).
- Zude M, Herold B, Roger JM, Bellon-Maurel V and Landahl S, Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. *J Food Eng* **77**:254–260 (2006).
- Sánchez M, De la Haba MJ, Guerrero JE, Garrido-Varo A and Pérez-Marín D, Testing of a local approach for the prediction of quality parameters in intact nectarines using a portable NIRS instrument. *Postharvest Biol Technol* **60**:130–135 (2011).
- Shao Y, Bao Y and He Y, Visible/near-infrared spectra for linear and nonlinear calibrations: A case to predict soluble solid contents and pH values in peach. *Food Bioprocess Technol* **4**:1376–1383 (2011).
- Clark CJ, McGlone VA, Requejo C, White A and Woolf AB, Dry matter determination in 'Hass' avocado by NIR spectroscopy. *Postharvest Biol Technol* **29**:300–307 (2003).
- Clark CJ, McGlone VA and Jordan RB, Detection of brownheart in 'Braeburn' apple by transmission NIR spectroscopy. *Postharvest Biol Technol* **28**:87–96 (2002).
- Hernández-Gómez A, He Y and García-Pereira A, Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using vis/NIR-spectroscopy techniques. *J Food Eng* **77**:313–319 (2006).
- Peirs A, Tirry J, Verlinden B, Darius P and Nicolai B, Effect of biological variability on the robustness of NIR models for soluble solids content of apples. *Postharvest Biol Technol* **28**:269–280 (2003).
- Liu F, Jiang Y and He Y, Variable selection in visible/near infrared spectra for linear and nonlinear calibrations: A case study to determine soluble solids content of beer. *Anal Chim Acta* **635**:45–52 (2009).
- Kim J, Mowat A, Poole P and Kasabov N, Linear and non-linear pattern recognition models for classification of fruit from visible–near infrared spectra. *Chemom Intell Lab Syst* **51**:201–216 (2000).
- Araújo M, Saldanha T, Galvão R, Yoneyama T, Chame H and Visani V, The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemom Intell Lab Syst* **57**:65–73 (2001).
- Xiaobo Z, Jiewen Z, Povey M, Holmes M and Hanpin M, Variables selection methods in near infrared spectroscopy. *Anal Chim Acta* **667**:14–32 (2010).
- Rossi F, Lendasse A, François D, Wertz V and Verleysen M, Mutual information for the selection of relevant variables in spectrometric nonlinear modelling. *Chemom Intell Lab Syst* **80**:215–226 (2006).
- Pent H, Long F and Ding C, Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *Trans Pattern Anal Mach Intell* **27**:1226–1238 (2005).
- De Ketelaere B, Howarth S, Crezee L, Lammertyn J, Viaene K, Bulens I, et al., Postharvest firmness changes as measured by acoustic and low-mass impact devices: A comparison of techniques. *Postharvest Biol Technol* **41**:275–284 (2006).
- Zude-Sasse M, Truppel I and Herold B, An approach to non-destructive apple fruit chlorophyll determination. *Postharvest Biol Technol* **25**:123–133 (2002).
- Valente M, Leardi R, Self G, Luciano G and Pain JP, Multivariate calibration of mango firmness using vis/NIR spectroscopy and acoustic impulse method. *J Food Eng* **94**:7–13 (2009).
- Suykens JAK, Van Gestel T, De Brabanter J, De Moor J and Vanderwalle B, *Least Squares Support Vector Machines*. World Scientific, Singapore (2002).
- Nicolai BM, Theron KI, Lammertyn J and Kernel PLS, Regression on wavelet transformed NIR spectra for prediction of sugar content of apple. *Chemom Intell Lab Syst* **85**:243–252 (2006).
- Lu R, Predicting firmness and sugar content of sweet cherries using near infrared diffuse reflectance spectroscopy. *Trans ASAE* **44**:1265–1271 (2001).
- Pérez-Marín D, Sánchez MT, Cano G and Garrido A, Prediction of texture in green asparagus by near infrared spectroscopy (NIRS). *J Food Qual* **25**:277–287 (2002).
- Martínez-Vega M, Sharifzadeh S, Wulfsohn D, Skov T, Clemmensen L and Toldam-Andersen T, A sampling approach for predicting the eating quality of apples using visible–near infrared spectroscopy. *J Sci Food Agric* **93**:3710–3719 (2013).
- Nicolai BM, Beullens K, Bobelyn E, Peirs A, Saey W, Theron K et al., Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol Technol* **46**:99–118 (2007).
- Suykens JAK and Vanderwalle J, Least squares support vector machine classifiers. *Neural Process Lett* **9**:293–300 (1999).
- Sun T, Lin H, Xu H and Ying Y, Effect of fruit moving speed on predicting soluble solids content of 'Cuiguan' pears (*Pomaceae pyrifolia* Nakai cv. Cuiguan) using PLS and LS-SVM regression. *Postharvest Biol Technol* **51**:86–90 (2009).
- Chauchard F, Cogdill R, Roussel S, Roger JM and Bellon-Maurel V, Application of LS-LVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemom Intell Lab Syst* **71**:141–150 (2004).
- Williams PC, Implementation of near-infrared technology, in *Near-Infrared Technology in the Agricultural and Food Industries*, ed. by Williams PC and Norris KH. American Association of Cereal Chemists, St Paul, MN, pp.145–169 (2001).
- Cover TM and Thomas JA, *Elementary Information Theory*. Wiley, New York (1991).
- Estévez PA, Tesmer A, Perez CA and Zurada JM, Normalized mutual information feature selection. *IEEE Trans Neural Networks* **20**:189–201 (2009).
- Guyon I and Elisseeff A, An introduction to variable and feature selection. *J Mach Learn Res* **3**:1157–1182 (2003).
- Koller D and Sahami M, Toward optimal feature selection, in *Proceedings of 13th International Conference on Machine Learning*, Bari, Italy, 3–6 July 1996, pp. 284–292 (1996).
- Herrera LJ, Pomares H, Rojas I, Verleysen M and Guillén A, Effective input variable selection for function approximation. Lecture notes in computer science. *International Conference on Artificial Neural Networks*, **4131**:41–50 (2006).
- Kraskov A, Stgbauer H and Grassberger P, Estimating mutual information. *Phys Rev E* **69**:066138 (2004).
- Golding JB, Spohr L, et al., Nondestructive assessment of peach and nectarine firmness. *Acta Hort* **687**:261–270 (2005).
- Herrero A, Diezma B, Lleó L, Valero C and Ruiz-Altisent M, Evaluación de técnicas acústicas para la determinación de firmeza en melocotón, in *V Congreso Nacional y II Congreso Ibérico Agroingeniería (2009)*, Lugo, España, 28–30 September (2009).
- Fu X, Ying Y, Zhou Y, Xie L and Xu H, Application of NIR spectroscopy for firmness evaluation of peaches. *J Zhejiang* **9**:552–557 (2008).
- Costa M, Noferini M, Fiori G, Miserocchi O and Bregoli AM, NIRS evaluation of peach and nectarine fruit quality in pre- and post- harvest

- conditions, in Proceedings of the 5th International Symposium on Peaches, ed. by Johnson RS and Chrisosto CH. *Acta Hort* **592**:593–599 (2002).
- 45 Qing Z, Ji B and Zude M, Wavelength selection for predicting physicochemical properties of apple fruit based on near-infrared spectroscopy. *J Food Qual* **30**:511–526 (2007).
- 46 Liu F, He Y and Wang L, Determination of effective wavelengths for discrimination of fruit vinegars using near infrared spectroscopy and multivariate analysis. *Anal Chim Acta* **615**:10–17 (2008).
- 47 Moons E, Dardenne P, Dubois A and Sindic M, Non-destructive visible and NIR spectroscopy measurement for the determination of apple internal quality. *Acta Hort* **517**:441–448 (1997).
- 48 Tu K, De Busscher R, De Baerdemaeker J and Schrevels E, Using a laser beam as the light source to study tomato and apple quality non-destructively, in *Proceedings of the Food Processing Automation IV Conference*, Chicago, 3–5 November 1995, pp. 3–5 and 528–536 (1995).

ANEXO III

'Determination of SSC and Firmness in *Prunus avium* cv *Chelan* by Vis/NIR spectroscopy using linear and nonlinear regression methods'

Victoria Lafuente, Luis Javier Herrera, Alberto Guillén, Razvan Guinea, Jesús Val, A. Ignacio Negueruela

DETERMINATION OF SSC AND FIRMNESS IN *Prunus Avium* cv Chelan BY Vis/NIR EQUIPMENT USING LINEAR AND NON-LINEAR REGRESSION METHODS

Lafuente^{1,2}, V., Herrera³, L.J., Guillén³, A., Guinea⁴, R., Val², J., Negueruela¹, A. I.

