UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

D O C T O R A L T H E S I S

Application of hyperspectral imaging combined with chemometrics for the non-destructive evaluation of the quality of fruit in postharvest

Sandra Munera Picazo

Supervisors:

José Blasco Ivars

Nuria Aleixos Borrás

Valencia, May 2019

ivia
instituto valenciano
de investigaciones agrarias

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

José Blasco Ivars, PhD in Computer Science and Head of the Agricultural Engineering Centre at the Instituto Valenciano de Investigaciones Agrarias and **Nuria Aleixos Borrás**, PhD in Computer Science and Full Professor at the Universitat Politècnica de València

Certify:

That the work '***Application of hyperspectral imaging combined with chemometrics for the non-destructive evaluation of the quality of fruit in postharvest***' has been developed by **Sandra Munera Picazo** under their supervision in the Instituto Valenciano de Investigaciones Agrarias, as a Thesis Project in order to obtain the degree of PhD at the Universitat Politècnica de València.

Valencia, May 2019

José Basco Ivars

Nuria Aleixos Borrás

Agradecimientos

En primer lugar, agradezco la financiación recibida por el **Instituto Nacional de Investigaciones y Tecnología Agraria y Alimentaria (INIA)**, a través del contrato predoctoral FPI-INIA (CPR2014-0082, #43), sin la cual la presente tesis no habría sido posible.

Muchas gracias a mis directores, **José Blasco** y **Nuria Aleixos,** por la cercanía que siempre he recibido y por haber confiado en mí. Porque durante este tiempo han respetado mis ideas y decisiones y me han motivado en todo momento para que dé lo mejor de mí.

Gracias a todos mis compañeros del centro de Agroingeniería del IVIA. Por considerarme una más desde el minuto uno y hacer que, cada uno a su manera, este trabajo haya sido más fácil. Mención especial a **Sergio Cubero**, **Carlos Bataller**, **Vicente Alegre** y **Santi López,** siempre atentos a resolver mis dudas y a ayudarme. Y también a **Vicky Cortés**, que a pesar de estar en la UPV, simpre ha sido un gran apoyo en todo momento.

Gracias a los profesores e investigadores **Pau Talens**, **Alejandra Salvador**, **Cristina Besada** y **Juan Gómez**, por su importante aporte y participación en el desarrollo de esta tesis.

Gracias a **José Manuel Amigo,** por su supervisión durante mis dos estancias en Dinamarca en la Universidad de Copenhague y por la impagable formación en quimiometría ofrecida. Gracias también a **Giancarlo Colleli** y **Maria Luisa Amodio,** por acogerme en su grupo como un miembro más, por su supervisión y ayuda durante mi estancia en Italia en la Universidad de Foggia.

Agradecimientos a la **Cooperativa Agrícola Nuestra Sra. del Oreto (La Alcudia, Valencia),** por el soporte técnico prestado y el suministro de los caquis utilizados para realizar los experimentos de esta tesis; a **Fruits de Ponent (Alcarràs, Lleida)** por el suministro de las nectarinas utilizadas para realizar los experimentos de esta tesis; a

Antonio Terrés por el suministro de las granadas utilizadas para realizar los experimentos de esta tesis; a **Ruchey y la Cooperativa Agrícola de Callosa d´en Sarriá (Callosa d'en Sarrià, Alicante)**, por el soporte técnico prestado y el suministro de los nísperos utilizados para realizar los experimentos de esta tesis. Mención especial a **Esteban Soler,** responsable de I+D+i de la Cooperativa Agrícola de Callosa d´en Sarriá, por su apoyo, sus amplios conocimientos y su interés en este proyecto.

Todo camino tiene un comienzo. Gracias a **Ángel Carbonell** por apostar por mí, introducirme en el mundo de la investigación y darme la oportunidad de ser miembro del grupo Calidad y Seguridad Alimentaria de la Universidad Miguel Hernández. A pesar del duro trabajo, me fui de allí cargada de conocimientos y experiencia.

Agradecer también a todas esas personas que he conocido durante todo este tiempo y que han sido un gran apoyo tanto fuera como dentro del laboratorio. En especial a mi mexicana favorita, **Nallely**, por ser buena amiga y compañera de piso, de laboratorio y de fatigas; a **Paqui Hernández** por su amable ayuda y consejo; a los perdidos en Copenhague, **Jerónimo**, **Ana** y **Santi** y a los italianos, **Alejandra**, **Carmen**, **Mudassir**, **Faramand**, **Danial**, **Alessia**, **Nadia** y **Eleonora** por hacerme sentir como en casa.

Gracias a mi **familia** y **amigos** porque siempre están ahí para apoyarme aunque no les quede claro lo que hago. A mis padres, **Paco** y **Julia**, por todo su esfuerzo, por quererme tanto y darme la oportunidad de ser libre en la elección de mi formación sin que el aspecto económico fuese un problema. Y a mi hermana **Mª Juli**, que a pesar de nuestras diferencias sé que puedo contar ella.

Por último gracias a ti, **Vicente**, por soportarme en mis peores momentos, apoyarme en mis decisiones y dejarlo todo por acompañarme en este camino. Te quiero muchísimo.

A mis padres

INDEX

ABSTRACT

Currently, the food industry, especially the post-harvest sector, has to face the challenge of satisfying the growing demand for products in a market without borders, while complying with established quality and safety standards. Most of the conventional analytical techniques that ensure these standards are invasive, contaminant, time-consuming and costly, and there is a high risk of human error. The objective of this doctoral thesis is to evaluate the potential of the hyperspectral imaging in the visible and near infrared range in combination with chemometrics for the assessment of the postharvest quality of fruit in a non-destructive, efficient and sustainable manner. To this end, different studies are presented in which the quality of some fruits is evaluated. Due to their economic, strategic or social value, the selected fruits are of special importance in the Valencian Community, such as Persimmon 'Rojo Brillante' (Protected Designation of Origin Kaki Ribera del Xúquer), the pomegranate 'Mollar de Elche' (Protected Designation of Origin Granada Mollar de Elche), the loquat 'Algerie' (Protected Designation of Origin Níspero from Callosa d'en Sarrià) or different nectarine cultivars.

First, the quality monitoring of ‘Big Top’ and ‘Magique’ nectarines was carried out using reflectance and transmittance images. At the same time, transmittance was evaluated for the detection of split pit. In addition, a classification was performed to distinguish the ‘Big Top’ and ‘Diamond Ray’ cultivars, which look very similar but have different flavour. Whereas that for the ‘Rojo Brillante’ persimmon, the hyperspectral imaging was studied on the one hand to monitor its maturity, and on the other hand to evaluate the astringency of this fruit, which must be completely eliminated before its commercialization. The physicochemical properties of the ‘Mollar de Elche’ pomegranate were evaluated by means of hyperspectral and colour imaging during its maturity using the information from the intact fruit and arils. Finally, this technique was used to characterise and identify the internal and external defects of the ‘Algerie’ loquat.

The results obtained for the monitoring of the postharvest quality of the nectarines showed that both the reflectance and transmittance images are precise techniques. In the prediction of the IQI and RPI quality indexes, R^2 values around 0.90 were obtained and in the discrimination according to firmness, accuracy around 95.0 % using selected wavelengths was obtained. Regarding the split pit detection, the use of the hyperspectral image in transmittance mode obtained a 93.5 % of fruits with normal bone correctly

classified and 100% with split pit using PLS-DA models and 7 wavelengths. The results obtained in the classification of 'Big Top' and 'Diamond Ray' fruits show accuracy higher than 96.0 % by using PLS-DA models and 14 selected wavelengths, higher than the obtained with colour image (56.9 %) and a trained panel (54.5 %).

According to persimmon, the results obtained indicated that it is possible to distinguish between three states of maturity with an accuracy of 96.0 % using QDA models and its firmness was predicted obtaining a R2 value of 0.80 using PLS-R. Regarding astringency, two similar studies were carried out. In the first study, the fruit was classified according to the time of treatment with high concentrations of CO_2 with a precision of around 95.0 % using QDA. In the second, the fruit was discriminated according to a threshold value of soluble tannins (0.04 %) and was determined what fruit area was better to perform this discrimination. Thus, an accuracy of 86.9 % was obtained using the middle area and 23 wavelengths.

The results obtained for the pomegranate indicated that the use of colour and hyperspectral images have a similar precision in the prediction of physicochemical properties using PLS-R and the intact fruit information. However, when the information from the arils was used, the hyperspectral image was more accurate. Regarding the discrimination by the state of maturity using PLS-DA, the hyperspectral image offered greater precision, of 95.0 % using the information from the intact fruit and 100 % using that from the arils.

Finally, the results obtained for the ‘Algerie’ loquat indicated that the hyperspectral image with the XGBOOST classification method could discriminate between sound samples and samples with defects with accuracy of 97.5 % and between sound samples or samples with internal or external defects with an accuracy of 96.7 %. It was also possible to distinguish between the different defects with an accuracy of 95.9 %.

Thus, it can be concluded that hyperspectral imaging combined with chemometrics can be an adequate tool for evaluating the quality of different fruits in postharvest. The results obtained could serve as a scientific basis for a future implementation of this technique in real lines of handling and packaging of fruits such as nectarine, persimmon, pomegranate or loquat.

RESUMEN

Actualmente la industria alimentaria, especialmente el sector poscosecha, tiene que enfrentarse al desafío de satisfacer la creciente demanda de productos en un mercado sin fronteras, mientras cumple con los requerimientos de calidad y seguridad establecidos por los reguladores. La mayoría de las técnicas analíticas convencionales que aseguran estos requerimientos son invasivas, contaminantes y consumen tiempo y dinero, presentando un alto riesgo de error humano. Por ello, el objetivo de esta tesis doctoral es evaluar una técnica de inspección no destructiva, como la imagen hiperespectral en el rango visible e infrarrojo cercano, en combinación con técnicas quimiométricas para la evaluación de la calidad de la fruta en poscosecha de manera eficaz y sostenible. Con este fin, se presentan diferentes estudios en los que se evalúa la calidad de algunas frutas que por su valor económico, estratégico o social, son de especial importancia en la Comunidad Valenciana como son el caqui 'Rojo Brillante' (Denominación de Origen Protegida Kaki Ribera del Xúquer), la granada 'Mollar de Elche' (Denominación de Origen Protegida Granada Mollar de Elche), el níspero 'Algerie' (Denominación de Origen Protegida Níspero de Callosa d'en Sarrià) o diferentes cultivares de nectarina.

En primer lugar se llevó a cabo la monitorización de la calidad poscosecha de nectarinas 'Big Top' y 'Magique' usando imagen hiperespectral en reflectancia y transmitancia. Al mismo tiempo se evaluó la transmitancia para la detección de huesos abiertos. Se llevó a cabo también un estudio para distinguir los cultivares 'Big Top' y "Diamond Ray", los cuales poseen un aspecto muy similar pero sabor diferente. En cuanto al caqui 'Rojo Brillante', la imagen hiperespectral fue estudiada por una parte para monitorear su madurez, y por otra parte para evaluar la astringencia de esta fruta, que debe ser completamente eliminada antes de su comercialización. Las propiedades físico-químicas de la granada 'Mollar de Elche' fueron evaluadas usando imagen de color e hiperespectral durante su madurez usando la información de la fruta intacta y de los arilos. Finalmente, esta técnica se usó para caracterizar e identificar los defectos internos y externos del níspero 'Algerie'.

Los resultados obtenidos para la monitorización de la calidad poscosecha de las nectarinas mostraron que tanto la imagen en reflectancia como en trasnmitancia son técnicas precisas. En la predicción de los índices de calidad IQI y RPI se obtuvieron valores de R^2 alrededor de 0,90 y en la discriminación por firmeza, una precisión entorno al 95 %

usando longitudes de onda seleccionadas. En cuanto a la detección de huesos abiertos, el uso de la imagen hiperespectral en transmitancia obtuvo un 93,5 % de clasificación correcta de frutas con hueso normal y 100 % con hueso abierto usando modelos PLS-DA y 7 longitudes de onda. Los resultados obtenidos en la clasificación de los cultivares 'Big Top' y 'Diamond Ray' mostraron una fiabilidad superior al 96,0 % mediante el uso de modelos PLS-DA y 14 longitudes de onda seleccionadas, superando a la imagen de color (56,9 %) y a un panel entrenado (54,5 %).

Con respecto al caqui, los resultados obtenidos indicaron que es posible distinguir entre tres estados de madurez con una precisión del 96,0 % usando modelos QDA y se predijo su firmeza obteniendo un valor de R^2 de 0,80 usando PLS-R. En cuanto a la astringencia, se llevaron a cabo dos estudios similares en los que en el primero se discriminó la fruta de acuerdo al tiempo de tratamiento con altas concentraciones de CO_2 con una precisión entorno al 95,0 % usando QDA. En el segundo se discriminó la fruta de acuerdo a un valor de contenido en taninos (0,04 %) y se determinó qué área de la fruta era mejor para realizar esta discriminación. Así se obtuvo una precisión del 86,9 % usando la zona media y 23 longitudes de onda.

Los resultados obtenidos para la granada indicaron que la imagen de color e hiperespectral poseen una precisión similar en la predicción de las propiedades fisicoquímicas usando PLS-R y la información de la fruta intacta. Sin embargo, cuando se usó la información de los arilos, la imagen hiperespectral fue más precisa. En cuanto a la discriminación del estado de madurez usando PLS-DA, la imagen hiperespectral ofreció mayor precisión, 95,0 %, usando la información de la fruta intacta y del 100 % usando la de los arilos.

Finalmente, los resultados obtenidos para el níspero indicaron que la imagen hiperespectral junto con el método de clasificación XGBOOST pudo discriminar entre muestras con y sin defectos con una precisión del 97,5 % y entre muestras sin defectos o con defectos internos o externos con una precisión del 96,7 %. Además fue posible distinguir entre los diferentes defectos con una precisión del 95,9 %.

Así se puede concluir que la imagen hiperespectral combinada con métodos quimiométricos puede llegar a ser una herramienta adecuada para la evaluación de la calidad de diferentes frutas en poscosecha. Los resultados aquí obtenidos podrían servir como base científica para una futura implementación de esta técnica en línea reales de manipulación y envasado de frutas, especialmente de nectarina, caqui, granada o níspero.

RESUM

Actualment, la indústria alimentària, especialment el sector post collita, s'ha de enfrontar al repte de satisfer la creixent demanda de productes en un mercat sense fronteres, mentre compleix amb els requeriments de qualitat i seguretat establits. La majoria de les tècniques analítiques convencionals que asseguren aquests requeriments són invasives, contaminants i consumeixen temps i diners, i hi ha al seu torn un alt risc d'error humà. Per això, l'objectiu de la present tesi doctoral se centra en avaluar la capacitat d'una tècnica innovadora i no destructiva com es la imatge hiperespectral en el rang visible i infraroig pròxim, en combinació amb tècniques quimiomètriques, per a l'avaluació de la qualitat de la fruita en post collita de manera eficaç i sostenible. A aquest efecte, es presenten diferents estudis en els quals s'avalua la qualitat d'algunes fruites que pel seu valor econòmic, estratègic o social, són d'especial importància a la Comunitat Valenciana com són el caqui 'Rojo Brillante' (Denominació d'Origen Protegida Kaki Ribera del Xúquer), la magrana 'Mollar de Elche' (Denominació d'Origen Protegida Granada Mollar de Elche), el nispro 'Algerie' (Denominació d'Origen Protegida Nispro de Callosa d'en Sarrià) o diferents cultivares de nectarina.

En primer lloc es va dur a terme la monitorització de la qualitat post collita de nectarines 'Big Top' i 'Magique' per mitjà d'imatge hiperespectral en reflectància i trasnmitancia. Així mateix es va avaluar la transmitància per a la detecció d'ossos oberts. Es va dur a terme també un estudi per distingir els cultivares 'Big Top' i 'Diamond Ray', els quals posseeixen un aspecte molt semblant però sabor diferent. Pel que fa al caqui 'Rojo Brillante', la imatge hiperespectral va ser estudiada d'una banda per a monitoritzar la seua maduresa, i per un altre costat per avaluar l'astringència d'aquesta fruita, que ha de ser completament eliminada abans de la seua comercialització. Les propietats fisicoquímiques de la magrana 'Mollar de Elche' van ser avaluades per la imatge de color i hiperespectral durant la seua maduresa usant la informació de la fruita intacta i els arils. Finalment, aquesta tècnica es va fer servir per caracteritzar i identificar els defectes interns i externs del nispro 'Algerie'.

Els resultats obtinguts per a la monitorització de la qualitat postcollita de les nectarines van mostrar que tant la imatge en reflectància com en trasnmitancia són tècniques precises. En la predicció dels índexs de qualitat IQI i RPI es van obtindre valors de R^2 al

voltant de 0,90 i en la discriminació per fermesa una precisió entorn del 95,0 % utilitzant longituds d'ona seleccionades. Pel que fa a la detecció d'ossos oberts, l'ús de la imatge hiperespectral en transmitància va obtindre un 93,5 % classificació correcta de fruites amb os normal i 100 % amb os obert usant models PLS-DA i 7 longituds d'ona. Els resultats obtinguts en la classificació dels cultivares 'Big Top' i 'Diamond Ray' van mostrar una fiabilitat superior al 96,0 % per mitjà de l'ús de models PLS-DA i 14 longituds d'ona, superant a la imatge de color (56,9 %) i a un panell sensorial entrenat (54,5 %).

Quant al caqui, els resultats obtinguts van indicar que és possible distingir entre tres estats de maduresa amb una precisió del 96,0 % usant models QDA i es va predir la seua fermesa obtenint un valor de R^2 de 0,80 usant PLS-R. Pel que fa a l'astringència, es van dur a terme dos estudis similars en què el primer es va discriminar la fruita d'acord al temps de tractament amb altes concentracions de CO_2 amb una precisió al voltant del 95,0 % usant QDA. En el segon, es va discriminar la fruita d'acord a un valor de contingut en tanins (0,04 %) i es va determinar quina part de la fruita era millor per a realitzar aquesta discriminació. Així es va obtindre una precisió del 86,9 % usant la zona mitjana i 23 longituds d'ona.

Els resultats obtinguts per la magrana van indicar que la imatge de color i hiperespectral posseïxen una precisió semblant a la predicció de les propietats fisicoquímiques usant PLS-R i la informació de la fruita intacta. No obstant això, quan es va usar la informació dels arils, la imatge hiperespectral va ser més precisa. Quant a la discriminació de l'estat de maduresa usant PLS-DA, la imatge hiperespectral va oferir major precisió (95,0 %) usant la informació de la fruita intacta i del 100 % usant la dels arils.

Finalment, els resultats obtinguts pel nispro indiquen que la imatge hiperespectral juntament amb el mètode de classificació XGBOOST va poder discriminar entre mostres amb i sense defectes amb una precisió del 97,5 % i entre mostres sense defectes o amb defectes interns o externs amb una precisió del 96,7 %. A més, va ser possible distingir entre els diferents defectes amb una precisió del 95,9 %.

D'aquesta manera, es pot concloure que la imatge hiperespectral combinada amb mètodes quimiomètriques pot arribar a ser una ferramenta adequada per a l'avaluació de la qualitat de diferents fruites en postcollita. Els resultats aquí obtinguts podrien servir com a base científica per a una futura implementació d'aquesta tècnica en línies reals de manipulació i envasat de fruites, especialment de nectarina, caqui, magrana o nispro.

PREFACE

1. Resarch freamwork of the doctoral thesis

This doctoral thesis is the result of the research work accomplished by the author during the period 2015-2019, as a member of the Centro de Agroingeniería of the Instituto Valenciano de Investigaciones Agrarias (IVIA) with a FPI-INIA grant (CPR2014-0082, #43), granted by the Instituto Nacional de Investigaciones y Tecnología Agraria y Alimentaria (INIA), partially supported by the European Regional Development Funds (FEDER). The thesis is part of three research projects funded by the INIA with the support of FEDER funds, in which I have participated:

- Sistemas no destructivos para la determinación automática de la calidad interna de frutas en línea utilizando métodos ópticos e información espectral. *Non-destructive systems for the automatic determination of the internal quality of fruit in-line using optical methods and spectral information.* (RTA2015-00078-00-00)
- Nuevas técnicas de inspección basadas en espectrometría para la estimación de propiedades y determinación automática de la calidad interna y sanidad de productos agroalimentarios aplicadas a líneas de inspección y manipulación. *Novel inspection techniques based on spectrometry for the estimation of properties and automatic determination of the internal quality and health of agri-food products applied to inspection and handling lines.* (SPEC-DACSA) (RTA2012-00062-C04-01).
- Nuevas técnicas de inspección interna basadas en visión por computador multiespectral para la estimación de propiedades y determinación automática de la calidad y sanidad de la producción agroalimentaria en líneas de inspección y manipulación (VIS-DACSA). *Novel techniques of internal inspection based on multispectral computer vision for the estimation of properties and automatic determination of the quality and health of agri-food production in inspection and handling lines.* (RTA2012-00062-C04-03).

These projects are focused on the evaluation of the potential of optical methods for the characterisation and inspection of the internal and external quality of the fruit both off-line and in real-time, in order to create automatic processes of inspection.

2. Dissemination of the results

Publications in international journals of JCR

Published

Munera, S., Blasco, J., Amigo, J.M., Cubero, S., Talens, P. & Aleixos, N. (2019). Use of hyperspectral transmittance imaging to evaluate the internal quality of nectarines. *Biosystems engineering 182*, 54-64.
DOI: 10.1016/j.biosystemseng.2019.04.001. Impact factor: 2,132 (Q1).

Munera, S., Amigo, J.M., Aleixos, N., Talens, P., Cubero, S. & Blasco, J. (2018). Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine. *Food Control 86*, 1-10.
DOI: 10.1016/j.foodcont.2017.10.037. Impact factor: 3,667 (Q1).

Munera, S., Amigo, J.M., Blasco, J., Cubero, S., Talens, P. & Aleixos, N. (2017). Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging. *Journal of Food Engineering 214*, 29-39.
DOI: 10.1016/j.jfoodeng.2017.06.031. Impact factor: 3,197 (Q1).

Munera, S., Besada, C., Aleixos, N., Talens, P., Cubero, S., Salvador, A., Sun, D.W. & Blasco, J. (2017). Non-destructive assessment of the internal quality of intact persimmon using colour and VIS/NIR hyperspectral imaging. *LWT - Food Science and Technology 77*, 241-248.
DOI: 10.1016/j.lwt.2016.11.063. Impact factor: 3,129 (Q1).

Submitted – Under review

Munera, S., Besada, C., Gómez-Sanchís, J., Aleixos, N., Salvador, A., Cubero, S., Talens, P. & Blasco, J. Detection of astringent and de-astringed hard 'Rojo Brillante' persimmon fruit using a sensory threshold by means of hyperspectral imaging. *Journal of Food Engineering*.

Munera, S., Hernández, F., Aleixos, N., Cubero, S. & Blasco, J. Computer vision and chemometrics for quality monitoring of intact 'Mollar de Elche' pomegranate fruit and arils during maturity. *Postharvest Biology and Technology*.

Pending submission

Munera, S., Gómez-Sanchís, J., Aleixos, N., Vila-Francés,J., Cubero, S., Soler, E., Colelli, G. & Blasco J. Discrimination of common defects on 'Algerie' loquat fruit using hyperspectral imaging and machine learning techniques.

Publications in professional journals

Munera, S., Amigo, J.M., Aleixos, N., Talens, P., Cubero, S., Blasco, J. (2019). Identificación de variedades de nectarina con apariencia similar y diferente sabor mediante imagen hiperespectral. *Agrícola Vergel* 414, 354-358.

Blasco, J., Lorente, D., Cortés, V., Talens, P., Cubero, S., **Munera, S**., Aleixos, N. (2016). Application of near infrared spectroscopy to the quality control of citrus fruits and mango. *NIR News 27*, 4-7.

Besada, C., **Munera, S.,** Gil, R., Cubero., Aleixos., Salvador, A. (2016). El análisis de imagen como herramienta para evaluar de forma objetiva la reacción de los taninos solubles del caqui con el cloruro férrico. *Levante Agrícola 432*, 181-185.

Munera, S., Cubero., S., Blasco, J., Besada, C., Gil, R., Salvador, A., Albert, F., Aleixos, N. (2016). Análisis de la distribución de la astringencia en caqui 'Rojo Brillante' usando imagen hiperespectral. *Phytoma 282*, 14-16.

Book chapters

Blasco, J., **Munera, S.,** Aleixos, N., Cubero, S., Molto, E. (2017). Machine vision-based measurement systems for fruit and vegetable quality control in postharvest. Chapter in: Advances in Biochemical Engineering/Biotechnology. Springer (Berlin, Germany). DOI: 10.1007/10_2016_51.

Communications in national and international conferences

Oral communications

Munera, S., Aleixos, N., Gómez-Sanchis, Besada, C., Cubero, S., Talens, A., Salvador, A., Blasco, J. (2018). Detection of astringent and deastringent persimmon fruits using hyperspectral imaging technology. EurAgEng 2018 conference. 8-11 July, Wageningen, The Netherlands.

Munera-Picazo, S., Cubero, S., Albert, F., Talens, P., Cortés, V., Blasco, J., Aleixos, N. (2015). Medición no destructiva del índice de madurez en mango 'Kent' usando tecnología de imagen hiperespectral. VIII Congreso Ibérico de Agroingeniería, 1-3 June, Orihuela, Spain.

Posters

Munera, S., Besada, C., Blasco, J., Cubero, S., Gil, R., Aleixos, N., Salvador, A. (2016). Analysis of astringency distribution on 'Rojo Brillante' persimmon using hyperspectral imaging. VI International Symposium on Persimmon. 16-20 October, Valencia, Spain.

Munera, S., Amigo, J. M., Blasco, J., Cubero, S., Talens, P., Aleixos, N. (2016). Hyperspectral reflectance imaging for flesh firmness monitoring of nectarines cv. `Magique'. Fruit Logistic. 10-13, October, Berlín (Alemania).

Munera, S., Blasco, J., Cubero, S., Talens, P., Aleixos, N. 2016. Hyperspectral transmittance imaging for ripeness determination of nectarine. 6th International Conference in Spectral Imaging (IASIM 2016). 06-08 July, Chamonix, France.

Munera, S., Blasco, J., Cubero, S., Besada, C., Salvador, A., Talens, P., Aleixos, N. (2016). Maturity assessment of `Rojo Brillante' persimmon by hyperspectral imaging. International Conference on Agricultural Engineering (CIGR AgEng 2016). 26-29 June, Aarhus (Dinamarca).

Munera. S., Besada, C., Blasco, J., Cubero, S., Salvador, A., Talens, P., Aleixos, N. (2016). Firmness prediction in 'Rojo Brillante' persimmon using hyperspectral imaging technology. VIII International Postharvest Symposium. 21-24 June, Cartagena, Spain.

Besada, C., **Munera, S**., Gil, R., Cubero, S., Aleixos, N., Salvador, A. (2016). Image analysis as a tool to evaluate the astringency level in persimmon fruit. VIII International Postharvest Symposium. 21-24 June, Cartagena, Spain.

Besada C., Gil R., **Munera S.,** Cubero S., Aleixos N., Salvador A. (2015). Image analysis of blue staining associated to soluble tannins-chloride ferric reaction as indicator of sensory astringency in persimmon. 11th Pangborn Sensory Science Symposium. 23-27 August. Gothenburg, Sweden.

Albert, F., Aleixos, N., Cubero, S., Bataller, C., **Munera-Picazo, S.,** Blasco, J. (2015). Análisis de algunas características externas de los alimentos mediante análisis de imágenes utilizando Food-ColorInspector. VIII Congreso Ibérico de Agroingeniería. 1-3 June, Orihuela, Spain.

Lorente, D., Cortés, V., **Munera-Picazo, S.,** Escandell-Montero, P., Cubero, S., Aleixos, N., Talens, P., Blasco, J. (2015). Detección de podredumbres en cítricos mediante espectroscopía VIS/NIR y métodos de aprendizaje automático. VIII Congreso Ibérico de Agroingeniería, 1-3 June, Orihuela, Spain.

Cubero, S., Besada, C., Alberti, F., Salvador, A., **Munera, S.,** Gil, R., Cortes, V., Talens, P., Blasco, J., Aleixos, N. (2015). Determination of the astringency of persimmon cv 'Rojo Brillante' using hyperspectral imaging. International Conference on Food and Biosystems Engineering, 28-31 May, Mykonos, Greece.

Pre-doctoral stages at a foreign institution

- **University of Copenhagen**. Spectroscopy and Chemometrics Group. Department of Food Science. Copenhagen (Denmark). From 1st May 2016 to 31st July 2016 (3 months)

  Supervisor: Dr. José Manuel Amigo.

- **University of Copenhagen**. Spectroscopy and Chemometrics Group. Department of Food Science. Copenhagen (Denmark). From 1st May 2017 to 16th June 2017 (1.5 months)
  Supervisor: Dr. José Manuel Amigo.
- **University of Foggia**. Department of Agricultural, Food and Environmental Sciences. Foggia (Italy). Form 1st May 2018 to 30th June 2018 (2 months)
  Supervisor: Prof. Giancarlo Colelli.

3. Structure of the doctoral thesis

This doctoral thesis has been organized in five sections: Introduction, Objectives, Chapters, General Discussion and Conclusions.

The **Introduction** section is focused on the discussion of the state of the recent research and applications of hyperspectral imaging as a valuable tool to assess the quality and safety of fruits and vegetables in postharvest.

The **Objectives** section presents the general and specific objectives of the thesis.

The results obtained are divided into seven **Chapters** as scientific publications presented in four sections according to the fruit studied.

The first section is focused in the application of hyperspectral imaging to evaluate the quality of **nectarines** cv. 'Big Top', 'Magique' and 'Diamond Ray'.

- **Chapter I** studies the feasibility of hyperspectral reflectance imaging as a tool to assess the physicochemical properties and sensory perception of 'Big Top' and 'Magique' nectarines during ripening using the Ripening Index (RPI) and the Internal Quality Index (IQI).
- **Chapter II** evaluates the capacity of hyperspectral imaging in transmittance to inspect the internal quality of 'Big Top' and 'Magique' nectarines.

- **Chapter III** evaluates the capacity of hyperspectral imaging to discriminate between 'Big Top' and 'Diamond Ray' nectarines which are very similar in appearance but different in taste.

The second section is focused in the application of hyperspectral imaging to evaluate the quality of **persimmon** cv. 'Rojo Brillante', which has its own Protected Designation of Origin, 'Kaki Ribera del Xúquer'.

- **Chapter IV** evaluates the feasibility of hyperspectral imaging as a non-destructive toot to predict the maturity and determine the level of astringency according to the time of de-astringent treatment.
- **Chapter V** studies the application of hyperspectral imaging to discriminate astringent and deastringed fruits non-destructively using the spectral information of from three different areas of the fruit.

The third section is focused in the application of hyperspectral imaging to evaluate the quality of **pomegranate** cv. 'Mollar de Elche', which which has its own Protected Designation of Origin, 'Granada Mollar de Elche'. This section is divided in one chapter.

- **Chapter VI** evaluates colour and hyperspectral imaging as a non-destructive method to predict the physicochemical properties and the maturity stage of this fruit using the spectral information of the intact fruit and arils.

The fourth section is focused in the application of hyperspectral imaging to evaluate the quality of **loquat** cv. 'Algerie', which has its own Protected Designation of Origin, 'Nísperos Callosa d'En Sarrià'. This section is divided in one chapter.

- **Chapter VII** investigates the capacity of hyperspectral imaging to identify common external and internal defects in this fruit.

The most relevant results obtained in the different chapters are analysed together in the **General Discussion** section. Finally, the last section compiled the most relevant **conclusions** of the thesis.

ABBREVIATIONS

A	Astringent	**IR**	Infrared
AA	Antioxidant Activity	**K-NN**	K-Nearest Neighbors
ANN	Artificial Neural Network	**L**	Luminosity
ANOVA	Analysis of Variance	**LCTF**	Liquid-Crystal Tuneable Filters
AOTF	Acousto-Optical Tuneable Filter	**LDA**	Linear Discriminant Analysis
AUC	Area Under the ROC Curve	**LS-SVM**	Least Squares-Support Vector Machines
BPNN	Back Propagation Neural Network	**LV**	Latent Variable
C	Chromaticity	**MI**	Maturity Index
CCD	Charge Coupled Device	**MLP**	Multi-Layer Perceptron
CI	Colour Index	**MSC**	Multiplicative Scatter Correction
CV	Cross Validation	**NIR**	Near Infrared Region
D	Diameter	**NMR**	Nuclear Magnetic Resonance
DA	Deastringed	**PC**	Principal Component
F	Firmness	**PCA**	Principal Component Analysis
h	Hue angle	**PCR**	Principal Component Regression
HSD	Honestly Significant Difference	**PLS**	Partial Least Square
HSI	Hue, Saturation, Intensity	**PLS-R**	Partial Least Square Regression
InGaAs	Indium Gallium Arsenide	**PLS-DA**	Partial Least Square - Discriminant Analysis
i-PLS	Interval Partial Least Squares	**QDA**	Quadratic Discriminant Analysis
IQI	Internal Quality Index	**R^2**	Coefficient of determination

RF	Random Forest
RGB	Red, Green, Blue
RH	Relative Humidity
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
RPD	Residual Predictive Deviation
RPI	Ripening Index
SE	Standard Error
SIMCA	Soft Independent Modeling of Class Analogy
SNV	Standard Normal variate
ST	Soluble Tannins
SVM	Support Vector Machine
TA	Titratable Acidity
TPC	Total Phenolic Compounds
TSS	Total Soluble Solids
UV	Ultraviolet
VIP	Variable Important in Projection
VIS	Visible
W	Weight
XGBOOST	eXtreme Gradient Boosting

INTRODUCTION

1. Hyperspectral imaging technique

Nowadays, the food industry has to face the challenge of satisfying the growing demands of consumers in a market without borders, while complying with the required quality and safety standards. Furthermore, to ensure these standards are met, most conventional techniques to assess the quality and safety of horticultural products employ destructive methods which are contaminating, time-consuming, costly and the few samples monitored are later corroborated with the whole batch.

One of the most successful non-destructive technique for quality assessment of food products is the measurement of their optical properties (ElMasry and Sun, 2010). These optical properties are based on reflectance, transmittance, absorbance, or scatter of polychromatic or monochromatic radiation in the UV, VIS, NIR regions of the electromagnetic spectrum (Figure 1).

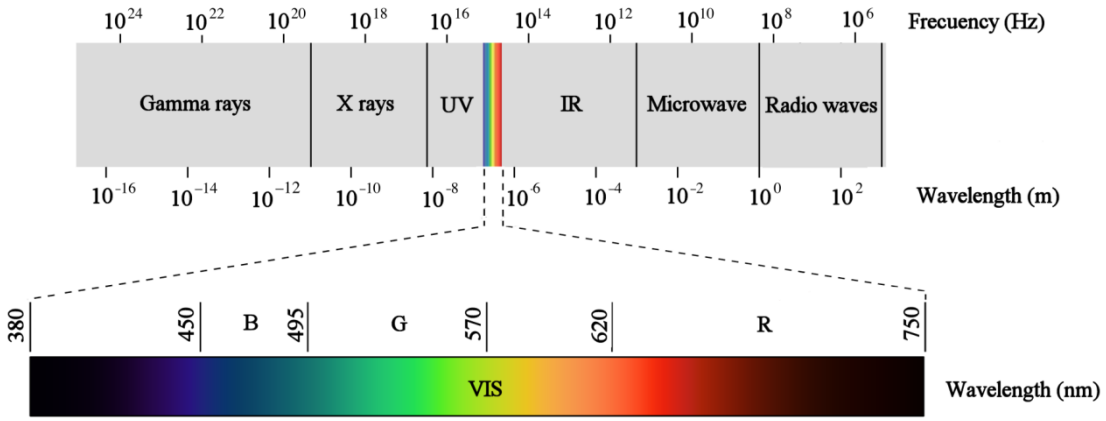


Figure 1. The electromagnetic spectrum.