¹Departamento de Física Aplicada, Universidad de Zaragoza, Spain

²Estación Experimental de Aula Dei, CSIC, Departamento de Nutrición Vegetal, Zaragoza, Spain

³Departamento de Arquitectura y Tecnología de los Computadores, Universidad de Granada, Granada, Spain

⁴Departamento de Óptica, Universidad de Granada, Granada, Spain

INTRODUCTION

The cherry tree (*Prunus avium*), one of the leading species of stone fruit worldwide and in Spain alone production reaches 96,946 tons. Cherries are not climacteric fruits and thus do not ripen once picked from the tree, so they should be harvested after displaying the desired physico-chemical and organoleptic characteristics. This leaves about 7 days for sale and consumption. As the cherry ripens on the tree, the content of soluble solids increases, while the acidity and firmness decrease. The indexes most commonly used to determine the time of cherry picking are the content in soluble solids, fruit firmness, and skin colour. The increase in the soluble-solid content, due to starch degradation, has traditionally been determined destructively through refractometry. The perception of firmness through manual compression remains the method most used by producers and consumers (Mitchman et al., 1998). Colour is considered the external quality factor that most influences the consumer (Francis, 1995), changing from green to bright red, and then to reddish-purple colours (depending on the variety). In the colour change undergone by the cherry during ripening, anthocyanin and carotenoids are accumulated as chlorophyll degrades (Webster and Looney, 1996). To specify the colour, some researchers have used colour charts to determine the optimal moment of harvest (Timm et al., 1993) while others have done so using the CIELAB coordinates (Gao et al., 1995).

The fruit and vegetable industry requires effective non-destructive quality-control systems, increasingly demanded by the consumer. For the study of the fruit maturity, optical techniques have been developed since the 1990s, such as near-infrared spectroscopy (NIRS), which offers a number of advantages over previous techniques: it avoids the destruction of the fruit, provides faster *in situ* measurement, reduces cost, and determines several quality parameters in a single measurement. This technique consists of a light beamed on the sample. One part undergoes specular reflection and is responsible for the shine; another part, after slightly penetrating the fruit, is absorbed selectively by the pigments; and the rest of the light is reflected by the sample in diffuse mode, producing the Vis/NIR spectra, from which the desired information is gathered. Many studies have been conducted with different fruits, using this technique: apples (Fan et al., 2009; Pissard et al., 2013), pears (Li et al., 2013), tomatoes (Tiwari et al., 2013), avocados (Clark et al., 2003), oranges (Cayuela, 2008), and cherries (Lu, 2001).

The current trend is to look for non-destructive maturity control methods, comparing the Vis-NIR spectral information of the samples with the reference values gained destructively. In this way, statistical models are created to predict the desired quality parameter. These models are constructed using linear multivariate calibration methods, such as multiple linear regression (MLR, Jha et al., 2014); principal-components regression (PCR, Hernández-Gomez et al., 2006), partial least squares (PLS, Liu et al., 2005). In many cases, this relationship is not strictly linear, so that non-linear methods must be used: least-squares support vector machines (LS-SVM, Chauchard et al., 2004) and artificial neural networks (ANN) (Shao et al., 2008; Pérez-Marín et al., 2007).

Normally, multivariate regression methods make use of all the variables of the spectrum to build the calibration models. These methods have the partial capacity to reduce the impact of problems such as

collinearity, band overlaps, and interactions. However, the influence of variables that are collinear or that do not contain relevant information may impair the construction of effective models (Xiaobo et al., 2010). Some variables provide useful information whereas others do not. Consequently, the prior selection of a small number of variables is necessary to achieve a better and simpler calibration. Many variable-selection methods are used in Vis/NIR spectroscopy: successive projections algorithm (SPA) (Costa-Pereira et al., 2008); genetic algorithm (GA) (Wu et al., 2014); uninformative variables elimination (EUV) (Wu et al., 2012); stability competitive adaptive reweighted sampling (SCARS) (Zhang et al., 2013); Monte Carlo uninformative variables elimination (MC-UVE) (Xu et al., 2014); mutual information (MI) (Rossi et al., 2006).

The aim of this work is to investigate the relationship between the parameters of quality, firmness, and soluble solids content, with the Vis/NIR reflection spectra in *Prunus avium* 'Chelan'. The calibration models were built by the regression methods of PLS and LS-SVM. In addition, a selection of variables is proposed in order to detect the most influential wavelengths in the physico-chemical parameters, and to obtain more stable and simple calibrations.

MATERIAL AND METHODS

Samples

A total of 282 samples of *Prunus avium* 'Chelan' were collected during the 2011 and 2012 seasons, cultivated on a commercial farm owned by Orchard, S.L. It is located at the boundary of La Almunia de Doña Godina (Zaragoza, Spain). The trees were planted in 2000, on a planting grid of 5x3m. Drip irrigation was applied, at a rate of 25,000L/ha/h. The samples were collected weekly during the harvest in the first season, and every other day for two weeks of harvest in the second season (47 samples per collection day). These samples were immediately transported to the laboratory, where the Vis/NIR spectra were measured at two opposite points on the fruit. Then, the firmness and soluble-solids content were determined at the same points on the fruit.

Vis/NIR spectroscopy

The equipment used was QualitySpec Pro 2600 modular reflectance equipment (Analytical Spectral Devices, INC. Colorado, USA). Fast-scanning spectrophotometer, with a measuring interval of 350-2500 nm ($\Delta\lambda=1$ nm). Tungsten halogen lamp (12V/45W) to provide light. The detection system consists of a monochromator, with a double InGaAs detector. The scanning speed is 10 scans/second. The light energy is collected through a bundle of specially formulated optical fibres. The fibre-optic cable has a conical view subtending a full angle of approximately 25 degrees.

The equipment measures the reflection spectra in units of optical density ($D.O=\log I/Reflectance$). Two spectral measurements were made per fruit at two opposing fixed positions at the equator of the fruit, using the mean of the two spectra for the calibration processes. The initial and final parts of the spectra were removed to eliminate useless information. The working range was 400-2420 nm for the QualitySpec 2600 instrument.

Firmness

The firmness was determined using a AGROSTA 100Field non-destructive digital durometer (Agro Technology, Forges Les Eaux, France). The equipment does not penetrate the fruit. Its measurement is based on quantifying the force necessary to retract a plunger a certain length to manually

apply it to the surface of the fruit, giving an idea of its firmness through its deformation. The results were expressed in DI_{10} (Durofel Index, where the durofel device was fitted with a 0.10 cm^2 probe). The measurements were made on each face of the cherry on which the Vis/NIR spectra were measured.

Soluble Solids

The total soluble solids were analysed using a digital refractometer (ATAGO PR-101 Co. Model, Tokyo, Japan). The measurement was performed on a drop of cherry juice, running the analysis twice.

Statistical treatment

The calibration equations were formulated using two methods of multivariate analysis: PLS and LS-SVM. The software used was Matlab R2009a (The MathWorks, Natick, United States), with its own programming.

Different pretreatments were tested on the study spectra before establishing the calibration models: standard normal variate (SNV), multiplicative scatter correction (MSC), derivative in first and second order, and normalization. Finally, normalization was applied to each of the variables of the spectrum, using mean zero and standard deviation one. This type of normalization is applied individually to each variable considered, whether it be spectral or otherwise (Rossi et al., 2006; Suykens et al., 2002).

The selection of calibration and validation samples of the entire set is a critical point in the construction of the calibration model. We conducted three distinct subdivisions for all the samples. For calibration, each subdivision was made up of 2/3 of the samples, and the remaining third was earmarked for validation samples (Flores et al., 2009; Paz et al., 2009).

Regression methods

Partial least-squares analysis (PLS)

This regression method developed by Wold (1982) is probably the most widely used today for treating multivariate data in chemometrics (Martens and Naes, 1989). This method is used to build predictive models of variables (y) when there are many factors (x), and when these are highly collinear. The X matrix is substituted by a matrix of latent variables (LVs), which are themselves linear combinations of the x vectors that maximize the covariance of the y matrix, using least squares to adjust both the latent variables as well as the regression coefficients, with a high (r^2) determination coefficient. The main characteristic of this composition, in our work, is to seek the maximum correlation between the spectra (x variables) and the characteristic to be determined (y variables).

This technique has been used in a large number of studies related to the forecasting of quality indexes in fruits and vegetables (Nicolai et al., 2006; Sánchez et al., 2012; Maniwaru et al., 2014; Martínez-Valdivieso et al., 2014).

L-fold cross-validation ($L = 20$) was used to calculate the optimum number of latent variables, and to avoid overfitting in the development of the calibration models (Xiaobo et al., 2007; Zhang et al., 2013).

Least-squares support-vector machine (LS-SVM)

This methodology started to be developed in 1990s. LS-SVM has the capability of dealing with linear and nonlinear multivariate calibration and solving these problems in a relatively fast way (Suykens et al., 1999; Vapnik et al., 1995). Previous research has demonstrated the potential of this non-linear regression model for several quantitative applications in agro-food products (Sun et al., 2009; Shao et al., 2011 and Torres-Mariani et al., 2014). In the LSSVM model development, the radial basis function (RBF) kernel was used. Grid search and cross validation (LS-SVMLab, Suykens et al., 2002) were conducted to

find the optimal values of gam (γ) and sig² (δ) parameters of the model; gam is used to maximize model performance and minimize model complexity while sig² is the width of the RBF kernel.

Performance metrics

Afterwards, in both methods, the calibration models were tested to predict SSC and firmness with a validation set. The best calibration models were selected based on the highest coefficient of determination for cross validation (R_{cv}^2) together with the lowest standard error of cross validation (SEVC) (Williams, 2001).

Another statistic used was residual predictive deviation (RPD), calculated as the ratio of the standard deviation (SD) of the reference data to the SECV (Williams, 2001). This latter statistic enables SECV to be standardized, facilitating the comparison of results found with sets of different means (Williams, 2001). Williams points out RPD values between 3 and 5 indicate Vis/NIR prediction efficiency.

With respect to the validation, the effect of the different settings on the performance of the model was evaluated by comparing the root mean square error of prediction (RMSEP), the coefficient of external validation (r^2), bias, and the standard error of prediction (SEP).

Variable selection

Mutual Information Backwards

In this work the backwards variable-selection criterion was based on mutual information (MI) (Herrera et al., 2006; Lafuente et al., 2014; Herrera et al., 2015). MI is a non-linear correlation measurement from Information Theory (Cover et al., 1991). For two sets of continuous variables, X and Y, it can be calculated by:

$$I(X, Y) = \iint \mu_{X,Y}(x, y) \log \frac{\mu_{X,Y}(x, y)}{\mu_X(x)\mu_Y(y)} dx dy$$

where $\mu_{X,Y}(x, y)$ is the joint probability density function (PDF) of X and Y, and $\mu_X(x)$ is the marginal density function of the set of variables X. Its advantage in comparison with other correlation criteria is that it is able to identify non-linear relations among the variables involved. The MI estimator used in this work is the k nearest-neighbours one (Kraskov et al., 2004)

The algorithm is based on the *Markov blanket* concept, which is very helpful in identifying collinearity among the input variables. This algorithm has also proved to perform better in feature selection processes for spectrometric data in chemical problems in comparison with similar algorithms eliminating collinear variables (Herrera et al., 2015).

Thus, given a set of input variables X and an output variable Y, a set of variables M_i in X is said to be a *Markov blanket* for a variable x_i in X with respect to Y, if $I(\{M_i \cup x_i\}, Y) = I(M_i, Y)$, provided that M_i itself has all the information that x_i has about Y. A Markov blanket is thus a group of variables that subsumes the mutual information content in a certain variable—that is, if a variable presents a Markov Blanket among the rest of available variables, it can thus be discarded, as it presents redundant information only with respect to the objective variable.

The algorithm operates as backwards variable-selection method that starts with the complete set of variables, and iteratively discards the ones detected to have a Markov Blanket in the remaining set X_G of variables.

The algorithm states the following steps:

1. Calculate the MI between each pair of input variables $I(x_i, x_j)$
2. Starting from the complete set of input variables $X_G = X$, iterate:

- a) For each variable x_i , let the candidate Markov blanket M_i be the set of p variables in X_G for which $I(x_i, x_j)$ is highest.
- b) Compute for each x_i

$$Loss_i = I(\{M_i \cup x_i\}, Y) - I(M_i, Y)$$

- c) Choose the x_i for which $Loss_i$ is lowest and eliminate x_i from X_G .

3. Continue with step 2 until no variables remain.

In this way, the relevance of variables (in reverse order) is ranked. Under this operation, it should be noted that variables having low influence with respect to the output variable (irrelevant variables) will soon be discarded, as the $Loss_i$ value should tend to 0. Similarly, redundant variables (variables with high collinearity) are iteratively discarded at earlier stages. Relevant variables with low redundancy will be the last ones to be considered in the iterative process, and thus easily identified. The p parameter of the algorithm (in step 2.a of the algorithm) will take the value $p = 1$, as recommended (Herrera et al., 2006)

RESULTS AND DISCUSSION

Table1. Samples number (n), range, mean, standard deviation (SD) and coefficient variation (%CV) for the two parameters studied in the calibration set (n=211) and in the validation set (n=70)

Parameter	Group	Set	n	Range	Mean	SD	%CV
Durofel	1	Calibration	211	18-58	41.1	10.0	24.4
		Validation	70	19-56	42.4	9.8	23.3
	2	Calibration	211	18-58	42.1	9.9	23.6
		Validation	70	21-55	41.3	10.1	24.6
	3	Calibration	211	21-58	41.2	9.9	24.1
		Validation	70	18-57	42.4	10.1	24
SSC (°Brix)	1	Calibration	211	9.5-30	17.7	4.8	27.1
		Validation	70	8.7-27.3	16.2	4.3	26.5
	2	Calibration	211	8.7-30	17.2	4.7	27.3
		Validation	70	9.8-28.1	17.5	4.7	26.9
	3	Calibration	211	8-7-28.1	17.4	4.7	27.0
		Validation	70	9.5-30	17.1	4.6	26.9

Table 1 shows mean values, standard deviation, range, and variation coefficient for the three subdivisions of samples used in calibration and validation. The range for the validation set was within the calibration set. This indicates that the selection of the groups was properly performed.

Figure 1 shows examples of cherry-absorption spectra found using the Vis/NIR equipment. In the visible region, the absorption peak of the chlorophyll (≈ 675 nm) is plotted, which decreases when the cherry ripens. In the infrared range (700-2000), several absorption bands appear: 980, 1470, and 1940 nm, related to the high water content of the fruit (Osborne et al., 1993; Lammertyn et al., 2000; Nicolai et al., 2007). Also, a band around 1200 nm appears, characteristic of the absorption of carbohydrates (Osborne et al., 1993). Spectra with similar bands for cherries were reported by Lu (2001).

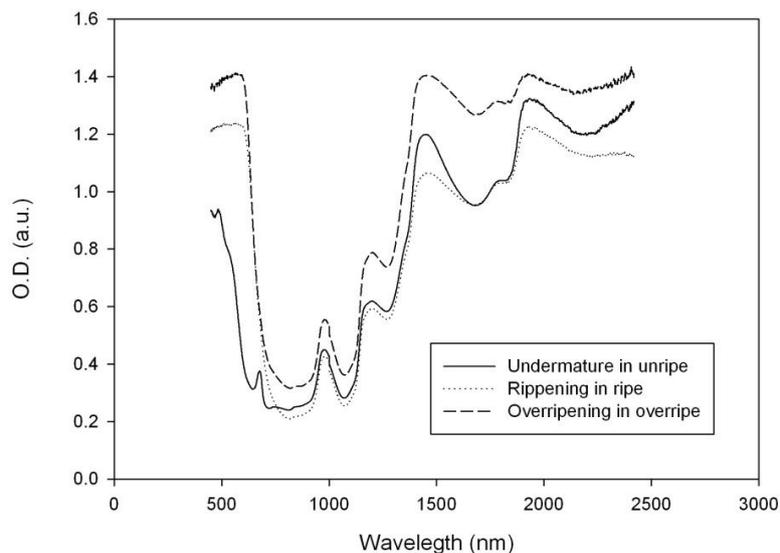


Figure 1. Reflected spectra of *Prunus avium* 'Chelan' with QualitySpec equipment (450-2500 nm).

Calibration Models

Tables 2 and 3 show calibration statistics for the three groups of samples selected to determine the firmness and SSC, respectively, using the ranges from 450-700 nm to determine firmness, and two intervals for SSC, i.e. 700-2420 and 700-1060 nm.

Regarding the method of regression, in the case of firmness, similar results were found with both methods, PLS and LS-SVM, respectively ($R_{cv}^2 = 0.82$ vs. 0.76 , $SECV = 4.78$ vs. 4.73). The RPD values found for the two regression methods are between 2.1-2.2. In SSC, very similar statistics resulted for both regression methods, LS-SVM and PLS, respectively ($R_{cv}^2 = 0.96$ vs. 0.97 , $SECV = 0.82$ and 0.83 °Brix). As for the interval of wavelengths used, better results were found for the interval 700-1060 nm ($R_{cv}^2 = 0.96$, $SECV = 0.83$) than for 700-2420 nm ($R_{cv}^2 = 0.92$, $SECV = 1.25$). The RPD values proved much higher for the range of 700-1060 nm, being between 5.5-5.8. These results are consistent with those of other studies. Firmness was related to the band of the visible spectrum, and the SSC were determined strictly in the infrared band close the spectrum (Carlini et al., 2000; McGlone et al., 2002; Martínez Vega et al., 2013).