When trained operators are employed in order to perform visual verifications, the decisions made by them could be affected by psychological factors such as fatigue or acquired habits. There is a high risk of human error in these processes, and this is one of the most important drawbacks that can be prevented by automated inspection systems based on computer vision (Cubero et al., 2011).

Systems based on VIS light have been widely used because such systems are designed with the intention of imitating the human eye, on the one hand, and due to

the fact that they are relatively inexpensive and faster, on the other. They are commonly used to detect external features such as size, shape, colour or the presence or absence of external defects in fruits (ElMasry et al., 2012; Font et al., 2014; Mohammadi et al., 2015; Benalia et al., 2016), meat (Muñoz et al., 2015; Zapotoczny et al., 2016; Barbin et al., 2016) or fish (Dutta et al., 2016; Sture et al., 2016; Coelho et al., 2016). However, some kinds of damage, the presence of contaminants or certain organoleptic characteristics cannot be seen with the naked eye, and therefore they are not detected by traditional cameras that only capture colour images (Nogales-Bueno et al., 2015a; Munera et al. 2018). In this regard, spectroscopy has been widely used to detect spectral features in other regions of the electromagnetic spectrum such as the UV, NIR or IR (Lorente et al., 2012) in food products, like meat (Alamprese et al., 2013), fish (Cascant et al., 2018) fruits (Lorente et al., 2015; Cortés et al., 2017), bakery (Li Vigni & Cocchi et al., 2013) or milk (Nuñez-Sanchez et al., 2016). This technology is rapid and inexpensive but in certain types of problems such as the assessment of samples like grains or small pieces of food, the detection of internal damage or the visualisation of different compounds in a sample, spatial information must be considered.

With the integration of the main advantages of spectroscopy and imaging, hyperspectral imaging technique can simultaneously acquire spectral and spatial information in one system that is critical for the quality prediction of agricultural and food products (Wu and Sun, 2013). A hyperspectral imaging system produces a two dimensional spatial array of vectors which represents the spectrum at each pixel location. The resulting three-dimensional dataset containing the two spatial dimensions and one spectral dimension is known as the datacube or hypercube (ElMasry and Sun, 2010) (Figure 2).

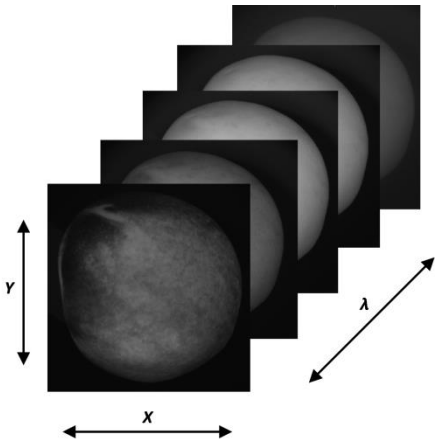


Figure 2. Hyperspectral image (hypercube) of a nectarine discomposed in their monochromatic images.

2. Hyperspectral imaging system

In recent years, a large amount of progress has been made in the technology used in hyperspectral imaging and the price of the equipment has gradually decreased. This fact has allowed incorporate this technique in many laboratories devoted to developing different hyperspectral applications. Nevertheless, this equipment is very different to those systems based on VIS sensors and it is essential to consider aspects such as the spatial and spectral distribution of the lighting, keeping the scene focused through the different wavelengths, the need to work with different filters or to move the sample as the image is acquired, and so on.

2.1 Wavelength dispersion devices

Wavelength dispersion devices are the core component of hyperspectral imaging systems. Their function is to disperse broadband light into different wavelengths and project the dispersed light to the camera (Qin et al., 2013). Many optical and electro-optical instruments can be used for this purpose; however, three types of technology are the most used nowadays: AOTF, LCTF and imaging spectrographs.

AOTF consist of a crystal in which selected wavelengths of light are separated from a broadband source using acoustic waves at specific radio frequencies. Alternative

compression and relaxation of the crystal lattice generates density changes that produce refractive index variations which act as a transmission diffraction grating. Unlike a classical diffraction grating, AOTF only diffract one specific wavelength of light, so they act more as a filter than a diffraction grating (Blasco et al., 2017).

LCTF use a stack of successively thicker, polariser birefringent liquid crystal plates which can generate a tuneable retardation of light transmission. Switching speed is limited by the relaxation time of the crystal and is of the order of 50 ms. Spectral resolution of LCTF is typically of the order of several nanometres (Blasco et al., 2017).

Imaging spectrographs are based on the characteristics of the scattering of electromagnetic wavelengths in materials and are characterised by the fact that they acquire spectral data about a scene line by line using the relative movement of that scene with respect to the instrument. These sensors offer a good spectral resolution, but they do not allow the complete image to be acquired without synchronising the image acquisition with the movement of the sample by means of an encoder or by using a mirror-scan (Lorente et al., 2012).

2.2 Camera

Regarding to the area detector or camera, it should have to be sensitive within the working range in which the images are going to be acquired. A standard CCD is sensitive to about 900 nm, which can impose restrictions on the system which thus fails to take full advantage of the possibilities of the filters (Blasco et al., 2017). This type of camera has been used by Gaston et al., (2010) to detect polyphenol oxidase in mushrooms and Chen et al. (2015) to predict anthocyanin in grapes. For systems sensitive beyond 1000 nm special cameras based on InGaAs sensors with stabilized temperature have to be normally employed (Blasco et al., 2017).

2.3 Lens

Another point in the system is the lens. A standard lens presents a high degree of chromatic dispersion in the infrared region of the electromagnetic spectrum because of the different optical paths taken by the infrared components of the light source.

This is due to the variation in the refraction index of the lens depending on the wavelength. This means that the focus plane can vary considerably between bands that are separated in the spectrum (for example, between the bands corresponding to the visible and infrared), thus giving rise to images that are focused in some bands and out of focus in others (Blasco et al., 2017).

2.4 Light

A proper lighting is crucial when using this kind of sensors. It must prevent unwanted bright spots, while also providing high-quality homogenous scene illumination. Emission should be ideally enough and uniform along the working spectrum and have adequate spatial homogeneity. Different kinds of light sources have different spectral emission. For example, daylight-type fluorescent tubes rarely go beyond 700 nm and therefore they have to be ruled out when it comes to designing a system that works in NIR region (Blasco et al., 2017). In contrast, incandescent lamps offer a high degree/grade/level of efficiency in NIR but produce inappropriate directional light. In addition, it is important to observe the geometry of the object to be analysed. The traditional systems based on the reflection of light at 45°, can be used to illuminate flat objects. If used to illuminate spherical objects, like some fruits, this type of illumination would produce bright spots on the object and it is more suitable to employ a diffuser in order to maximise its reflectivity and minimise the directional reflections that cause the bright spots on the spherical surface (Figure 3).

2.5 Sensing modes

Depending on the objective of the analysis and the position of the detector, lighting and sample, there are three common sensing modes for hyperspectral imaging. External features are typically detected using reflectance mode, being the most common mode in quality and safety assessment in postharvest. In this mode, the detector captures the reflected light from the illuminated sample avoiding specular reflection (Figure 3A). Internal defects can be detected using transmittance mode, in which the detector is located in the opposite side of the light source and captures the

transmitted light through the sample (Figure 3B). If is necessary a dipper information into the sample and less surface effects and the influence of thickness, interactance mode may be selected. In this mode, both light source and the detector are located in the same side of sample and parallel to each other (Nicolai et al., 2007).

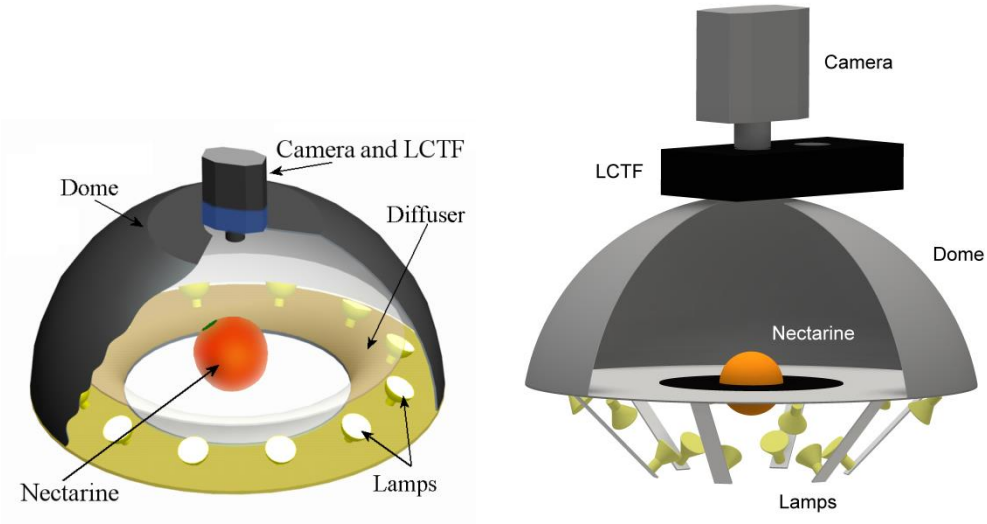


Figure 3. Example of a hyperspectral imaging system to acquire images in reflectance (A) and transmittance (B).

2.6 Hardware and software

Due to the hyperspectral image datasets size, an efficient use of hardware and software is required to enable fast and reliable streaming of data from camera detectors to disk drive storage and visualization displays. This includes, on the one hand, the proper selection of hardware components such as memory, processors and storage devices; on the other, the design and implementation of extremely efficient numerical algorithms ranging from simple vector and matrix operations to complex operations (Burger and Gowen, 2011).

3. Image analysis

Previous of pre-processing of the hypercube is essential to obtain uncontaminated data and to warranty the correct application of any multivariate data analysis (Vidal and Amigo, 2012). The techniques for image pre-processing will depend on the type of image measured, the device used and the information expected to obtain with the analysis.

The first step of image analysis is the correction of the radiance due to the differences in camera efficiency and the configuration of the systems. This kind of conditioning is called radiometric calibration, which converts the digital intensity values registered by the sensor to real or relative reflectance values (ElMasry & Nakauchi, 2016). This can be performed using equation (1) (Gat, 2000):

$$\rho_{xy}(x,y,\lambda) = \frac{\mathrm{R^{abs}}}{\mathrm{R^{abs}_{white}}} = \rho^{\mathrm{Ref}}(\lambda)\frac{\mathrm{R(x,y,\lambda)-R_{black}(x,y,\lambda)}}{\mathrm{R_{white}(x,y,\lambda)-R_{black}(x,y,\lambda)}} \quad (1)$$

where $\rho^{Ref}(\lambda)$ is the reflectance value of the standard surface (Teflon® standard with a 99% reflectance), $R(x,y,\lambda)$ is the reflectance of the sample captured by the sensor of the camera, $R_{white}(x,y,\lambda)$ is the reflectance captured of the standard surface to obtain the maximum intensity of each pixel in each wavelength, and $R_{black(}x,y,\lambda)$ is the reflectance captured without any illumination source and with the lens of the camera covered.

The morphology of the horticultural products, during calibration, has to be considered because it can affect the scattering and reflectance patterns. Due to the curvature of many fruits, vegetables or grains surface, light reflected is typically not uniform as seen by a camera. It makes difficult to detect surface anomalies or to evaluate the quality of the sample surface (ElMasry & Nakauchi, 2016). To mitigate this problem, Tao & Wen (1999) proposed an adaptive spherical image transform and Gomez-Sanchis et al. (2008) a Lambertian transform. However, applying some pre-treatments as SNV or MSC is possible to compensate the morphology effect in the spectra. Furthermore, other pre-treatments as smoothing or de-noising are used to

avoid the influence of artefacts as the instrumental noise, being Savitzky-Golay method one of the most accepted (Rinnan et al., 2009).

As is known, hyperspectral images are usually composed of thousands or even millions of data and it requires much storage space, large transmission bandwidths and long transmission times (Vidal and Amigo, 2012). Therefore, the compression of the image is useful to retain only the needed information. Data binning can be applied in both spatial and spectral dimensions by replacing the original data values which fall in a given small interval, a bin, by a value representative of that interval (Vidal and Amigo, 2012). For spectral reduction, variable selection using different mathematical methods could be an alternative. These methods are explained in 'Chemometrics' section.

Another important step in the image analysis is the segmentation. This process divides an image in ROI. Thresholding is widely used for image segmentation because is useful for images containing objects against a contrasting background. Thus, all pixels at or above the threshold set to 1 correspond to ROI whereas all pixels set to 0 correspond to the background. A set of morphological operations as neighbourhood operations or dilation and erosion may be utilized if the initial segmentation by thresholding is not satisfactory. Image segmentation can be also implemented by identifying the edge pixels located at the boundaries, due to the dramatically change in the grey level of those points Ngadi and Lu (2010)

4. Chemometrics

Multivariate analysis or chemometrics is a key tool in the analysis of the immense amount of spatial and spectral information that hyperspectral imaging offers. These mathematical and statistical methods are considered as a standard procedure for establishment of quantitative and qualitative models, allowing a reduction in the dimensionality of the data and retaining essential spectral information.

4.1 Exploratory analysis

The first step in the analysis of the spectral data is to summarize and to visualize the data as well as possible and in an unsupervised way (Szymanska et al., 2015). One of the most known exploratory methods and widely used is PCA. This technique transforms the original measured variables into new uncorrelated variables called PCs to reduce the data dimensionality, allowing its visualization and retaining as much as possible the information present in the original data. Each PC is a linear combination of the original measured variables. This technique affords a group of orthogonal axes that represent the directions of greatest variance in the data. The first PC accounts for the maximum of the total variance, the second is uncorrelated with the first and accounts for the maximum of the residual variance, and so on, until the total variance is accounted for. The loadings are the correlation coefficients between the original variables and the principal components. The values that represent the samples in the space defined by the principal components are the component scores (Berrueta et al., 2007).

4.2 Quantitative and qualitative analysis

The goal of using hyperspectral imaging and chemometric methods in the assessment of horticultural products is to create models to quantify different properties present in the samples or classify the samples according to these properties, i.e. to replace slow, expensive measurement of the property of interest, by a spectral feature that is cheaper or faster, but is still sufficiently accurate (Xiaobo et al., 2010).

First, when generating a model, a calibration or training set of samples must contain many representative examples that include both common and rare types of the target population with known properties or class to which they belong. At the same time, an independent, second sampling of the target population must be carried out to produce a test set to be used exclusively for validating of the model. However, there are situations when the use of a test set to validate the calibration model is not possible because its sampling is difficult, expensive or limited. For this situation, the

viable alternative is the use of CV, but it can never substitute for proper test set validation (Westad and Marini, 2015). In CV, the prediction ability of the model is determined by developing a model with part of the data set and using another part for testing the model. Both, training and test sets, contain samples representative of each class. This procedure, consisting of model development and model testing, is repeated several times so that the same samples have the probability to be used as training and as test objects (Berrueta et al., 2007). One of the most used methods is *k*-fold-cross-validation, which consists in assigning 1/*k* of the samples randomly to the test set.

Numerous chemometric methods have been established in order to perform quantification and qualification models, however, PLS, LDA, SIMCA, SVM or ANN are the most used in the multivariate analyisis of horticultural products data.

PLS-R is a widely used method for modelling the linear relationship between dependent variables Y or results of reference analysis and independent variables X or spectral data. The principle of PLS is to find the components in the input matrix X that describe as much as possible the relevant variations in the input variables and at the same time have maximal correlation with the target value in Y, given less weight to the variations that are irrelevant or noisy. PLS maximizes the covariance between matrices X and Y (Berrueta et al., 2007). On the contrary, PLS-DA aims to find the variables and directions in the multivariate space which discriminate the established classes in the calibration set. In contrast to PCA that only uses the information of matrix X, PLS-DA also takes into account the information in matrix Y (Berrueta et al., 2007).

LDA is another supervised pattern recognition method, which maximizes the ratio between both variances compared to the within-group variance (Berrueta et al., 2007). It searches for a linear function of the variable in multivariate space. When the number of variables is larger than the number of samples, LDA cannot be used directly. In that case, first PCA is employed for data compression to transform the original data set comprising of large number of inter-correlated variables into a reduced new set of variables.

SIMCA uses PCA for classification by creating confidence region around each class using residuals of the samples in the calibration set. The new objects are projected as a

member of a particular class based on their Euclidian distance which does not exceed a confidence limit from a particular principle component (Kumar et al., 2014).

Other methods can be used for defining linear or non-linear separations by integrating non-linear functions. SVM (Cortes & Vapnik, 1995) form a set of supervised learning methods for classification and regression tasks. SVM classifies the data in high-dimensional space with a separator described by a hyperplane. This hyperplane is expressed in terms of a linear combination of functions parameterized by so-called support vectors ANN imitate the structure and functioning of the human nervous system, to build parallel, distributed and adaptive information-processing systems, able to solve large-scale complex problems such as pattern recognition, non-linear modelling, classification and control (Pérez-Marín et al., 2007).

The performance of a quantitative model is usually evaluated in terms of standard error, RMSE and R^2 of calibration, CV and prediction in which is also used the RPD defined as the ratio between the standard deviation of the reference data and RMSEP (Williams, 1987). Generally, a good model should have higher values of R^2_C, R^2_{CV}, R^2_P and RPD and lower values of $RMSE_C$, $RMSE_{CV}$ and $RMSE_P$.

In qualitative models, confusion matrix, a form of contingency table, is used to show the counts of the correct and incorrect classifications from each class. Although it shows all of the information about the classifier's performance, more meaningful measures can be extracted from sensitivity, which measures the proportion of positives that are correctly identified, specificity, which measures the proportion of negatives that are correctly identified, class error and accuracy (Eq. 3, 4, 5 and 6):

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$Class\ error = 1 - \left(\frac{Sensitivity+Specificity}{2}\right) \quad (5)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

where TP and TN stand for true positive and true negative, respectively, accounting for the samples that have been correctly assigned as belonging (TP), or not belonging (TN),

to a specific class. FP and FN stand for false positive and false negative, respectively, accounting for the samples that have been wrongly assigned as belonging (FP), or not belonging (FN), to a specific class (Amigo et al., 2015).

The ROC curve is also used as a method to describe the performance of a classification. This curve is generated by plotting the 1-specificity versus sensitivity as *X* and *Y* axes, respectively. The best classification is when the coordinates of the ROC space are 0 and 1, representing 100% sensitivity (no false negative) and 100% also specificity (no false positive). The AUC integrates the sensitivity over the specificity. This area has a value between 0.5 and 1, where a value of 1 represents a perfect classification while 0.5 represents a test without discriminatory capacity.

4.3 Optimal wavelength selection

Spectral wavelengths in hyperspectral images are characterized by their large degree of dimensionality with collinearity and redundancy. Only the most important wavelengths having the great influence in prediction should be kept in the model in order to accelerate the process. Wavelengths selection may also be based on the model constructed, but variable selection is then applied uniquely as a post-processing step of the data.

Often, the selection of the optimal wavelengths is performed from a fitted PLS model, which is optimized for some number of components using CV. Loadings can be used as a measure of importance to select wavelengths. For each component the wavelengths with a loading weight above a certain threshold in absolute value may be selected. A second possibility is to use the vector of regression coefficients which is a single measure of association between each wavelength and the response. Again, wavelength having small absolute value of this filter measure can be eliminated (Mehmood et al., 2012).

VIP scores are calculated as a weighted sum of the squared correlations between the PLS components and the original variables. The weights correspond to the percentage variation explained by the PLS component in the model. Wavelengths with

a VIP score close to or greater than 1 can be considered the ones making the highest contribution in the PLS model (ElMasry et al., 2008).

The iPLS, introduced by Nørgaard et al. (2000), is especially designed for wavelength selection, where spectra are split into smaller subintervals with equal distance. Then a PLS-R is fitted to each sub-interval. The sub-interval having the smallest prediction error is selected.

Other methods used to select optimal wavelengths need a fitted ANN model which can be also combined with PLS.

5. Application of hyperspectral imaging for quality assessment of horticultural products

Quality is one of the major positioning tools of the producer for marketability, profitability and for consumer satisfaction. The 'optimal' quality of a product is related to a determined development of ripening degree, where the composition or combination of physical attributes and chemical components has maximum acceptance by consumers and producers. This section focuses on the recent applications of hyperspectral imaging to assess these attributes of quality in the horticultural products.

5.1 Evaluation of maturity and physicochemical properties

Greater knowledge of the chemical composition of fruit and vegetables can increase their added value. Consumers are willing to pay higher prices for products with an optimal maturity and health-stimulating properties such as bioactive compounds or antioxidant properties. In addition, the increase in exports and the longer distances involved mean that it becomes necessary to deliver higher quality and more consistent fresh products in the country of origin in order to meet the quality standards upon arrival at the destination. TSS, TA, pigments such as chlorophylls or carotenoids or ascorbic acid and TPC are the parameters that are most widely used to determine the maturity and the AA of fresh products. Their traditional assessment in

many fruit or vegetables is inefficient and incompatible with large-scale production and trading.

Amodio et al. (2017) predicted TSS, individual sugars and organic acids, phenols, and AA of fennel heads in relation to different sheath layers and harvest times using a VIS-NIR and NIR system. The calibration models were developed by PLS-R and TSS, AA and sucrose could be predicted with satisfactory accuracy, R^2 = 0.77, 0.78 and 0.77, respectively using VIS-NIR. In addition, they mapped the TSS from the external to the internal leaves and classified the fennels according to harvest time using PLS-DA, with a rate of non-error of 88.6%. Baiano et al., (2012) used VIS-NIR hyperspectral imaging for the prediction of parameters such as TSS, TA, pH and sensory characteristics of seven variety of table grape. Good correlations by using PLS-R were found between each of the quality parameters and the spectra information. Nevertheless, spectra information was not correlated with the sensory data, making difficult the prediction of the consumer liking. In the case of tomato, the most important bioactive compounds are lycopene and phenolic compounds. Liu et al., (2015) assessed the application of multispectral imaging with 19 wavelengths in UV, VIS and NIR region for predicting the content of these compounds in tomato. PLS, LS-SVM and BPNN were applied to develop quantitative models. BPNN model considerably improved the performance with R^2 = 0.94 and 0.97, for lycopene and TPC prediction, respectively.

According to the requirement of high quality in transoceanic shipment, Hu et al., (2015) investigated the use of VIS-NIR reflectance and transmittance as well as their combined modes, as a potential to be a non-destructive and non-contact measurement tool to predict blueberry mechanical properties. They used a selection algorithm, called random frog, to select optimal wavelengths and obtained similar results with full and selected spectra. Leiva-Valenzuela et al. (2013 & 2014) also used VIS-NIR reflectance and transmittance hyperspectral imaging to acquire images of blueberries in three different orientations (i.e., stem end, equator and calyx). They developed calibration models using PLS, coupled with iPLS for selection of wavelengths, to predict TSS and F, and assess the effect of fruit orientation in the image acquisition in order to guarantee the durability of the fruit during transportation. Results showed that the orientation did not have a significant effect on

F prediction, and hence there is no need to orient this type of fruit for hyperspectral imaging. Zhu et al (2016) used VIS-NIR and PLS-R to obtain the best model for determining the F of peach pulp, yielding prediction results with a R^2 0.85. The F distributions for different cross sections of peach pulp were quantitatively visualized at the pixel level. Furthermore, the histogram revealed the existence of a wide range of F inside peach pulp.

5.2 Discrimination of varieties

Due the large number of different cultivars that are currently available to growers, fruits of the same appearance but different organoleptic properties (and hence different marketing values) are grown at the same time and have close harvesting times. This is the case of cereals such as rice or maize, different cultivars can show variations in size, shape, colour and constitution, which cannot be identified by the human eye. Kong et al., (2013) classified four cultivars of rice using NIR hyperspectral imaging. They used SIMCA, kNN and SVM, a novel machine learning algorithm called Random Forest (RF) on full spectra and on selected wavelengths. SIMCA, SVM and RF models showed 100 % classification rates in full spectra. The best models on optimal wavelengths were RF, KNN and SVM but obtained less accuracy than full spectra. Moreover, the price and quality of some seeds in countries like China, strongly depends on geographic origin, on which climate and soil conditions are crucial influencing factors. Huang et al., (2016) tried to identify maize seeds of different year using VIS-NIR. They developed models using least squares support vector machine (LS-SVM) with 94.4% correct classification.

For the preparation of coffee beverages there are two main species used, Arabica and Robusta coffee. Due to its better taste and aroma, Arabica coffee is of higher quality than Robusta coffee, but it is more difficult to grow, even in function of its lower resistance to plant diseases, and therefore it is more expensive. Calvini, Ulrici and Amigo, (2015) classified these species of coffee with NIR hyperspectral imaging and by using sparse methods, such as sPCA + kNN or sPLS-DA. They compared classical

classification methods with the new methods. The last ones led to the analogous or even better classification results, between 80 % and 90 % accuracy.

In the production of wine, grape variety is typically determined by means of visual methods based on the staff experience and knowledge but it would be appropriate to have rapid and inexpensive analytical methods to classify grapes according to their variety. Nogales-Bueno et al., (2015a) discriminated between Tempranillo, Graciano, Garnacha and Mazuelo grapes with the use of NIR hyperspectral imaging. A step wise LDA) was developed for the data set in order to discriminate grapes according to the variety. As a result, 100 % of the samples were correctly classified in the internal validation process and 86 % in the leave-one-out CV process. Rodríguez-Pulido et al., (2013) classified wine grapes by varieties according to the seeds. They used two red varieties (Tempranillo and Syrah) and one autochthonous white variety (Zalema) cultivated in two kinds of soil. A general discriminant analysis was carried out using only a six selected wavelengths with PCA, and the results were compared with the classification obtained by using the whole spectra. Using the full NIR spectra, it was possible to classify 100 % of the samples according to their variety. The result using only six selected wavelengths was lower, 96.0 %.

5.3 Evaluation of industrial properties and control of the process

The importance of the measurement of the chemical or physical properties of raw materials and the correlation with the ensuing industrial processes suggests the need to develop rapid, accurate and non-invasive systems that can be used as a trusted technique to monitor and help detecting these properties. For instance, Rady et al., (2014) studied the potential of VIS-NIR spectroscopy and hyperspectral imaging to estimate internal or external constituents in two potato cultivars that are important to the processing industries like glucose, sucrose, specific gravity, primordial leaf count, and TSS. PLS-R was used to obtain the prediction models with the full spectra of slices and whole potato and with numerous pre-treatments. Spectroscopy method obtained the best correlations.

Kernel hardness is an important characteristic that influences the processing and end-use quality of maize products, and is of great importance to producers, processors and workers in the grain trade. Although the processes used during their transformation are well known, there are some factors that greatly influence processing and lead to repercussions in final products. In maize, for example, hardness is a significant factor related to losses during dry-milling, where softer kernels yield smaller amounts of large grits than harder kernels, and extremely hard kernels require more energy input and are more prone to stress cracks and breakage. For this reason, Williams and Kucheryavskiy (2016) used NIR hyperspectral imaging to classify maize kernels of three hardness categories: hard, medium and soft. Furthermore, they used PLS-DA to perform a pixel-wise classification (no acceptable results) and an object-wise classification using two methods for feature extraction — score histograms and mean spectra (sensitivities and specificities higher than 0.93). Serranti et al. (2013) verified the possibility of recognition of oat and groat before and after de-hulling. Classification models were built using PLS-DA and allowed to obtain a predictive accuracy near to 100 % with a reduced set of three wavelengths in the NIR region.

Regarding to wine, in a cellar it is really important to know the characteristics of grapes that are taken by the vine growers. Grape variety, maturity or phenolics compounds content are typically analyzed in order to determine grape quality, set grape price and classify grapes for the various wines produced. Nogales-Bueno et al., (2014) developed a non-destructive hyperspectral method for the determination of the principal parameters that compose technological and phenolic maturity (i.e. pH, TA, TSS, and TPC) in white and red grapes. In this fruit, anthocyanins are a group of the most crucial phenolic components of red wine grapes and are a key factor in the quality of the wine. Thereby, Fernandes et al., (2015) estimated anthocyanin concentration, TSS and pH in in whole Port wine grapes using VIS-NIR hyperspectral imaging and neural networks. Correlation results presented a R^2 of 0.95, 0.92 and 0.73 for anthocyanin content, TSS, pH, respectively. Chen et al., (2015) developed a model to measure anthocyanin content in wine grape skins using NIR hyperspectral imaging and applied it for predicting the phenolic maturation stage after veraison to guide selecting the best harvest time. In order to control the features of wines, the condition

of seeds is becoming an important factor for deciding the moment of harvesting by winemakers. Nogales-Bueno et al (2015b) also screened the extractable polyphenolic compounds in intact grape skin by modified PLS. The results were values of 0.82, 0.79 and 0.82 of R2 for total phenolic, anthocyanin and flavanols, respectively. Rodíiguez-Pulido et al., (2014) used NIR hyperspectral imaging to determine flavanols in seeds of red and white because these are the most representative phenolic compounds in this part of the berry and the phenolic maturity is decisive for the production of quality red wines. Calibrations were performed by PLS-R and provided R^2 0.73 for total flavanol content and R^2 0.85 for predicting flavanols extracted with a model solution. Delwiche, Souza and Kim (2013) tried to predict milling quality in soft red and white wheat with NIR hyperspectral imaging using flour yield, softness equivalent and sucrose solvent retention capacity as reference parameters. However, results concluded that hyperspectral imaging will not be sufficient so as to replace actual pilot milling procedures.

The consumption of fresh ready-to-eat products and minimally processed foods has increased in recent years. This sector is asking for fast, cheap and objective techniques to evaluate the overall quality and safety of these products in order to obtain decision tools for implementing new packaging procedures. Diezma et al. (2013) and Lara et al. (2013) monitored the evolution of leafy vegetables during storage using VIS-NIR hyperspectral imaging and acquiring the images through packing films and without these. In the first study, the tests performed showed the ability in discriminating between different storage periods of virtual images resulting from the application of three analytical techniques (spectral angle mapper distance, PLS-DA and a nonlinear index) to the hyperspectral images combined with the wavelengths selected. In the second one, the effect of the variation of the transmittance of the plastic on the leaves spectra was removed applying radiometric correction. They didn't compare the difference between the plastic covers used. Chaudhry et al. (2018) monitored the evolution of rocket leaves during storage by acquiring images both through packing films and without them.

5.4 Detection of physical damages and defects

Some damage (for example, mechanical) is difficult to detect in fruits and vegetables at early stages until external symptoms become visible. The development of browning is the consequence of a series of biochemical reactions in which polyphenol oxidase enzymes are present, and is the major factor leading to loss of quality that results in a reduction in their market value. In these cases, hyperspectral imaging can be used to detect early damage that can be seen in spectral ranges other than visible or that alters the composition in such a way that it can be detected by spectral imaging. Baranowski et al., (2013) examined the applicability of hyperspectral imaging in VIS and NIR ranges for classification of apple bruising with respect to the time after damage of five selected cultivars. They used numerous classifiers such as SVM, linear logistic regression, neural networks and decision trees. Results showed that the best accuracy to distinguish between days after bruising for apples of all five cultivars was observed for the sequential minimal optimization model (99.4 %) with second derivative and the linear logistic regression model (97.7 %). Huang et al., (2015) developed an online multispectral imaging system to classify bruise in apple by using the wavelengths 780, 850 and 960 nm selected with PCA. The classification accuracy of bruises in static tests was 91.5 % but in online tests was 74.6% due to the poor quality of 850 and 960 nm images. Lee et al., (2014a) investigated the extended range of NIR using hyperspectral imaging to detect bruises on pears. A simple F-value statistics was sufficient to find the optimal waveband ratio (R1074/R1016) and threshold for maximizing the bruise detection accuracy (92.0 %). Lü et al., (2011) developed a VIS-NIR hyperspectral imaging system covering the spectral region from 408 nm to 1117 nm for the bruise detection in kiwifruit. Selecting five important wavelengths with PCA, they developed SVM algorithms to identify and segregate the bruised tissue from the normal tissue of kiwifruit with accuracy of 87.5 %. Vélez-Rivera et al., (2014) studied the possibility of early detecting mechanical damages in 'Manila' mangoes, which is a very soft fruit and it is therefore very susceptible to mechanical damage during postharvest, by identifying the spectral bands that best categorise whether a mango is damaged or not. They obtained 97.9 % rate of correct classification on the third day

after the damage had been caused using K-NN and the whole spectra and 91.4% when only the most discriminant bands were used. López-Maestresalas et al., (2016) detected blackspots in potatoes using two hyperspectral imaging setups, one ranging from 400 to 1000 nm, and another covering the 1000-2500 nm range. Samples were analysed 1, 5, 9 and 24 h after bruising. PCA, SIMCA and PLS-DA were used to build classifiers, being PLS-DA the best model achieving an overall correct classification rate above 94% for both hyperspectral setups.

Low temperature storage is an effective way to slow these decay processes and maintain crop quality. However, it can also cause physiological disorders in some fruit and vegetables either during cold storage, at market or at home. The symptoms of this damage include internal browning and a dry mealy texture, together with a lack of taste and aroma. Simko et al. (2015) developed some indices to estimate decay and freezing injury in different cultivars of lettuces, based on ratios of particular wavelengths obtained from VIS-NIR hyperspectral imaging and chlorophyll fluorescence imaging with accuracy near to 97 %. Pan et al. (2016) detected cold injury in peaches. A multilayer perception ANN model was developed to select eight optimal wavelengths.

Bruised and cold injures can depreciate the value of the fruits or can incline the consumer to purchase a different fruit. However, other damages must be detected during postharvest quality control because they prevent the fruit to be marketed or can be used by fungal pathogens to contaminate the fruit. Cracking is one of the main reasons for rejection of the fruit by retailers, and serves as a potential vector for pathogenic penetration and contamination of the fruit. The feasibility of hyperspectral fluorescence imaging to detect cuticle crack defects on cherry tomatoes was investigated by Cho et al. (2013). Fluorescence intensity in the area of cracked cuticle was significantly higher in the blue-green spectral region than that of the sound surfaces. ANOVA and PCA were employed to investigate optimal fluorescence wavebands. A pair of selected wavebands was able to detect defective cherry tomatoes with > 99 % accuracy. Yu et al. (2014) investigated the potential of hyperspectral imaging for crack identification in fresh jujube. They employed PLS-R, spatial PCA and spatial independent component analysis to extract characteristic

wavelengths, and then LS-SVM qualitative discrimination models were established based on characteristic wavelengths and evaluated by ROC curve. Results demonstrated that the PLSR–LS-SVM model was the best to finish the qualitative discrimination of crack features in fresh jujube with a high accuracy of 100 %.

Zhang et al. (2015) used hyperspectral imaging combined with multivariate analysis and band math methods to detect common defects such as scars, insect damages, indentation, and spots in peaches. The overall classification accuracy of 93.3 % indicated that the selected wavelengths and proposed method were suitable and efficient for the common defect detection. Li, Rao and Ying (2011) detected various common defects on oranges: insect damage, wind scarring, thrips scarring, scale infestation, canker spot, copper burn, phytotoxicity, heterochromatic stripe, and normal surface. They reduced images data to few optimal wavelengths to form multispectral images by using PCA method. The two-band ratio and PCA coupled with a simple threshold method achieved the best 93.7 % orange surface defects identification accuracy and no false positives. Nakariyakul and Casasent (2011) discriminated internally damaged almond nuts from normal ones. They used the ratios 850/1210 nm and 1160/1335 nm, selected by sequential forward selection and SVM, which achieved 91.2 % accuracy.

Moreover, this technology can be used to obtain information about the presence of certain substances that can be the cause of some damages related with the evolution of the fruit. Gaston et al., (2010) investigated the potential of VIS-NIR hyperspectral imaging for the prediction of polyphenol oxidase enzyme activity on mushroom caps, causing the browning that is the major quality loss that accounts for a reduction in their market value. PCR models built on raw reflectance and MSC reflectance data were found to be the best model approach with R^2 of 0.78. In a similar study, Yang et al., (2015) studied the pericarp browning of lychees since this is regarded as the major problem of postharvest lychee. It is closely related to degradation of red pigments (anthocyanins) and the formation of brown coloured by-products. They selected two sets of optimal wavelengths using SPA and stepwise regression algorithms. Then, they built calibration models based on spatial and spectral information using the radial basis function support vector regression algorithm to

generate a map of anthocyanin distribution. Results showed the fused model showed higher R^2 of 0.89 and 0.87 for the training and the testing sets.

6. Application of hyperspectral imaging for safety assessment of horticultural products

The assessment of quality in postharvest is not only based on terms of appearance, flavour or nutritional value, but also they should not risk the consumers's health. They have no way to detect the presence of harmful substances and depend entirely on the responsibility of all members of the production and distribution chain. Due to numerous horticultural products may be consumed fresh, any pathogenic organism or toxic substance for human beings that can be transported on its surface constitute a potential hazard. This section focuses on the recent applications of hyperspectral imaging that proving the ability of this technology to detect the presence of harmful organism or substances to health.

6.1 Detection of biological damage

Insects can cause damage to fruit and significant economic losses for grower, processers and exporters in regions where it occurs. Haff et al., (2013) detected fruit fly larvae infested mango in grey scale images previously generated based on absorbance at particular NIR wavelengths (which are not indicated). Their algorithm incorporated background removal, application of a Gaussian blur, thresholding, and particle count analysis to identify locations of infestations. For heavily infested samples, the lowest overall error rate achieved was 2.0 %, with 1.0 % false positive and 3.0 % false negative. For samples with lower infestation rates, the error rates were much higher, the lowest overall error being 12.3 %. Wang et al. (2011) identified the effective wavelengths that have the maximum discriminatory capability in jujube fruits affected by insect damage and created a discriminant function to identify the stem-end/calyx-end, the sound cheek, and insect damage as well as to distinguish insect-damaged fruits from those free of infestation. According to the results, over 98.0 % of

the intact jujubes and 94.0 % of the insect-infested jujubes represented in the images were correctly recognised, and the overall classification accuracy was about 97.0 %. Huang et al. (2013) detected insect-damaged green soybean using VIS-NIR hyperspectral transmission imaging. Four statistical image features (minimum, maximum, mean, and standard deviation) were extracted from the images for classification and given as input to a discriminant classifier. The support vector data description classifier achieved 100% calibration accuracy. The model achieved 97.3 % and 87.5 % accuracies for normal and insect-damaged samples, respectively, with a 95.6 % overall classification accuracy, for the investigated independent test samples. Chelladurai et al., (2014) used soft X-ray and NIR hyperspectral imaging techniques to acquire images of soybeans infested by egg, larval, and pupal stages of *C. maculatus* along with uninfected and completely damaged soybeans. The LDA classifier for soft X-ray images correctly identified more than 86.0% of uninfested soybeans and 83 % of soybeans infested with all developmental stages of *C. maculatus* except the egg stage. Pair-wise LDA classification models developed from NIR hyperspectral data selected with PCA (960 nm, 1030 nm and 1440 nm) yielded more than 86.0 % and 87.0 % classification accuracy for uninfected and infested seeds, respectively. Singh et al., (2010) identify insect-damaged wheat kernels using NIR hyperspectral and digital colour imaging. They used three statistical discriminant classifiers (LDA, QDA, and Mahalanobis) and BPNN classifier. The QDA classifier using combined NIR image features and top 10 features from 230 colour image features gave the highest classification accuracy and classified 96.3% healthy and 91 % - 100 % insect-damaged wheat kernels. In other cases, the presence of a few fruits affected by fungus or bacteria in a shipment can render the entire shipment unmarketable.

Early detection of fungal infections in postharvest is especially important because only a few infected fruits can spread the infection to a whole batch during operations such as storage or exportation, thus causing great economic losses. Green and blue moulds, caused by *Penicillium digitatum* and *Penicillium italicum*, are the most economically important postharvest diseases of citrus in all production areas. Gómez-Sanchís et al. (2013) detected mandarins affected by green blue mould. They used ANN and classification and regression trees for the segmentation and classification of

images. Feature selection methods were used in order to reduce the dimensionality of the hyperspectral images and determine the 10 most relevant. The classification trees showed the best results of classification, an accuracy of around 93 %. Lorente et al. (2013) proposed a methodology to select features in multi-class classification problems using the ROC curve, in order to detect rottenness in citrus fruits. They select a reduced set of features (wavelengths at 550, 670, 690, 720 and 950 nm, CI and normalised difference vegetation index) with which obtained a classification success rate of around 89 %. Folch-Fortuny et al., (2016) detected symptoms of diseases caused by *Penicillium digitatum* in citrus fruits using N-way PLS-DA. A double cross-validation strategy was used to validate the discriminant models and a permutation testing on VIP values was used to select five bands (650, 660, 700, 750 and 760 nm) offering 91.0 % correct classification rate.

Kong et al., (2014) developed calibration models based on hyperspectral imaging data to fast detection of peroxidase activity in tomato leaves which infected with *Botrytis cinerea* and compare the performance of different calibration models. Five pre-treatment methods were investigated. 21 optimal wavelengths were selected by genetic algorithm-PLS and used as inputs of three calibration models. The optimal prediction result was achieved by a new fast learning neural algorithm named extreme learning machine with selected wavelengths, and it obtained a $R^2 = 0.87$.

Wang et al., (2012) compared the spectral characteristics of good onions and onions affected by bacterial diseases (sour skin) in the spectral region of 950 nm – 1650 nm to determine the optimal bands with PCA for discriminating the two classes and develop classification models to detect infected onions. The best classification approach used three parameters (maximum, contrast and homogeneity) of the log-ratio images as the input features of SVM, which discriminated 87.1 % healthy and sour skin-infected onions.

Citrus canker is also a severe disease of citrus causing enormous socioeconomic loses. Kim et al., (2014) detected citrus black spots symptoms or other potentially confounding peel conditions such as greasy spot, wind scar, or melanose. Spectral angle mapper and spectral information divergence hyperspectral analysis approaches

were used to classify fruit samples obtaining and accuracy of 97.9 % and 97.1 % respectively.

6.2 Detection of biological and chemical contaminants

Foodborne illnesses result in high economic and social costs and hence, the interest in methods and technologies for detecting contaminating food and preventing foodborne illness have grown significantly in the food and agricultural industries and also in regulatory agencies. For instance, Lee et al., (2014b) used hyperspectral fluorescence imaging with UV-A excitation to detect faecal contamination on leafy greens. They determined two ratios with the most significant wavelengths, F665.6/F680.0 for Romaine lettuce which obtained a R^2 of 0.98 and F660.8/F680.0 for baby spinach which obtained a R^2 of 0.96. On the other hand, Everard et al. (2014) studied three techniques, i.e. UV induced fluorescence, violet induced fluorescence, and VIS/NIR reflectance using hyperspectral imaging in combination with multivariate image analysis as well for detection of faecal contamination on spinach leaves. PLS-DA and two band ratio analysis techniques were used to compare these techniques. High detection accuracy was found for the two fluorescence configurations compared to the VIS/NIR. Both fluorescence configurations had 100 % detection rates for faecal contamination up to 1:10 dilution level and violet had 99.0 % and 87.0 % detection rates for 1:20 and 1:30 levels, respectively. Tomato hornworm is one of the several types of large caterpillars that attack tomatoes in the US whose faecal matter is closely related with the presence of *Escherichia coli* and *Salmonella*. Yang et al., (2014) developed a multispectral fluorescence imaging algorithm to detect aqueous frass contamination on mature tomatoes. The fluorescence intensities at five wavelengths (515 nm, 640 nm, 664 nm, 690 nm, and 724 nm) were used to compute three simple ratio functions to detect frass contamination. The algorithms detected over 99 % of the 0.2 kg/L and 0.1 kg/L frass contamination spots and successfully differentiated these spots from tomato skin surfaces, stem scars, and stems. In a previous work, Yang et al. (2012) worked on the development of a simple multispectral algorithm to detect four concentrations of aqueous faecal dilutions that was applied to apple

surfaces. The algorithm utilized the fluorescence intensities at four wavebands, 680, 684, 720, and 780 nm, and detected more than 99.0 % of the faecal spots.

The presence of toxigenic fungi in agricultural products is not only a loss of value, but concerns about food security, due to the possibility of which producing mycotoxins. Aflatoxins are toxic compounds produced by many species of *Aspergillus*, especially by *A. flavus* and *A. parasiticus* and fumonisins and trichothecenes by *Fusarium spp.* Cereals are the most susceptible to be contaminated with this type of organism. Del Fiore et al. (2010) and then Yao et al. (2013), tried to detect toxigenic fungi in maize kernels. The first study used VIS-NIR spectral range (400–1000 nm), and the results showed that hyperspectral imaging is able to rapidly discriminate from 48 h after inoculation with *A. niger* or *A. flavus*. In the second one, fluorescence emission was used to discriminate between toxigenic and untoxigenic fungi in both sides of the kernel when the infection was mild. The separation ability achieved between strains for the adjacent kernels on the germ side was 100 %.Then, they classified by healthy and contaminated kernel with a classification accuracy for the 100 ppb threshold on the germ side of 94.4 %. Recent works have studied the aflatoxin detection in this cereal. Kandpal et al., (2015) used three different varieties of maize (yellow, white and purple) which were inoculated with four different aflatoxin concentrations (10, 100, 500 and 1000 ppm). A PLS-DA model was developed to categorize control and infected kernels and the highest overall classification accuracy was 96.9 % in purple variety. They generated a contamination map with the PLS-DA model provided the visual appearance of infected samples.

Teena et al., (2014) studied the presence of *Aspergillus flavus* damages in date fruits using NIR hyperspectral imaging. Four wavelengths (1120, 1300, 1610 and 1650 nm) were identified as the most significant to classify by using PCA. The classification accuracies of infected date samples were 91.0-99.0 % and 92.0-100 % while comparing with sterile control and untreated control, respectively in various approaches. Atas, Yardimci and Temizel (2012) tried to detect aflatoxin in contaminated chili pepper by using UV and halogen lighting. 83.3 % accuracy rate was achieved for the under halogen. The most frequently selected spectral bands were 540, 550, 560, 590, 640 and 650 nm. UV excitation achieved 87.5 % of classification accuracy. 400 and 420 nm

spectral bands were selected as the most discriminative spectral bands. Hierarchical bottleneck backward elimination feature selection method and MLP classifier were used for get these results.

Gluten is found in processed foods made from wheat, barley and rye. It produces a mucous inflammation and bad-absorption syndrome, which is characterised by inappropriate absorption of nutrients in the bowel. Treatment of patients with celiac disease is based mainly on them following a gluten-free diet (Munera et al., 2014). Oats are considered a good addition to the gluten-free diet, but it is a challenge to keep them segregated from other gluten-rich grains. The demands for better detection tools for identifying and screening oat grain by the oat-processing industry led Erkinbaev, Henderson and Paliwal (2017) to discriminate oats from barley, wheat, and rye. A procedure was developed to classify six grains (oat, dehulled oat, barley, dehulled barley, wheat and rye) using NIR hyperspectral imaging in the wavelength range of 900-1700 nm coupled with multivariate data analysis. The reflectance spectra were analysed using PCA and PLS-DA. Good results of de-hulled oats grain prediction (99.0 %) were achieved using only few selected key wavelengths (1069, 1126, 1189, 1243, and 1413 nm).

7. Importance of persimmon, pomegranate, loquat and nectarine

In this doctoral thesis, persimmon, pomegranate, loquat and nectarine have been selected due to the special interest for the Valencian Community. This interest is mainly due to the fact that these fruits have a high strategic value as energizers of the some growing areas of this region, but also because they are, in the most cases, the main producers of them in Spain and in Europe, generating a high economic and social value.

In the case of persimmon (*Diospyros kaki* L.), the biggest producer in the world after China is Spain. In the last twenty years, the land area devoted to cultivating this crop has risen from 2,000 to 18,500 ha, and production has increased from 33 to 404 thousand tons (FAOSTAT, 2017). Part of this growth is due to the increase in the

production of the 'Rojo Brillante' cultivar in the Valencian Community. This cultivar is astringent at harvest and must be subjected to post-harvest treatments to remove astringency. This has been traditionally a handicap for the commercialization of this fruit since once the fruit losses the astringency by overripe, it acquires a soft jelly-like consistency being difficult to handle and eat. The development of the de-astringency methods based on high CO_2 concentrations allowed removal of the astringency while preserving high flesh F (Arnal and Del Río, 2003), which has facilitated a rapid commercial expansion of this crop. Nowadays 'Rojo Brillante' persimmon is one of the most appreciated persimmon cultivars worldwide because it good aspect, high size, flavour and absence of seeds. This fact was decisive for obtaining the 'Kaki Ribera del Xúquer' Protected Designation of Origin by the EU in 2001.

Regarding to pomegranate (*Punica granatum* L.), this fruit is a promising source of bioactive phytochemicals with a wide range of biological properties related to protection against oxidative stress, which is related to pathologies such as cardiovascular or neurodegenerative diseases and cancer (Mena et al., 2011). In Europe, Spain is the first producer of this fruit with a total production of 76 thousand tons (MAPA, 2018) coming mainly from the Valencian Community (60 thousand tons) where the cultivar 'Mollar de Elche' is very appreciable for consumers and has been granted Protected Designation of Origin status by the EU in 2015.

Loquat fruit (*Eriobotrya japonica* L.) is native from China, which is the first producing country in the world. Although loquat is a minor crop in Spain, this is the main loquat-producing country in the Mediterranean region with 29 thousand tons (MAPA, 2018) and the main exporter in the world (Besada et al., 2017). The production is concentrated in Andalusia and the Valencian Community, where loquat from Callosa d'en Sarria has been granted Protected Designation of Origin status by the EU in 1992. The most important cultivar is 'Algerie', which accounts for more than 80 % of total production. The interest of this fruit lies in the fact that loquat trees are harvested during a short period (from mid-April to the end of April), when there is low competition with other fruit on the market (Ballester et al., 2018).

Peaches and nectarines (*Prunus persica* L. Batsch) are, after apples, the most economically important fruit crop in EU (Elsadr and Sherif, 2016). The surface area of

the land devoted to the planting of these fruits is around 228 000 ha in 2018/19, with an estimation of production of 3.5 million tons of fruit (USDA, 2018). Over the last five years, Spain has become the largest peach and nectarine producer in EU with around 1.4 million tons, nectarine production being 547 thousand tons (MAPA, 2018). The most important regions are Aragón, Cataluña and Murcia and the Valencian Community presents a significant increase in the production (USDA, 2018). Due to the importance of nectarine production, it is one of the fruits to which most effort has been devoted by plant breeders in recent years in order to improve agronomic performance, and enhanced fruit appearance and quality (Reig et al., 2013). This fact has resulted in a significant increase in the number of new cultivars available to fruit growers and nowadays, the market offers sweet, semi-sweet, balanced, acidic and highly acidic cultivars, of which the first two cultivars are the most widely accepted by consumers (Iglesias, 2012).

References

Alamprese, C., Casale, M., Sinelli, N., Lanteri, S. & Casiraghi, E. (2013). Detection of minced beef adulteration with turkey meat by UV–vis, NIR and MIR spectroscopy. *LWT - Food Science and Technology 53*, 225-232.

Amigo, J.M., Babamoradia, H. & Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta 896*, 34–51.

Amodio, M.L., Capotorto, I., Chaudhry, M.M.A. & Colelli, G. (2017). The use of hyperspectral imaging to predict the distribution of internal constituents and to classify edible fennel heads based on the harvest time. *Computers and Electronics in Agriculture 134*, 1-10.

Arnal, L. & Del Río, M.A. (2003). Removing astringency by carbon dioxide and nitrogen-enriched atmospheres in persimmon fruit cv. 'Rojo brillante'. *Journal of Food Science 68*, 1516-1518.

Atas, M., Yardimci, Y. & Temizel, A. (2012). A new approach to aflatoxin detection in chili pepper by machine vision. *Computers and Electronics in Agriculture 87*, 129–141.

Ballester, C., Buesa, I., Soler, E., Besada, C., Salvador, A., Bonet, L. & Intrigliolo, D.S. (2018). Postharvest regulated deficit irrigation in early- and intermediate-maturing loquat trees. *Agricultural Water Management 205*, 1-8.

Baiano, A., Terracone, C., Peri, G. & Romaniello, R (2012). Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Computers and Electronics in Agriculture 87,* 142–151.

Baranowski, P., Mazurek, W. & Pastuszka-Wozniak, J. (2013). Supervised classification of bruised apples with respect to the time after bruising on the basis of hyperspectral imaging data. *Postharvest Biology and Technology 86*, 249–258.

Barbin, D.F., Mastelini, S.M., Barbon, S., Campos, G.F.C., Barbon, A.P. & Shimokomaki, M. (2016). Digital image analyses as an alternative tool for chicken quality assessment. *Biosystems engineering 144*, 85-93.

Benalia, S., Cubero, S, Prats-Montalbán, J.M., Bernardia, B., Zimbalattia, G. & Blasco, J. (2016). Computer vision for automatic quality inspection of dried figs (Ficus carica L.) in real-time. *Computers and Electronics in Agriculture 120*, 17-25.

Besada, C., Sanchez, G., Gil, R., Granell, A. & Salvadora, A. (2017). Volatile metabolite profiling reveals the changes in the volatile compounds of new spontaneously generated loquat cultivars. *Food Research International 100*, 234-243.

Berrueta, L.A., Alonso-Salces, R.M. & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography 1158*, 196–214.

Blasco, J., Munera, S., Aleixos, N., Cubero, S. & Moltó, E. (2017). Machine vision-based measurement systems for fruit and vegetable quality control in postharvest. In: Bernd Hitzmann (Ed.). Measurement, modeling and automation in advanced food processing (Pp 71-91). Advances in Biochemical Engineering/Biotechnology 161. Springer.

Burger, J. & Gowen, G. (2011). Data handling in hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems 108*, 13-22.

Calvini, R., Ulrici, A. & Amigo, J. M. (2015). Practical comparison of sparse methods for classification of Arabica and Robusta coffee species using near infrared hyperspectral imaging. *Chemometrics and Intelligent Laboratory Systems 146*, 503–511.

Cascant, M.M., Breil, C., Silvie, A., Tixier, F., Chemat, F., Garrigues, S. & de la Guardia, M. (2018). Determination of fatty acids and lipid classes in salmon oil by near infrared spectroscopy. *Food Chemistry 239*, 865-871.

Chaudhry, M.M.M., Amodio, M.L., Babellahi, F., de Chiara, M.L.V., Amigo, J.M. & Colelli, G. (2018). Hyperspectral imaging and multivariate accelerated shelf life testing (MASLT) approach for determining shelf life of rocket leaves. Journal of Food Engineering 238, 122-133.

Chelladurai, V., Karuppiah, K., Jayas, D.S., Fields, P.G. & White, N.D.G. (2014). Detection of Callosobruchus maculatus (F.) infestation in soybean using soft X-ray and NIR hyperspectral imaging techniques. *Journal of Stored Products Research 57*, 43-48.

Chen, S., Zhang, F., Ning, J., Liu, X., Zhang, Z. & Yang S. (2015). Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging. Food Chemistry 172, 788–793.

Cho, B.K., Kim, M.S., Baek, I.S., Kim, D.Y., Lee, W.H., Kim, J., Bae, H. & Kim, Y.S. (2013). Detection of cuticle defects on cherry tomatoes using hyperspectral fluorescence imagery. *Postharvest Biology and Technology 76*, 40–49.

Coelho, P.A., Torres, S.N., Ramírez, W.E., Gutierrez, P.A., Toro, C.A., Soto, J.G., Sbarbaro, D.G. & Pezoa, J.E. (2016). A machine vision system for automatic detection of parasites *Edotea magellanica* in shell-off cooked clam *Mulinia edulis*. *Journal of Food Engineering 181*, 84-91.

Cortés, V., Rodríguez, A., Blasco, J., Rey, B., Besada C., Cubero, S., Salvador, A., Talens, P. & Aleixos, N. (2017). Prediction of the level of astringency in persimmon using visible and near-infrared spectroscopy. *Journal of Food Engineering 204*, 27-37.

Cortes, C. & Vapnik, V. (1995). Support Vector Networks. *Machine Learning 20,* 273 – 297.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. & Blasco, J. (2011). Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables. *Food Bioprocess Technology 4*, 487–504.

Del Fiore, A., Reverberi, M., Ricelli, A., Pinzari, F., Serranti, S., Fabbri, A.A., Bonifazi, G. & Fanelli, C. (2010). Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. *International Journal of Food Microbiology 144*, 64–71.

Delwiche, S. R., Souza, E. J. & Kim, M.S. (2013). Limitations of single kernel near-infrared hyperspectral imaging of soft wheat for milling quality. *Biosystems engineering 115*, 260-273.

Diezma, B., Lleó, L. Roger, J.M., Herrero-Langreo, A., Lunadeic, L. & Ruiz-Altisent, M. (2013). Examination of the quality of spinach leaves using hyperspectral imaging. *Postharvest Biology and Technology 85*, 8–17.

Dutta, M.K., Isaac, A., Minhas, N. & Sarkar, B. (2016). Image processing based method to assess fish quality and freshness. *Journal of Food Engineering 177*, 50-58.

Elsadr, H.T. & Sherif, S. (2016). Peaches and Nectarines. In: Caballero, B., Finglas, P.M., Toldrá, F. (Ed). Encyclopedia of Food and Health (pp 270-276). Elsevier.

ElMasry, G. & Nakauchi, S. (2016). Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality - A comprehensive review. *Biosystems Engineering 142*, 53-82.

ElMasry, G., Cubero, S. Moltó, E. & Blasco, J. (2012). In-line sorting of irregular potatoes by using automated computer-based machine vision system. *Journal of Food Engineering 112*, 60-68.

ElMasry, G. & Sun, D.W. (2010). Principles of hyperspectral imaging technology. In Sun, D.W. (Ed). Hyperspectral Imaging for Food Quality Analysis and Control (Pp 3-43). Elsevier.

ElMasry, G. Wang, N., Vigneault, C., Qiao, J. & ElSayed, A. (2008). Early detection of apple bruises on different background colors using hyperspectral imaging. *LWT - Food Science and Technology 41*, 337–345.

Erkinbaev, C., Henderson, K., & Paliwal, J. (2017). Discrimination of gluten-free oats from contaminants using near infrared hyperspectral imaging technique, *Food Control 80*, 197-203.

Everard, C. D., Kim, M. S. & Lee, H. (2014). A comparison of hyperspectral reflectance and fluorescence imaging techniques for detection of contaminants on spinach leaves. *Journal of Food Engineering 143*, 139–145.

FAOSTAT. http://www.fao.org/faostat/en/#data/QC - Accessed May, 2019.

Fernandes, A.M., Franco, C., Mendes-Ferreira, A., Mendes-Faia, A., Leal da Costa, P. & Melo-Pinto, P. (2015). Brix, pH and anthocyanin content determination in whole

Port wine grape berries by hyperspectral imaging and neural networks. *Computers and Electronics in Agriculture 115*, 88–96.

Folch-Fortuny, A., Prats-Montalbán, J.M., Cubero, S., Blasco, J. & Ferrer, A. (2016). VIS/NIR hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus fruits. *Chemometrics and Intelligent Laboratory Systems 156,* 241-248.

Font, D., Tresanchez, M., Pallejà, T., Teixidó, M., Martinez, D., Moreno, J. & Palacín, J. (2014). An image processing method for in-line nectarine variety verification based on the comparison of skin feature histogram vectors. *Computers and Electronics in Agriculture 102*, 112–119.

Gaston, E., Frías, J.M., Cullen, P.J, O´Donnell, C.P. & Gowen, A. (2010). Prediction of polyphenol oxidase activity using visible near-infrared hyperspectral imaging on mushroom (Agaricus bisporus) caps. *Journal of Agricultural and Food Chemistry 58*, 6226–6233.

Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report, Opto- Knowledge Systems Inc. OKSI.

Gómez-Sanchis, J., Blasco, J., Soria-Olivas, E., Lorente, D., Escandell-Montero, P., Martínez-Martínez, J.M., Martínez-Sober, M. & Aleixos, N. (2013). Hyperspectral LCTF-based system for classification of decay in mandarins caused by Penicillium digitatum and Penicillium italicum using the most relevant bands and non-linear classifiers. *Postharvest Biology and Technology 82*, 76–86.

Gómez-Sanchis, J., Moltó, E., Camps-Valls, G., Gómez-Chova, L., Aleixos, N. & Blasco, J. (2008). Automatic correction of the effects of the light source on spherical objects. An application to the analysis of hyperspectral images of citrus fruits. *Journal of Food Engineering 85*, 191-200.

Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G. & Frias, J.M. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology 18,* 590-598 .

Haff, R.P., Saranwong, S., Thanapase, W., Janhiran, A., Kasemsumran & S., Kawano, S. (2013). Automatic image analysis and spot classification for detection of fruit fly infestation in hyperspectral images of mangoes. *Postharvest Biology and*

Technology 86, 23–28.

Hu, M.H., Dong, Q.L., Liu, B.L., Opara, L. & Chen, U.L. (2015). Estimating blueberry mechanical properties based on random frog selected hyperspectral data. *Postharvest Biology and Technology 106*, 1–10.

Hu, M.H., Dong, Q.L. Liu, B.L. & Linus, U. (2016). Prediction of mechanical properties of blueberry using hyperspectral interactance imaging. *Postharvest Biology and Technology 115*, 122-131.

Huang, M., Tang, J., Yang, B. & Zhu, Q. (2016). Classification of maize seeds of different years based on hyperspectral imaging and model updating. *Computers and Electronics in Agriculture 122*, 139–145.

Huang, W., Li, J., Wang, Q. & Chen, L. (2015). Development of a multispectral imaging system for online detection of bruises on apples. *Journal of Food Engineering 146*, 62–71.

Huang, M., Wana, X., Zhang, M. & Zhu, Q. (2013). Detection of insect-damaged vegetable soybeans using hyperspectral transmittance image. *Journal of Food Engineering 116*, 45–49.

Iglesias, I. (2012). Producción, consumo e innovación varietal en el melocotón. *Alimentación, Equipos y Tecnología 268*, 25-31.

Kandpal, L., M., Lee, S., Kim, M.S., Bae, H. & Cho, B.K. (2015). Short wave infrared (SWIR) hyperspectral imaging technique for examination of aflatoxin B1 (AFB1) on corn kernels. *Food Control 51*, 171-176.

Kim, D., Burks, T.F., Ritenour, M.A. & Qin, J. (2014). Citrus black spot detection using hyperspectral imaging. *International Journal Agricultural & Biology Engineering 7,* 20-27.

Kong, W., Liu, F., Zhang, C., Bao, Y. ,Yu, J. & He, Y. (2014). Fast detection of peroxidase (POD) activity in tomato leaves which infected with Botrytis cinerea using hyperspectral imaging. Spectrochimica Acta Part A: *Molecular and Biomolecular Spectroscopy 118,* 498–502.

Kong, W., Zhang, C., Liu, F., Nie, P. & He, Y. (2013). Rice seed cultivar identification using near-infrared hyperspectral imaging and multivariate data analysis. *Sensors 13*, 8916-8927.

Kumar, N., Bansal, A., Sarma, G.S. & Rawal, R.K. (2014). Chemometrics tools used in analytical chemistry: An overview. *Talanta 123*, 186-199.

Lara, M.A., Lleó, L., Diezma-Iglesias, B., Roger, J.M. & Ruiz-Altisent, M. (2013). Monitoring spinach shelf-life with hyperspectral image through packaging films. *Journal of Food Engineering 119*, 353–361.

Lee, W.H., Kim, M.S., Lee, H., Delwiche, S.R., Bae, H., Kim, D.Y. & Cho, B.K. (2014a). Hyperspectral near-infrared imaging for the detection of physical damages of pear. *Journal of Food Engineering 130*, 1–7.

Lee, H., Everard, C. D., Kang, S., Cho, B.K., Chao, K., Chan, D. E. & Kim, M. S. (2014b). Multispectral fluorescence imaging for detection of bovine faeces on Romaine lettuce and baby spinach leaves. *Biosystems engineering 127*, 125 – 134.

Leiva-Valenzuela, G.A., Lu, R. & Aguilera, J. M. (2014). Assessment of internal quality of blueberries using hyperspectral transmittance and reflectance images with whole spectra or selected wavelengths. *Innovative Food Science and Emerging Technologies 24*, 2–13.

Leiva-Valenzuela, G.A., Lu, R. & Aguilera, J. M. (2013). Prediction of firmness and soluble solids content of blueberries using hyperspectral reflectance imaging. *Journal of Food Engineering 115*, 91–98.

Li, J., Rao, X. & Ying, Y. (2011). Detection of common defects on oranges using hyperspectral reflectance imaging. *Computers and Electronics in Agriculture 78*, 38–48.

Li Vigni, M. & Cocci, M. (2013). Near infrared spectroscopy and multivariate analysis to evaluate wheat flour doughs leavening and bread properties. *Analytica Chimica Acta 764*, 17-23.

Liu, C., Liu, W., Chen, W., Yang, J. & Zheng L. (2015). Feasibility in multispectral imaging for predicting the content of bioactive compounds in intact tomato fruit. *Food Chemistry 173*, 482–488.

López-Maestresalas, A., Keresztes, J.C., Goodarzi, M., Arazuri, S., Jaren, C., & Saeys W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control 70*, 229-241.

Lorente, D., Escandell-Montero, P., Cubero, S., Gómez-Sanchis, J. & Blasco, J. (2015).

Visible–NIR reflectance spectroscopy and manifold learning methods applied to the detection of fungal infections on citrus fruit. *Journal of Food Engineering 163*, 17-24.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S. & Blasco, J. (2013). Selection of optimal wavelength features for decay detection in citrus fruit using the ROC curve and neural networks. *Food and Bioprocess Technology 6*, 530–541.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology 5*, 1121–1142.

Lü, Q., Tang, M., Cai, J., Zhao, J. & Vittayapadung, S. (2011). Vis/NIR hyperspectral imaging for detection of hidden bruises on kiwifruits. *Czech Journal Food Science 29*, 595–602.

MAPA. (2018). Datos Avances de Frutales No Cítricos y Frutales Secos años 2018. Available on-line at https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/superficies-producciones-anuales-cultivos/ Accesed in May 2019.

Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. Chemometrics and Intelligent Laboratory Systems 118, 62–69.

Mena, P, Gironés-Vilaplana, A., Moreno, D.A. & García-Viguera, C. (2011). Pomegranate fruit for health promotion: myths and realities. In: A. Jiménez, C. García-Viguera (Eds.), Antioxidant Properties of Crops III, 5, Functional Plant Science and Biotechnology (pp. 33-42). Global Science Books.

Mohammadi, V., Kheralipour, K. & Ghasemi-Varnamkhastia, M. (2015). Detecting maturity of persimmon fruit based on image processing technique. *Scientia Horticulturae 184*, 123-128.

Muñoz, I., Rubio-Celorio, M., Garcia-Gil, N., Guàrdia, M.D. & Fulladosa, E. (2015). Computer image analysis as a tool for classifying marbling: A case study in dry-cured ham. *Journal of Food Engineering 166*, 148–155.

Nakariyakul, S. & Casasent D. P. (2011). Classification of internally damaged almond nuts using hyperspectral imagery. *Journal of Food Engineering 103*, 62–67.

Ngadi, M.O. & Liu, L. (2010). Hyperspectral Image Processing Techniques. In: Sun, D.W. (Ed). Hyperspectral Imaging for Food Quality Analysis and Control (Pp 99-127). Elsevier.

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W, Theron, K.I. & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology 46*, 99-118 .

Nogales-Bueno, J., Rodríguez-Pulido, F.J., Heredia, F.J. & Hernández-Hierro, J.M. (2015a). Comparative study on the use of anthocyanin profile, color image analysis and near-infrared hyperspectral imaging as tools to discriminate between four autochthonous red grape cultivars from La Rioja (Spain). *Talanta 131*, 412–416.

Nogales-Bueno, J., Baca-Bocanegra, B., Rodríguez-Pulido, F. J. Heredia, F.J. & Hernández-Hierro, J. M. (2015b). Use of near infrared hyperspectral tools for the screening of extractable polyphenols in red grape skins. *Food Chemistry 172*, 559–564.

Nogales-Bueno, J., Hernández-Hierro, J.M., Rodríguez-Pulido, F. J. & Heredia, F.J. (2014). Determination of technological maturity of grapes and total phenolic compounds of grape skins in red and white cultivars during ripening by near infrared hyperspectral image: A preliminary approach. *Food Chemistry 152*, 586–591.

Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L. & Engelsen, S.B. (2000). Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. Society for Applied Spectroscopy 54, 413-419.

Nuñez-Sánchez, N., Martínez-Marín, A.L., Polvillo, O., Fernández-Cabanás, V.M., Carrizosa, J., Urrutia, B., & Serradilla, J.M. (2016). Near Infrared Spectroscopy (NIRS) for the determination of the milk fat fatty acid profile of goats. *Food Chemistry 190*, 244-252.

Pan, L., Zhang, Q., Zhang, W., Sun, Y., Hu, P. & Tu, K. (2016). Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chemistry 192*, 134–141.

Pérez-Marín, D., Garrido-Varo, A., & Guerrero, J. E. (2007). Non-linear regression

methods in NIRS quantitative analysis. *Talanta 72*, 28-42.

Qin, J., Chao, K., Kim, M.S., Lu, R. & Burks, T.F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering 118*, 157–171.

Rady, A., M., Guyer, D.E., Kirk, W. & Donis-González, I.R. (2014). The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. *Journal of Food Engineering 135*, 11–25.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G.S.V. & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194–200.

Reig, G., Alegre, S., Gatius, F., & Iglesias, I. (2013). Agronomical performance under Mediterranean climatic conditions among peach [Prunus persica (L.) Batsch] cultivars originated from different breeding programs. *Scientia Horticulturae 150*, 267–277.

Rinnan, Å., van den Berg, F. & Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry 28*, 1201-1222.

Rodríguez-Pulido, F.J., Hernández-Hierro, J.M., Nogales-Bueno, J., Gordillo, B., González-Miret, M. L. & Heredia, F. J. (2014). A novel method for evaluating flavanols in grape seeds by near infrared hyperspectral imaging. *Talanta 122*, 145–150.

Rodríguez-Pulido, F.J., Barbin, D.F., Sun, D.W., Gordillo, B., González-Miret, M.L. & Heredia, F.J. (2013). Grape seed characterization by NIR hyperspectral imaging. *Postharvest Biology and Technology 76,* 74-82.

Schmilovitch, Z., Ignat, T., Alchanatis, V., Gatker, J., Ostrovsky, V. & Felföldi, J. (2014). Hyperspectral imaging of intact bell peppers. *Biosystems Engineering 117*, 83-93.

Serranti, S., Cesare, D., Marini, F. & Bonifazi, G. (2013). Classification of oat and groat kernels using NIR hyperspectral imaging. *Talanta 103*, 276–284.

Simko, I., Jimenez-Berni, J. A. & Furbank, R.T. (2015). Detection of decay in fresh-cut lettuce using hyperspectral imaging and chlorophyll fluorescence imaging. *Postharvest Biology and Technology 106*, 44–52.

Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2010). Identification of insect-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging. *Computers and Electronics in Agriculture 73,* 118–125.

Sture, Ø., Øye, E.R., Skavhaug, A. & Mathiassen, J.R. (2016). A 3D machine vision system for quality grading of Atlantic salmon. *Computers and Electronics in Agriculture 123*, 142–148.

Szymanska, E., Gerretzen, J., Engel, J., Geurts, B., Blanchet, L. & Buydens, L.M.C (2015). Chemometrics and qualitative analysis have a vibrant relationship. *Trends in Analytical Chemistry 69*, 34–51

Tao, Y. & Wen, Z. (1999). An adaptive spherical image transform for high-speed fruit defect detection. *Transactions of the ASAE 42,* 241-246.