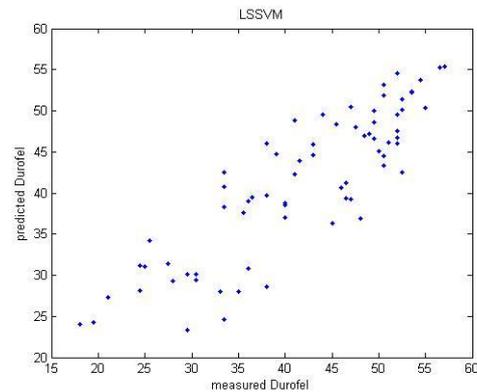
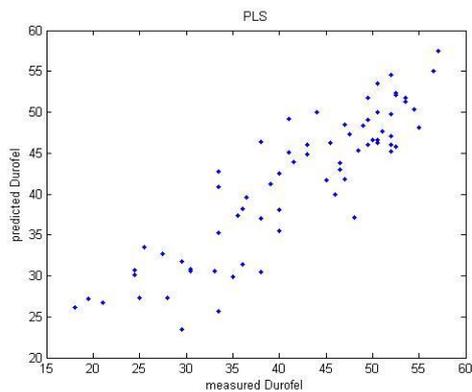
Table 2. Calibration statistics for three calibration groups, using PLS and LSSVM regression methods for firmness in 450-750 nm range and, for SST, the 700-2420 nm and 700-1060 nm range.

		PLS			LS-SVM		
		Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
Firmness (450-700 nm)	N	211	211	211	211	211	211
	LVs	10	10	10			
	SECV	4.58	4.78	4.71	4.51	4.73	4.69
	R_{cv}²	0.83	0.82	0.82	0.76	0.76	0.74
	RPD	2.18	2.07	2.10	2.22	2.09	2.11
SSC (700-2420 nm)	N	211	211	211	211	211	211
	LVs	17	13	15			
	SECV	1.32	1.28	1.3	1.25	1.22	1.21
	R_{cv}²	0.98	0.96	0.97	0.92	0.92	0.92
	RPD	3.64	3.67	3.61	3.84	3.85	3.88
SSC (700-1060 nm)	N	211	211	211	211	211	211
	LVs	11	11	11			
	SECV	0.82	0.83	0.83	0.83	0.86	0.82
	R_{cv}²	0.97	0.97	0.98	0.96	0.96	0.96
	RPD	5.85	5.66	5.66	5.78	5.46	5.73

Validation

Validation is necessary to ensure an independent measurement of the precision for the calibration models through RMSEP and r_p^2 . Figure 2 shows the prediction statistics for one of the prediction groups in the intervals of wavelengths in which the best results have been achieved in the calibration equations, for firmness (450-700 nm) and soluble solids (700-1060 nm).

The SECV and RMSEP statistics were similar to those found in calibration, indicating that the use of these calibration models is viable for predicting firmness and SSC. For the two parameters, the prediction statistics were similar for both regression methods, though somewhat better for the linear method (PLS) for the total soluble solids content (PLS, RMSEP=4.61°Brix, r_p^2 =0.79; LSSVM, RMSEP=5.03°Brix, r_p^2 =0.75).



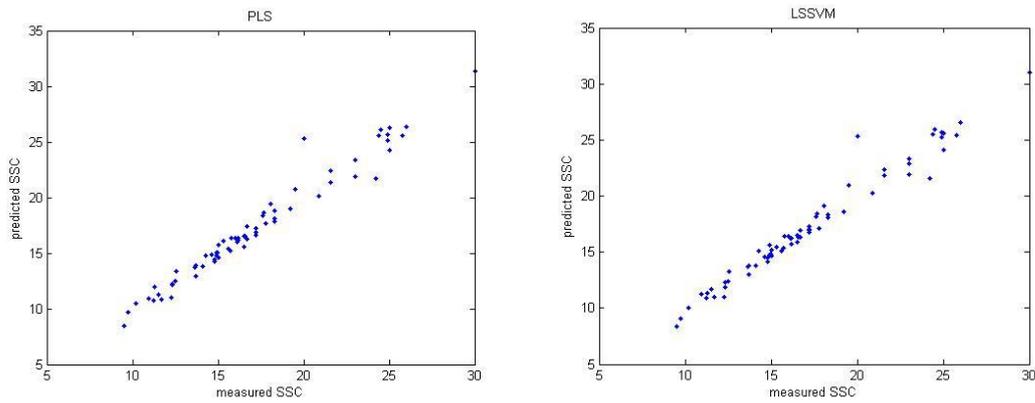


Figure 2. Prediction statistics for firmness and SSC using PLS y LS-SVM regression methods

VARIABLE SELECTION

Figure 3 shows for both parameters the yields of the models found with respect to the number of variables used. In both cases, the yield of the model stabilized through the use of approximately 40 wavelengths. The R_{cv}^2 values were 0.74 and 0.97 for firmness and SSC, respectively, i.e. similar to those found with all the variables of the work intervals. Figure 4 shows the wavelengths that correlate best with firmness and soluble-solids content. In the case of firmness, the first variables selected correspond to the absorption band of the chlorophyll. This again confirms that firmness is related to chlorophyll content (Lafuente et al., 2014; Zude et al., 2011; Zude et al., 2002). For SSC, the prominent wavelengths correspond to the spectral region between 700-820 nm, which is one of the areas presenting the least variation in the absorbance spectrum (Qing et al., 2007). Variables also appear (980 nm) related to the water content of the cherry (Osborne et al., 1993; Lammertyn et al., 2000; Nicolai et al., 2007).

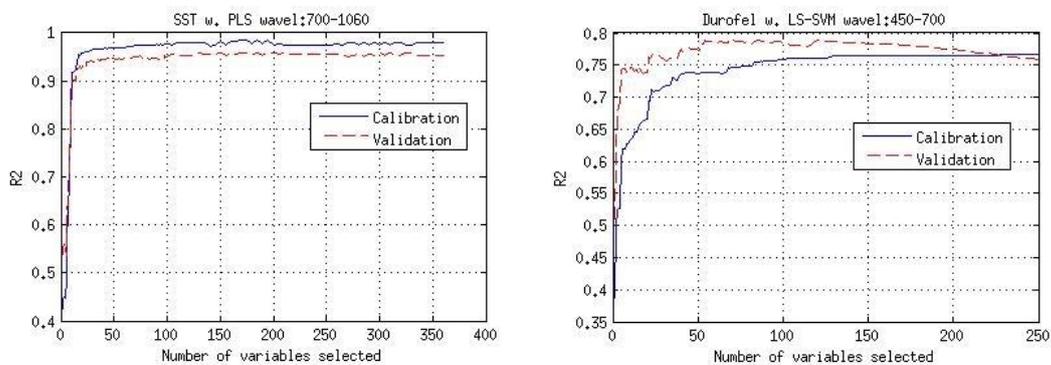


Figure 3. Example of variable selection for SSC and firmness

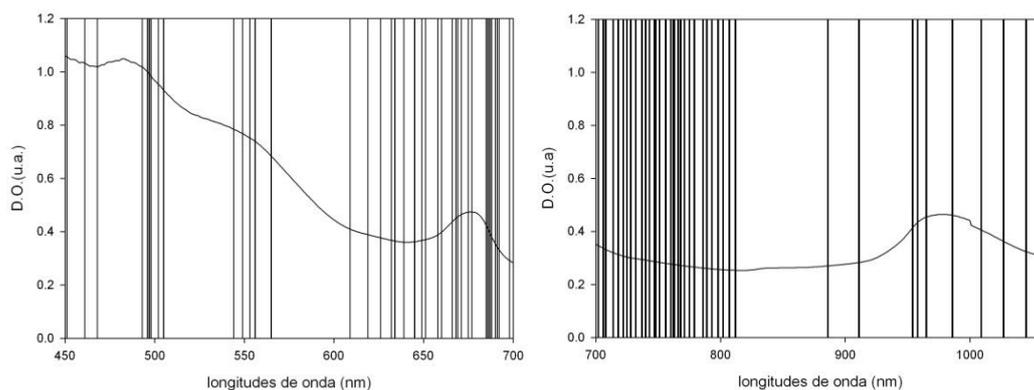


Figure 4. Wavelength selection to determine firmness (a) and SSC (b)

CONCLUSIONS

In this paper, predictive models to determine SSC and firmness in *Prunus Avium* cv. 'Chelan' via Vis/NIR spectroscopy are developed using linear (PLS) and non-linear (LS-SVM) regression methods. The development of this methodology is of great interest to the fruit sector in the area, since this is a major export fruit.

Initially, the spectral-range selection for determining firmness (450-700 nm) and SSC (700-1060 nm) supported the reliability of the prediction. With this, the r_p^2 and RMSEP values improved with respect to the use of the complete spectrum.

The use of non-linear regression methods (LS-SVM) did not improve the statistics of the calibration models for firmness (PLS, $r_p^2=0.79$, RMSEP=4.69; LS-SVM, $r_p^2=0.75$, RMSEP=5.03). In the case of SSC, the results were also similar for both regression methods (PLS, $r_p^2=0.97$, RMSEP=0.83°Brix; LS-SVM, $r_p^2=0.97$, RMSEP=0.86 °Brix).

Finally, the selection of variables by the MI method reduced the number of wavelengths, from a measurement of 350 wavelengths to approximately 40, giving results similar to those corresponding to the entire interval. This enables the development of much simpler equipment for determining the maturity index of the cherry.

Further work is needed on the development of the NIRS methodology and on new calibration equations for other varieties of cherry and other species.