Teena, M.A., Manickavasagan, A., Ravikanth, L. & Jayas, D.S. (2014). Near infrared (NIR) hyperspectral imaging to classify fungal infected date fruits. *Journal of Stored Products Research 59,* 306-313.

USDA. (2018). Stone Fruit annual report. Available on-line at https://gain.fas.usda.gov/. Last visit May 2019.

Vélez-Rivera, N., Gómez-Sanchis, J., Chanona-Pérez, J., Carrasco, J.J., Millán-Giraldo, M., Lorente, D., Cubero, S. & Blasco, J. (2014). Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. *Biosystems engineering 122*, 91-98.

Vidal, M. & Amigo, J.M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems 117*, 138–148.

Wang, W., Li, C., Tollner, E.W., Gitaitis, R.D. & Rains, G.C. (2012). Shortwave infrared hyperspectral imaging for detecting sour skin (Burkholderia cepacia) infected onions. *Journal of Food Engineering 109*, 38–48.

Wang, J., Nakano, K., Ohashi, S., Kubota, Y., Takizawa, K. & Sasaki, Y. (2011). Detection of external insect infestations in jujube fruit using hyperspectral reflectance imaging. *Biosystems engineering 108*, 345-351.

Westad, F. & Marini, F. (2015). Validation of chemometric models – A tutorial. *Analytica Chimica Acta 893*, 14-24.

Williams, P. J., & Kucheryavskiy, S. (2016). Classification of maize kernels using NIR hyperspectral imaging. *Food Chemistry 209*, 131-138.

Williams, P.C. (1987). Variables affecting near-infrared reflectance spectroscopic analysis. In: Williams, P., Norris, K. (Eds). Near-infrared Technology in the Agricultural and Food Industries. (pp. 143–166). St. Paul, MN: American Association of Cereal Chemists.

Wu, D. & Sun, D.W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part I: Fundamentals. *Innovative Food Science and Emerging Technologies 19*, 1–14.

Xiaobo, Z., Jiewen, Z., Povey, M.J.W., Holmes, M. & Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. Analytica Chimica Acta 667, 14-32.

Yang, Y.C., Sun, D.W., Pu, H., Wang, N.N. & Zhu, Z. (2015). Rapid detection of anthocyanin content in lychee pericarp during storage using hyperspectral imaging coupled with model fusion. *Postharvest Biology and Technology 103*, 55–65.

Yang, C.C., Kim, M.S., Millner, P., Chao, K., Cho, B.K., Mo, C., Lee, H. & Chan, D.E. (2014). Development of multispectral imaging algorithm for detection of frass on mature red tomatoes. *Postharvest Biology and Technology 93,* 1–8.

Yang, C.C., Kim, M.S., Kang, S., Cho, B.K., Chao, K., Lefcourt, A.M. & Chan, D.E. (2012). Red to far-red multispectral fluorescence image fusion for detection of fecal contamination on apples. *Journal of Food Engineering 108*, 312–319.

Yao, H., Hruska, Z., Kincaid, R., Brown, R.L. Bhatnagar, D. & Cleveland, T.E. (2013). Detecting maize inoculated with toxigenic and atoxigenic fungal strains with fluorescence hyperspectral imagery. *Biosystems engineering 115*, 125-135.

Yu, K., Zhao, Y., Li, X., Shao, Y., Zhu, F. & He, Y. (2014). Identification of crack features in fresh jujube using Vis/NIR hyperspectral imaging combined with image processing. *Computers and Electronics in Agriculture 103*, 1–10.

Zapotoczny, P., Szczypinski, P.M. & Daszkiewicz, T. (2016). Evaluation of the quality of cold meats by computer-assisted image analysis. *LWT - Food Science and Technology 67*, 37-49.

Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C. & Huang, D. (2015). Hyperspectral

imaging combined with multivariate analysis and band math for detection of common defects on peaches (Prunus persica). *Computers and Electronics in Agriculture 114*, 14–24.

Zhu, N., Lin, M., Nie, Y., Wu, D. & Chen, K. (2016). Study on the quantitative measurement of firmness distribution maps at the pixel level inside peach pulp. *Computers and Electronics in Agriculture 130*, 48–5.

OBJECTIVES

General objective

The main objective of this doctoral thesis is to investigate the potential of hyperspectral imaging combined with chemometrics to be applied as a non-destructive postharvest tool for the evaluation of the quality of fruit such as nectarine, persimmon, pomegranate and loquat.

Specific objectives

1. To investigate the potential of hyperspectral reflectance imaging to predict the internal quality of nectarine 'Big Top' and 'Magique' by means of new ripening indices (RPI and IQI) through models based on PLS (**Chapter I**).
    - To create a tool to visualise the ripeness of each fruit by projecting the models on the pixels of the fruits in the images.
    - To test the performance of the indices in two cultivars with different physicochemical properties.
2. To investigate the possible use of hyperspectral imaging in transmittance mode as a non-destructive tool to estimate ripeness in nectarines 'Big Top' and 'Magique' and, at the same time, to detect split pit defect (**Chapter II**).
3. To develop statistical predictive models capable of distinguishing cultivars of nectarines, 'Big Top' and 'Diamond Ray', with a very similar appearance but different taste (**Chapter III**).
    - To investigate two approaches, based on the analysis of the individual spectrum of each pixel and on the mean spectrum of each fruit.
    - To visualise the result over the images of nectarines.
4. To propose a new non-destructive approach based on hyperspectral imaging and multivariate analysis to determine the F, ripeness state and astringency level of intact persimmon cv. 'Rojo Brillante' as alternative to the current destructive and/or subjective techniques (**Chapter IV**).

5. To study the application of hyperspectral imaging to predict the ST content in persimmon fruits cv. 'Rojo Brillante' in order to correctly discriminate A from DA persimmons using 0.04 % of ST as the threshold (**Chapter V**).
    - To determine which part of the fruit is the most appropriate to measure and obtain the ST content to be able to make this prediction.
    - To reduce the amount of spectral information generated and to speed up this process.
6. To evaluate the capability of both machine vision techniques, colour and hyperspectral imaging, to predict the physicochemical properties and the maturity stage of pomegranate fruits cv. 'Mollar de Elche' using the information of the intact fruit and arils (**Chapter VI**).
7. To develop classification models to discriminate common defects of loquat cv. 'Algerie' by using hyperspectral imaging combined with two robust machine learning techniques, RF and XGBOOST (**Chapter VII**).

I. NECTARINE

CHAPTER I

Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging

Sandra Munera [a], José Manuel Amigo [b], Sergio Cubero [a], Pau Talens [c], José Blasco [a] and Nuria Aleixos [d]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Department of Food Sciences, Faculty of Science, University of Copenhagen, Rolighedsvej 30, Frederikberg C DK-1958, Denmark

[c] Departamento de Tecnología de Alimentos, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[d] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

Journal of Food Engineering 214 (2017), 29-39

Abstract

Visible-near-infrared (450-1040 nm) hyperspectral reflectance imaging was studied in order to assess the internal physicochemical properties and sensory perception of 'Big Top' and 'Magique' nectarines (*Prunus persica* L. Batsch var. *nucipersica*) (yellow and white-flesh cultivar, respectively) during ripening using the RPI and the IQI. Hyperspectral images of the intact fruits were acquired during the ripeness under controlled conditions, and their physicochemical properties (flesh F, TSS, TA and flesh colour) were analysed. IQI and RPI were used to relate the spectral information obtained from nectarines with the physicochemical properties and the sensory perception of their maturity using PLS-R with proper variable selection. Optimal results were obtained with R^2 values higher than 0.87 for the two indices and the two cultivars. The ripeness of each fruit could be visualised by projecting the PLS-R models of the IQI on the pixels of the fruits in the images, showing great potential for further monitoring of the evolution of intact nectarine ripeness in industrial setups.

1. Introduction

During recent years the production of nectarines (*Prunus persica* L. Batsch var. nucipersica) has increased progressively. Currently, Europe produces almost 1.5 million tonnes of nectarines, Spain being the second producer after Italy, with an annual production of over 0.5 million tonnes (Europ^ech, 2016). As the market has become more demanding, the improvements in nectarine production have been focused on fruit presentation, with special emphasis on the colour, size and shape, on the ease with which they can be handled, and on increasing the sensory attributes. Nowadays, the market offers sweet, semi-sweet, balanced, acidic and highly acidic cultivars, of which the first two cultivars are the most widely accepted by consumers (Iglesias, 2012).

As a climacteric fruit, nectarine undergoes a 'ripening phase' that transforms the mature fruit into 'ready to eat'. This phase is associated with increased respiration and ethylene production, softening, loss of green colour and development of yellow or red, and the production of its characteristic aroma (Ritenour et al., 1997). Although to date no minimum quality level has been established, it is essential to understand the changes occurring in this fruit for its successful manipulation, transport and marketing. Nectarine ripeness and quality in general have been traditionally measured using the physical and chemical properties that best describe this progress: F, colour, TS) and/or TA. However, when quality is measured from the perspective of consumers, these parameters do not always match the consumer's preferences (Echeverría et al., 2015). Available data indicate that only measuring F, TSS or TA in the flesh give no satisfactory minimum ripening index; those properties change from one cultivar to another and for a given cultivar in relation to fruit size, climatic conditions, and cultural practices (Crisosto, 1994). For this reason, single physicochemical parameters are not always satisfactory ripening estimators and could be more useful to combine them in indices (Crisosto, 1994). On the other hand, changes in the flesh colour are not affected by sunlight and, thus, can be more reliable to estimate maturity. In this context, Vásquez-Caicedo et al. (2005) proposed a RPI for mangoes which was based on the combination of different properties of the fruit: F, TA and TSS; while Cortés et al (2016) proposed an

IQI that avoids the use of TA and introduced colour parameters such as *L**, *C**, and *h** in CIELCh coordinates.

Any index giving a comprehensive account of the quality should preferably be objective (a measurement) rather than subjective (an evaluation) and, ideally, non-destructive. When destructive measurements are used, the tendency is to use as few samples as possible, which often results in increased lot-to-lot variability in the quality index. Thus, sample variability becomes a factor to consider during laboratory studies and/or commercial applications (Valero et al., 2007). Numerous studies have examined the application of non-destructive technologies for quality assessment in stone fruits. Sonego et al (1995) detected nectarine woolliness using NMR imaging and X-ray computed tomography. Later, Arana et al. (2005) estimated this property using mechanical impacts. Wang et al (2006) and Diezma- Iglesias et al. (2006) assessed the mechanical properties of peaches by the excitation dynamic characteristics also through mechanical impact. X-rays were also used for the detection of changes in internal quality in peaches by Barcelon et al., 2009, and Pereira et al. (2013) used an MNR spectrometer to classify plums according to sweetness. In the search for a method of non-destructive analysis, NIR spectroscopy is especially widespread (Cortés et al., 2017; Lorente et al., 2015) and has been used to classify intact peaches or nectarines according to their degree of ripeness or different irrigation strategies (Carlomagno et al., 2004; Pérez-Marín et al., 2011). The flesh colour of clingstone peaches was assessed by Slaughter et al. (2013) using interactance spectroscopy without cutting the skin. Time-resolved reflectance spectroscopy was used by Zerbini et al. (2006) and Tijskens et al. (2007) to obtain a model of softening and assessing harvest maturity of nectarines. On the other hand, colour image analysis has been applied for in-line verification of nectarine cultivar (Font et al., 2014) or for assessing quality and marketability of fresh-cut nectarines (Pace et al., 2011).

Hyperspectral imaging is a non-destructive technology that integrates spectroscopy and conventional imaging to obtain both spatial and spectral information from an object simultaneously. The resulting spectrum for each pixel acts like a fingerprint, which can be used to characterise the composition of that particular pixel in the image, something that is not possible with conventional spectroscopy. It allows visualisation

of the biochemical constituents of a sample, separated into particular areas of the image, since regions of a sample with similar spectral properties have similar chemical composition (Gowen et al., 2007). VIS-NIR hyperspectral imaging has been applied as a powerful process analytical tool for rapid, non-destructive inspection of the internal and external quality attributes in fruits such as banana (Rajkumar et al., 2012), pear (Li et al., 2016a), persimmon (Munera et al., 2017), citrus fruits (Gómez-Sanchis et al., 2013), grapes (Baiano et al., 2012), blueberries (Leiva-Valenzuela et al., 2013) or apples (Baranowski et al., 2013), as well as in pepper (Schmilovitch et al., 2014), tomato (Liu et al., 2015) or potato (López-Maestresalas et al., 2016). In regards to stone fruit, peaches are the most studied. Several studies have been performed with the aim of detecting different types of defects. Zhang et al. (2015) and Li et al. (2016b) detected common defects in skin while Pan et al. (2016) and Sun et al. (2017) used this technology to detect chilling injury. In nectarines, Huang et al. (2015) used this technology for the detection of defective features. Apart from defects, few works have been carried out to estimate other properties. Lu and Peng (2006) presented one of the first works to detect firmness in peaches using hyperspectral scattering, later, Lleó et al. (2011) classified peaches by maturity using multispectral indices.

This technique offers an immense amount of spectral and spatial information for each sample. For this reason, chemometrics is an indispensable tool for reducing the dimensionality of the data, retaining essential spectral information and classifying and quantifying important areas of the scene (Amigo et al., 2013; Lorente et al., 2012). PCA is one of the chemometric methods which have been specially designed as a tool for obtaining an overview of the main source of variance in individual and set samples (Amigo et al., 2015). PLS is another method used for constructing predictive models. Unlike PCA, PLS is a reliable form of analysis, directed towards factor spaces that are associated with high variation in the responses but biased towards directions that are accurately predicted. PLS is widely used in hyperspectral imaging to extract and summarise spectral information from hyperspectral images, to reduce the high dimensionality of the spectral data and to overcome the problem of multi-collinearity (Vinzi et al., 2010).

The objectives of this work are: i) to investigate the potential of hyperspectral reflectance imaging in the VIS-NIR (450-1040 nm) to predict for the first time the internal quality of nectarine by means of new ripening indices (RPI and IQI) through models based on PLS, ii) to create a new tool to visualise the ripeness of each fruit by projecting the models on the pixels of the fruits in the images, and iii) to test the performance of the indices in two cultivars with different physicochemical properties.

2. Material and methods

2.1. Fruit samples

In this study two cultivars of nectarine were used. 'Big Top', with yellow flesh, and 'Magique', a white-flesh cultivar, were harvested in a commercial orchard in Lerida (Spain) in the commercial maturity period during the 2015 season. Fruits without any defects or bruises were selected and grouped in batches of 25 samples.

Then the fruits were stored under controlled conditions (15ºC – 90 % relative humidity) until senescence. Due to the fact that they are different cultivars (as can be seen in Fig. 1), the ripening times for each of them were also different. For this reason both cultivars were considered separately and the analyses of 'Big Top' nectarines were performed before storage and then on the 1st, 2nd, 3rd, 5th and 8th day (150 fruits in total), and for 'Magique' nectarines the analyses were performed before storage and then on the 2nd, 4th, 7th, 10th and 14th day (150 fruits in total).

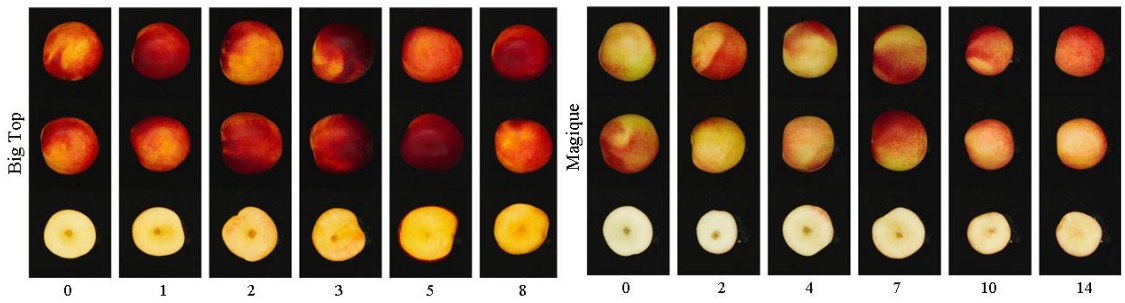


Figure 1. External and internal appearance of both cultivars of nectarine during postharvest storage.

2.2. Hyperspectral imaging acquisition

The hyperspectral system (Fig. 2) was composed of an industrial camera (CoolSNAP ES, Photometrics, AZ, USA), coupled to two LCTF (Varispec VIS-07 and NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA), capable of acquiring images at 60 different wavelengths every 10 nm in the working spectral range of 450 nme1040 nm. The system was configured to capture images of 1392 x 1040 pixels with a spatial resolution of 0.14 mm/pixel. To avoid problems of unfocused images due to the refraction of light across this wide spectral range, the focus was adjusted on the central band of the acquisition interval (740 nm) and the images were captured using lenses capable of covering the whole spectral range without going out of focus (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA). To optimise the dynamic range of the camera, prevent saturated images and correct the spectral sensitivity of the different elements of the system, a calibration of the integration time of each band was performed. For the reflectance mode, the integration time of each band was calibrated to capture the averaged grey level of a white reference target (Spectralon 99%, Labsphere, Inc, NH, USA) corresponding to 90% of the dynamic range of the camera.

The illumination system consisted of 12 37-watt halogen spotlights (Eurostar IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current (12 V). The scene was lit indirectly by means of diffuse reflection inside a hemispherical dome, where the fruits were introduced manually upon a fruit holder, with the stem-apex axis lying horizontal. The inner surface of the aluminium dome was painted white in order to maximise its reflectivity, and given a rough texture in order to minimise directional reflections, which could cause bright spots, thus providing highly homogeneous light.

Two hyperspectral images per fruit (side A and B) were acquired in reflectance mode using customised software developed at IVIA.

2.3. Reference analysis

In this work, RPI and IQI are calculated as shown in equations (1) and (2) and are used to relate the spectral information obtained from nectarines with the physicochemical properties and the sensory perception of their maturity.

$$RPI = \ln \frac{100 \times F \times TA}{TSS} \quad (1)$$

$$IQI = \ln \frac{100 \times F \times L^* \times H^*}{TSS \times C^*} \quad (2)$$

Both indices give an estimate of fruit maturity, but the advantage of IQI over RPI is that the first one replaces the TA measure, which is complex and time-consuming, for colour parameters that can be more easily obtained with a colorimeter or even a colour camera. All parameters listed in Equations (1) and (2) were analysed immediately after hyperspectral imaging acquisition. First, colour images were acquired to observe the external and internal (after F analysis) appearance of each fruit (Fig. 1). The image acquisition system consisted of a digital camera (EOS 550D, Canon Inc, Japan) introduced into a square inspection chamber that included a calibrated and uniform illumination system composed of four lamps, each containing two BIOLUX 18W/965 fluorescent tubes (Osram GmbH, Germany) with a colour temperature of 6500 K. The angle between the axis of the lens and the sources of illumination was approximately 45° and polarising filters were placed in front of the lamps and in the camera lenses to avoid direct reflections on the camera.

The analysis of F was performed on two opposite sides of each fruit using a XT2 Stable texturometer (MicroSystems Haslemere, UK) provided with a 6mmflat plunger. The crosshead speed during the puncture test was 1 mm/s. The maximum force, expressed in Newton (N), was registered on opposite sides of the fruits.

Flesh colour was analysed using a MINOLTACM-700d colorimeter (Minolta Co. Tokyo, Japan) with the standard illuminant D65 and the observer 10º. L*, C* and h* parameters were obtained at the CIELCh space. Then the juice of each nectarine was used to analyse TA and TSS compounds.

TA was determined using a Crison pH-Burette 24 automatic titrator (Crison, Barcelona, Spain) and NaOH 0.5 N, according to standard UNE34211:1981 (AENOR, 1981). Results were expressed as % of citric acid. TSS was determined using a digital refractometer RFM330 þ VWR (Internacional Eurolab S.L., Barcelona, Spain) at 20 ºC and results were expressed as % of TSS.

2.4. Image processing

A total of 300 images of each cultivar were obtained and imported into the commercial software MATLAB R2015a (The MathWorks, Inc. MA, USA) to be pre-processed using the customised toolbox, called HYPER-Tools (Amigo et al., 2015).

The first step of the image processing consisted in the correction of the relative reflectance using Eq. (3) (Gat, 2000):

$$\rho_{xy}(x,y,\lambda) = \frac{R^{abs}}{R^{abs}_{white}} = \rho^{Ref}(\lambda)\frac{R(x,y,\lambda)-R_{black}(x,y,\lambda)}{R_{white}(x,y,\lambda)-R_{black}(x,y,\lambda)} \quad (3)$$

where $\rho^{Ref}(\lambda)$ is the standard reflectance of the white reference target (99% in this work), $R(x,y,\lambda)$ is the reflectance of the fruit captured by the CCD sensor of the camera, $R_{white}(x,y,\lambda)$ is the reflectance captured by the CCD of the white reference target, and $R_{black}(x,y,\lambda)$ is the reflectance captured by the CCD while avoiding any light source in order to quantify the electronic noise of the CCD.

The corrected images were clipped and compressed to obtain images with a new dimension of 256 x 163 pixels and a spatial resolution of 0.56 mm/pixel in order to reduce the computation time. With these new dimensions, a hypercube was generated by joining images of a certain number of samples in the image depending on the calibration or prediction set. After proper removal of the background using K-means, SNV pre-treatment was applied to the NIR region due to the scattering effects, which are one of the main drawbacks of this region (Vidal and Amigo, 2012).

Finally, the average reflectance spectrum was determined by averaging the relative reflectance spectra of a central square region of interest on each side of each fruit. Altogether, 150 average spectra representing all the nectarines of each cultivar were

obtained for their ripeness prediction using PLS-R models. With the model thus obtained, the prediction was performed on the whole surface of the nectarine.

2.5. Data analysis

After the pre-processing steps indicated in the previous section (2.5), the whole dataset of each cultivar was randomly partitioned into two separate folds. 68% of the samples (204) were used for the calibration (C) of the models and the remaining 32% (96) were used for independent testing or prediction. The PLS models were then used to predict the independent set of the samples.

Before the quantitative analysis, PCA was performed to explore the image of nectarines in an unsupervised manner. In this case, the images from the calibration set were used with the aim of visualising the changes on the surface of each fruit during storage.

PLS regressions were performed using the PLS_Toolbox (Eigenvector Research Inc., USA) working under MATLAB in order to correlate the spectral response and each of the quality indices considered (RPI and IQI). The input spectra (NIR region previously pre-treated using SNV) were normalised using mean-centering. A single 10-fold venetian blind cross-validation (i.e. splitting evenly the data into 10 sets and leaving each one of the sets out in each iteration of the validation procedure) was used on the calibration set to choose the optimal number of LVs as well as to obtain an estimate of the error rate of the PLS models. The accuracy of the PLS models and predictive capability were evaluated by the RMSE and R^2 between the predicted and the measured values of the quality indices for calibration, CV and prediction. Furthermore, the RPD, defined as the ratio between the standard deviation of the reference data and $RMSE_P$, was also used (Williams, 1987).

In order to visualise the ripeness in the fruit surface of the prediction set, the index of each pixel within nectarine was calculated by inputting the extracted spectrum of the corresponding pixel into the previously established PLS model. Then the positions of all corresponding pixels were used to visualise the distribution of the predicted

value in each sample. In this case, each side of each fruit was considered a sample. All the steps for visualisation were implemented using the HYPER-Tools toolbox.

ANOVA, followed by Tukey's HSD test, was conducted to determine significant differences in the reference properties of the fruit during ripeness using the software Statgraphics (Manugistics Corp., Rockville, USA).

3. Results and discussion

3.1. Physicochemical analysis

Table 1 summarises the means and standard deviations of the physicochemical properties measured on nectarine samples of each cultivar during storage. F ranged from 47.3 to 8.8 and 57.9 - 6.1 N for 'Big Top' and 'Magique', respectively. These changes are due to pectin solubilisation and degradation by enzymes acting on cell walls, whose activity is related to ethylene biosynthesis and action. Cellulases are active in the first stage of slow softening. Later, in the so-called 'melting' phase, the combined action of pectinmethylesterase and polygalacturonase resulted in a large decline in firmness in a few days at room temperature (Zerbini et al., 2006). F thresholds have been defined to describe bruising thresholds and identification of important ripening stages (Valero et al., 2007), finding that fruit above 35 N are significantly less susceptible to be bruised, between 18 N and 35 N were described as 'ready to buy' and below 18 N as 'ready to eat'.

The TSS of the 'Big Top' cultivar increased from 9.8 % to 15.0 % until the last day of storage, when it decreased until 12.0%. In the case of 'Magique', TSS values were stable during ripening and significant differences were only found between day 0 and the last day of storage (10.7-11.4 %). Something similar happened for TA. Both cultivars maintained the same value until the last day of storage, when this parameter changed. According to Iglesias and Echeverría (2009), TSS below 10 % is generally unacceptable to consumers. But to reach good consumer acceptance (good flavour) it is important to achieve a minimum level of TA (which has not yet been established) and surpass a certain level of TSS, which the authors set at 12 %, otherwise the taste is

judged as lacking in flavour, insipid or flat. In this case, both cultivars achieved higher values than the specified minimum value of acceptability, especially 'Big Top'.

Table 1. Results of physicochemical properties of both cultivars of nectarine during postharvest storage.

	Day	F (N)	TSS (%)	TA (%)	Flesh colour parameters		
					*L**	*h**	*C**
'Big Top'	0	47.3 ± 5.1 ^{a}	9.8 ± 1.2 ^{d}	0.4 ± 0.1 ^{a}	72.1 ± 1.9 ^{a}	79.2 ± 2.1 ^{a}	42.4 ± 2.4 ^{d}
	1	43.3 ± 5.3 ^{a}	10.6 ± 1.3 ^{cd}	0.4 ± 0.1 ^{a}	70.5 ± 1.8 ^{ab}	78.3 ± 2.5 ^{a}	45.2 ± 2.9 ^{c}
	2	31.4 ± 8.3 ^{b}	10.8 ± 1.4 ^{cd}	0.4 ± 0.1 ^{a}	69.2 ± 2.4 ^{b}	75.4 ± 3.4 ^{b}	44.7 ± 2.4 ^{cd}
	3	15.5 ± 4.2 ^{c}	12.7 ± 1.8 ^{b}	0.4 ± 0.1 ^{a}	65.3 ± 2.5 ^{c}	78.1 ± 2.8 ^{a}	56.0 ± 4.9 ^{b}
	5	12.5 ± 2.8 ^{cd}	15.0 ± 2.6 ^{a}	0.4 ± 0.1 ^{a}	66.3 ± 2.5 ^{c}	72.5 ± 2.3 ^{c}	46.1 ± 3.1 ^{c}
	8	8.8 ± 1.2 ^{d}	12.0 ± 1.6 ^{bc}	0.3 ± 0.1 ^{b}	62.2 ± 1.7 ^{d}	74.1 ± 1.2 ^{bc}	59.9 ± 3.8 ^{a}
'Magique'	0	57.9 ± 3.9 ^{a}	10.7 ± 0.8 ^{b}	0.5 ± 0.1 ^{ab}	72.1 ± 2.0 ^{ab}	102.1 ± 1.0 ^{a}	23.7 ± 1.5 ^{d}
	2	51.3 ± 5.2 b	10.5 ± 0.8 ^{ab}	0.5 ± 0.1 ^{ab}	73.4 ± 1.9 ^{a}	100.2 ± 2.1 ^{b}	20.5 ± 1.8 ^{e}
	4	32.7 ± 9.4 ^{c}	10.3 ± 1.0 ^{ab}	0.6 ± 0.1 ^{a}	73.1 ± 1.8 ^{ab}	98.1 ± 2.1 ^{c}	21.3 ± 1.8 ^{e}
	7	11.8 ± 3.2 ^{d}	11.3 ± 1.5 ^{ab}	0.5 ± 0.1 ^{ab}	70.9 ± 2.6 ^{bc}	91.6 ± 3.0 ^{d}	25.4 ± 1.7 ^{c}
	10	7.4 ± 1.5 ^{e}	10.8 ± 1.5 ^{ab}	0.5 ± 0.1 ^{b}	69.3 ± 3.9 ^{c}	87.9 ± 1.4 ^{e}	33.1 ± 2.7 ^{a}
	14	6.1 ± 0.9 ^{e}	11.4 ± 1.4 ^{a}	0.4 ± 0.1 ^{c}	68.9 ± 2.7 ^{c}	84.8 ± 1.2 ^{f}	27.8 ± 2.6 ^{b}

Values are mean ± standard deviation. Different superscript letters in the same column and nectarine cultivar indicate significant differences between groups (p-value<0.05), according to Tukey's HSD test

Regarding the flesh colour, 'Big Top' nectarines are characterised by a high presence of yellow components or carotenoids, unlike 'Magique', where these are present in a very low proportion (Gil et al., 2002), as shown in Figure 1. Both cultivars obtained similar values of L* at day 0 of storage but 'Big Top' underwent a big reduction of this parameter. 'Magique' presented high values of h* (greenish-yellow above 85º) and a large reduction of this parameter. However, 'Big Top' presented high values of C* and underwent a higher increase in this value than the white-flesh cultivar. The loss in visual appearance is related to a reduction in lightness and variation of flesh colour described by those parameters (Pace et al., 2011).

3.2. Spectral analysis

The pre-processed (SNV) average spectra of the calibration set of both nectarine cultivars on different days of postharvest storage are illustrated in Figure 2. Both cultivars followed the same spectral pattern.

The difference in the VIS range was due to the colour feature of the samples over the entire visible colour spectral range. In this case the differences between 650 and 700 nm belong to chlorophylls, as reported by Lleó et al. (2011) or Rajkumar et al. (2012). Riper fruits reflected more light because these molecules are degraded during ripeness. Furthermore, the 'Magique' cultivar obtained lower values of reflectance in these bands, possibly due to higher chlorophyll content than the yellow-flesh cultivar. Carotenoids are present in the 450-600 nm range, but in this case no large differences were observed.

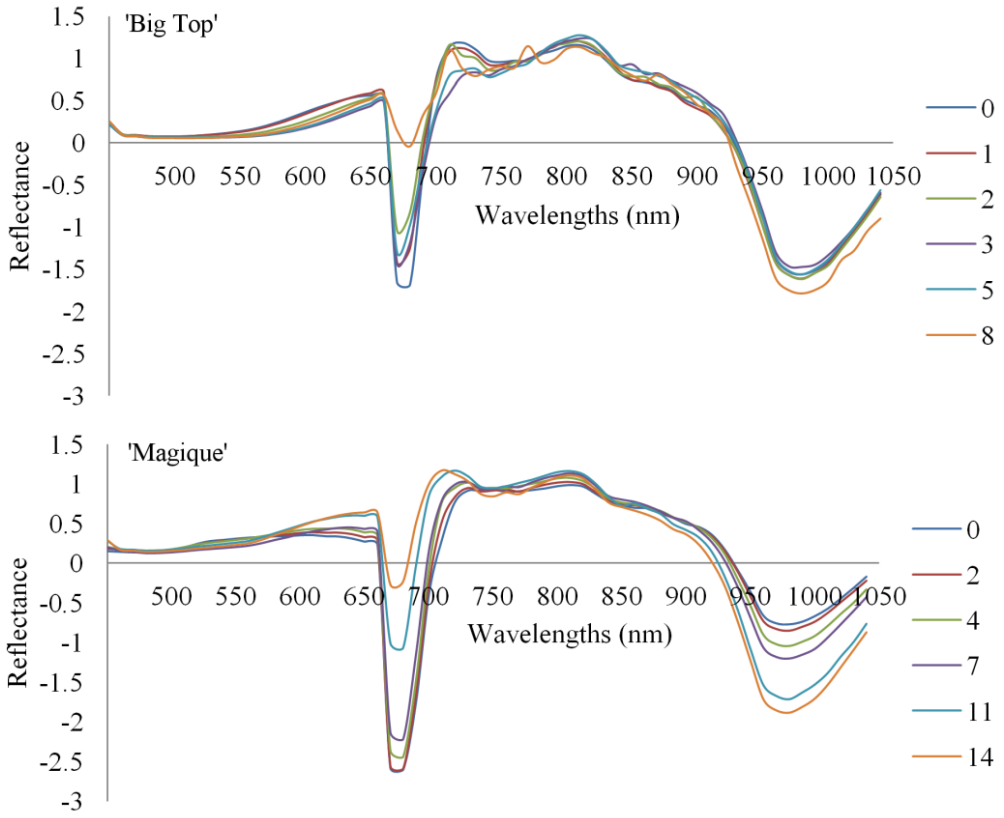


Figure 2. Average spectra of calibration samples of 'Big Top' and 'Magique' cultivars on each day of analysis.

Differences in the NIR region could be attributed to the chemical differences among nectarines at different maturity stages (Liu et al., 2015). In this case, the main differences were localised in the reflection valley around 950-1000 nm, primarily assigned to water absorption bands (Lu and Peng, 2006; Lleó et al., 2011). This valley was more pronounced in riper fruits because the water content increases in the flesh during the onset of ripening, due to cell breakage and osmotic movement of water from the flesh to the peel (Rajkumar et al., 2012). The 'Magique' cultivar presented more differences in water content between days of analysis than 'Big Top'. This indicates that external and internal colour is not the only difference between these cultivars.

3.3. Principal components analysis

Results of PCA for the two cultivars are shown in Figure 3. In these figures, the colour of each pixel in the nectarine represents the score obtained by this pixel based on the colour scale at the right part of the images. The score maps were obtained by refolding the score vectors obtained for each factor, whereas the loadings were related to the spectral variability.

In the model for 'Big Top', the first two PCs were necessary to explain 88.9 % of the variance (72.5 % and 16.4 %, respectively). However, the first PC of the 'Magique' cultivar explained 90.7% of the total variance, whereas the second PC only explained 3.8 %. The images of the scores, the first and second PC for 'Big Top' and the first PC for 'Magique', seem to show an evolution of the ripening process of the fruit during postharvest storage. The corresponding loadings show that wavelengths between 650 nm and 730 nm and between 940 nm and 1040 nm (especially for 'Magique') could be important for the internal quality prediction of both cultivars. As commented before, these zones in the spectrum belong to chlorophylls and water absorbance, respectively. However, predictive models need to be investigated for use in quantitative analyses capable of identifying nectarine ripeness.

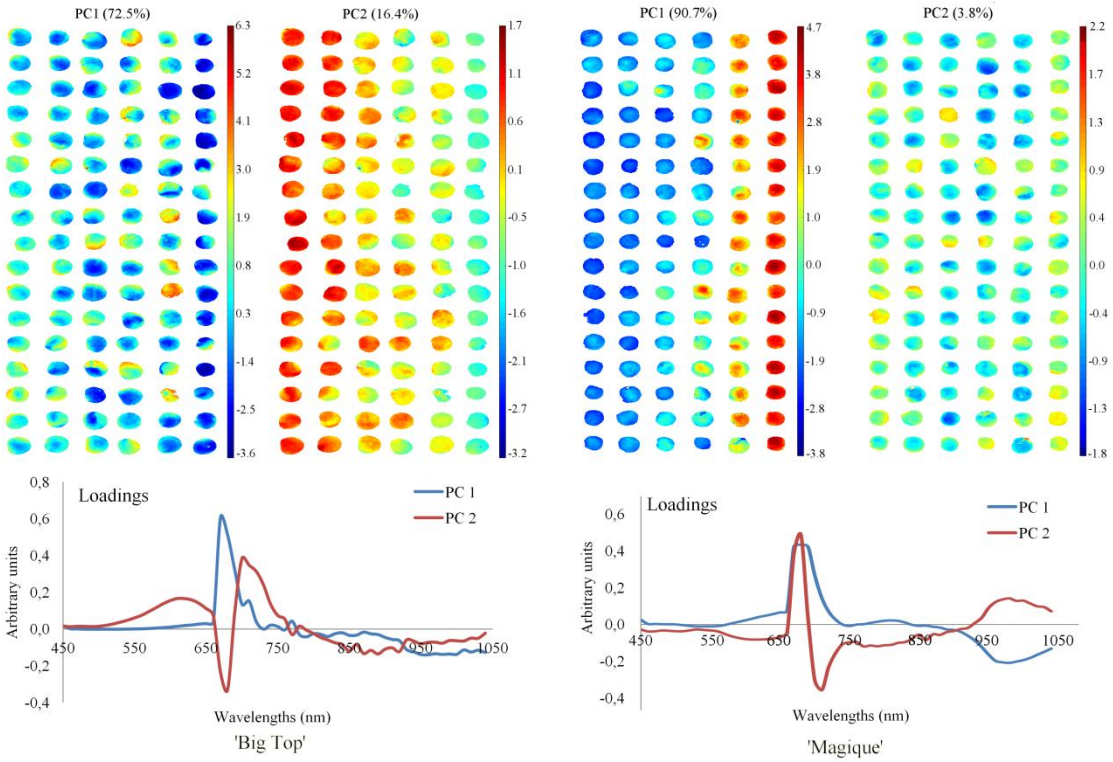


Figure 3. PCA model. Score surface and loadings for PC1 and PC2 of both cultivars.