REFERENCES

- Carlini, P., Massantini, R., Mencarelli, F. 2000. Vis-NIR measurement of soluble solids in cherry and apricot by PLS regression and wavelength selection. *Journal Agricultural Food Chemistry*, 48, pp. 5236-5242.
- Camps, C., Christen, D. 2009. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. *LWT- Food Science and Technology* 42, 1125-1131.
- Cayuela, J.A. 2008. Vis/NIR soluble solids prediction in intact oranges (*Citrus sinensis* L.). *Postharvest Biology and Technology*. 47, 75-80
- Cayuela, J. A. Weiland, C. 2010. Intact orange quality prediction with two portable NIR spectrometers. *Postharvest Biology and Technology*. 58, 113-120
- Cho, R. K., Lee, K. H., Iwamoto, M. 1992. Determination of firmness of intact apple fruits by near infrared reflectance spectroscopy. In: Taylor, R. A. (Ed.), *Proc 6th Int. Diffuse Reflect. Spectr. Conf. Counc. Dor Near IR Spectr.*, Chambersburg, PA., pp. 71-76.
- Clark, C.J., McGlone, V.A., Requejo, C., White, A. and Woolf, A.B. 2003. Dry matter determination in Hass avocado by NIR spectroscopy. *Postharvest Biology and Technology* 29. Pp: 300-307.
- Costa-Pereira, A. F., Coelho-Pontes, M. J., Gambarra-Neto, F. F., Bezerra-Santos, S. R., Harrop-Galvao, R. K., Ugulino-Araújo, M. C. 2008. NIR spectrometric determination of quality parameters in vegetable oils using iPLS and variable selection. *Food Research International*, 41: 341-348.
- Cover, T.M., Thomas, J.A. 1991. *Elementary Information Theory*. Wiley, New York.
- Gao, L., Mazza, G., 1995. Characterization, quantization and distribution of anthocyanins and colorless phenolics in sweet cherries. *Journal of Agricultural and Food Chemistry* 43, 343-346.
- Fan, G., Zha, J., Du, R. and Gao, L. 2009. Determination of soluble solids and firmness of apples by Vis/NIR transmittance. *Journal of Food Engineering*, 93, 416-420
- Flores, K., Sánchez, M.T., Pérez-Marín, D., Guerrero, J.E., Garrido-Varo, A. 2009. Feasibility in NIRS instruments for predicting internal quality in intact tomato. *Journal of Food Engineering*. 91, 311-318.
- Francis, F.J. 1995. Quality as influenced by colour. *Food Quality Prefer.* 6, 149-155
- Hernández-Gómez, A., He, Y., García-Pereira, A. 2006. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis-NIR spectroscopy techniques. *Journal of Food Engineering*, 77, 313-319.
- Herrera, L. J., Pomares, H., Rojas, I., Verleysen, M., Guillén, A. 2006. Effective input variable selection for function approximation. *Springer Verlag Part 1 LNCS* 4131. Pp. 41-60.
- Herrera, L.J., Lafuente, V., Ghinea, R., Pérez, M.M., Negueruela, I., Pomares, H., Rojas, I., Guillén, A. 2015. Mutual information-based feature selection in spectrometric data for agricultura applications. *Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol I, IMECS 2015, March 18-20, 2015, Honk Kong.*
- Jha, S. N., Narsaiah, K., Jaiswal, R., Bhardwaj, M., Gupta, R., Kumar, R., Sharma, R. 2014. Nondestructive prediction of maturity of mango using near infrared spectroscopy. *Journal of Food Engineering*. 124. 152-157
- Kraskov, A., Stgbauer, H., Grassberger P. 2004. Estimating mutual information. *Phys Rev E*. 69: 066138

- Lafuente, V., Herrera, L.J., Pérez, M.M., Val, J., Negueruela, I. 2014. Firmness prediction in *Prunus Persica* 'Calrico' peaches by visible/short wave near infrared spectroscopy and acoustic measurements using optimised linear and non-linear chemometric models. *J Sci Food Agric*
- Lammertyn, J., Peirs, A., De Baerdemaeker, J., Nicolai, B. 2000. Light penetration properties of NIR radiation in fruit respect to non-destructive quality assessment. *Postharvest Biology and Technology*, 1, 121-132.
- Li, J., Huang, W., Zhao, C and Zhang, B. 2013. A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy. *Journal of Food Engineering*. Volume 116, Issue 2, pp. 324-332.
- Liu, Y and Ying, Y. 2005. Use of FT-NIR spectrometry in non-invasive measurements of internal quality of 'Fuji' apples. *Postharvest Biology and Technology*, 37, 65-71.
- Liu, F., Jiang, Y. and He, Y. 2009. Variable selection in visible/near infrared spectra for linear and nonlinear calibrations: A case study to determinate soluble solids content of beer. *Analytica Chimica Acta* 635, 45-52
- Lu, R. 2001. Predicting firmness and sugar content of sweet cherries using near-infrared diffuse reflectance spectroscopy. *Transactions of the ASAE. American society of agricultural engineers* ISSN 0001-2351. Vol.44, pp. 1265-1271.
- McGlone, V.A. and Kawano, S. 1998. Firmness, dry-matter and soluble solids assessment of postharvest kiwifruit. *Postharvest Biology and Technology* 13:131-141
- McGlone, V. A., Jordan, R. B., Martinsen, P. J. 2002. Vis/NIR estimation at harvest of pre- and post-storage quality indices for Royal Gala apple. *Postharvest Biology and Technology* 25 135-144.
- Maniwaru, P., Nakano, K., Boonyakiat, D., Ohashi, D., Hiroi, M., Tohyama, T. 2014. The use of visible and near infrared spectroscopy for evaluating passion fruit postharvest quality. *Journal of Food Engineering*: 143. 33-43
- Martens, H and Naes, T. 1989. *Multivariate Calibration*. John Wiley & Sons Inc., Chichester, UK
- Martínez-Valdivieso, D., Font, R., Blanco-Díaz, M.T., Moreno-Rojas, J. M., Gómez, P. Alonso-Moraga, A., Del Río Celestino, M. 2014. Application of near-infrared reflectance spectroscopy for predicting carotenoid content in summer squash fruit. *Computers and Electronics in Agriculture*. 108, 71-79.
- Mitcham, E.J., Clayton, M. and Biasi, W. V. 1998. Comparison of devices for measuring cherry fruit firmness. *Hort. Sci.*, 33: 723-727.
- Moons, E., Dardenne P., Dubois, A., Sindic, M. 1997. Non-destructive visible and NIR spectroscopy measurement for the determination of apple internal quality. *Acta Horticulturae* 517: 441-448.
- Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K., Lammertyn, J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*, 46, pp. 99-118.
- Osborne, B. G., Feam, T., Hindle, P. H. 1993. *Practical NIR spectroscopy with applications*. In: Osborne, B. G., Feam, T., Hindle, P. H. (Eds.) *Food and Beverage Analysis*. Longman, Essex, pp. 11-35.
- Paz, P., Sánchez, M.T., Pérez-Marín, D., Guerrero, J.E., Garrido-Varo, A. 2009. Evaluating NIR instruments for quantitative and qualitative assesment of intact Apple quality. *Journal Science Food and Agriculture* 89 781-790

- Pérez-Marín, D., Garrido-Varo, A., Guerrero, J.E. 2007. Non-linear regression methods in NIRS quantitative analysis. *Talanta* 72: 28-42.
- Pissard, A., Fernández-Pierna, J.A., Baeten, V., Sinnaeve, G., Lognay, G., Mouteau, A., Dupont, P., Rondia, A., Lateur, M. 2013. Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near infrared spectroscopy. *J. Sci Food Agric*, 93: 238-244.
- Qing, Z., Ji, B., Zude, M. 2007. Wavelength selection for predicting physicochemical properties of apple fruit based on near-infrared spectroscopy. *Journal of Food Quality* 30: 511-526.
- Roger, J. M and Bellon-Maurel, V. 2000. Using genetic algorithms to select wavelengths in near-infrared spectra: application to sugar content prediction in cherries. *Applied Spectroscopy*. Volume 54, 9.
- Rossi, F., Lendasse, A., François, D., Wertz, V., Verleysen, M. 2006. Mutual Information for the selection of relevant variables in spectrometric nonlinear modelling. *Chemometrics Intelligent Laboratory Systems*. 80: 215-226
- Sánchez, M.T., De la Haba, M.J., Benítez-López, M., Fernández-Novales, J., Garrido-Varo, A., Pérez-Marín, D. 2012. Non-destructive characterization and quality control of intact strawberries based on NIR spectral data. *Journal of Food Engineering* 110: 102-108
- Shao, Y., He, Y., Bao, Y and Mao, J. 2008. Near- Infrared spectroscopy for classification of oranges and prediction of the sugar content. *International Journal of Food Properties*, 12, 644-658.
- Sun T., Lin H., Xu H., Ying Y. 2009. Effect of fruit moving speed on predicting soluble solids content 'Cuiguan' pears (*Pomaceae pyrifolia* Nakai cv. Cuiguan) using PLS and LS-SVM regression. *Postharvest Biology and Technology* 51 pp. 86-90
- Suykens J.A.K, Vandewalle J. *Neural Process. Lett.* 9.1999. 293.
- Suykens J.A.K, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. *Least Squares Support Vector Machines*, World Scientific, Singapore, 2002.
- Timm, E. J., G. K. Brown, P. R. Armstrong, and R. M. Beaudry. 1993. A portable instrument for measuring firmness of cherries and berries. *ASAE Paper No. 93-6539*. St. Joseph, Mich.: ASAE.
- Tiwari, G., Slaughter, D. C., Cantwell, M. 2013. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biology and Technology*. 86: 221-229.
- Torres-Mariani, N., Camara da Costa, R., Gomes de Lima, K. M., Nardini, V., Cunha Junior, L. C., Almeida Texeira, G. H. 2014. Predicting soluble solid content in intact jacobicaba [*Myrciria jacobicaba* (Vell.) O. Berg] fruit using near infrared spectroscopy and chemometrics. *Food Chemistry*. 159. 458-462.
- Travers, S., Bertelsen, B., Kucheryavskiy, S. 2013. Predicting apple (cv. Elshof) postharvest dry matter and soluble solids content with near infrared spectroscopy. *Journal of the Science Food and Agriculture*. 94 (5), 955-962.
- Vapnik, V. N. 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- Webster, A. D. and Looney, N.E. 1996. *The cherries: crop physiology. Production and uses*. CAB International. Wallingford, Oxon, Reino Unido.
- Wu, D., Nie, P., He, Y., Bao, Y. 2012. Determination of calcium content in powdered milk using near and mid-infrared spectroscopy with variable selection and chemometrics. *Food and Bioprocess Technology*. Volume 5, Issue 4, pp. 1402-1410.

Wu, Z., Wu, E., Wang, F., Long, J., Xu, X., Jiao, A., Jin, Z. 2014. Rapid determination of process variables of Chinese rice wine using FT-NIR spectroscopy and efficient wavelengths selection methods. *Food Analytical Methods* DOI 10.1007/s12161-014-0021-6.

Xiaobo, Z., Jiewen, Z., Yanxiao, L. 2007. Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models. *Vibrational Spectroscopy* 44: 220-227.

Xiaobo, Z., Jiewen, Z., Povey, M., Holmes, M., Hanpin, M. 2010. Variables selection methods in near-infrared spectroscopy. *Analytica Chimica Acta*, 667: 14-32

Xu, F., Huang, X., Dai, H., Chen, W., Ding, R., Teye, E. 2014. Nondestructive determination of bamboo shoots lignifications using FT-NIR with efficient variables selection algorithms. *Analytical Methods* 6: 1090-1095.

Zhang, X., Li, W., Yin, B., Chen, W., Kelly, D. P., Wang, X., Zheng, K., Du, Y. 2013. Improvement of near infrared spectroscopy (NIRS) analysis of caffeine in roasted Arabica coffee by variable selection method of stability competitive adaptive reweighted sampling (SCARS). *Spectrochimica Acta parta A: Molecular and Biomolecular Spectroscopy*. 114: 350-356

Zude-Sasse, M., Truppel, I., Herold, B. 2002. An approach to non-destructive apple fruit chlorophyll determination. *Postharvest Biology and Technology* 25: 123-133

Zude-Sasse, M., Pflanz, M., Spinelli, Dosche, C., Torricelli, A. 2011. Non-destructive analysis of anthocyanins in cherries by means of Lambert-Beer and multivariate regression based on spectroscopy and scatter correction using time-resolved analysis. *Journal of Food Engineering* 103: 68-75

ANEXO IV

Parámetros de calidad de las distintas especies y variedades de fruta usados para elaborar la memoria de esta Tesis Doctoral.

Fruta	Nº de muestras
Manzana	365
Melocotón	260
Cereza	282

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
8,5	88,8	27	14,7	173	14,3	49,5	11,3	68,5	20,1	10,6
9,3	97,8	33	13,8	173	14,7	43	12,1	65,7	23,2	9,3
8,3	98,6	28	14,1	200	13,6	51,5	11,5	62,0	23,9	9,5
8,45	84,1	29	15	200	13,9	53	11,6	71,6	19,6	11,7
10,05	105,5	29	13,8	186	13,2	43	14,9	62,7	24,3	7,7
9	98,0	6	14,5	185	9,8	51,5	12,7	69,2	19,9	11,6
8,7	90,5	36	15	202	14,6	47	13,3	77,4	13,1	11,0
8,6	98,3	37	13,8	203	15,8	47	12,5	88,1	14,8	5,5
8,5	92,5	31	14,4	177	14,5	51,5	13,3	75,2	8,2	15,2
9	103,1	53	14,9	176	18,1	48	14,4	63,6	23,0	10,6
8,55	86,1	27	14,9	175	12,4	52	12,3	71,5	16,6	12,9
8,35	85,4	31	14,5	175	14,0	52	12,1	66,9	21,3	13,8
8	91,8	20	15,7	193	9,7	43	12,4	69,0	20,9	10,5
8,8	95,6	55	15,1	194	16,0	54	10,3	75,0	13,2	13,2
9,4	98,5	72	14,4	169	22,9	50,5	12,9	71,4	17,3	14,4
8,75	91,5	13	15	169	8,7	49,5	14,6	66,1	26,8	10,3
9,4	98,5	54	14,1	182	13,8	43	14,8	62,3	28,6	9,5
8,35	81,7	40	14,1	182	16,6	48	11,8	76,9	13,2	11,1
8,5	88,2	37	14,8	199	14,5	51	8,7	75,5	6,6	13,2
9,4	107,0	43	13,6	198	14,9	47,5	12,3	76,4	16,8	8,2
8,95	88,8	25	14,2	212	11,1	56,5	12,3	81,0	8,1	13,0
8,85	97,4	32	14,8	212	13,4	52	11,7	72,6	14,5	14,4
8,2	97,0	31	14,2	211	14,2	46	14,1	63,4	25,8	9,6

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
9	96,1	35	14,2	211	13,1	51,5	11,8	64,2	22,5	10,2
8,7	102,8	36	14,2	171	18,2	43	13,1	70,4	17,8	9,0
8,7	77,1	25	14,6	170	11,7	45,5	13,0	69,2	22,0	9,9
8,4	72,6	63	13,8	186	19,1	53	12,9	69,9	19,8	10,6
9,05	82,2	74	13,4	185	16,3	52	14,3	67,5	26,1	10,2
9,45	86,4	54	14,6	161	16,2	40	15,4	66,5	18,5	10,7
9,1	80,7	45	14,4	161	14,0	43	12,4	63,9	24,4	12,3
10	81,8	31	15	176	13,2	49	12,9	75,9	8,2	14,0
9,95	89,9	55	12	175	15,6	45,5	12,3	69,2	18,8	11,4
9,1	75,2	39	15,3	156	14,5	47,5	13,0	63,9	21,0	9,5
9,6	88,5	41	13,5	156	14,1	51	13,0	76,2	16,5	10,0
8,9	81,8	48	13,6	175	13,9	50,5	14,5	64,1	22,4	10,6
8,25	72,8	61	13,8	175	16,6	44	11,9	64,0	25,4	9,9
8,7	76,3	24	13,8	173	14,1	50,5	11,6	88,4	3,0	16,1
8,35	64,8	47	14	173	14,3	51	13,1	69,1	21,3	11,0
9,05	86,9	37	14,2	185	13,5	44	14,1	62,4	29,1	10,9
8,75	74,2	45	14,5	187	15,3	47,5	14,8	66,2	25,1	10,7
8,85	76,8	46	14,9	200	16,7	56,5	9,5	74,3	-1,5	18,7
8,65	74,7	31	13,3	200	10,3	49,5	12,4	43,1	49,3	31,0
8,55	75,4	47	14,4	220	14,8	46,5	10,9	44,5	45,3	32,6
9,05	70,7	38	14,2	221	13,8	50,5	12,4	42,5	53,1	31,8
8,3	82,2	49	13,2	200	15,9	49	13,9	41,5	49,1	30,9
9,2	85,2	43	12,8	200	16,9	52,5	11,6	49,7	42,6	35,3