3.4. PLS prediction of ripening indices

Figure 4 shows the evolution of the indices IQI and RPI for each cultivar during storage. As expected, in both cases the values of the indices decreased as the fruit matures, mainly due to the progressive decrease of F and increase of TSS (Table 1). The IQI performed better for 'Big Top', as it could discriminate all maturity stages measured. On the other hand, RPI could not properly separate all the stages for 'Big Top', but behaved better for 'Magique', which could be explained because the TA did not change for 'Big Top' throughout the experiments, while it decreased slightly for 'Magique', and also because the changes in the flesh colour showed that luminosity and chroma were higher for 'Big Top' than for 'Magique'.

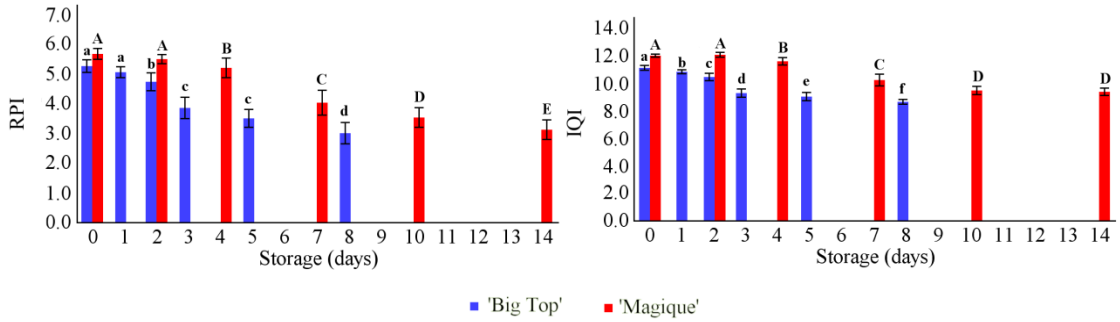


Figure 4. Evolution of the different ripening indices during storage of 'Big Top' and 'Magique' nectarines. *Columns are mean and bars are standard deviation. Different letters in each nectarine cultivar set indicate significant differences between groups (p-value<0.05), according to Tukey's (HSD) test.*

PLS-R models were performed to evaluate the internal quality of nectarines with the spectral range of 450-1040 nm. Table 2 shows the RPI and IQI prediction results for 'Big Top' and 'Magique' cultivars.

Table 2. Prediction of RPI and IQI by using PLS-R and all wavelengths for each cultivar.

		#W	#LV	R^2_C	$RMSE_C$	R^2_{CV}	$RMSE_{CV}$	R^2_P	$RMSE_P$	RPD
'Big Top'	RPI	60	7	0.91	0.26	0.87	0.31	0.87	0.37	2.4
	IQI	60	7	0.92	0.28	0.89	0.33	0.89	0.33	3.0
'Magique'	RPI	60	6	0.95	0.24	0.93	0.27	0.91	0.35	2.9
	IQI	60	6	0.93	0.31	0.91	0.35	0.89	0.44	2.7

W = wavelengths; LV = latent variables; C = calibration; CV = cross validation; P = prediction

The optimal model was chosen when the number of LV yields the lowest RMSE of calibration and CV ($RMSE_C$ and $RMSE_{CV}$). When the full spectral data were correlated with the ripeness indices, 7 LVs were determined for 'Big Top' and 6 LVs were determined for 'Magique'. In general, optimal predictions were obtained for both indices and cultivars. In 'Big Top', the R^2 and RMSE values were 0.87 and 0.37 for RPI and 0.89 and 0.33 for IQI. As in the case of 'Magique', these values were 0.91 and 0.35 and 0.89 and 0.44, respectively. Furthermore, according to Williams (1987), values of RPD between 2 and 2.5 indicate that coarse quantitative predictions are possible and a

value above 2.5 means good to excellent prediction accuracy. Taking these values into consideration, IQI was better predicted than RPI for the yellow flesh cultivar and RPI was somewhat better predicted than IQI for the white-flesh cultivar.

3.5. Variable selection

Due to the large amount of information contained in the hyperspectral images, much of which is redundant, time consuming and unsuitable for in-line inspection (Lorente et al., 2012), it is necessary to select the best-known wavelengths from the entire spectral data. The use of average spectra of the samples could significantly reduce the amount of data. However, the high dimensional full spectra suffered from co-linearity and redundancy of wavelengths, resulting in complex models and poor performances (Zhang et al., 2016).

VIP scores were used to select the best variables for predicting the proposed indices (Figure 5).

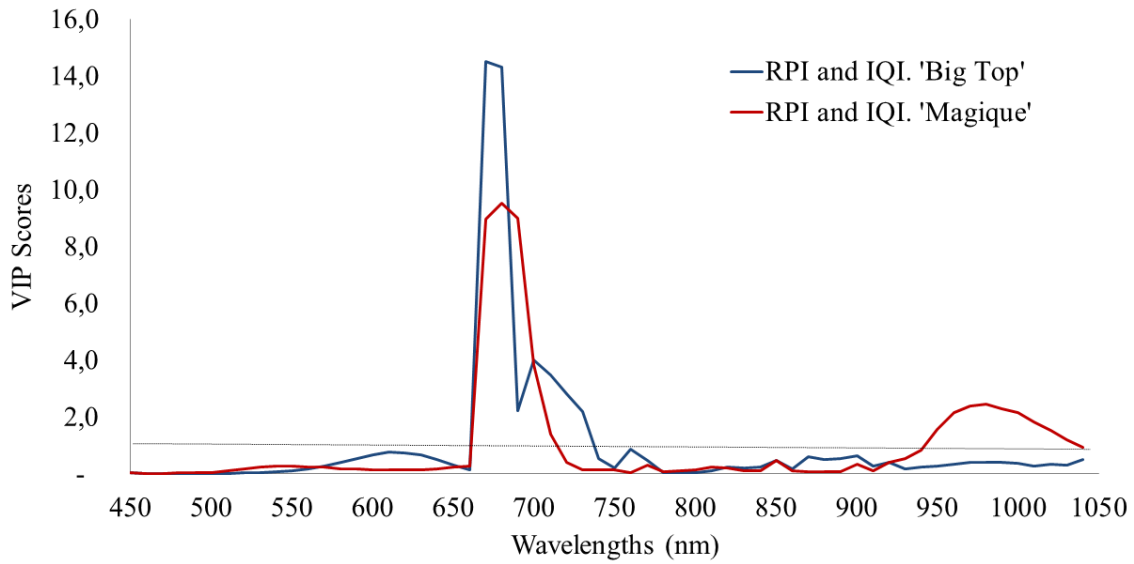


Figure 5. Optimal wavelength selection for prediction of the indices in both cultivars using VIP scores.

These are calculated as a weighted sum of the squared correlations between the PLS components and the original variables. The weights correspond to the percentage

variation explained by the PLS component in the model. The number of terms in the sum depends on the number of PLS components found to be significant in predicting. Variables with a VIP score close to or greater than 1 can be considered the ones making the highest contribution to prediction in the PLS-R model. While for 'Big Top' the optimal wavelengths to predict both RPI and IQI were located only in the VIS region (670-730 nm and 760 nm), for 'Magique' several wavelengths in the NIR region (670-700 nm and 970-990 nm) were also important.

As shown in Table 3, the PLS-R models using the optimal wavelengths maintained a similar performance to PLS-R models created with the full spectrum (Table 2). Likewise, their calibration and prediction errors do not worsen and both indices remain within the same range as in nectarine samples.

Table 3. Prediction of RPI and IQI by using PLS-R and optimal wavelengths for each cultivar

		#W	#LV	R^2_C	$RMSE_C$	R^2_{CV}	$RMSE_{CV}$	R^2_P	$RMSE_P$	RPD
'Big Top'	RPI	8	5	0.87	0.31	0.85	0.34	0.89	0.33	2.7
	IQI	8	5	0.87	0.35	0.86	0.37	0.90	0.32	3.1
'Magique'	RPI	7	5	0.88	0.34	0.87	0.36	0.90	0.36	2.8
	IQI	7	5	0.87	0.43	0.85	0.45	0.88	0.44	2.7

W = wavelengths; LV = latent variables; C = calibration; CV = cross validation; P = prediction

Values of RPD indicated that IQI was the best index predicted for the yellow-flesh cultivar. However, similar results were obtained for IQI and RPI in 'Magique'. The main differences between these two indices reside in the requirements of time and costs to obtain them. For this reason IQI is more suitable to be used as a standard index on an inspection line. In general, the prediction results using the optimal wavelengths were acceptable and revealed the potentiality of hyperspectral imaging as a rapid and non-destructive method to obtain the estimation of ripeness in nectarines.

One of the main advantages of hyperspectral imaging is being able to map the spatial distribution of different properties of the samples under study. In order to visualise the ripeness of the fruits throughout postharvest storage, the PLS-R model of RPI with optimum wavelengths was used to transfer the calibrated results of

multivariate analysis to each pixel of the image. Fig. 6 shows the predicted IQI value for each pixel in the nectarine according to the colour scale bar on the right side of the image. The colour scale ranges from red colours to show relatively high IQI values (lower ripeness) to blue colours to show lower IQI values indicating a higher degree of ripeness. This way, the difference in colour (and thus in ripeness) from sample to sample was easy to distinguish from the resulting images. Moreover, samples showed a variation of colours which indicates that ripeness in each fruit had a heterogeneous distribution, as also discussed Herrero-Langreo et al. (2011) for peaches.

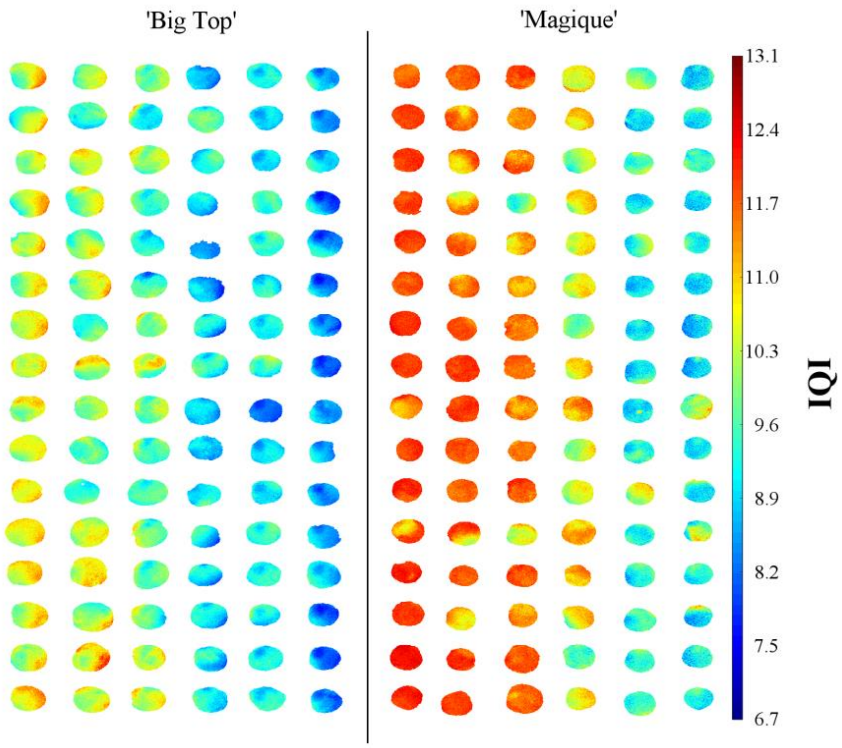


Figure 6. Visualisation of IQI prediction using PLS-R and optimal wavelengths for each cultivar of nectarine.

Towards an optimal variable selection study and implementation, the proposed methodology could be adapted into in-line sorting equipment. Hyperspectral systems are still expensive, and both the acquisition of images and the processing of spectral and spatial information require a great computational cost. Instead, the emergence in the market of new multispectral cameras that use custom Bayer-like matrix filters to

capture a reduced set of 9-15 wavelengths simultaneously at high speed, will allow the progressive incorporation of these systems for in-line operations in the industry. This will entail the challenge of finding the proper setup of discrete wavelengths towards a fast acquisition imaging system that gives us the ideal descriptive or quantitative results, which will require the design and development of new experiments with more fruits.

4. Conclusions

This work demonstrates the capability of hyperspectral imaging to monitor the ripeness of two cultivars of nectarine using ripening indices. The ripening evolution could be observed during storage through PCA of the spectral information in a non-supervised manner. On the other hand, PLS models produced optimal prediction for both cultivars of around R^2 = 0.90 of RPI and IQI indices, which were obtained by measuring the physicochemical properties destructively. In the case of the 'Big Top' cultivar, the R^2 values were 0.87 and 0.89 for RPI and IQI, respectively; while for the 'Magique' cultivar these values were 0.91 and 0.89, respectively.

However, hyperspectral systems capture a huge amount of data, most of which are redundant, which makes the process slow and hence it is important to obtain a reduced set of wavelengths that preserves the most variability of the problem. In our study, optimal wavelength selection was performed by means of VIP scores. A total of eight wavelengths in the VIS region were selected for 'Big Top' and seven for 'Magique', but in this second case some of them were also located in the NIR region. The simplified models also yielded good performance in prediction with R^2 values of around 0.90 and RPD higher than 2.5 for both indices and both cultivars. However, each cultivar needed a particular set of wavelengths.

The predictions of the IQI of each individual pixel in the spectral images were used to create new images that allowed the ripeness distribution (ripening maps) to be visualised within and between fruits using the calibrated PLS model, thus facilitating the visual observation of the state of fruit ripening.

Acknowledgements

This work has been partially funded by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria de España (INIA) through research project RTA2015-00078-00-00 with the support of European FEDER funds. Sandra Munera thanks INIA for the grant FPI-INIA num. 43 (CPR2014-0082), partially supported by European Union FSE funds.

References

AENOR. (1981). Productos derivados de frutas y verduras, determinación de la acidez valorable. UNE 34211: 1981

Amigo, J.M., Babamoradia, H. & Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta 896*, 34–51.

Amigo, J.M., Martí-Aluja, I. & Gowen, A. (2013). Hyperspectral Imaging and Chemometrics: A Perfect Combination for the Analysis of Food Structure, Composition and Quality. In: Marini, F (Ed.), *Data Handling in Science and Technology* (pp 343-370). Elsevier.

Arana, I., Jarén, C. & Arazuri, S. (2005). Nectarine woolliness detection by non-destructive mechanical impact. *Biosystems Engineering 90*, 37–45.

Baiano, A., Terracone, C., Peri, G. & Romaniello, R. (2012). Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Computers and Electronics in Agriculture 87*, 142–151.

Baranowski, P., Mazurek, W. & Pastuszka-Wózniak, J. (2013). Supervised classification of bruised apples with respect to the timeafter bruising on the basis of hyperspectral imaging data. *Postharvest Biology and Technology 86*, 249–258.

Barcelon, E. G., Tojo, S. & Watanabe, K. (2009). X-ray computed tomography for internal quality evaluation of peaches. *Journal of Agricultural Engineering Research 73*, 323-330.

Carlomagno, G., Capozzo, L., Attolico, G. & Distante, A. (2004). Non-destructive grading of peaches by near-infrared spectrometry. *Infrared Physics & Technology 46*, 23–29.

Cortés, V., Ortiz, C., Aleixos, N., Blasco, J., Cubero, S. & Talens, P. (2016). A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy. *Postharvest Biology and Technology 118*, 148–158.

Crisosto, C.H. (1994). Stone fruit maturity indices: a descriptive. *Postharvest News and Information 6*, 65-68.

Diezma-Iglesias, B., Valero, C., García-Ramos, F.J. & Ruiz-Altisent, M. (2006). Monitoring of firmness evolution of peaches during storage by combining acoustic and impact methods. *Journal of Food Engineering 77*, 926–935.

Echeverría, G., Cantín, C.M., Ortiz, A., López, M.L., Lara, I. & Graell, J. (2015). The impact of maturity, storage temperature and storage duration onsensory quality and consumer satisfaction of 'Big Top®' nectarines. *Scientia Horticulturae 190*, 179–186.

Europêch. Chambre d'Agriculture des Pyrénées-Orientales. (2016). Synthese de la recolte Europeenne 2015. Previsions de recolte 2016: Pêche , nectarine, pavie, pêches plates. http://www.agro-alimentarias.coop/ficheros/doc/04992.pdf . Accessed 20.08.16.

Font, D., Tresanchez, M., Pallejà, T., Teixidó, M., Martinez, D., Moreno, J., Palacín, J. 2014. An image processing method for in-line nectarine variety verification based on the comparison of skin feature histogram vectors. Computers and Electronics in Agriculture 102, 112–119.

Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report, Opto- Knowledge Systems Inc. OKSI.

Gil, M., Tomás-Barberán, F. A., Hess-Pierce, B., Kader, A. A. (2002). Antioxidant capacities, phenolic compounds, carotenoids, and vitamin C contents of nectarine, peach, and plum cultivars from California. *Journal of Agricultural and Food Chemistry 50*, 4976-4982.

Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology 18*, 590–598.

Herrero-Langreo, A., Lunadei, L., Lleó L., Diezma, B. & Ruiz-Altisent, M. (2011). Multispectral Vision for Monitoring Peach Ripeness. *Journal of Food Science 2*, 178-187.

Huang, F., Zhang, S., Yang, Y., Man, Z., Zhang, X. & Wu, Y. (2015). Application of hyperspectral imaging for detection of defective features in nectarine fruit. *Transactions of the Chinese Society for Agricultural Machinery 11,* 252-259.

Iglesias, I. & Echeverría, G. (2009). Differential effect of cultivar and harvest date on nectarine colour, quality and consumer acceptance. *Scientia Horticulturae 120*, 41–50.

Iglesias, I. (2012). Producción, consumo e innovación varietal en el melocotón. *Alimentación, Equipos y Tecnología 268*, 25-31.

Leiva-Valenzuela, G. A., Lu, R. & Aguilera, J. M. (2013). Prediction of firmness and soluble solids content of blueberries using hyperspectral reflectance imaging. *Journal of Food Engineering 115*, 91–98.

Li, B., Hou, B., Zhang, D., Zhou, Y., Zhao, M., Hong, R. & Huang, Y. (2016). Pears characteristics (soluble solids content and firmness prediction,varieties) testing methods based on visible-near infraredhyperspectral imaging. *Optik 127*, 2624–2630.

Li, J., Chen, L., Huang, W., Wang, O., Zhang, B., Tian, X., Fan, S. & Li, B. (2016). Multispectral detection of skin defects of bi-colored peaches based on VIS–NIR hyperspectral imaging. *Postharvest Biology and Technology 112*, 121–133.

Liu, C., Liu, W., Chen, W., Yang, J. & Zheng, L. (2015). Feasibility in multispectral imaging for predicting the content of bioactive compounds in intact tomato fruit. *Food Chemistry 173*, 482–488.

Lleó, L., Roger, J.M., Herrero-Langreo, A., Diezma-Iglesias, B. & Barreiro, P. (2011). Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening. *Journal of Food Engineering 104*, 612–620.

López-Maestresalas, A., Keresztes, J.C., Goodarzi, M., Arazuri, S., Jarén, C. & Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control 70*, 229–241.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology 5*, 1121–1142.

Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering 93*, 161–171.

Munera, S., Besada, C., Blasco, J., Cubero, S., Salvador, A., Talens, P. & Aleixos, N. (2017). Astringency assessment of persimmon by hyperspectral imaging. *Postharvest Biology and Technology 125*, 35–41.

Pace, B., Cefola, M., Renna, F. & Attolico, G. (2011). Relationship between visual appearance and browning as evaluated by image analysis and chemical traits in fresh-cut nectarines. *Postharvest Biology and Technology 61*, 178–183.

Pan, L., Zhang, Q., Zhang, W., Sun, Y., Hua, P. & Tu, K. (2016). Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chemistry 192*, 134–141.

Pereira, F. M. V., Carvalho, A. S., Cabeça, L. F. & Colnago, L. A. (2013). Classification of intact fresh plums according to sweetness using time-domain nuclear magnetic resonance and chemometrics. *Microchemical Journal 108*, 14–17.

Pérez-Marín, D., Sánchez, M. T., Paz, P., González-Dugo, V. & Soriano, M. A. (2011). Postharvest shelf-life discrimination of nectarines produced under different irrigation strategies using NIR-spectroscopy. *LWT - Food Science and Technology 44,* 1405-1414.

Qin, J. & Lu, R. (2008). Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique. *Postharvest Biology and Technology 49*, 355–365.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G. S. V. & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194–200.

Ritenour, M. A., Mangrich, M. E., Beaulieu, J. C., Rab, A. & Saltveit, M. E. (1997). Ethanol effects on the ripening of climacteric fruit. *Postharvest Biology and Technology 12*, 35–42.

Rodríguez-Pulido, F. J., Barbin, D. F., Sun, D. W., Gordillo, B., González-Miret, M. L. & Heredia, F. J. (2013). Grape seed characterization by NIR hyperspectral imaging. *Postharvest Biology and Technology 76*, 74-82.

Slaughter, D. C., Crisosto, C. H. & Tiwari, G. (2013). Nondestructive determination of flesh color in clingstone peaches. *Journal of Food Engineering 116*, 920–925.

Schmilovitch, Z., Ignat, T., Alchanatis, V., Gatker, J., Ostrovsky, V. & Felföldi, J. (2014). Hyperspectral imaging of intact bell peppers . *Biosystems Engineering 117*, 83-93.

Sonego, L., Ben-Arie, R., Raynal, J. & Pech, J. C. 1995. Biochemical and physical evaluation of textural characteristics of nectarines exhibiting woolly breakdown: NMR imaging, X-ray computed tomography and pectin composition. *Postharvest Biology and Technology 5*, 187-198.

Tijskens, L. M. M., Zerbini, P. E., Schouten, R. E., Vanoli, M., Jacob, S., Grassi, M., Cubeddu R., Spinelli, L. & Torricelli, A. (2007). Assessing harvest maturity in nectarines. *Postharvest Biology and Technology 45*, 204–213.

Valero, C., Crisosto, C. H. & Slaughter, D. (2007). Relationship between nondestructive firmness measurements and commercially important ripening fruit stages for peaches, nectarines and plums. *Postharvest Biology and Technology 44*, 248–253.

Vásquez-Caicedo, A. L., Sruamsiri, P., Carle, R. & Neidhart, S. (2005). Accumulation of all-trans-β-carotene and its 9-cis and 13-cis stereoisomers during postharvest ripening of nine Thai mango cultivars. *Journal of Agricultural and Food Chemistry. 53*, 4827–4835.

Vidal, M. & Amigo, J. M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems 117*, 138–148.

Vinzi, V., Chin, W. W., Henseler, J., & Wang, H. (2010). Handbook of partial least squares. H. Wang (Ed.), Berlin: Springer.

Wang, J., Teng, B. & Yu, Y. (2006). The firmness detection by excitation dynamic characteristics for peach. *Food Control 17*, 353–358.

Williams, P.C. (1987). Variables affecting near-infrared reflectance spectroscopic analysis. In: Williams, P., Norris, K. (Eds). Near-infrared Technology in the Agricultural and Food Industries. (pp. 143–166). St. Paul, MN: American Association of Cereal Chemists.

Zerbini, P. E., Vanoli, M., Grassi, M., Rizzolo, A., Fibiani, M., Cubeddub, R., Pifferi, A., Spinelli, L. & Torricelli, A. (2006). A model for the softening of nectarines based on sorting fruit at harvest by time-resolved reflectance spectroscopy. *Postharvest Biology and Technology 39,* 223–232.

Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C. & Huang, D. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (*Prunus persica*). Computers and Electronics in Agriculture 114, 14–24.

Zhang, C., Guo, C., Liu, F., Kong, W., He, Y. & Lou, B. (2016). Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. *Journal of Food Engineering 179*, 11-18.

CHAPTER II

Use of hyperspectral transmittance imaging to evaluate the internal quality of nectarines

Sandra Munera[a], José Blasco[a], José Manuel Amigo[b], Sergio Cubero[a], Pau Talens[c] and Nuria Aleixos[d]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Department of Food Sciences, Faculty of Science, University of Copenhagen, Rolighedsvej 30, Frederikberg C DK-1958, Denmark

[c] Departamento de Tecnología de Alimentos, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[d] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

Biosystems Engineering 182 (2019), 54-64

Abstract

The internal quality of nectarines (*Prunus persica* L. Batsch var. *nucipersica*) cv. 'Big Top' (yellow flesh) and 'Magique' (white flesh) has been inspected using hyperspectral trans-mittance imaging. Hyperspectral images of intact fruits were acquired in the spectral rangeof 630-900 nm using transmittance mode during their ripening under controlled conditions. The detection of split pit disorder and classification according to an establishedfirmness threshold were performed using PLS-DA. The prediction of the IQI related to ripeness was performed using PLS-R. The most important variableswere selected using interval-PLS. As a result, an accuracy of 94.7 % was obtained in thedetection of fruits with split pit of the 'Big Top'cultivar. Accuracies of 95.7 % and 94.6 % were achieved in the classification of the 'Big Top'and 'Magique'cultivars, respectively, according to the F threshold. The internal quality was predicted through the IQI with R^2 values of 0.88 and 0.86 for the two cultivars. The results obtained indicate the greatpotential of hyperspectral transmittance imaging for the assessment of the internal quality of intact nectarines.

1. Introduction

Nectarine (*Prunus persica* L. Batsch var. *nucipersica*) is one of the fruits to which plant breeders have devoted the most effort in recent years in order to improve agronomic performance and enhance their appearance and quality (Iglesias & Echeverría, 2009; Munera et al., 2017; Reig, Alegre, Gatius & Iglesias, 2013). However, this effort has not resulted in an increase in consumption due to the fruit being harvesting too early, which means that the products often lack flavour and have excessive firmness, irregular quality and a lack of product identification (Iglesias & Echeverría, 2009; Munera et al., 2018). Therefore, a prior evaluation of quality would be necessary to offer consumers fruits that best match their preferences. Some of these preferences are related to the ripeness of the fruit when consumed. But the skin colour of red cultivars makes it virtually impossible to visually determine the exact stage of maturity. On the other hand, ripening of peaches and nectarines is related with changes during storage that transform a mature fruit into one that is ready to be eaten (Crisosto, 1994). Therefore, maturity at harvest determines the quality of fruit when it reaches the consumer (Jacob et al., 2006).

Hyperspectral imaging has emerged as a potential and powerful tool for safety and quality inspection of agricultural products (Lorente et al., 2012). This non-destructive technique integrates conventional imaging and spectroscopy to obtain both spatial and spectral information from an object simultaneously, thus making it a useful tool for evaluating individual fruits, vegetables or grains (Qin, Chao, Kim, Lu, & Burks, 2013). Most of the hyperspectral imaging systems found in the literature have been implemented to capture images of the samples illuminated by appropriate lighting systems that make it possible to capture the light reflected by the sample. The differences found between the light emitted by the lamps and the radiation reflected by the samples allows certain attributes related to the composition or the quality to be estimated. Examples are found in vegetables, such as pepper (Schmilovitch et al., 2014), tomato (Liu, Liu, Chen, Yang, & Zheng, 2015) or rocket leaves (Chaudhry et al., 2018), cereals, like maize (Williams & Kucheryavskiy, 2016), or rice (Kong, Zhang, Liu, Nie, & He, 2013), and fruits such as bananas (Rajkumar, Wang, EImasry, Raghavan, &

Gariepy, 2012), pears (Li et al., 2016), grapes (Baiano, Terracone, Peri, & Romaniello, 2012), strawberries (Zhang et al., 2016) or apples (Baranowski, Mazurek, & Pastuszka-Wozniak, 2013). In the case of stone fruit, Herrero-Langreo, Lunadei, Lleó, Diezma, and Ruiz-Altisent (2011) assessed the ripeness of peaches by using multispectral indexes; Lu and Peng (2006) assessed the F of peaches; Zhu, Lin, Nie, Wu and Chen (2016) obtained F distribution maps inside the peach pulp. This technology was also used to monitor the ripeness of two cultivars of nectarines (Munera et al., 2017) and to discriminate between similar cultivars with precision (Munera et al., 2018).

On the contrary, hyperspectral imaging in transmittance mode is more effective in detecting internal defects and concentrations in translucent materials, as is the case of some fruits. When a fruit is illuminated with a strong light, the incident radiation may be reflected, absorbed or transmitted, and the relative contribution of each phenomenon depends on the chemical constitution and physical parameters of the sample (Nicolaï et al., 2007). The transmission mode may be less susceptible to surface properties and hence better for detecting composition or internal disorders than the reflectance mode (Schaare & Fraser, 2000). When this mode is used in hyperspectral imaging, the camera is located on the opposite side to the light source and captures the light transmitted through the sample. Transmittance has already been used to analyse the mechanical properties of blueberries (Hu, Dong, Liu, Opara & Chen, 2015; Leiva-Valenzuela, Lu, & Aguilera, 2014), and to detect pits in cherries (Qin & Lu, 2005; Siedliska, Baranowski, Zubik & Mazurek, 2017), defects in pickling cucumbers (Cen, Lu, Ariana & Mendoza, 2014) and damage in soybeans (Huang, Wan, Zhang & Zhu, 2013). However, to our knowledge, no previous works have been undertaken to study the application of hyperspectral imaging in transmittance mode in stone fruit such as nectarines. This technique could be an interesting alternative to evaluate their physicochemical properties but also important disorders such as split pit (Figure 1). This phenomenon consists in the splitting of the pit along the suture/seam of the endocarp, resulting in the two halves of the endocarp being detached from each other inside the mesocarp.

When this disorder happens, the fruit generally develops rot problems far more quickly than sound fruit, and there is a higher risk of the disease spreading more

rapidly from split pit fruit to other fruit during the postharvest operations of storage or marketing (Tani, Polidoros & Tsaftaris, 2007).

Figure 1. Example of nectarine with split pit defect.

In most cases, even in the most advanced cases, no visual symptoms of pit splitting or breakage can be observed, and it is only detected when the fruit is opened (Kritzinger, Lötze & Jooste, 2017). This can be a big problem in nectarines because it can affect 45 % of the fruits, depending on the cultivar and the season (IRTA, 2016). Therefore, non-destructive techniques such as computed tomography (Kritzinger, Lötze & Jooste, 2017), X-ray (Han, Bowers & Dodd, 1992) or, more recently, acoustic vibration methods (Nakano et al., 2018) have been used in an attempt to detect this problem in plums and peaches.

The aim of this work is to investigate the potential use of hyperspectral imaging in transmittance mode as a tool for the non-destructive evaluation of the internal quality of two cultivars of nectarine. This quality evaluation is related to the detection of fruit with split pit and to the ripeness monitoring determined by two indicators, the internal quality index, IQI, and a firmness threshold (35 N).

2. Material and methods

2.1. Fruit samples

This study was performed in parallel to a previous work in which the ripeness of 'Big Top' (yellow flesh cultivar) and 'Magique' (white flesh cultivar) nectarines was monitored using hyperspectral imaging in the reflectance mode (Munera et al., 2017).

In this case, a total of 168 fruits of each cultivar (336 in total), 'Big Top' and 'Magique', were harvested in a commercial orchard in Lerida (Spain) in the commercial maturity period and grouped in 6 batches of 28 fruits, where 5 of which were stored under controlled conditions (15 °C, 90 % relative humidity) until senescence. The image acquisition and the analyses of the 'Big Top' cultivar were performed before storage (for one set) and after the 1st, 2nd, 3rd, 5th and 8th days (for the remaining five sets), collecting a total of 168 mean spectra; for 'Magique' nectarines they were performed before storage (for one set) and after the 2nd, 4th, 7th, 10th and 14th days (for the remaining five sets), also collecting a total of 168 mean spectra. Different days were selected for the analyses due to different ripening speeds for each cultivar (Munera et al., 2017).

Initially, all of the fruits presented a sound appearance and there were no external signs of split pit in any of them. The experiments to detect this disorder were carried out after the image acquisition. A total of 137 'Big Top' fruits out of 168 (81.5 %) presented a normal pit and 31 (18.5 %) were identified as split pit (Figure 1). In the case of the 'Magique' cultivar, no fruit presented split pit.

2.2. Hyperspectral imaging acquisition and processing

The hyperspectral imaging system used to acquire the images in transmittance mode (Figure 2) was composed of an industrial camera (CoolSNAP ES, Photometrics, AZ, USA), coupled to two LCTF (Varispec VIS-07 and NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). A lens capable of maintaining the focus across the full

spectral range (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA) was also used.

The camera was configured to acquire images with a size of 1392 × 1040 pixels and a spatial resolution of 0.14 mm/pixel. The camera and the filters are sensitive in the range from 400 to 1100 nm. However, little light crosses the nectarines and the images appeared very dark when the time of the light exposition was limited to no more than 10 s per wavelength in order to avoid any damage in the fruit. Therefore, a calibration was carried out so that the integration time was increased as much as possible while ensuring that the maximum intensity (saturation) was not reached for any wavelength in any region of the image. To avoid the low sensitivity of the sensors close to the edges of this range, the images were captured at every 10 nm in the working spectral range of 630–900 nm, resulting in 28 images obtained at different wavelengths. This is in accordance with Qin and Lu (2005), who selected the spectral range from 692 to 856 nm to detect pits in cherries using transmittance.

The fruit was placed manually in a holder with a foam foil located between the camera and the illumination system in which the fruit was inserted to ensure that only the light that was transmitted through the fruit reached the camera (Figure 2). The nectarines were oriented so that the pedicel was pointing downwards and directly illuminated by the twelve halogen spotlights (37 W) (Eurostar IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current (12 V). The lamps were arranged equidistant from each other outside a hemispherical aluminium diffuser (Figure 2).

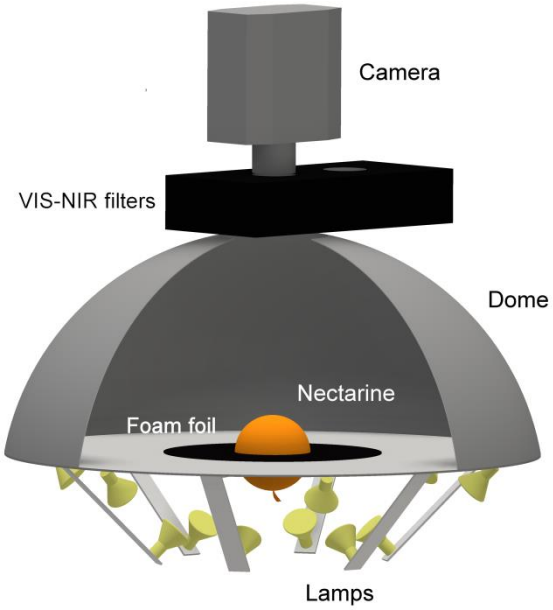


Figure 2. Hyperspectral acquisition system.

In order to extract the actual response of the samples at each wavelength, while avoiding light-dependent intensities, a correction was applied. Several methods have been described to correct the effect of the spectrum of the light source in transmittance mode, from no correction (Siedliska et al., 2017), which is clearly wrong, to the use of different materials, such as opal glass, or measuring the light source directly with no samples (Ariana & Lu, 2008; Cogdill, Hurburgh, & Rippke, 2004). This last option is actually equivalent to correcting the images using the reflectance of a standard white reference. A correction was then performed using the image of a standard white reference (Spectralon 99 %, Labsphere, Inc, NH, USA) captured with a reduction in the integration time to prevent saturation (Gómez-Sanchis et al., 2014). The influence of the minimum dark current of the camera was also captured by switching off the lamps and placing a cap in the lens to prevent the light from getting inside the camera. The correction was performed using the correction in Equation (1):

$$I = \frac{I_0 - I_{black}}{I_{white} - I_{black}} \quad (1)$$

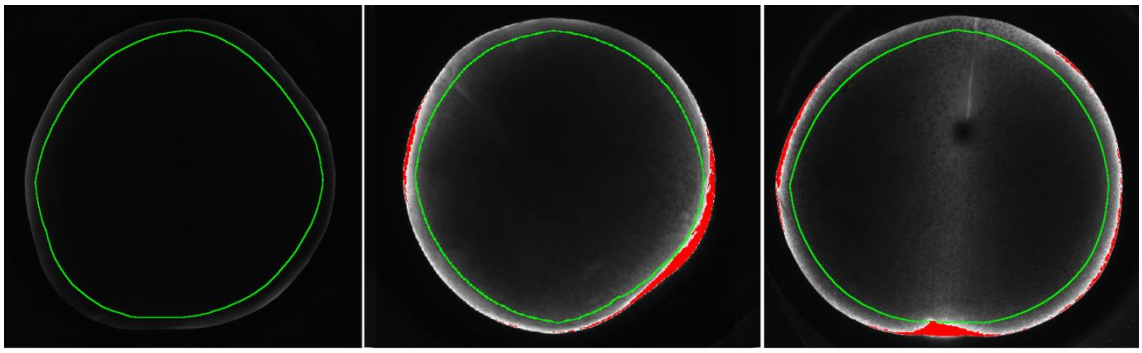
where I_0 is the raw acquired image of the fruit, I_{white} is the image of the standard white reference, and I_{black} is the image acquired while avoiding any light source. The images obtained were processed using the toolbox HYPER-Tools (Mobaraki & Amigo, 2018) for MATLAB R2017b (The MathWorks, Inc. MA, USA).

As Ariana and Lu (2008) pointed out, transmittance is affected by the diameter of the fruit, and therefore the effect of the fruit size was corrected using Equation 2:

$$I_d = \frac{I \, x \, d_n}{d_t} \quad (2)$$

where I is the corrected image obtained previously, d_n is the diameter of the individual fruit and d_t is the average of the diameters of all the fruits of each cultivar.

Finally, the mean transmittance spectrum was obtained by averaging the relative transmittance spectra without including the possible saturated pixels on the edge of the fruit (Figure 3). A total of 168 mean spectra representing the 'Big Top' fruits and 168 mean spectra representing 'Magique' fruits (28 mean spectra of each cultivar in each day of analysis) were obtained for assessment of their internal quality by means of multivariate data analysis methods. In the case of 'Big Top' cultivar, 137 mean spectra corresponded to fruits with normal pit and 31 with split pit.



A B C

Figure 3. Image processing to select the ROI of each type of fruit: less ripe fruit (A), riper fruit (B) and split pit fruit (C). *Green line = limit of the ROI (analysed area); red pixels = saturated pixels*

2.3. Reference quality parameters

The determination of reference quality parameters was performed after image acquisition on each day of analysis in order to monitor the ripening of both cultivars of nectarines. The analysis of the flesh firmness (F) was performed using a texturometer (XT2 Stable, MicroSystems Haslemere, UK) equipped with a 6 mm flat plunger. The crosshead speed during the puncture test was 1 mm s^{-1}. The maximum force, expressed in N, was registered on opposite sides of the fruits. The colour of the flesh was obtained using a colourimeter (MINOLTA CM-700D, Minolta Co. Tokyo, Japan) with the standard illuminant D65 and the CIE standard observer 10°. L*, C* and h* parameters were obtained in the CIELCh colour space. The TSS value was analysed from the juice of each nectarine with a digital refractometer (RFM330 + VWR, Internacional Eurolab S.L., Barcelona, Spain) at 20 °C and the results were expressed as a percentage of the TSS. The IQI was calculated using Equation (3) (Cortés et al., 2016). This index relates internal physicochemical properties to a sensory perception of its ripeness.

$$IQI = \ln \frac{100 \times F \times L^* \times h^*}{TSS \times C^*} \quad (3)$$

The ANOVA, followed by Tukey's Honestly Significant Difference (HSD) test was conducted to determine significant differences (significance defined at p-value ≤ 0.05) in the reference properties of the fruit during the ripening process using the software Statgraphics (Manugistics Corp., Rockville, USA).

2.4. Multivariate analysis

In this work, the prediction of the ripeness properties by means of the IQI was performed using models based on PLS-R and the discrimination between split and normal pit and the corresponding F was carried out by means of models based on PLS-DA.

PLS-R searches for a linear regression model of latent variables by projecting prediction variables X and response variables Y into a new latent space where the

covariance between these latent variables is maximised. In this work, the goal is to find the latent multidimensional direction in the wavelengths space that explains the direction of the maximum multidimensional covariance in the reference parameter space (Lorente et al., 2012).

In PLS-DA the Y variable is categorical, expressing the class membership of the samples. It is performed in order to sharpen the separation between groups of observations by maximising the covariance between the wavelengths and the classes, such that a maximum separation among these classes is obtained (Lorente et al., 2012).

All models were calibrated using the mean transmittance spectra of two thirds of the fruit and later validated using the remaining third. For the detection of split pit, the mean transmittance spectra of 92 fruits with normal pit and 20 with split pit were used as a training set to calibrate the model, and the remaining spectra of 45 fruits with normal pit and 11 with split pit were used as a test set to validate the model. Both the fruits in the calibration and the validation sets were selected with different degrees of ripeness. In the case of ripeness monitoring, the models for 'Big Top' were calibrated using the mean transmittance spectrum of 92 fruits and validated using 45 (removing split pit fruits). The models for 'Magique' were calibrated using the mean spectra of 112 fruits and validated using 56.

All spectra were previously pre-processed using SNV to remove the scatter and then normalised using mean-centring (Rinnan, van den Berg, & Engelsen, 2009). A 10-fold cross-validation was used to choose the optimal number of LVs as well as to obtain an estimation of the error rate of the models. The accuracy of the PLS-R models and predictive capability were evaluated by the R^2 and the RMSE between the predicted and the measured values of the reference parameter for calibration, CV and prediction. Furthermore, the RPD, defined as the ratio between the standard deviation of the reference data and RMSEP, was used (Williams, 1987). The results of the PLS-DA models were expressed as a percentage of correct classification and total accuracy for calibration, CV and prediction.

2.4.1 Selection of optimal wavelengths

Since hyperspectral images have a high dimensionality, which makes it almost impossible to develop automatic inspection systems capable of working in-line or in real time, it is necessary retain the most original information in a few bands, while preserving the greatest amount of variability and the most significant information (Du & Sun, 2006). The i-PLS algorithm was performed to select the optimal wavelengths in order to detect normal and split pit fruits, classify them according to the firmness threshold and predict the IQI. This is a method proposed by Nørgaard et al. (2000), in which the whole spectrum is split into equidistant subintervals and models are calculated for each of these intervals (spectral regions). This method performs a sequential search for the best wavelength or combination of wavelengths. It can be performed in either forward or reverse mode, where intervals are successively included or removed from the analysis, respectively. In this case, the forward i-PLS was applied to the training set automatically using the same number of LV as the PLS-R and PLS-DA models, and each interval corresponded to an individual wavelength. The multivariate data analysis was performed using the PLS_Toolbox (Eigenvector Research Inc., USA) working under MATLAB (R2017b, The MathWorks, Inc. MA, USA).

3. Results and discussion

3.1 Detection of split pit fruit

The presence of split pit allows the light to cross through the stone fruit without any interference along the suture of the fruit (Figure 3-C). Therefore, the SNV pre-treated mean spectra of both types of fruit followed a very different pattern, as Fig. 4 shows.

The discrimination between normal and split pit fruit was performed by means of PLS-DA. The model was built using all of the 28 wavelengths in the spectral range 630–900 nm and calibrated using three LV. In the calibration of the model, a total accuracy of 94.6 % was obtained, 95.0 % of normal pit and 93.4 % of split pit fruits being classified correctly. In the prediction of the test set, a total accuracy of 93.0 % was

obtained, 91.3 % of normal pit and 100 % split pit being classified correctly (Table 1).

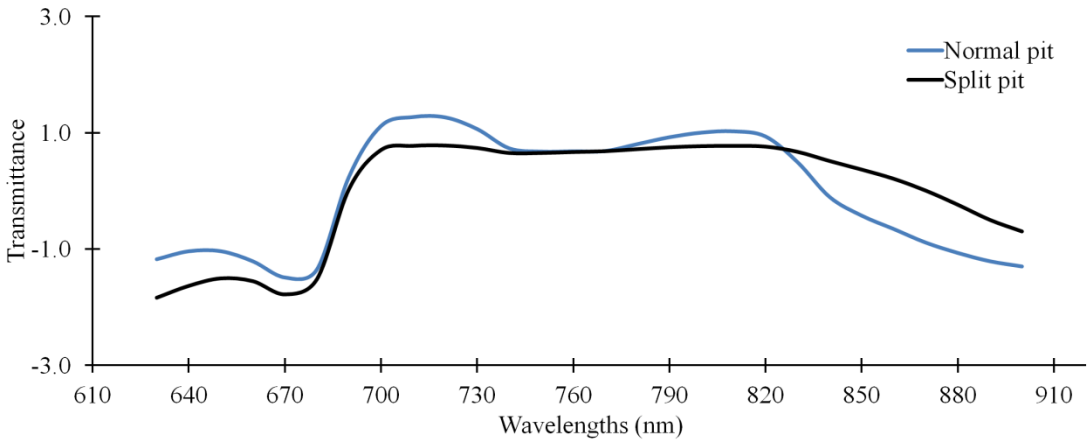


Figure 4. Mean spectra of 'Big Top' fruits with normal and split pit

As stated previously, no study has been performed to detect split pit using hyperspectral imaging. However, Qin and Lu (2005) used this technology to detect the presence of pits in cherries and achieved similar results, an accuracy of 96.5 %. Other techniques have already been used with the aim of detecting split pit disorder. Han, Bowers and Dodd (1992) used X-ray images and obtained a total classification accuracy of 95.5 % using 94 normal pit fruits, 5 cracked and 99 split pit of different cultivars of peach. An acoustic vibration method developed by Nakano et al. (2018) obtained a total classification accuracy of 97.8 % using 256 normal pit fruits and 57 split pit in the same cultivar and stage of ripeness. Comparing these results with hyperspectral transmittance imaging, it can be stated that this technology is a feasible alternative for the detection of split pit, especially taking into account the high accuracy in identifying fruits with split pit that was achieved regardless of the stage of ripeness.

To select the optimal wavelengths, the forward i-PLS method was used. This method has been used previously to select the optimal wavelengths in the detection of early bruise on apples (Ferrari, Foca, Calvini & Ulrici, 2015) or to assess the internal quality of blueberries (Leiva-Valenzuela et al., 2014). Usually, the selection of these wavelengths would be based on the physicochemical properties of the fruit, however, in this case it is based on those wavelengths that transmit more or less light due to the

presence of normal or split pit. In this study, 7 optimal wavelengths were selected (630, 670, 680, 700, 740, 800 and 870 nm) which are those that present more differences along the transmittance mean spectrum of both types of fruit (Figure 4). Therefore, a new PLS-DA model was developed with these wavelengths, also calibrated using 3 LVs. However, the results (Table 1) were better than those obtained using the full spectrum for all the testing sets. Thus, the total accuracy in the calibration rose from 94.6 % using all the wavelengths to 97.3 % and in the classification of the test set, it increased from 93.0 % to 94.7 %.

Table 1. Results of the detection of split and normal pit fruits of the 'Big Top' cultivar using all the selected wavelengths.

#V	#LV	Class	Calibration				Cross validation				Prediction			
			NP	SP	CC (%)	A (%)	NP	SP	CC (%)	A (%)	NP	SP	CC (%)	A (%)
28	3	NP	90	1	98.9	98.2	85	6	93.4	94.6	42	4	91.3	93.0
		SP	1	19	95.0		1	19	95.0		0	11	100	
7	3	NP	91	0	100	99.1	89	2	97.8	97.3	43	3	93.5	94.7
		SP	1	19	95.0		1	19	95.0		0	11	100	

#V=number of variables; #LV=number of latent variables; NP = normal pit; SP = split pit; CC = correct classification; A = accuracy.

3.2 Ripeness monitoring

3.2.1 Analysis of the reference parameters and spectral information

Figure 5 shows the evolution of the physicochemical properties measured in fruits of the 'Big Top' and 'Magique' cultivars throughout the experiment.

In the case of 'Big Top', these properties were measured only in fruits with a normal pit. The F decreased from 46.3 N to 10.1 N for 'Big Top' and from 57.9 N to 6.1 N for 'Magique'. As stated by Munera et al. (2017), these changes are due to pectin solubilisation and degradation by enzymes acting on the cell walls, whose activity results in a large decline in firmness. Valero, Crisosto, and Slaughter (2007) found that fruits below 35 N could be considered as 'ready to buy' because they are susceptible to damage during postharvest handling, while fruits above this firmness were less

susceptible to bruising but could be either mature or immature. This F threshold was therefore selected to classify the fruit because it indicates changes during postharvest ripening and the susceptibility to damage by bruising (Crisosto, Slaughter, Garner, & Boyd, 2001).

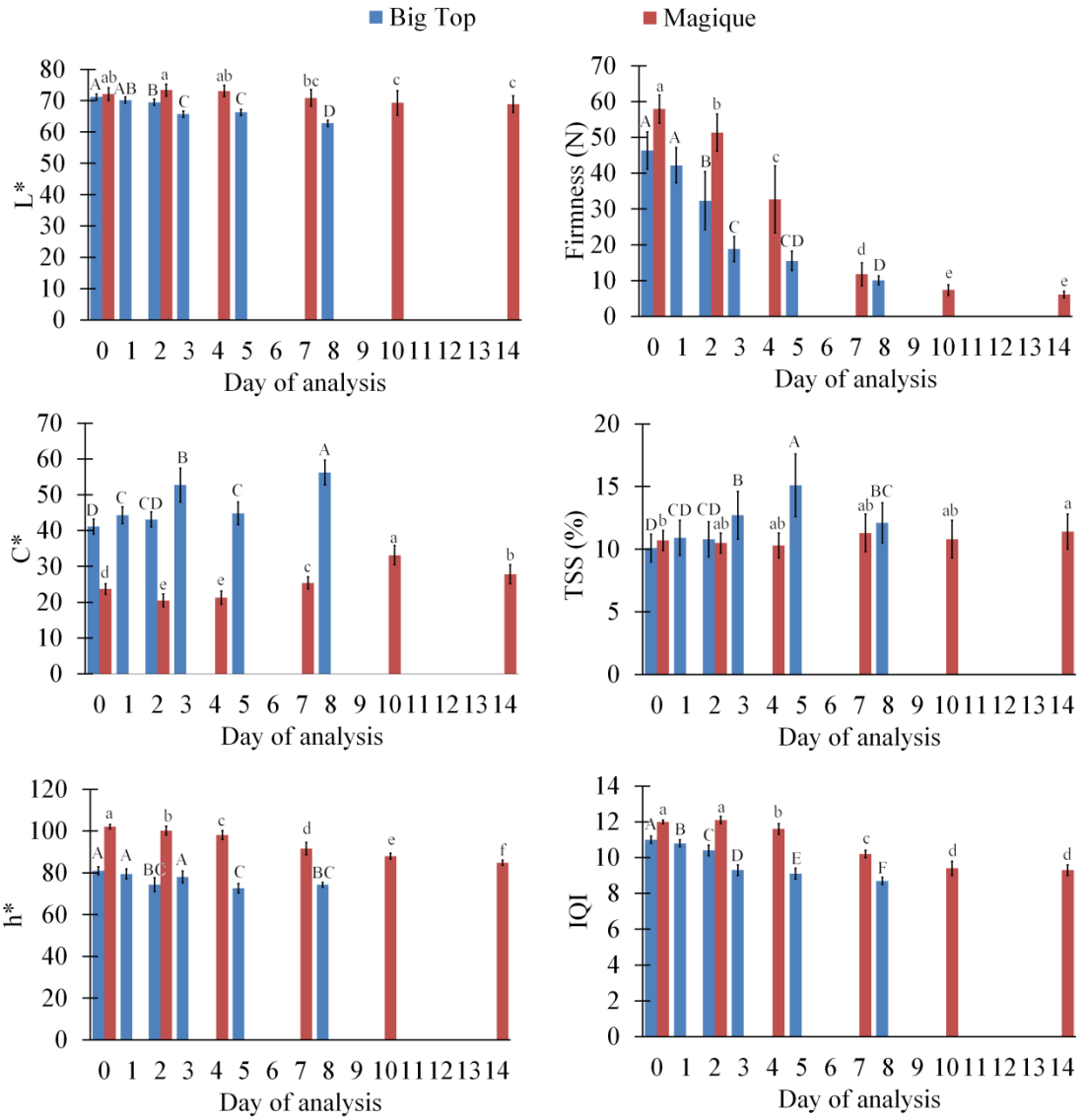


Figure 5. Results of the analysis of the reference quality parameters.

Columns are mean and bars are standard deviation. Different letters in each nectarine cultivar set indicate significant differences between groups (p-value<0.05), according to Tukey's (HSD) test.

Regarding the colour of the flesh, both cultivars obtained similar L* values at the beginning of the experiments, but 'Big Top' underwent a higher reduction in this parameter as the fruit ripened, which is related to a reduction in the brightness perceived during the maturation process. In contrast, 'Magique' presented higher values of h*, starting with a green colour and eventually reaching a greenish-yellow colour. On the other hand, the fruits from cv. 'Big Top' changed from greenish yellow at the beginning to yellow. In the case of C*, no progressive evolution was observed in either cultivar, but 'Big Top' presented higher values, which means that the colouration was more intense in this cultivar.

The TSS obtained for the 'Big Top' cultivar increased from 10.1 % to 15.1 % on the fifth day, and then dropped to 12.1 % due to over-ripeness. In the case of 'Magique', these values did not change significantly until the last day, when the fruits could be considered over-ripe.

The IQI decreased during fruit ripening for both cultivars, mainly due to the progressive decrease in F and the colour parameters L* and h* and the increase in TSS (Figure 5), which is in agreement with Munera et al. (2017). As they pointed out, IQI is more suitable for use as a standard index on an inspection line because obtaining the reference parameters requires less time and costs.

Figure 6 shows the average transmission spectra of both nectarine cultivars pre-processed using SNV on the different days of postharvest storage. Both cultivars followed a similar spectral pattern during ripeness. The main differences between the days of analysis are observed in the region 630–750 nm for 'Big Top' and also 820–900 nm for 'Magique'. In both cultivars, as the fruits ripen more light is transmitted in the VIS region around 670 nm because the chlorophyll content decreases. In contrast, in the NIR region, the transmission of light is lower in the ripest fruits, probably because the effective absorption bands related to water (OH) and sugar (CH) bonds are relatively wide, partially covering this range (Golic, Walsh, & Lawson, 2003).

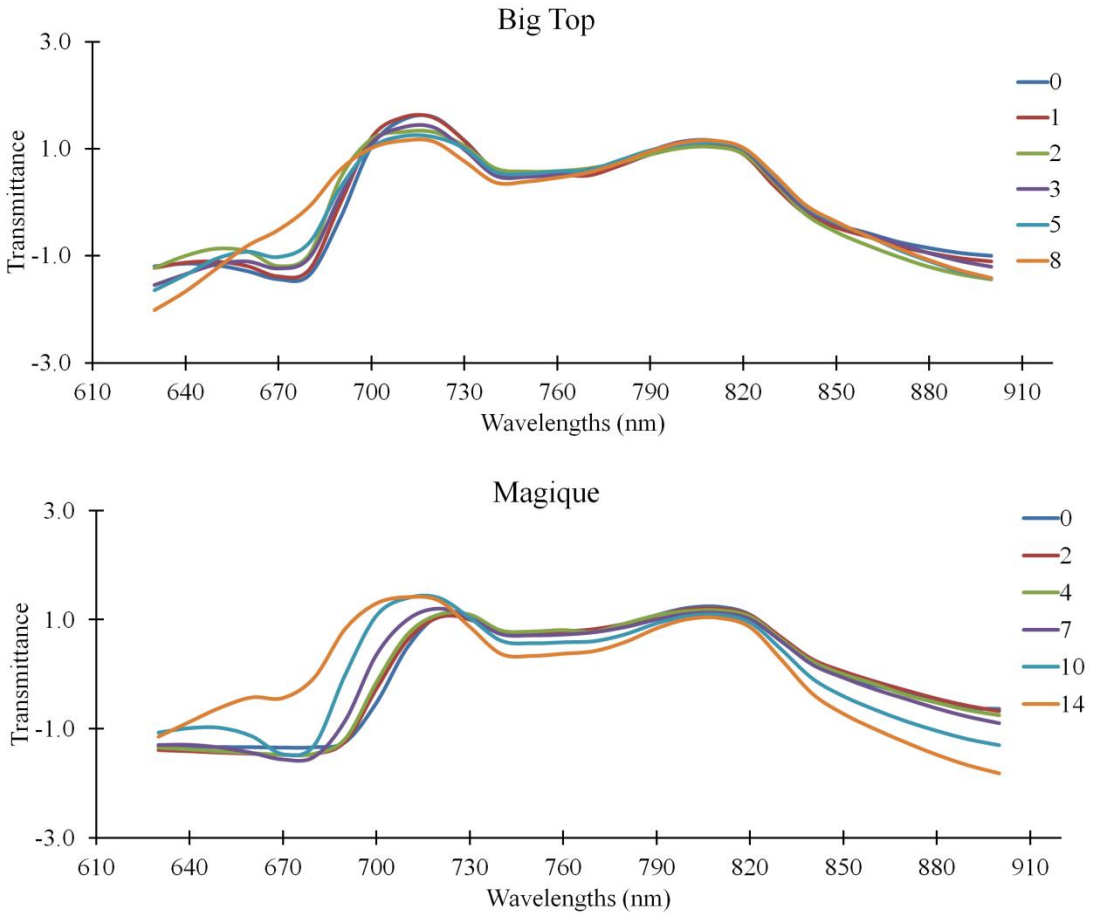


Figure 6. Mean spectra of the fruits of the 'Big Top' and 'Magique' cultivars on each day of analysis.

3.2.2 Prediction of Internal Quality Index (IQI)

With the aim of predicting the IQI and monitoring the ripeness of both cultivars, a PLS-R model was performed for each cultivar using all 28 wavelengths in the spectral range 630–900 nm.

The optimal model was chosen when the number of LV yields the lowest RMSE for calibration and CV. As Table 2 shows, the calibration of the prediction models was performed using 9 LVs and 7 LVs for the 'Big Top' and 'Magique' cultivars, respectively.

In 'Big Top', the R^2 and RMSE values in the calibration were 0.88 and 0.33, and for 'Magique' 0.88 and 0.44, respectively. Regarding the prediction of the test set, the R^2 and RMSE values for 'Big Top' were 0.89 and 0.34, and for 'Magique' 0.88 and 0.43, respectively.

The value of RPD was 2.7 for 'Big Top' and 2.8 for the 'Magique' cultivar. According to Williams (1987), RPD values between 2 and 2.5 indicate that coarse quantitative predictions are possible and a value above 2.5 means good to excellent prediction accuracy. Taking into consideration these values, IQI prediction was excellent for both cultivars (Table 2).

Table 2. Results of prediction of IQI using all the wavelengths.

Cultivar	#LV	Calibration		Cross validation		Prediction		RPD
		R^2	$RMSE_C$	R^2	$RMSE_{CV}$	R^2	$RMSE_P$	
'Big Top'	9	0.93	0.25	0.88	0.33	0.89	0.34	2.7
'Magique'	7	0.90	0.38	0.88	0.44	0.88	0.43	2.8

#LV = number of latent variables

Munera et al. (2017) achieved an R^2 of 0.89 for both cultivars to estimate the IQI using hyperspectral reflectance imaging on the same sets of fruits. The RMSE in the prediction was 0.33 and 0.44 for 'Big Top' and 'Magique', while the RPD achieved was 3.0 and 2.7, respectively. Therefore, transmittance imaging also has a great potential to obtain and estimate the stage of ripeness of nectarines, but it is not greater than reflectance imaging. The selection of one or the other mode would therefore depend on the application (i.e. split pit can only be detected by transmittance).

3.2.3 Classification according to firmness

In order to discriminate the fruits using the selected F threshold (35 N) between 'ready to buy' (F < 35 N) and 'hard' fruit (F > 35 N), a PLS-DA model was performed for each cultivar. The models were built using all captured wavelengths of the spectral range 630–900 nm.

The model for the 'Big Top' cultivar was calibrated using 4 LVs, obtaining a total accuracy of 95.7 % in the prediction set. The correct classification of 'ready to buy' fruit was 100% while 93.1 % of 'hard' fruits were classified correctly. In the case of the 'Magique' cultivar, the model was calibrated using 5 LVs, obtaining an overall classification of 94.5% which is slightly lower than for 'Big Top'. For this cultivar, 90.9 %

of ‘ready to buy’ fruits and 95.7 % of ‘hard’ fruits were classified correctly. Complete results for all sets are described in Table 3.

Table 3. Results of classification of both cultivars of nectarine by firmness using all wavelengths.

Cultivar	#LV	Class	Calibration				Cross validation				Prediction			
			H	RB	CC (%)	A (%)	H	RB	CC (%)	A (%)	H	RB	CC (%)	A (%)
BT	4	H	32	2	94.1	96.7	32	2	94.1	94.5	17	0	100	95.7
		RB	1	56	98.3		3	54	94.7		2	27	93.1	
M	5	H	35	3	92.1	90.2	35	3	92.1	89.3	22	1	95.7	94.5
		RB	8	66	89.2		9	65	87.8		3	30	90.9	

BT = ‘Big Top’; M = ‘Magique’; #LV=number of latent variables; H *= ‘hard’, RB = ‘ready to buy’; CC = correct classification; A = accuracy.*

3.2.4 Selection of the optimal wavelengths

The i-PLS algorithm was also applied to the models created to predict both IQI and F. Since most of the wavelengths selected by i-PLS were common for the two quality indicators, only one set of wavelengths per variety was selected to estimate both. Therefore, 13 optimal wavelengths were used to build the models of ‘Big Top’ (630, 640, 660–690, 710–730, 800, 810, 890 and 900 nm) and 9 for the ‘Magique’ cultivar (630–690, 890 and 900 nm). Despite the two cultivars analysed in this study are different in the colour of the flesh and in the ripeness pattern, most of selected wavelengths for both cultivars are located in the VIS region (630–690) nm) related to carotenoids, chlorophylls and other pigments responsible for fruit colour (Rajkumar et al., 2012). In the case water absorption, several wavelengths were select around 750 nm (first overtone of OH) (710–730 nm) for ‘Big Top’ and others were selected at the beginning of the spectral valley around 970 nm (third overtone of OH) (890–900 nm) for both cultivars. The wavelengths selected around 850 nm (800–810 nm) are assigned usually to the absorption of acids and sugars (Yang et al., 2015).

In the previous work which uses reflectance mode in the spectral range 450–1050 nm to predict the IQI of nectarines (Munera et al., 2017), the optimal wavelengths

selected for 'Big Top' were 670–730 nm and 760 nm, and for 'Magique', 670–700 nm and 970–990 nm. Several wavelengths were the same or close for both modes, reflectance and transmittance. However, other wavelengths were different because the spectral ranges and the selection methods used were different. Furthermore, while in reflectance mode the penetration depth can be of few millimetres in the fruit obtaining the information from the external layers, in transmittance mode the information was obtained from the interior of the fruit.

For the evaluation of IQI, the PLS-R models were calibrated using 8 and 5 LVs for 'Big Top' and 'Magique', respectively (Table 4). The results obtained in the calibration of the model and prediction of the test set were similar to those using all the wavelengths for 'Big Top' cultivar but were improved in the case of 'Magique' cultivar (Table 2). The values of R^2 of 0.91 and 0.89 and RMSE of 0.29 and 0.41 were obtained in the calibration (CV) of 'Big Top' and 'Magique'. For the prediction of the test set, values were 2.7 and 3.0, respectively.

Table 4. Results of prediction of IQI using the selected wavelengths.

Cultivar	#V	#LV	Calibration		Cross validation		Prediction		RPD
			R^2	$RMSE_C$	R^2	$RMSE_{CV}$	R^2	$RMSE_P$	
BT	13	8	0.93	0.25	0.91	0.29	0.88	0.35	2.7
M	9	5	0.90	0.37	0.89	0.41	0.89	0.40	3.0

BT = 'Big Top'; M = 'Magique'; #LV=number of latent variables; #V = number of variables

To classify the fruit by F, the PLS-DA models created using the selected wavelengths were calibrated using 5 LVs for 'Big Top' and 2 LV for 'Magique' (Table 5).

As in the case of using all the wavelengths (Table 3), the model for 'Big Top' obtained a total accuracy of 95.7 % in the prediction set. The correct classification of fruits as 'ready to buy' was 100 % while 93.1 % of 'hard' fruits were classified correctly. In the case of the 'Magique' cultivar, the model achieved an overall classification of 94.6 %. For this cultivar, 90.9 % 'ready to buy' and 100 % 'hard' fruits were classified correctly. The results obtained using the selected set of wavelengths was very similar

to those obtained with all the captured wavelengths. Complete results for all sets are described in Table 5.

Table 5. Results of classification of both cultivars of nectarine by firmness using a threshold and the selected set of wavelengths.

	#V	#LV	Class	Calibration				Cross validation				Prediction			
				H	RB	CC (%)	A (%)	H	RB	CC (%)	A (%)	H	RB	CC (%)	A (%)
BT	13	5	H	31	3	91.2	94.5	32	2	94.1	93.4	17	0	100	95.7
			RB	2	55	96.5		4	53	93.0		2	27	93.1	
M	9	2	H	37	1	97.4	91.1	36	2	94.7	90.2	23	0	100	94.6
			RB	9	65	87.8		10	64	86.5		3	30	90.9	

BT = 'Big Top'; M = 'Magique'; #V=number of variables; #LV=number of latent variables; H *= 'hard', RB = 'ready to buy'; CC = correct classification; A = accuracy.*

3.3 Hierarchical classification

Hierarchical classification allows recognising different classes under study in a single step. This approach has been successfully applied to determine the geographical origin of green coffee beans using spectroscopy (Giraudo et al., 2019), to classify the roasted coffee by cup quality using spectroscopy (Craig et al., 2018) or to identify defective hazelnuts using RGB image analysis (Giraudo et al., 2018).

With the aim of obtaining the estimation of the total internal quality of the 'Big Top' nectarines at the same time, including both the detection of split pit disorder and the stage of ripeness, a hierarchical model of two levels was built using the PLS-DA models previously calibrated with the optimal wavelengths. The class of each fruit in the test set was predicted by introducing the mean spectrum measured into the hierarchical model. The result can be seen in Figure 7, which shows the fruit coloured in black if the mean value was assigned by the model to the split pit class, dark blue if it was assigned to 'hard' fruit with normal pit or light blue if it was assigned to the 'ready to eat' and normal pit class.

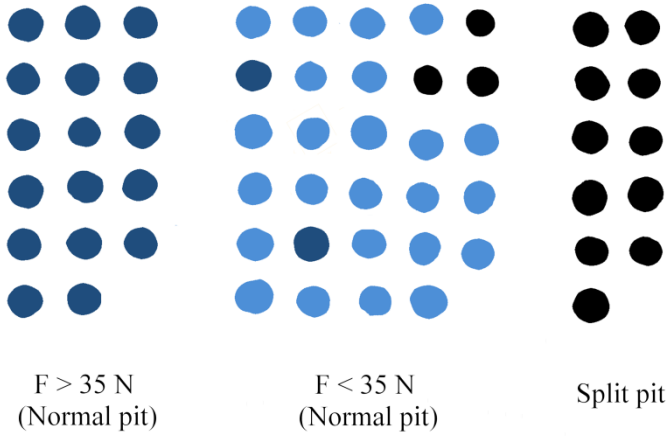


Figure 7. Visual verification of the hierarchical classification of the test set of 'Big Top' nectarines.

The results obtained using this approach were the same as individual models (Tables 1 and 5). All the split pit and 'hard' fruits with normal pit were correctly classified. Three 'ready to buy' fruits with normal pit were classified as defective (10.3 %) may be due to the fact that riper fruit can transmit more light than less ripe fruits, and two other 'ready to buy' fruits with normal pit were classified as 'hard' (6.9 %). These results indicate that it is possible to detect split pits and estimate the ripeness of the nectarines 'Big Top' in only one step, which makes hyperspectral imaging an even more practical tool for quality control of nectarines.

4. Conclusions

This paper presents a new approach for the evaluation of the internal quality of nectarines by means of hyperspectral imaging. The transmittance mode was evaluated as a potential non-destructive method to detect split pit fruits and to monitor their ripeness using two quality indicators. The detection of split pit fruits of the 'Big Top' cultivar using PLS-DA was successful, achieving a 100 % correct classification for split pit fruit and 91.3 % for normal pit using all the captured wavelengths. The ripeness of the 'Big Top' and 'Magique' cultivars was determined by two indicators: the ripening index, IQI, and an F threshold (35 N) that is based on the susceptibility to suffer

damage by bruising. The prediction of the IQI was performed by means of PLS-R models, obtaining an R^2 of 0.89 and 0.88 and an RPD of 2.7 and 2.8 for the 'Big Top' and 'Magique' cultivars, respectively. The classification of the fruits by F was performed by PLS-DA, which correctly classified 95.7 % of the 'Big Top' fruits and 94.5 % of the 'Magique' fruits.

To reduce the huge amount of data captured by the hyperspectral imaging system, an optimal wavelength selection was performed by means of forward i-PLS. Thus, the simplified models obtained similar results to those models that used all the wavelengths. Finally, a hierarchical model was built to evaluate the total internal quality of the 'Big Top' cultivar in one step. The prediction was visualised on the fruit surface, indicating that 10.3 % of 'ready to buy' fruits were classified as split pit and 6.9 % as 'hard'.

These results confirm the great potential of this technique to evaluate the internal quality of these two cultivars of nectarine, especially to detect internal defects such as split pit disorder. Nevertheless, this method should be tested in other cultivars and on a larger sample set of fruits grown in different areas and seasons before it can be implemented in an in-line system. Furthermore, the development of a transmission system must take into account the fact that, in order to detect split pit fruits, the fruit must be oriented such that light penetrates through the fruit from the pedicel to the back and the time of the light exposure must be limited in order to avoid any damage to the fruit.

Acknowledgements

This work has been partially funded by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria de España (INIA) through research project RTA2015-00078-00-00 with the support of European FEDER funds. Sandra Munera thanks INIA for the grant FPI-INIA num. 43 (CPR2014-0082), partially supported by European Union FSE funds.

References

Ariana, D.P. & Lu, R. (2008) Quality evaluation of pickling cucumbers using hyperspectral reflectance and transmittance imaging: Part I. Development of a prototype. *Sensing and Instrumentation for Food Quality and Safety 2*, 144–151.

Baiano, A., Terracone, C., Peri, G. & Romaniello, R. (2012). Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Computetrs and Electronics in Agriculture 87,* 142-151.

Baranowski, P., Mazurek, W. & Pastuszka-Wozniak, J. (2013). Supervised classification of bruised apples with respect to the timeafter bruising on the basis of hyperspectral imaging data. *Postharvest Biology and Technology 86*, 249- 258.