Manzana cv. <i>Golden Smoothie Delicious</i>		Melocotón vs. <i>Calrico</i>				Cereza cv. <i>Chelan</i>				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
8,65	75,1	56	12,9	185	16,3	50	12,5	31,0	57,1	39,8
8,85	78,5	60	13	185	18,6	52,5	10,2	46,3	50,1	36,2
8,9	84,4	66	13,8	208	17,6	48,5	12,4	39,1	49,2	33,0
8,7	81,8	48	13,7	207	16,3	47,5	12,3	42,2	49,7	32,2
9,95	69,9	46	12,4	695	20,6	52	11,3	49,6	40,4	35,2
10,25	72,6	36	16,3	516	10,6	48	10,6	39,8	54,4	38,5
9,95	70,1	39	16,2	566	13,9	46,5	10,7	43,8	53,7	34,6
9,35	64,8	65	13,6	717	21,6	51,5	12,7	31,3	50,1	30,1
8,9	78,0	26	15,2	445	6,6	45,5	13,6	40,3	52,9	33,1
9,75	77,1	23	15,9	625	13,5	50	13,9	30,1	58,2	44,0
9,05	64,4	17	16,3	520	9,2	45,5	13,4	32,8	56,1	33,9
8,7	68,9	23	17,2	428	8,4	49,5	13,4	35,6	53,0	31,8
9,25	69,3	21	14,9	518	8,3	53	12,6	49,3	42,6	32,2
9,7	70,6	32	16,2	707	22,6	52,5	13,7	33,6	54,8	35,7
9,9	77,1	27	16	516	9,3	54	13,4	42,8	46,6	31,2
9,4	75,0	26	16,6	574	12,4	50,5	15,8	30,7	58,5	38,1
9,55	71,4	26	16,4	619	13,1	52	15,7	44,7	43,2	31,5
9,95	77,5	27	13,9	566	10,6	48	16,6	29,4	56,6	37,5
9,2	70,8	31	16,6	580	11,0	52	11,9	54,8	36,7	35,3
9,5	67,8	25	18	535	9,6	53	13,6	48,3	42,8	26,6
8,85	53,2	49	14,4	668	13,8	52	10,3	46,7	51,2	35,5
8,75	72,9	34	15,4	723	20,5	49,5	17,0	27,5	58,0	40,8
9,25	63,8	29	17	770	23,8	52,5	12,6	48,3	46,2	32,7

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
10,7	68,9	18	15,2	580	11,8	47,5	15,6	29,3	59,8	42,6
9,25	62,9	23	16,7	566	11,0	47,5	16,6	35,6	57,6	33,3
9,8	69,0	26	15,8	576	12,0	52	10,8	62,1	21,1	44,2
8,4	67,5	23	16,2	525	9,6	52,5	9,8	54,6	38,3	37,3
9,75	73,9	21	16,3	332	16,9	50	14,9	44,6	48,5	32,6
9,15	74,3	30	16,2	248	14,3	49,5	15,8	28,4	57,3	40,5
9,65	47,3	16	15,5	206	8,7	53,5	17,1	29,5	54,7	35,1
8,9	51,3	31	15,8	331	17,9	52	15,3	45,9	41,6	33,8
9	43,0	31	14,8	226	15,3	58	10,9	64,0	20,1	39,3
10,1	55,8	31	15,4	198	12,0	53,5	12,6	38,0	55,7	35,1
9,5	43,6	25	16,3	271	18,4	54,5	11,7	55,2	24,9	40,2
10,25	48,3	30	17,4	319	13,1	54,5	15,9	32,1	58,9	41,3
10,05	40,6	26	16,4	165	10,5	56	13,0	62,0	25,5	36,9
9,5	35,0	26	15,6	194	11,8	50	12,9	38,4	55,9	37,5
10,5	52,8	22	17	220	13,4	56,5	11,3	65,1	14,8	37,6
9,5	50,5	28	16,3	185	10,4	55,5	9,7	58,4	33,8	36,2
10,95	61,4	28	16	224	12,6	53	15,1	35,2	60,0	36,8
9,05	44,7	34	16,3	153	7,3	46,5	8,8	60,6	23,0	40,8
9,65	47,1	37	16,2	269	21,4	49,5	16,0	44,6	47,6	28,2
10,3	50,2	23	17,9	181	7,4	57	12,3	69,2	10,9	42,2
10	45,8	29	17,8	206	11,5	52	15,7	31,7	54,5	30,4
11,4	46,1	24	14,1	180	10,0	56,5	12,4	52,0	39,5	36,7
10,8	57,9	33	16,8	313	20,4	53,5	14,8	43,2	50,5	30,7

Manzana cv. <i>Golden Smoothee Delicious</i>		Melocotón vs. <i>Calrico</i>				Cereza cv. <i>Chelan</i>				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
10,05	42,7	37	14	263	19,0	55	14,6	43,8	49,8	31,0
9,4	49,6	20	16,7	246	11,5	53,5	10,9	54,0	38,2	32,8
9,45	53,0	33	19,1	207	14,8	56	10,4	60,2	26,9	38,9
10	45,3	43	10,5	230	14,6	52	14,9	39,7	52,3	31,8
10,2	38,4	20	16,4	219	10,1	47	16,6	15,9	44,9	27,3
9,7	49,8	28	17,7	212	14,6	42,5	15,9	17,8	46,0	30,2
10,85	54,4	42	16,7	264	19,7	48	19,2	10,8	39,8	18,5
9,75	44,2	41	15,7	278	17,8	47	15,5	14,4	43,7	24,7
9,65	40,2	46	16,5	248	16,4	48	13,0	24,9	52,6	41,8
11	44,5	38	15,9	213	16,1	41	16,9	11,8	41,1	20,2
10,9	44,4	33	16,7	262	19,9	50,5	15,9	25,2	50,7	33,1
10,6	42,7	31	15	280	20,8	39	16,7	27,1	48,7	29,2
11,8	60,6	33	17,7	296	15,0	39,5	15,1	26,2	48,5	28,2
12,05	53,2	40	14,3	285	18,4	46,5	11,4	30,6	55,9	46,4
10,35	47,0	28	17,7	220	11,9	42	14,7	30,2	54,0	39,0
10,15	42,3	38	15,7	259	20,1	38	14,1	16,2	45,7	27,9
10,8	42,4	23	15,3	230	12,0	42	15,8	17,5	47,1	30,1
10,25	45,9	22	18,6	250	11,6	40	14,3	20,2	49,3	34,4
11,5	49,0	35	16,5	292	19,3	40	15,6	18,2	46,4	29,5
10,95	53,0	13	16,9	257	15,8	34	15,8	19,6	47,8	31,6
10,4	40,8	36	15,7	225	15,3	42	14,6	30,3	55,1	39,3
10,45	40,7	34	11,5	208	10,6	42,5	15,2	29,6	55,4	47,1
12,65	47,5	28	17	243	19,1	33,5	15,7	24,1	48,7	30,6

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
10,35	41,5	39	15,7	277	23,0	44,5	18,0	18,5	48,0	31,6
10,95	40,2	39	16,3	288	16,5	39,5	17,3	18,7	45,2	26,9
12,9	60,4	38	15,3	255	19,9	46	17,2	21,2	48,6	31,1
9,75	39,7	33	14,5	236	18,6	40	14,1	15,0	44,6	25,9
10,9	46,8	45	15,9	252	17,3	42,5	16,4	16,7	46,4	28,9
10,35	44,8	38	15,1	177	13,3	36	17,2	13,2	41,3	21,2
10,1	49,7	30	15,7	278	15,5	41,5	16,2	27,2	55,4	46,1
11,85	42,4	43	15	254	21,5	44	18,0	20,1	47,3	29,1
11,4	40,6	19	17,3	231	11,0	33,5	13,7	22,9	51,1	38,8
11,75	45,3	30	15	301	20,8	38	15,3	35,5	56,1	39,7
12,1	48,2	41	15,1	245	18,7	36,5	16,3	22,9	50,8	36,5
11,2	39,0	39	14,6	281	16,9	33,5	14,8	18,6	48,1	31,8
12,05	39,5	14	17	269	10,1	29	15,5	13,2	43,3	22,8
11,75	35,9	38	16,7	240	14,7	39,5	17,6	10,3	39,5	17,8
12,45	46,3	55	14,9	206	15,3	35	17,0	9,0	38,2	15,6
11,65	41,1	28	15,1	261	16,4	36,5	14,2	29,7	57,0	41,6
12	42,0	53	13,2	195	18,5	51	20,2	8,0	36,8	13,8
11,65	38,8	50	14,7	248	16,0	36,5	14,9	14,8	44,8	25,6
11,4	41,5	22	16	238	17,9	44	16,1	20,6	49,8	34,3
11,1	39,5	44	13,9	280	22,1	35,5	18,3	14,0	44,2	24,1
11,25	40,1	33	16,4	230	17,8	38	12,0	23,5	51,9	37,7
11,6	43,2	41	14	213	14,7	43,5	13,8	14,1	43,9	24,4
11,85	38,8	32	16,9	258	11,5	33,5	15,0	15,9	46,2	27,4