Cen, H., Lu, R., Ariana, D.P. & Mendoza, F. (2014). Hyperspectral imaging-based classification and wavebands selection for internal defect detection of pickling cucumbers. *Food Bioprocess Technology 7*, 1689–1700.

Chaudhry, M.M.M., Amodio, M.L., Baballahi, F., de Chiara, M.L.V., Amigo, J.M., Colelli, G. (2018). Hyperspectral imaging and multivariate accelerated shelf life testing (MASLT) approach for determining shelf life of rocket leaves. *Journal of Food Engineering 238*, 122-133.

Cogdill, R., Hurburgh, C. & Rippke G. (2004). Single-kernel maize analysis by near-infrared hyperspectral imaging. *Transactions of the ASAE*, 47, 311-320

Cortés, V., Ortiz, C., Aleixos, N., Blasco, J., Cubero, S. & Talens, P. (2016). A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy. *Postharvest Biology and Technology 118,* 148-158.

Craig, A.P., Botelho, B.G., Oliveira, L.S. & Franca, A.S. (2018). Mid infrared spectroscopy and chemometrics as tools for the classification of roasted coffees by cup quality. Food Chemistry 245, 1052–1061.

Crisosto C.H. (1994). Stone fruit maturity indices: a descriptive review. *Postharvest News and Information*, 5, 65-68.

Crisosto, C.H., Slaughter, D., Garner, D., & Boyd J. (2001). Stone fruit critical bruising thresholds. *Journal of the American Pomological Society 55*, 76-81

Du, C. J. & Sun, D. W. (2006). Learning techniques used in computer vision for food quality evaluation: A review. *Journal of Food Engineering 72*, 39-55.

Giraudo, A., Grassi, S., Savorani, F., Gavoci, G., Casiraghi, E. & Geobaldo, F. (2019). Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis. Food Control 99, 137-145.

Giraudo, A., Calvini, R., Orlandi, G, Ulrici, A., Geobaldo, F. & Savorani, F. (2018). Development of an automated method for the identification of defective hazelnuts based on RGB image analysis and colourgrams. Food Control 94, 233-240.

Golic, M., Walsh, K., & Lawson, P. (2003) Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. *Applied Spectroscopy*, 57, 139-145

Gómez-Sanchis, J., Lorente, D., Soria-Olivas, E., Aleixos, N., Cubero, S., & Blasco, J. (2014) Development of a hyperspectral computer vision system based on two liquid crystal tuneable filters for fruit inspection. Application to detect citrus fruits decay. *Food and Bioprocess Technology*, 7, 1047-1056.

Han, Y.J., Bowers & S. V., Dodd, R.B. (1992). Nondestructive detection of Split-pit peaches. *Transactions of the ASAE 35,* 2063-2067.

Herrero-Langreo, A., Lunadei, L., Lleó, L., Diezma, B. & Ruiz-Altisent, M. (2011). Multispectral vision for monitoring peach ripeness. *Journal of Food Science 2*, 178-187

Hu, M.H., Dong, K.L., Liu, B. L., Opara, U.L. & Chen, L. (2015) Estimating blueberry mechanical properties based on random frog selected hyperspectral data. *Postharvest Biology and Technology 106*, 1-10

Huang, M., Wana, X., Zhang, M. & Zhu, Q. (2013). Detection of insect-damaged vegetable soybeans using hyperspectral transmittance image. *Journal of Food Engineering 116,* 45–49

Iglesias, I. & Echeverría, G. (2009). Differential effect of cultivar and harvest date on nectarine colour, quality and consumer acceptance. *Scientia Horticulturae 120*, 41-50.

Institut de Recerca i Tecnologia Agroalimentàries (IRTA). (2016). XX Exposición de variedades de melocotón y nectarina. http://kp.eufrin.eu/ Accessed 03/11/18

Jacob, S., Vanoli, M., Grassi, M., Rizzolo, A., Eccher Zerbini, P., Cubeddu, R., Pifferi, A., Spinelli, L., & Torricelli A. (2006). Changes in sugar and acid composition of 'Ambra' nectarines during shelf life based on non-destructive assessment of maturity by time-resolved reflectance spectroscopy. *Journal of Fruit and Ornamental Plant Research*, 14, 183-194

Kong, W., Zhang, C., Liu, F., Nie, P. & He, Y. (2013). Rice seed cultivar identification using near-infrared hyperspectral imaging and multivariate data analysis. *Sensors 13*, 8916-8927.

Kritzinger, I., Lötze, E. & Jooste, M. (2017). Stone hardening and broken stones in Japanese plums (*Prunus salicina* Lindl.) evaluated by means of computed tomography scans. *Scientia Horticulturae 221*, 1–9

Leiva-Valenzuela, G.A., Lu, R. & Aguilera, J.M. (2014). Assessment of internal quality of blueberries using hyperspectral transmittance and reflectance images with whole spectra or selected wavelengths. *Innovative Food Science and Emerging Technologies 24*, 2–13

Li, B., Hou, B., Zhang, D., Zhou, Y., Zhao, M., Hong, R. & Huang, Y. (2016). Pears characteristics (soluble solids content and firmness prediction, varieties) testing methods based on visible-near infrared hyperspectral imaging. *Optik 127,* 2624-2630.

Liu, C., Liu, W., Chen,W., Yang, J. & Zheng, L. (2015). Feasibility in multispectral imaging for predicting the content of bioactive compounds in intact tomato fruit. *Food Chemistry 173*, 482-488.

Lorente, D., Aleixos, N., Gomez-Sanchis, J., Cubero, S., García-Navarrete, O.L., Blasco, J., (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess and Technology 5*, 1121-1142.

Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering 93*, 161-171.

Vidal, M. & Amigo, J.M. (2018). HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems 172*, 174-187.

Munera, S., Amigo, J.M., Aleixos, N., Talens, P., Cubero, S. & Blasco, J. (2018). Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine. *Food Control 86*, 1-10.

Munera, S., Amigo, J.M., Blasco, J., Cubero, S., Talens, P. & Aleixos, N. (2017). Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging. *Journal of Food Engineering 214,* 29-39.

Nakano, R., Akimoto, H., Fukuda, F., Kawai, T., Ushijima, K., Fukamatsu, Y., Kubo, Y., 1, Fujii, Y., Hirano, K., Morinaga, K. & Sakurai, N. (2018). Nondestructive detection of split pit in peaches using an acoustic vibration method. *The Horticulture Journal 87*, 281–287.

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, I.K., & Lammertyn, J., (2007). Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biology and Technology* 46, 99–118.

Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L. & Engelsen, S.B. (2000). Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared *Spectroscopy. Society for Applied Spectroscopy 54*, 413-419.

Qin, J. & Lu, R. (2005). Detection of pits in tart cherries by hyperspectral transmission imaging. *Transactions of the ASAE*, 48, 1963-1970.

Qin, J., Chao, K., Kim, M.S., Lu, R. & Burks, T.F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering 118*, 157-171.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G.S.V. & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194-200.

Reig, G., Alegre, S., Gatius, F. & Iglesias, I. (2013). Agronomical performance under Mediterranean climatic conditions among peach [*Prunus persica* (L.) Batsch] cultivars originated from different breeding programs. *Scientia Horticulturae 150*, 267-277.

Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry 28*, 1201-1222.

Schaare, P.N. & Fraser, D.G. (2000) Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit (*Actinidia chinensis*). *Postharvest Biology and Technology 20*, 175 – 184.

Schmilovitch, Z., Ignat, T., Alchanatis, V., Gatker, J., Ostrovsky, V. & Felfoldi, J. (2014). Hyperspectral imaging of intact bell peppers. *Biosystems Engineering 117,* 83-93.

Siedliska, A., Baranowski, P., Zubik, M. & Mazurek, W. (2017). Detection of pits in fresh and frozen cherries using a hyperspectral system in transmittance mode. *Journal of Food Engineering 215*, 61-71.

Tani, E., Polidoros, A.N., & Tsaftaris, A.S. (2007). Characterization and expression analysis of FRUITFULL-and SHATTER-PROOF-like genes from peach (*Prunus persica*) and their role in split-pit formation. Tree physiology, 27, 649-659.

Valero, C., Crisosto, C.H. & Slaughter, D. (2007). Relationship between non-destructive firmness measurements and commercially important ripening fruit stages for peaches, nectarines and plums. *Postharvest Biology and Technology 44*, 248-253.

Williams, P. J., & Kucheryavskiy, S. (2016). Classification of maize kernels using NIR hyperspectral imaging. *Food Chemistry 209*, 131-138.

Williams, P.C. (1987). Variables affecting near-infrared reflectance spectroscopic analysis. In: Williams, P., Norris, K. (Eds.), Near-infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemists, St. Paul, MN, pp. 143-166.

Zhang, C., Guo, C., Liu, F., Kong, W., He, Y. & Lou, B. (2016). Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. *Journal of Food Engineering 179,* 11-18.

Zhu, N., Lin, M., Nie, Y., Wu, D. & Chen, K. (2016). Study on the quantitative measurement of firmness distribution maps at the pixel level inside peach pulp. *Computers and Electronics in Agriculture 130*, 48-56.

CHAPTER III

Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine

Sandra Munera[a], José Manuel Amigo[b], Nuria Aleixos[c], Pau Talens[d], Sergio Cubero[a] and José Blasco[a]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Department of Food Sciences, Faculty of Science, University of Copenhagen, Rolighedsvej 30, Frederikberg C DK-1958, Denmark

[c] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[d] Departamento de Tecnología de Alimentos, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

Food Control 86 (2018), 1-10

Abstract

Product inspection is essential to ensure good quality and to avoid fraud. New nectarine cultivars with similar external appearance but different physicochemical properties may be mixed in the market, causing confusion and rejection among consumers, and consequently affecting sales and prices. Hyperspectral reflectance imaging in the range of 450-1040 nm was studied as a non-destructive method to differentiate two cultivars of nectarines with a very similar appearance but different taste. PLS-DA was used to develop a prediction model to distinguish intact fruits of the cultivars using pixel-wise and mean spectrum approaches, and then the model was projected onto the complete surface of fruits allowing visual inspection. The results indicated that mean spectrum of the fruit was the most accurate method, a correct discrimination rate of 94.0 % being achieved. Wavelength selection reduced the dimensionality of the hyperspectral images using the regression coefficients of the PLS-DA model. An accuracy of 96.0% was obtained by using 14 optimal wavelengths, whereas colour imaging and a trained inspection panel achieved a rate of correct classification of only 57.0 % of the fruits.

1. Introduction

The surface area of the land devoted to the planting of peaches and nectarines (*Prunus persica* L. Batsch) in the EU was around 232 000 ha in 2015/16, with a production of nearly 3.7 million tons of fruit. Spain is the main producer with around 1.4 tons, which accounts for almost 40 % of the total EU peach and nectarine production (USDA, 2016). Due to the importance of nectarine (*Prunus persica* L. Batsch var. *nucipersica*) production, it is one of the fruits to which most effort has been devoted by plant breeders in recent years in order to improve agronomic performance, and enhanced fruit appearance and quality (Reig, Alegre, Gatius, & Iglesias, 2013). This fact has resulted in a significant increase in the number of new cultivars available to fruit growers. These cultivars are similar in appearance but present different sensory properties and therefore different acceptance by the consumer (Iglesias & Echeverría, 2009). In this context, one of the most widely accepted and cultivated nectarine cultivars in Europe is 'Big Top' due to its presentation, size, sweet taste and low acidity (Echeverría, Cantín, Ortiz, Lopez, & Graell, 2015). However, a stagnation of nectarine consumption is occurring owing to early harvesting, which leads to flavourless fruits being offered with excessive F or irregular quality (Iglesias & Echeverría, 2009). These authors also point out the lack of an adequate identification of the product in the market. The mixture of sweet and acid cultivars on the shelf could lead to consumer rejection, which in turn might affect sales and prices.

The internal quality assessment of stone fruits has traditionally been performed by destructive methods, which are contaminating, time-consuming and only a few samples per batch can be monitored (Pérez-Marín et al., 2011). Moreover, there is an important lack of classification tools for differentiating cultivars that are very similar to one another. There is therefore a strong need to develop non-destructive and instantaneous methodologies that allow the correct identification of the cultivar in the postharvest stage.

Hyperspectral imaging is a computer vision technique which combines conventional two-dimensional digital imagery with spectroscopy to detect spectral features in regions of the electromagnetic spectrum such as the UV, NIR or IR (Lorente

et al., 2012). This technique is starting to be used as a scientific tool for quality assurance of a wide range of food including bakery products (Erkinbaev, Henderson & Paliwal, 2017; Verdú et al., 2016), meat (Feng et al., 2018; Iqbal, Sun, & Allen, 2014), or vegetables (López- Maestresalas et al., 2016). Fruits are of major interest for the use of this technology in the food industry (Keresztes, Goodarzi, & Saeys, 2016; Munera et al., 2017). However, due to the high importance of other fruits such as citrus or apples, few scientific studies have been done for quality control of stone fruit quality assessment using hyperspectral imaging. Herrero-Langreo et al. (2011) assessed the ripeness of peaches by using multispectral indexes. Lu and Peng (2006) assessed the F of peaches and Zhu et al. (2016) obtained F distribution maps inside the peach pulp, while Zhang et al. (2015), Li et al. (2016), Pan et al. (2016) and Sun et al. (2017) detected different types of defects and injuries, including decay. Regarding nectarine, Huang et al. (2015) used the same technique to detect defective features and Munera et al. (2017) to monitor its ripeness.

Hyperspectral imaging generates a huge amount of redundant and frequently highly correlated data that need to be processed (Vélez-Rivera et al., 2014). To handle such an amount of data and extract the useful information, it must be assisted by chemometric methods. These methods connect chemical measurements with the essential spectral information in order to classify and/or quantify important characteristics. PCA is one of the most popular methods commonly used both to reduce the dimensionality of data and to obtain an overview of all the relevant information in the dataset. It is an unsupervised projection method which summarises data by forming new independent linear combinations of the original variables (Jolliffe, 2002).

PLS-DA is a variant of PLS-R in which the independent variable is categorical, expressing the class membership of the samples. It is performed in order to sharpen the separation between groups of observations by maximising the covariance between the spectra and the independent variable such that a maximum separation among classes is obtained. Furthermore, it is commonly used to understand which variables contain the discriminating information (Lorente et al., 2012). Some examples of the

use of this method include the detection of decay lesions in citrus fruits (Folch-Fortuny et al., 2016), classification of oat kernels (Serranti et al., 2013), the classification of edible fennel heads based on the harvest time (Amodio et al.,, 2017), and the examination of aflatoxin on corn kernels (Kandpal et al., 2015).

In this paper, we put forward a novel approach based on VIS-NIR hyperspectral imaging and chemometric methods to develop statistical predictive models capable of distinguishing cultivars of nectarines with a very similar appearance but different taste. Previous studies have been conducted to differentiate among nectarine cultivars using colour images (Font et al., 2014). However, they use fruits with clearly different appearance. In this work, 'Diamond Ray' and 'Big Top' cultivars have been used due to their similar skin and flesh appearance. Furthermore, these cultivars are grown and marketed at the same time and become a problem for producers when they are mixed, either accidentally or intentionally, in the market.

In addition, using the spectral and spatial information provided by the hyperspectral images, two approaches are further investigated: the first based on the analysis of the individual spectrum of each pixel and the second based on the mean spectrum of each fruit. Finally, visualisation of the result of the classification model over the images of nectarines is proposed to establish a practical tool for nectarine classification in the packing houses.

2. Material and methods

2.1. Fruit samples

Nectarines cv. 'Diamond Ray' and 'Big Top' were selected as reference cultivars of sweet and acid cultivars, respectively (Reig, Iglesias, & Echeverría, 2009), due to their similar skin and flesh appearance (Figure 1).

These two cultivars are difficult to distinguish by the naked eye, which is problematic for producers when they are mixed in the market.

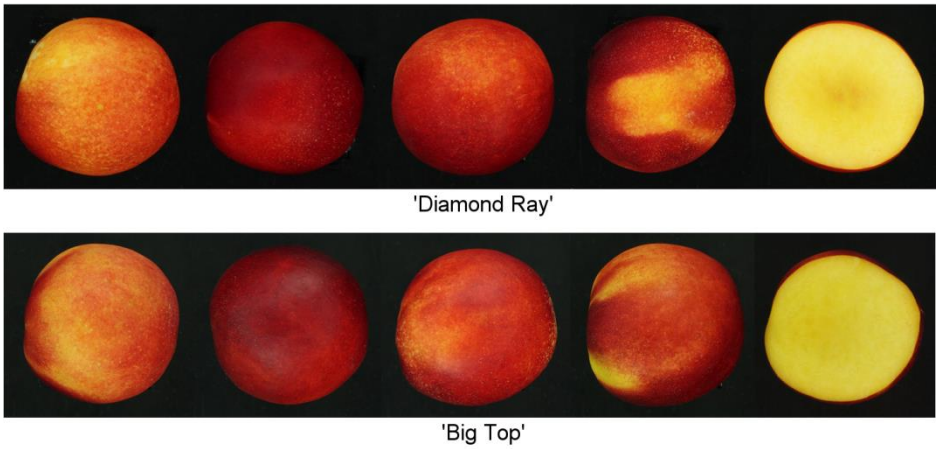


Figure 1. External and internal appearance of both cultivars of nectarine.

Fruits were harvested in a commercial orchard in Lerida (Spain) at the commercial maturity stage in the summer season of 2016. A total of 125 fruits of each cultivar without defects or bruises were selected and stored under controlled conditions (1 ^{o}C; 90 % RH) in order to avoid the further ripening of either cultivar during the experiment.

2.2. Hyperspectral image acquisition and processing

The hyperspectral imaging system consisted of an industrial camera (CoolSNAP ES, Photometrics, AZ, USA), coupled to two LCTF (Varispec VIS-07 and NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). The camerawas configured to acquire images with a size of 1392 x 1040 pixels and a spatial resolution of 0.14 mm/pixel at 60 different wavelengths every 10 nm, in the working spectral range 450-1040 nm. In order to avoid problems of unfocused images due to the refraction of light across this wide spectral range, the focus was adjusted on the central band of the acquisition interval (740 nm) and the images were captured using lenses capable of covering the whole spectral range without going out of focus (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA). To optimise the dynamic range of the camera, prevent saturated images and correct the spectral sensitivity of the different elements of the system, a calibration of the integration time of each band was performed by capturing

the averaged grey level of a white reference target (Spectralon 99%, Labsphere, Inc, NH, USA) corresponding to 90 % of the dynamic range of the camera.

The scene was illuminated by indirect light from twelve halogen spotlights (37 W) (Eurostar IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current (12 V) and arranged equidistant from each other inside a hemispherical aluminium diffuser. The inner surface of the aluminium diffuser was painted white with a rough texture to maximise its reflectivity, the rough texture being applied in order to minimise directional reflections, which could cause bright spots, thus resulting in highly homogeneous light.

The fruits were introduced manually into a fruit holder, with the stem-apex axis lying horizontal. Two images of each fruit were acquired using customised software developed at IVIA. A total of 250 images of each cultivar were imported into MATLAB R2015a (The MathWorks, Inc. MA, USA) to be pre-processed using the customised toolbox HYPER-Tools (Amigo, Babamoradia, & Elcoroaristizabal, 2015).

The image processing started with the correction of the relative reflectance by using equation (1) (Gat, 2000):

$$\rho_{xy}(x,y,\lambda) = \frac{R^{abs}}{R^{abs}_{white}} = \rho^{Ref}(\lambda)\frac{R(x,y,\lambda)-R_{black}(x,y,\lambda)}{R_{white}(x,y,\lambda)-R_{black}(x,y,\lambda)} \quad (3)$$

where $\rho^{Ref}(\lambda)$ is the standard reflectance of the white reference target (99% in this work), $R(x,y,\lambda)$ is the reflectance of the fruit captured by the CCD sensor of the camera, $R_{white}(x,y,\lambda)$ is the reflectance captured by the CCD of the white reference target, and $R_{black}(x,y,\lambda)$ is the reflectance captured by the CCD while avoiding any light source in order to quantify the electronic noise of the CCD.

The images were then clipped and spatially compressed to reduce the computation time, and a proper removal of the backgroundwas performed using K-means clustering. Thus, the relative reflectance spectrum of all the pixels in each fruit image was extracted.

2.3. Colour image acquisition and processing

Before image acquisition, the SCwas analysed to obtain the L*, a* and b* colour coordinates (CIELAB colour space) of each fruit, also using a colorimeter (MINOLTACM-700d, Minolta Co. Tokyo, Japan) configured with the standard illuminant D65 and the observer 10 º. The skin colour was obtained as the average of the values of two measurements, one in the blush zone (reddish colour) and another in the ground zone (yellowish colour).

The colour imaging system consisted of a digital camera (EOS 550D, Canon Inc, Japan) arranged inside a square inspection chamber that included a calibrated and uniform illumination system composed of four lamps, each containing two fluorescent tubes BIOLUX 18W/965 (Osram GmbH, Germany) with a colour temperature of 6500 K. The angle between the axis of the lens and the sources of illumination was approximately 45 º, and polarising filters were placed in front of the lamps and in the camera lenses to eliminate specular bright spots that could alter the true colour.

The fruits were introduced manually upon a fruit holder, with the stem-apex axis lying horizontal. Two images were acquired for each fruit, corresponding to each of the two sides delimited by the suture of the fruit. Then, a total of 250 images of each cultivar were imported into customised software developed at IVIA (FoodImage-Inspector v4.0, freely available at http//www.cofilab.com, Spain) to analyse the SC and to obtain the percentage of the reddish and yellowish zones on the fruit. This segmentation was based on the Bayes theorem to assign all the pixels in the image to the two classes used in a previous training. The RGB colour coordinates of the acquired images were converted to the L*, a*, b* coordinates and then corrected using a colour reference target (ColorChecker Digital SG, X-Rite, MI, USA).

2.4. Visual analysis with trained panel

The panel was composed of five panellists, ages 29-50 years (three male and two female), with expertise in fruit quality and marketing. The panellists were trained using 20 colour images of nectarines of the calibration set (10 from each cultivar chosen at

random). A total of 40 colour images of fruits of the validation set (20 from each cultivar chosen at random) were presented with randomised order to each panellist to be classified as belonging to the 'Diamond Ray' or 'Big Top' cultivar.

2.5. Reference analysis

The characterisation of the physicochemical properties of the samples using reference methodswas performed immediately after the acquisition of the images. F was registered on opposite sides of the fruits using an XT2 Stable texturometer (MicroSystems Haslemere, UK) equipped with a 6 mm flat plunger. The crosshead speed during the puncture test was 1 mm s^{-1}. The maximum force was expressed in Newton (N). Immediately after SC and F measurements, a juice sample was taken from each fruit for TSS and TA measurements. TSS were determined using a digital refractometer RFM330 þ VWR (Internacional Eurolab S.L., Barcelona, Spain) at 20 ^{o}C and results were expressed as percentage of TSS. TA was determined using a Crison pH-Burette 24 automatic titrator (Crison, Barcelona, Spain) and NaOH 0.5 N, according to standard UNE34211:1981 (AENOR, 1981). The results were expressed as the percentage of malic acid.

The ANOVA was conducted using the software Statgraphics (Manugistics Corp., Rockville, USA) in order to determine significant differences in the physicochemical properties (F, TSS, TA and L*, a* and b* colour coordinates) between cultivars.

2.6. Chemometric methods

To identify both nectarine cultivars with high precision, two approaches were studied for setting up the classification models: i) including in the model the individual spectrum of each pixel in the nectarine image, and ii) using only the mean spectrum of all the pixels corresponding to each fruit. Thus, 512 828 pixel spectra were used in the first approach, and the mean spectra of 500 fruits were used for the second. The data of all the fruits of both cultivars were collected and randomly partitioned into two sets: two thirds of the samples were used to calibrate the models (calibration set) and for

cross-validation, while the remaining third was used for independent test prediction (validation set).

Both the directly acquired spectrum of each pixel and that obtained as an average for each fruit were pre-treated using SNV in order to reduce the physical variability between samples due to light scatter (Rinnan, van den Berg, & Engelsen, 2009). This correction was performed using equation (2):

$$X_{corr} = \frac{X_{org} - a_0}{a_1} \quad (2)$$

where x_{corr} and x_{org} are the corrected and raw spectra, respectively, a_0 is the average value of the sample spectrum to be corrected and a_1 is the standard deviation of the sample spectrum.

Later, mean centring was applied to normalise the full spectrum. Multivariate analyses were then performed using the PLS_Toolbox (Eigenvector Research Inc., USA) and the HYPER-Tools toolbox (Amigo et al., 2015) both working under MATLAB R2015a.

PCA was used to explore the differences between the two cultivars using the pixel and mean spectra of the calibration set previously pre-processed by means of SNV and mean centring. Later, PLS-DA models were built to sort the fruits into one of the two studied cultivars. The models were also calibrated using the preprocessed pixel and mean spectra of the calibration set and tested using only samples of the validation or prediction set.

In order to compare the performance of the hyperspectral imaging with the colour imaging system, a PLS model was also built using the mean value of the L*a*b* colour coordinates. A single 10-fold venetian blind CV was used to choose the optimal number of LVs as well as to obtain an estimation of the error rate of the models. All models were statistically validated by using the sensitivity, specificity, class error and accuracy (Eqs. (3)-(6)):

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$Class\ error = 1 - \left(\frac{Sensitivity+Specificity}{2}\right) \quad (5)$$

$$Accuracy\ (\%) = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (6)$$

where TP and TN stand for true positive and true negative, respectively, accounting for the samples that have been correctly assigned as belonging (TP), or not belonging (TN), to a specific class. FP and FN stand for false positive and false negative, respectively, accounting for the samples that have been wrongly assigned as belonging (FP), or not belonging (FN), to a specific class.

The ANOVA, followed by Tukey's HSD test was also conducted in order to determine significant differences in the accuracy of the models using the software Statgraphics.

3. Results and discussion

3.1. Cultivar characterisation

3.1.1. Physicochemical properties

Table 1 shows the results obtained from the reference analysis of the physicochemical properties. F is one of the physicochemical properties commonly used to assess ripeness. In this work, the measures of F obtained for both cultivars showed no statistical differences, which means that they were in a similar stage of ripeness. According to the mean value of F measured for each cultivar, these fruits were considered as being within the group that Valero, Crisosto and Slaughter (2007) described as 'ready to buy'.

As noted above, the principal difference between these two cultivars is the flavour; i.e. the typical TSS values for 'Big Top' being higher than in 'Diamond Ray' and vice versa for TA. The measured values (Table 1) agreed with Crisosto et al. (2006), who found that 'Diamond Ray' had 0.8 % TA and 10.3 % TSS. The difference in TSS content

between these cultivars may be attributable to the stage of maturity, the season or the production area (Crisosto, 1994). Regarding the 'Big Top' cultivar, Giné-Bordonaba et al. (2014) reported results similar to those in the present study, i.e. 0.3 % TA and TSS between 12.2 % and 13.5 %.

Table 1. Results of analysis of physicochemical properties of both cultivars of nectarine.

Property		'Diamond Ray'		'Big Top'	
		Mean	SD	Mean	SD
F (N)		33.8 [a]	9.5	34.8 [a]	7.1
TSS (%)		11.9 [b]	1.6	12.7 [a]	2.3
TA (%)		0.7 [a]	0.1	0.4 [b]	0.1
Skin colour colorimeter	L*	36.9 [a]	6.6	36.5 [a]	6.0
	a*	27.0 [a]	4.2	26.2 [a]	3.9
	b*	13.3 [a]	5.1	13.4 [a]	4.9
Skin colour imaging	L*	28.0 [a]	8.6	27.0 [a]	8.2
	a*	44.9 [a]	5.4	41.0 [b]	5.5
	b*	27.2 [a]	8.8	24.8 [b]	8.6
	Blush zone (%)	67.0 [a]	21.4	66.3 [a]	18.4
	Ground zone (%)	33.0 [a]	21.4	33.7 [a]	18.4

Different superscript letters in the same row indicate significant differences between cultivars (p-value<0.05). SD = standard deviation; F = firmness; TA = tritratable acidity; TSS = total soluble solids

The mean L*, a* and b* colour coordinates of the skin colour using the colorimeter were not statistically different between cultivars (Table 1). However, colorimeters measure small regions only, which can be a major limitation in applications where distinguishing the colours all over the sample is of interest. This means that they are not well suited to measuring objects with a heterogeneous colour (Gardner, 2007), such as nectarines of these cultivars. However, a colour camera provides images in which the colours of the pixels are determined individually (Cubero, Aleixos, Moltó, Gómez-Sanchis, & Blasco, 2011), along with their spatial distribution. The analysis of the colour of the nectarines using imaging enable the evaluation of the skin colour of

the different colour zones separately and calculation of the relative distribution (percentage) of reddish or yellowish colour in the whole fruit.

Using this percentage, a mean value of the L*, a* and b* coordinates was calculated from the images. On average, a reddish colour was present on 67 % of the fruit surface and a yellowish colour on 33% in both cultivars (Table 1). Even so, the mean colour using imaging indicated that the a* and b* scores were statistically different in the two cultivars, i.e. both were higher in 'Diamond Ray'. However, the differences were too small to be detected visually by the human eye, especially during a rapid fruit-sorting process.

3.1.2. Spectral analysis

Differences between cultivars were observed in their hyperspectral spectra (Figure 2). The pre-processed (SNV) mean spectra of the two cultivars followed a similar spectral pattern but had clear differences at specific wavelengths.

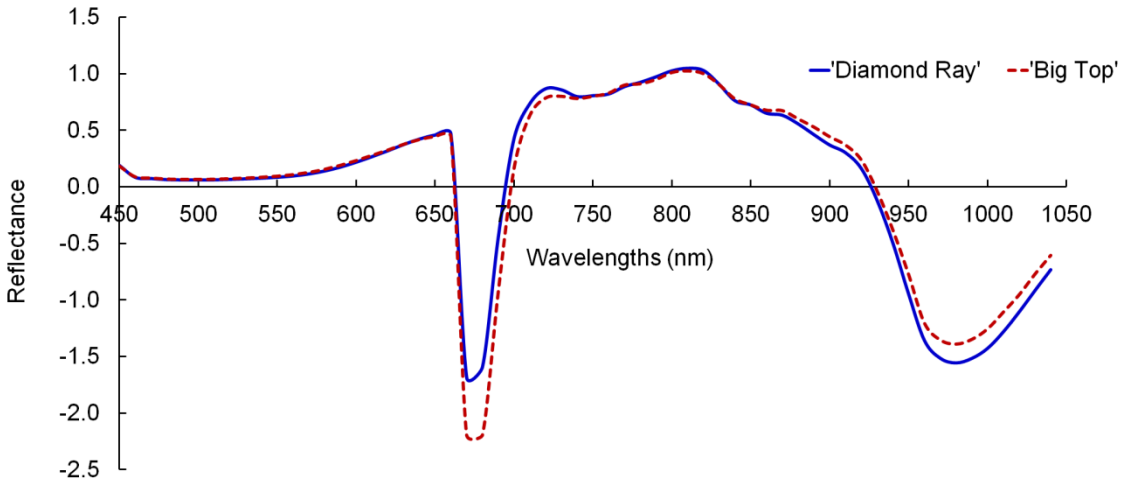


Figure 2. Mean hyperspectral image spectra of 'Diamond Ray' and 'Big Top' cultivars.

In the VIS region, no apparent differences could be visualised in the range between 400 and 600 nm where carotenoids are present. In contrast, the 'Big Top' cultivar had lower reflectance (higher absorbance) than 'Diamond Ray' near 680 nm, which is associated with chlorophylls (Rajkumar et al., 2012), suggesting a higher content of this

molecule. This agrees with the differences in the values of a* and b* found in the colour analysis (Table 1).

In the NIR region, the absorption bands for acids and sugars are usually found around 800 nmand 840 nm respectively, attributable to the hydroxyl groups of these compounds (Malegori et al., 2017; Yang et al., 2015). However, only small differences are usually observable due to the water absorption bands which dominate the spectrum (Nicolaï et al., 2007). In this region, the main differences observed in the spectra were at wavelengths above 850 nm and, in particular, around 970 nm, where Lu and Peng (2006) described a peak associated with water absorption, which in this case was more pronounced in 'Magique' nectarines (Figure 2).

3.2. Overview of the spectral data

A PCA was performed in order to obtain an overview of the distribution of the spectral data information from the samples of both cultivars. The PCA results from the individual pixel spectra and the mean spectra of each fruit are shown in Figures 3 and 4, respectively.

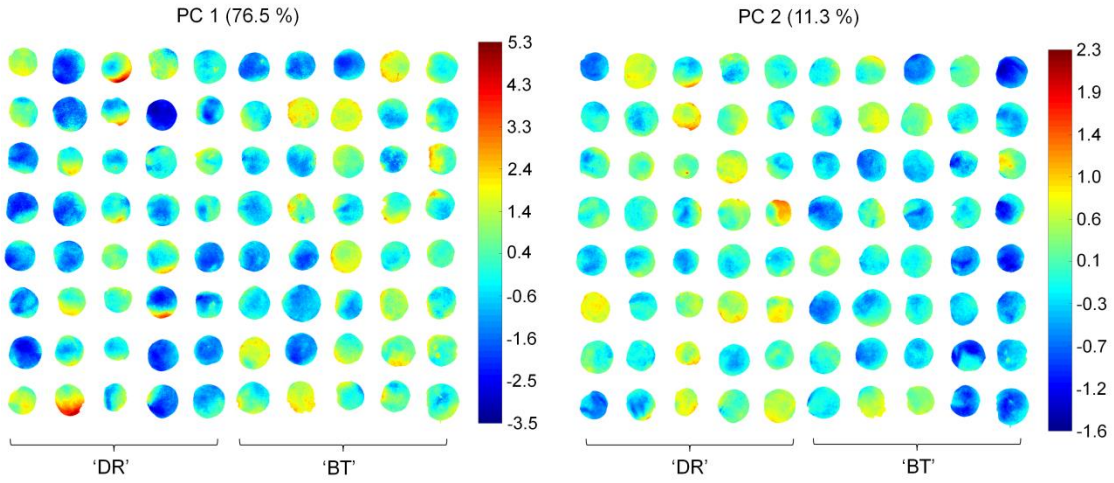


Figure 3. Score image of the two first PC of the PCA model using pixel spectra of 40 fruits of each cultivar from the calibration set.
The percentages indicate the explained variance (87.8 % of the total variance). The variations in the colour in both score plots show features linked to the distribution or content of the biochemical constituents in each fruit and cultivar. 'DR'='Diamond Ray'; 'BT'='Big Top'.

Forty samples of each cultivar were randomly selected to provide individual pixel spectra and this data was used to generate a score image plot. The first two PCs explained 87.8% of the total variance (76.5 % and 11.3 %, respectively). The variations in the colour within each fruit showed the distribution or content of the biochemical constituents. A possible trend was discerned in PC2, where pixels with low values (dark blue) were found mostly in 'Big Top' samples; however, there was little difference in individual fruit spectra of the 'Diamond Ray' and 'Big Top' cultivars.