Manzana cv. <i>Golden Smoothee Delicious</i>		Melocotón vs. <i>Calrico</i>				Cereza cv. <i>Chelan</i>				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
12,65	44,4	25	15,4	287	15,7	38	15,0	13,9	43,2	24,0
14,2	44,5	19	17,2	254	9,8	38	12,8	16,4	46,3	28,2
12,95	38,1	18	17	237	10,7	48,5	13,0	22,0	51,5	37,9
11,5	36,2	85	14,65	65	34,6	40	16,5	63,3	3,1	1,8
12,45	37,5	86	15,4	109	34,3	40,5	20,7	49,3	5,7	2,5
11,1	51,1	86	15,5	103	38,2	43	18,7	47,0	4,8	1,9
11,2	38,4	66	14,1	156	24,9	34	20,0	52,7	3,5	2,0
10,5	33,9	85	14,65	106	40,9	44	18,3	41,9	4,5	2,4
13,1	46,2	86	15,5	116	40,1	41	16,7	35,1	5,8	1,4
11,95	41,7	53	15,15	176	24,0	39,5	18,8	41,0	3,4	1,3
13,8	41,6	68	13,95	188	20,5	43	17,7	51,6	2,6	1,8
14,5	35,5	65	13,65	174	25,1	40	15,7	47,2	5,5	2,0
11,75	37,4	64	13,5	131	23,7	50,5	17,8	53,3	4,4	3,4
12,8	39,4	86	14,35	110	46,6	41	17,2	59,0	5,4	1,7
16	44,6	83	15,15	152	23,4	46,5	17,4	42,4	5,2	2,3
11,3	40,5	85	14,75	159	32,4	47	17,8	25,6	6,3	2,6
13,75	42,8	64	14,05	128	25,1	42	17,3	54,9	6,0	2,9
12,65	38,9	82	14,6	133	31,0	51	17,2	49,7	5,0	2,7
11,85	37,8	86	16,3	113	35,7	52	18,3	55,6	5,3	4,1
14,15	30,7	86	16,3	91	41,5	49	18,1	50,4	6,3	2,8
12,95	41,7	87	14,95	83	37,7	49	14,6	51,8	6,2	3,8
12,5	42,6	78	14,35	137	29,5	42	15,9	46,0	4,3	1,9
13,2	37,0	77	14,6	147	27,5	45	19,2	25,8	7,3	3,3

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
11,75	36,0	21	14,65	192	20,3	46,5	16,5	40,8	9,4	3,4
11,75	34,6	20	15,55	232	21,5	49	19,6	35,2	7,0	2,3
12,3	31,0	19	14,35	216	16,4	55,5	21,5	44,7	6,3	2,2
13,4	41,9	30	14,1	175	25,8	50,5	16,1	42,0	7,8	3,3
13	38,0	27	14,25	217	18,6	50,5	17,6	42,2	3,5	2,0
12,75	39,1	23	19,05	139	12,3	46,5	14,3	44,4	4,9	3,0
12,85	39,3	24	14,6	293	17,1	52,5	19,7	48,8	3,4	1,7
13,1	35,5	24	15,95	138	11,4	50,5	18,2	43,7	4,8	1,5
14,2	31,1	30	14,4	162	25,2	42,5	18,4	41,8	5,3	2,8
12,2	39,2	21	14,4	194	15,2	50	18,5	48,0	4,1	2,2
14,45	34,1	44	14,05	111	32,1	50,5	16,5	58,0	2,8	2,4
12,85	28,9	19	15,15	218	12,5	44,5	16,4	32,9	7,6	3,0
13,1	36,1	27	14,3	130	22,0	44	17,2	33,0	11,6	2,9
12,5	39,7	21	14,65	183	12,7	41	15,6	47,9	7,4	2,6
13,75	35,2	29	15,55	196	16,7	45	17,8	44,6	7,4	2,1
14,55	30,2	31	14,4	209	18,2	50,5	14,9	6,6	36,9	11,3
15,2	33,8	47	14,3	203	18,7	52,5	17,0	1,4	11,3	2,5
13,4	39,1	28	13,25	199	14,1	42	12,6	5,2	34,0	9,0
12,85	22,3	31	14,55	214	14,8	38	11,6	5,3	34,0	9,2
12,5	37,8	32	13,65	186	17,2	50	14,8	17,3	45,2	24,0
14	38,3	28	14,15	256	17,0	51,5	21,3	4,2	27,5	7,3
13,25	38,7	23	14,65	210	15,2	47	14,8	4,4	29,4	7,6
11,15	33,8	26	14,05	204	16,6	43,5	15,4	2,6	19,9	4,5

Manzana cv. <i>Golden Smoothee Delicious</i>		Melocotón vs. <i>Calrico</i>				Cereza cv. <i>Chelan</i>				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
14,55	34,5	27	14,05	202	19,6	53	17,1	2,5	18,4	4,3
13,75	27,4	23	15,65	193	15,1	48,5	15,8	1,5	11,8	2,6
13,45	33,3	22	14,2	266	16,8	48,5	16,2	3,0	22,9	5,2
13,2	32,8	40	14,5	215	23,3	49,5	16,2	4,1	29,0	7,1
13,45	38,2	28	13,4	198	15,6	52	14,2	2,7	21,2	4,7
12,4	36,4	36	14,9	250	19,4	50	17,5	2,8	21,0	4,9
12,9	29,1	33	14,95	240	15,0	45	15,0	3,5	25,5	6,1
13,65	38,0	46	14,85	232	22,5	36,5	9,6	9,7	41,4	16,7
13,4	26,3	38	14,3	180	14,8	50	17,5	2,2	16,4	3,7
13,1	26,8	46	13,5	222	24,7	48,5	17,3	2,6	18,4	4,5
14,55	30,0	58	14,4	230	22,3	45	13,2	4,7	32,2	8,2
14,25	21,6	37	13,9	219	16,9	25,5	20,6			
15,3	28,3	19	15	286	18,4	37	24,3			
15,1	24,2	23	16,25	248	10,8	25,5	24,1			
15,1	29,4	33	13,5	249	17,5	32	23,7			
14,1	32,1	28	13,15	160	14,6	30	22,4			
13,8	30,8	26	15	195	14,2	21,5	23,6			
13,4	34,5	21	13,7	220	15,0	27,5	22,3			
14,2	27,4	16	14,45	241	13,5	32	20,5			
15,4	34,0	18	13,8	207	15,2	33	23,0			
15,25	31,9	26	15,6	187	17,3	36	20,4			
15,45	32,7	14	13,7	164	8,4	30,5	21,0			
14,9	27,8	19	14,2	259	13,4	26,5	24,2			

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
14,95	25,5	19	14,8	222	15,2	32,5	22,5			
15,6	28,5	11	12,85	160	8,7	32	20,8			
15,45	31,5	38	13,3	247	16,5	31	25,1			
14,25	31,8	35	13,8	209	14,7	21,5	21,4			
14,95	36,5	29	14,35	271	16,1	25,5	22,4			
16,3	27,7	59	15,2	268	19,6	23	26,4			
15,3	34,0	56	13,45	215	18,5	27,5	27,3			
15,7	30,4	52	16,7	193	15,0	34,5	23,9			
16,3	15,6	54	17,9	178	8,8	30,5	22,8			
15,7	13,8	62	13,45	222	25,0	30,5	21,0			
15,6	13,5	59	16,95	164	19,7	25	20,7			
17,1	12,1	64	14,25	152	22,5	19,5	24,5			
12,6	13,8	53	14,05	235	18,4	28	21,6			
13,8	11,4	63	14,15	200	21,5	31	23,8			
14	18,7	55	13,75	217	21,0	37	18,6			
17,7	16,7	53	16,6	228	14,8	33	24,8			
17,2	14,7	55	15	225	16,5	35	25,5			
18,7	20,2	47	14,9	223	17,2	33	23,7			
17	14,3	56	13,8	215	21,8	30	28,1			
17,6	13,2	54	15,25	263	17,9	30,5	24,0			
18,8	16,8	37	13,65	206	14,7	38	21,6			
19,2	16,2	44	16,8	230	14,8	28	22,7			
20,3	16,2	54	14,7	252	21,2	25,5	25,6			

Manzana cv. Golden Smoothie Delicious		Melocotón vs. Calrico				Cereza cv. Chelan				
SST (°Brix)	Firmeza (N)	Firmeza (N)	SST (°Brix)	Peso (g)	AWETA	Durofel	SST (°Brix)	L	a	b
17,2	13,5	54	13,75	198	17,5	35	23,0			
19,4	10,0	49	16	206	14,2	33,5	24,9			
17,4	15,5	70	13,2	155	22,4	36	24,2			
13,6	15,2	42	13,7	222	18,4	28	27,7			
13,2	19,8	54	15,2	233	17,7	29,5	24,0			
13	15,9	44	15,25	221	18,7	21	25,8			
14,5	14,5	30	16,7	224	9,0	28,5	24,3			
13,3	12,2	21	16	223	9,3	25	24,2			
14	17,9	15	16,2	216	6,1	24	24,3			
14,9	14,1	25	15,75	207	7,1	25	19,5			
14,8	12,8	29	15,8	240	11,8	35,5	21,5			
14,2	16,0	28	14,8	241	10,6	23,5	22,1			
13,7	13,0	25	19,2	212	6,8	33,5	25,3			
14,2	16,7	40	14,2	183	5,6	29,5	21,8			
12,8	13,5	23	16,05	207	3,7	25,5	26,0			
13,3	17,4	12	14,65	165	3,8	24,5	23,3			
13,4	14,0	32	14,9	201	6,4	29	21,5			
13	19,9	17	13,45	199	6,3	28,5	22,9			
13	15,6	10	13,1	202	3,9	25	18,4			
15	19,1	24	14,65	180	4,5	34,5	21,7			
14,7	14,6	17	15,55	131	3,1	27,5	20,8			
13,1	4,3	19	14	191	3,8	33	23,8			
16,1	12,5	29	15,35	157	5,1	25,5	26,7			