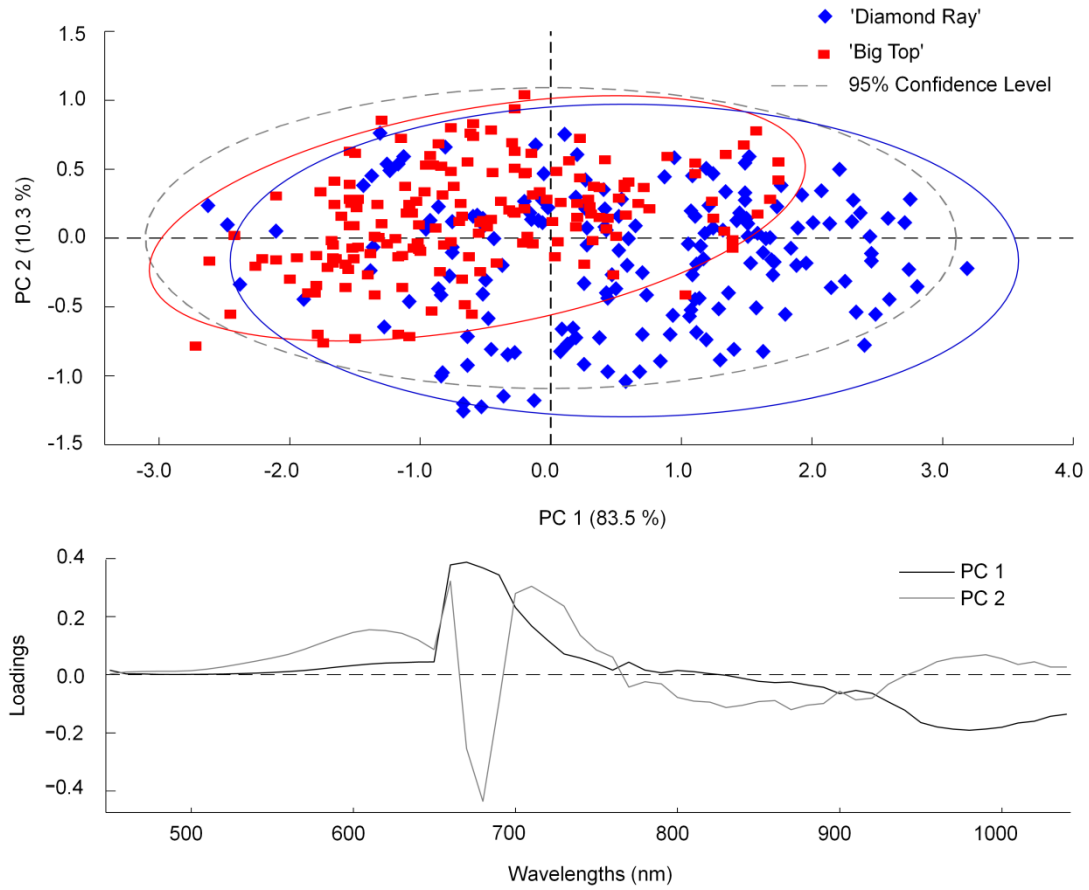


Figure 4. Score (top) and loadings plot (bottom) of the PCA of the mean spectra of the calibration set.

In the PCA of the mean spectra of the calibration set, the first two PCs (Fig. 4) explained 93.3 % of the variance (81.4 % and 11.9 %, respectively). The ellipses for the two cultivars appeared distinct, but discrimination between them was not possible because of overlap.

Although the loadings obtained for PC1 and PC2 (Figure 4) might have offered information on the most important wavelengths to distinguish the cultivars, this was not useful because separation was not evident in the preceding plot (Figure 4). PCA maximises the variance in the first components, which may or may not be related to the segregation of the classes; this does not guarantee the class separability of data due to its unsupervised nature (Jolliffe, 2002).

3.3. Cultivar classification using individual pixel spectrum

A PLS-DA model was performed using the spectral range of 450-1040 nm and the spectrum of the individual pixels of each fruit of the calibration set. The values obtained for sensitivity and specificity (Table 2) indicated that the number of samples correctly identified as belonging to a specific cultivar, or not, was above 0.80 in the CV set, using five LVs. Sensitivity of 0.83 and 0.86 was determined for 'Diamond Ray' and 'Big Top' respectively being the accuracy of classification 84.8 % and error 0.15.

Table 2. Cultivar discrimination using the pixel spectrum approach.

#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
60	5	Calibration	'DR'	0.83	0.86	0.15	84.8
			'BT'	0.86	0.83		
		CV	'DR'	0.83	0.86	0.15	84.8
			'BT'	0.86	0.83		
		Validation pixel	'DR'	0.79	0.89	0.16	83.8
			'BT'	0.89	0.79		
		Validation object	'DR'	0.78	0.91	0.16	84.4
			'BT'	0.91	0.78		

V=Variables; LV=Latent variables; CV=Cross validation; 'DR'=Diamond Ray'; 'BT'='Big Top'

Using the spatial data collected by the imaging system the combined results were applied to the calibration set. The predicted class of each pixel was obtained by introducing the spectrum measured for those pixels into the previously built model, and visualising the result. Each pixel was coloured blue if it was assigned to 'Diamond

Ray' or red if it was assigned to 'Big Top', as shown in Figure 5A. The accuracy of this classification was 83.8 % and error 0.16.

To classify each fruit using this approach, the whole fruit was assigned to the class found in the majority of its pixels (Fig. 5B). In this case, the accuracy and the classification error were 84.4 % and 0.16. In both cases, 'Big Top' was also the best discriminated, with a sensitivity of about 0.90.

3.4. Cultivar classification using mean fruit spectrum

The sensitivity and specificity in the results of calibration using the mean fruit spectrum approach giving values above 0.90 using six LVs (Table 3). In this case, both cultivars were discriminated similarly and the accuracy of classification of the CV was 93.2 % and error 0.07.

Table 3. Cultivar discrimination using the mean spectrum approach.

#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
60	6	Calibration	'DR'	0.94	0.94	0.06	93.8
			'BT'	0.94	0.94		
		CV	'DR'	0.93	0.94	0.07	93.2
			'BT'	0.94	0.93		
		Validation	'DR'	0.94	0.94	0.06	94.4
			'BT'	0.94	0.94		

V=Variables; LV=Latent variables; CV=Cross validation; 'DR'=Diamond Ray'; 'BT'='Big Top'

In order to get a graphical view of the veracity of the classification obtained using the validation set, the class for each fruit was predicted by introducing the mean spectrum measured into the previously built model. The result was visualised showing the fruit coloured blue if the mean value was assigned by the model to 'Diamond Ray' or red if it was assigned to 'Big Top' (Fig. 6A). The results for the validation set were similar to those obtained in the calibration, showing an accuracy of 94.4 % with a classification error of 0.06. The ANOVA results indicated that the mean spectra model was significantly better than the pixel model ($p < 0.05$) to classify the fruits.

As Williams and Kucheryavskiy (2016) pointed out, using properly computed object features as the mean spectrum decreases the amount of data, leading to more stable classification models. Furthermore, this approach avoids classifying by pixels when objects from different classes contain many similar pixels and are easily miss-assigned to the opposite class, such as for the cultivars studied in this work. On the other hand, it is important to include the negative influence of the sphericity of the fruits on the reflectance of the light. As seen in Figure 5A, most errors occur at the borders of the fruit, since the centres are usually well illuminated. The pixels near the borders are therefore more likely to be wrongly classified, thus affecting the overall result. In contrast when using the mean fruit spectrum, the averaging minimises these errors.

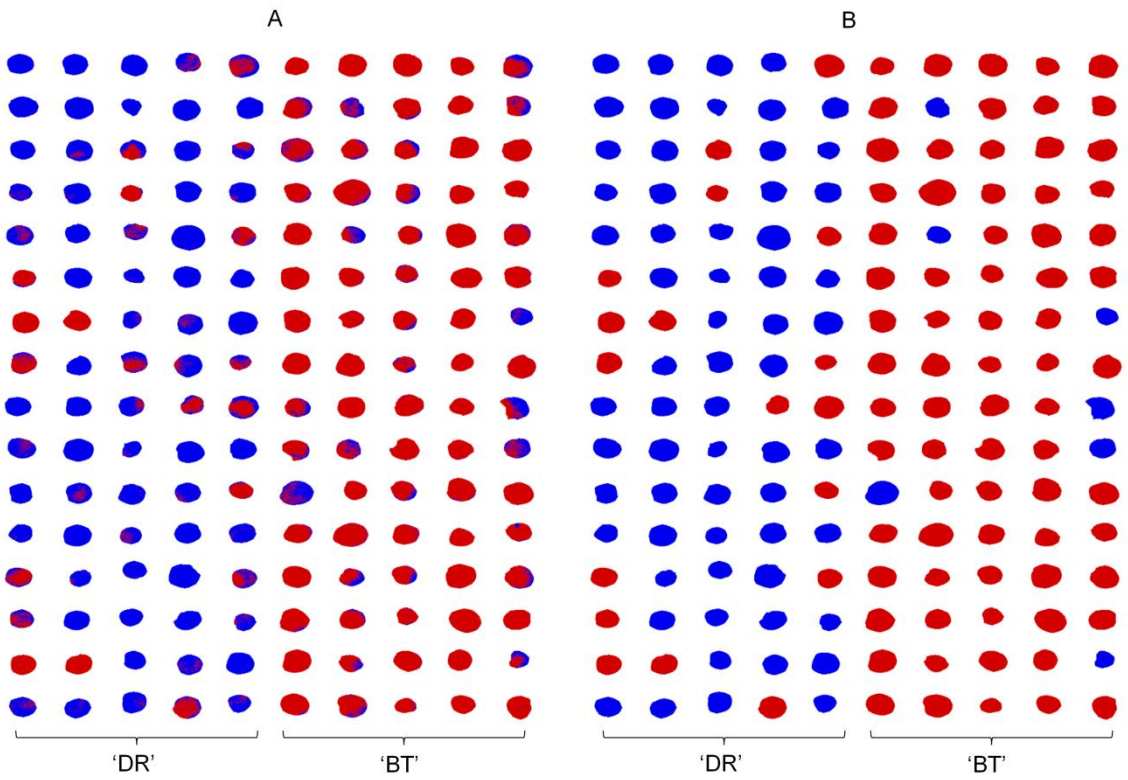


Figure 5. Visualisation of cultivar classification using individual pixel spectrum: A) Pixel classification method; B) Object classification method.

Blue='Diamond Ray'; Red='Big Top'.

3.5. Selection of the optimal wavelengths

In order to optimise the algorithms for an automatic in-line sorting system working at high speed, it is important to reduce the computational complexity generated by the huge amount of data obtained by hyperspectral imaging systems. This problem is commonly alleviated by techniques that retain the information in the few bands that reveal the most variability and therefore most significant information in the hyperspectral image (Du & Sun, 2006). The method used in this study was the vector of the regression coefficients. This measures the association between each variable and the response and selects variables in two steps: (i) the PLS-DA model is fitted to the data, and (ii) the variable selection is based on a threshold (Mehmood et al., 2012). Variables with a high absolute value can be selected because they make the highest contribution to cultivar classification and those with a small absolute value can be ignored. In this study, the regression coefficients were obtained from the PLS-DA model using the mean fruit spectrum approach, due to its higher accuracy in the classification of both cultivars.

Figure 7 shows the vector of regression coefficients. Those peaks where the absolute value was highest were selected as important wavelengths. In the VIS region the selected wavelengths were at 630, 650, 680 and 720 nm while in the NIR region they were 750-770, 790, 810-840, 860 and 900 nm.

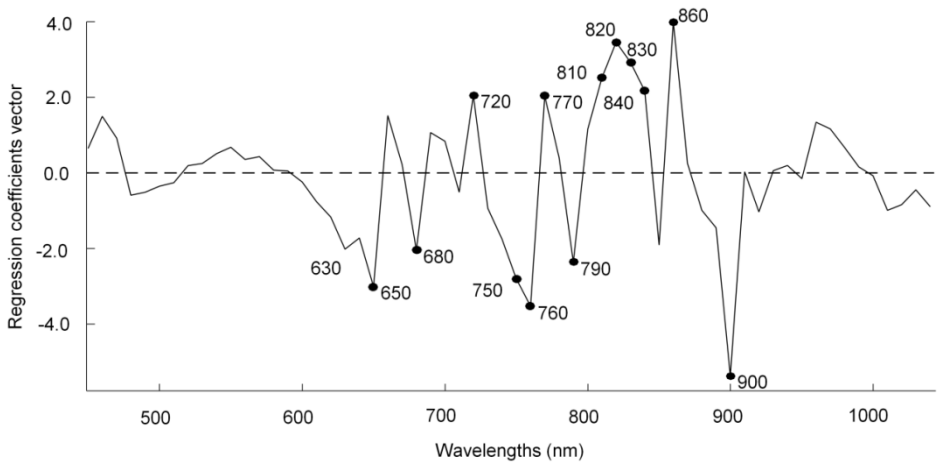


Figure 7. Vector of regression coefficients of the PLS-DA model using mean spectra and with the optimal wavelengths selected.

The optimised PLS-DA model was performed using the 14 selected wavelengths as input (Table 4). The sensitivities and specificities in the CV were similar to the full model using six LV (Table 3).

Table 4. Cultivar discrimination using the mean spectrum and the optimal wavelengths selected.

#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
14	6	Calibration	'DR'	0.94	0.94	0.06	93.8
			'BT'	0.94	0.94		
		CV	'DR'	0.93	0.94	0.07	93.2
			'BT'	0.94	0.93		
		Validation	'DR'	0.95	0.98	0.04	96.3
			'BT'	0.98	0.95		

V=Variables; LV=Latent variables; CV=Cross validation; 'DR'=Diamond Ray'; 'BT'='Big Top'

In the prediction set, using only the 14 wavelengths, the sensitivity for the two cultivars increased from 0.94 for both to 0.95 and 0.98, in 'Diamond Ray' and 'Big Top' respectively. Fig. 6 shows the results of both classifications, using the full spectrum (Figure 6A) and the optimal wavelengths (Figure 6B) in which more fruits were coloured as they should be when the wavelengths selected as the most important. However, the accuracy obtained, 96.3 %, was not statistically different ($p > 0.05$) from the accuracy of the full model (96.3 % and 94.4 %, respectively).

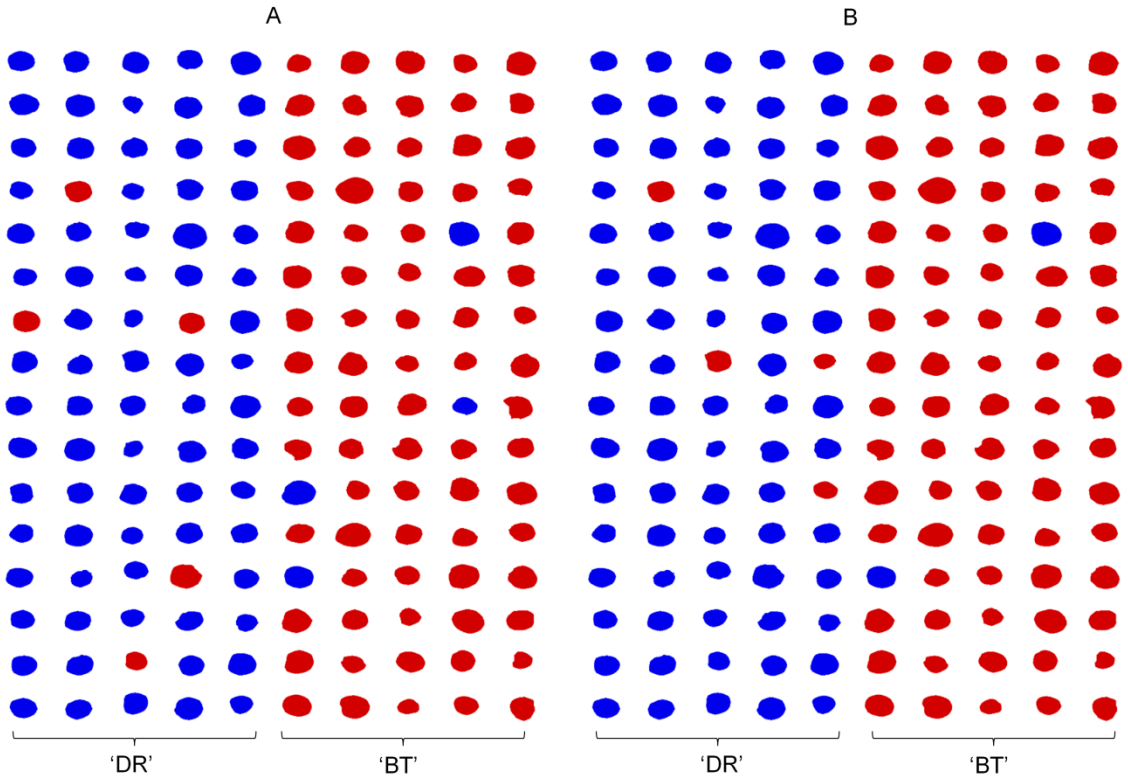


Figure 6. Visualisation of cultivar classification using mean spectrum: A) Classification using the full range; B) Classification using 14 optimal wavelengths.
Blue='Diamond Ray'; Red= 'Big Top'.

3.6. Hyperspectral imaging vs. colour and visual analysis

When the validation set was classified visually by the trained panel, the same fraction of each cultivar was identified correctly (Table 5). However, the accuracy was very low, i.e. 54.5 % with a classification error of 0.46. This demonstrates difficulty of the human eye to distinguish between the similar external appearances of these cultivars.

Classification by the colour data had similar accuracy (p-value > 0.05) to that achieved by the trained panel (Table 5), i.e. 56.9 % accuracy and error of 0.43. This is especially poor in comparison with the results of the hyperspectral imaging using 14 wavelengths, i.e. 96.3 %, error 0.04 (Table 4).

Table 5. Results of cultivar classification using colour imaging and a trained panel

	#V	#LV		Class	Sensibility	Specificity	Error	Accuracy (%)
Colour imaging *PLS-DA*	3	2	Calibration	'DR'	0.75	0.61	0.32	68.0
				'BT'	0.61	0.75		
			CV	'DR'	0.75	0.62	0.32	68.3
				'BT'	0.62	0.75		
			Validation	'DR'	0.65	0.49	0.43	56.9
				'BT'	0.49	0.65		
Trained panel	-	-	Validation	'DR'	0.54	0.55	0.46	54.5
				'BT'	0.55	0.54		

V=Variables; LV=Latent variables; CV=Cross validation; 'DR'=Diamond Ray'; 'BT'='Big Top'

These results are in agreement with the work carried out by Nogales-Bueno, Rodríguez-Pulido, Heredia, and Hernández-Hierro (2015) that used NIR hyperspectral and colour imaging to discriminate between four red grape cultivars. Only 52.0 % of the samples were correctly classified using colour imaging but this figure increased to 86.0 % using hyperspectral imaging. Furthermore, Font et al. (2014) described an in-line system for verification of nectarine cultivars with close harvest times using different colour space layers of the skin colour histogram. The success of their technique was 100 % in comparing fruits of three cultivars with a single cultivar for reference. In the same experiments, human classification achieved 86 % accuracy, likely attributable to the large differences in the skin colour of the cultivars tested.

The high rate of accuracy in classification of these cultivars using hyperspectral imaging was important because of the external similarity of the cultivars studied. This makes it difficult to accurately identify the cultivars by colour features, although they appear very different to consumers at the table. This is a genuine problem for the industry. Although colour imaging is a rapid and inexpensive tool, it has lower discrimination power for cultivars with very similar appearance, which necessitates the use of more VIS wavelengths and optimal wavelengths in the NIR region.

4. Conclusions

The capability of VIS-NIR hyperspectral imaging to discriminate very similar cultivars of nectarine has been demonstrated in this work.

The classification of these two cultivars by colour imaging or by a trained panel was very poor, achieving an accuracy of only 56.9 % and 54.5 % respectively. However, hyperspectral imaging supported by chemometric methods and optimised by reduction of the spectral and spatial information enabled classification more accurately than by traditional manual or colour-based systems, and it is also faster than destructive methods.

The use of the mean spectrum of the fruit as input of the predictive models provided classification accuracy of 94.4 %. To cope with the huge amount of data captured by the hyperspectral systems, the vector of the regression coefficients of a PLS-DA model identified 14 wavelengths which were selected as optimal, producing the best classification model with a classification accuracy of 96.3 %.

This technique may have potential as a tool for rapid and non-destructive cultivar discrimination, allowing the selection of fruit that is better suited to the consumer's preferences. Nevertheless, the results of this study should be confirmed on a larger sample set of fruits grown in different areas and harvested at different stages of ripeness before they can be implemented in an in-line system.

Acknowledgements

This work was partially funded by INIA and FEDER funds through project RTA2015-00078-00-00. Sandra Munera thanks INIA for the FPI-INIA grant num. 43 (CPR2014-0082), partially supported by European Union FSE funds. The authors wish to thank Fruits de Ponent (Lleida) for providing the fruit.

References

AENOR. (1981). Productos derivados de frutas y verduras, determinación de la acidez valorable. UNE 34211: 1981

Amigo, J.M., Babamoradia, H. & Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta 896*, 34–51.

Amodio, M.L., Capotorto, I., Chaudhry, M.M.A. & Colelli, G. (2017). The use of hyperspectral imaging to predict the distribution of internal constituents and to classify edible fennel heads based on the harvest time. *Computers and Electronics in Agriculture 134*, 1-10

Crisosto, C.H. (1994). Stone fruit maturity indices: a descriptive. *Postharvest News and Information 6*, 65-68.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. & Blasco, J. (2011). Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables. *Food Bioprocess Technology 4*, 487–504

Du, C.J. & Sun, D.W. (2006). Learning techniques used in computer vision for food quality evaluation: a review. *Journal of Food Engineering, 72*, 39–55.

Echeverría, G., Cantín, C.M., Ortiz, A., López, M.L. & Graell, J. (2015).The impact of maturity, storage temperature and storage duration on sensory quality and consumer satisfaction of 'Big Top' nectarines. *Scientia Horticulturae 190*, 179-186.

Erkinbaev, C., Henderson, K., & Paliwal, J. (2017). Discrimination of gluten-free oats from contaminants using near infrared hyperspectral imaging technique, *Food Control 80*, 197-203.

Feng, C.H., Makino, Y., Oshita, S., & García-Martín J.F. (2017). Hyperspectral Imaging and Multispectral Imaging as the Novel Techniques for Detecting Defects in Raw and Processed Meat Products: Current State-of-the-Art Research Advances. *Food Control*, In press, DOI: /10.1016/j.foodcont.2017.07.013.

Folch-Fortuny, A., Prats-Montalbán, J.M., Cubero, S., Blasco, J. & Ferrer, A. (2016). VIS/NIR hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus fruits. *Chemometrics and Intelligent Laboratory Systems 156*, 241-248

Font, D., Tresanchez, M., Pallejà, T., Teixidó, M., Martinez, D., Moreno, J. & Palacín, J. 2014. An image processing method for in-line nectarine variety verification based on the comparison of skin feature histogram vectors. *Computers and Electronics in Agriculture 102*, 112–119.

Gardner, J.L. (2007). Comparison of calibration methods for tristimulus colorimeters. *Journal of Research of the National Institute of Standards and Technology 112*, 129–138.

Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report, Opto- Knowledge Systems Inc. OKSI.

Giné-Bordonaba, J., Cantina, C.M., Larrigaudière, C., López, L., López, R. & Echeverría, G. (2014). Suitability of nectarine cultivars for minimal processing: The role of genotype, harvest season and maturity at harvest on quality and sensory attributes. *Postharvest Biology and Technology 93*, 49-60.

Herrero-Langreo, A., Lunadei, L., Lleó L., Diezma, B. & Ruiz-Altisent, M. (2011). Multispectral Vision for Monitoring Peach Ripeness. *Journal of Food Science 2*, 178-187.

Huang, F., Zhang, S., Yang, Y., Man, Z., Zhang, X. & Wu, Y. (2015). Application of hyperspectral imaging for detection of defective features in nectarine fruit. *Transactions of the Chinese Society for Agricultural Machinery 11,* 252-259.

Iglesias, I. & Echeverría, G. (2009). Differential effect of cultivar and harvest date on nectarine colour, quality and consumer acceptance. *Scientia Horticulturae 120*, 41–50.

Iqbal A., Sun, D.-W., & Allen, P. (2014). An overview on principle, techniques and application of hyperspectral imaging with special reference to ham quality evaluation and control. *Food Control 46*, 242-254.

Jolliffe, I.T. (2002). Principal component analysis (2nd ed.). New York: Springer.

Kandpal, L. M., Lee, S. , Kim, M.S., Bae, H. & Cho, B.K. (2015). Short wave infrared (SWIR) hyperspectral imaging technique for examination of aflatoxin B1 (AFB1) on corn kernel. *Food Control 51*, 171-176.

Keresztes, J.C., Goodarzi, M., & Saeys, W. (2016). Real-time pixel based early apple bruise detection using short wave infrared hyperspectral imaging in combination with calibration and glare correction techniques. *Food Control 66*, 215-226.

Li, J., Chen, L., Huang, W., Wang, O., Zhang, B., Tian, X., Fan, S. & Li, B. (2016). Multispectral detection of skin defects of bi-colored peaches based on VIS–NIR hyperspectral imaging. *Postharvest Biology and Technology 112*, 121–133.

López-Maestresalas, A., Keresztes, J.C., Goodarzi, M., Arazuri, S., Jaren, C., & Saeys W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control 70*, 229-241.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology 5*, 1121–1142.

Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering 93*, 161–171.

Malegori, C., Nascimento, E. J., Tonetto de Freitas, S., Pimentel, M. F. & Casiraghi, E. (2017). Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms. *Talanta, 165*, 112-116.

Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems 118*, 62–69.

Munera, S., Amigo, J.M., Blasco, J., Cubero, S., Talens, P. & Aleixos, N. (2017). Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging. *Journal of Food Engineering 214*, 29-39.

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys,W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biology and Technology 46*, 99–118.

Nogales-Bueno, J., Rodríguez-Pulido, F. J., Heredia, F. J. & Hernández-Hierro J. M. (2015). Comparative study on the use of anthocyanin profile, color image analysis

and near-infrared hyperspectral imaging as tools to discriminate between four autochthonous red grape cultivars from La Rioja (Spain). *Talanta 131*, 412–416.

Pan, L., Zhang, Q., Zhang, W., Sun, Y., Hua, P. & Tu, K. (2016). Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chemistry 192*, 134–141.

Pérez-Marín, D., Sánchez, M. T., Paz, P., González-Dugo, V. & Soriano, M. A. (2011). Postharvest shelf-life discrimination of nectarines produced under different irrigation strategies using NIR-spectroscopy. *LWT - Food Science and Technology 44*, 1405-1414.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G.S.V. & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194–200.

Reig, G., Alegre, S., Gatius, F., & Iglesias, I. (2013). Agronomical performance under Mediterranean climatic conditions among peach [*Prunus persica* (L.) Batsch] cultivars originated from different breeding programs. *Scientia Horticulturae 150*, 267–277.

Reig, G., Iglesias, I., & Echeverría, G. (2009). Agronomical performance, fruit quality and sensory attributes of several flat peach and flat nectarine cultivars. VII International Peach Symposium 962, 563-569.

Rinnan, Å., van den Berg, F., Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry 28*, 1201-1222.

Serranti, S., Cesare, D., Marini, F. & Bonifazi, G. (2013). Classification of oat and groat kernels using NIR hyperspectral imaging. *Talanta 103*, 276–284.

Sun, Y., Wang, Y., Xiao, Y., Gu, X., Pan, L. & Tu, K. (2017). Hyperspectral imaging detection of decayed honey peaches based on their chlorophyll content. *Food Chemistry 235*, 194–202

USDA. (2016). EU-28 Stone Fruit Annual. https://gain.fas.usda.gov/ Accessed 15.04.17.

Valero, C., Crisosto, C. H. & Slaughter, D. (2007). Relationship between nondestructive firmness measurements and commercially important ripening fruit stages for peaches, nectarines and plums. *Postharvest Biology and Technology 44*, 248–253.

Vélez-Rivera, N., Gómez-Sanchis, J., Chanona-Pérez, J.J., Carrasco, J.J., Millán-Giraldo, M., Lorente, D., Cubero, S. & Blasco, J. (2014). Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. *Biosystems Engineering 122,* 91-98.

Verdú, S., Vásquez, F., Grau, R., Ivorra, E., Sanchez, A.J., & Barat J.M. (2016). Detection of adulterations with different grains in wheat products based on the hyperspectral image technique: The specific cases of flour and bread. *Food Control 62*, 373-380.

Yang, C.H., Sun, D.W., Pu, H., Wang, N.N. & Zhu, Z. (2015). Rapid detection of anthocyanin content in lychee pericarp during storage using hyperspectral imaging coupled with model fusion. *Postharvest Biology and Technology 103*, 55–65.

Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C. & Huang, D. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (*Prunus persica*). *Computers and Electronics in Agriculture 114*, 14–24.

Zhu, N., Lin, M., Nie, Y., Wu, D. & Chen, K. (2016). Study on the quantitative measurement of firmness distribution maps at the pixel level inside peach pulp. *Computers and Electronics in Agriculture 130*, 48–56.

II. PERSIMMON

CHAPTER IV

Non-destructive assessment of the internal quality of intact persimmon using colour and VIS/NIR hyperspectral imaging

Sandra Munera[a], Cristina Besada[b], Nuria Aleixos[c], Pau Talens[d], Sergio Cubero[a], Alejandra Salvador[b], Da-Wen Sun[e f], José Blasco[a]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Centro de Tecnología Postcosecha, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[c] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[d] Departamento de Tecnología de Alimentos, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[e] School of Food Science and Engineering, South China University of Technology, Guangzhou, 510641, PR China

[f] Food Refrigeration and Computerised Food Technology (FRCFT), Agriculture and Food Science Centre, University College Dublin, National University of Ireland, Belfield, Dublin 4, Ireland

LWT - Food Science and Technology 77 (2017), 241-248

Abstract

The internal quality of intact persimmon cv. 'Rojo Brillante' was assessed trough VIS-NIR hyperspectral imaging. Fruits at three stages of commercial maturity were exposed to different treatments with CO_2 to obtain fruit with different ripeness and level of astringency (ST content). Spectral and spatial information were used for building classification models to predict ripeness and astringency trough multivariate analysis techniques like LDA, QDA and SVM. Additionally, flesh F was predicted by PLS-R. The full spectrum was used to determine the internal properties and later PCA was used to select optimal wavelengths (580, 680 and 1050 nm). The correct classification was above 92.0 % for the three classifiers in the case of ripeness and 95.0 % for QDA in the case of astringency. A value of R^2 0.80 and a ratio of prediction deviation (RPD) of 1.9 were obtained with the selected wavelengths for the prediction of firmness which demonstrated the potential of hyperspectral imaging as a non-destructive tool in the assessment of the F, ripeness state and astringency level of 'Rojo Brillante' persimmon.

1. Introduction

Spain is one of the major producers of persimmon (*Diospyros kaki* L.) among European countries (Plaza et al., 2012). The principal variety grown in Spain is 'Rojo Brillante', mostly located in the region of Ribera de Xuquer Valley near Valencia (Spain) with more then 100.000 T per year. This cultivar is very appreciated by consumers because it good aspect, high size, flavour and absence of seeds. However, this cultivar is astringent at harvest and the fruit cannot be consumed until a high degree of over-ripeness when allowed to rest and soften for a long period after harvest.

This has been traditionally a handicap for the commercialization of this fruit since once the fruit losses the astringency by overripe, it acquires a soft jelly-like consistency being difficult to handle and eat. Now, some methods have been developed to eliminate quickly the astringency without losing the F, as exposing fruit to high CO_2 concentrations (95-100 %) during 18-24 h. This method is based on promoting anaerobic respiration in the fruit, giving rise to an accumulation of acetaldehyde which reacts with the ST that are the responsible for the astringency (Matsuo, Ito, & Ben-Arie, 1991).

In Figure 1 can be appreciated the differences between a persimmon naturally de-astringed by over-ripeness and another de-astringed using a CO_2 treatment. Since the success of the treatment was demonstrated (Besada et al., 2010; Salvador et al., 2007), it has been adopted by industry as the standard deastringency method, and utilized to give the fruit in addition a sweet taste and firm texture similar to the apple, highly appreciated by the consumers.

However, the effectiveness depends on the fruit F at harvest, since maturation process is accompanied by a gradual decrease of firmness (Salvador et al., 2008). A problem is that the stage of maturity at harvest is currently determined based on the visual inspection of experienced growers or using colorimeters due the relationship between the changes in external colour and the internal changes (Salvador et al., 2007).

Figure 1. Persimmon de-astringed using a CO_2 treatment (left) and persimmon naturally de-astringed by over-ripeness (right). The first shows firm and crisp flesh while the second present a very soft texture.

The current way to know the level of astringency in the fruit after CO_2-treating is by destructive measurement of ST content in random fruits by means of the tannin print method (Matsuo & Ito, 1982) which consists of using a $FeCl_3$-impregnated filter paper to obtain a print of the content and distribution of the tannins trough the reaction with the $FeCl_3$ in the paper. Then, this print is visually assessed by trained workers being this method subjective and destructive and therefore the development of other new non-destructive and accurate methods is needed.

Computer vision systems have been traditionally used to create tools for the objective estimation of the quality of intact fruit production (Cubero et al., 2011) and have already been explored to assess quality of persimmon. Mohammadi et al. (2015) used colour information to determine the maturity of this fruit through colour analysis and classify the fruit into three commercial maturity stages.

Standard computer vision systems tend to mimic the human eye and hence are based on sensors sensible to visible wavelengths. But to analyse internal composition it is necessary the use of technology sensible to non-visible wavelengths related with chemical compounds. This can be achieved by using hyperspectral imaging (Lorente et al., 2012) that is a powerful non-invasive technology that allows obtaining the spatial distribution of the spectral information and it is being used from recent in the internal quality inspection of food (Cheng & Sun 2015; Cheng, Sun, & Cheng, 2016, Cheng, Sun, Pu, & Liu, 2016; Gómez-Sanchis et al., 2013) or to assess some properties of fruits like the ripeness in apples (ElMasry et al., 2008), citrus fruits (Folch- Fortuny et al., 2016), pepper (Schmilovitch et al., 2014), or mango (Vélez-Rivera et al., 2014).

Hyperspectral imaging in persimmon has been used by Munera et al., (2017) to create images showing the distribution of the predicted astringency of each pixel in the fruit, and by Wei et al., (2014) to predict F. However, in this work, the authors claimed that more research is needed to include more samples as well as different regions and different postharvest treatments to ascertain the discrimination power of this method and it is therefore necessary to investigate new methods especially to discriminate among fruits with slightly different stages of ripeness or levels of astringency as those exposed to a CO_2 treatment, to achieve a demand from both the industry and the consumers.

This work proposes a new non-destructive approach based on VIS-NIR hyperspectral imaging and multivariate analysis to determine the F, ripeness state and astringency level of intact persimmon 'Rojo Brillante' as alternative to the current destructive and/or subjective techniques.

2. Material and methods

2.1. Fruit samples and internal quality assessments

A total of 90 persimmon (*Diospyros kaki* cv. 'Rojo Brillante') fruit were harvested in L'Alcudia (Valencia, Spain) at three different stages of commercial maturity (M1, M2 and M3) (Figure 2) corresponding to different moments of the season (early November, end November, and mid December). A total of 30 fruits, with apparently similar size and colour were collected for each maturity stage. In order to obtain three different levels of astringency, the fruits in each maturity stage were equally divided into three sets. The first set (control fruits with high astringency) consisted of fruits not treated, the second set (medium astringency fruits) consisted of fruits treated in closed containers at 20 ^{o}C with 90% RH and 95 % of CO_2 for a period of 12 h, and the remaining set (deastringed fruits) were fruits treated under the same conditions for 24 h.

Figure 2. Images of persimmon at each maturity stage.

After each treatment, all the fruits were measured using a colorimeter, a digital camera, and a hyperspectral imaging system. Later, flesh F of all fruits was determined by means of a universal testing machine (4301, Instron Engineering Corp., MA, USA) equipped with an 8 mmpuncture probe. The crosshead speed during the F testing was 10 mm/min. During the test, the force increased smoothly until it drastically decreased when the flesh was broken and the maximum peak force was registered. Results were expressed as the mean of the load (in N) required for breaking the flesh of the fruit on the two sides after peel removal. To analyse the astringency of the fruits, they were sliced and frozen at -20 °C to determine ST using the Folin-Denis method (Taira, 1995), as described by Arnal and Del Río (2004). This method is based on the reduction of the Folin-Ciocalteu reagent by soluble tannins in alkaline solution. Calibration curve was made with gallic acid. ST were extracted by homogenization of 5 g of flesh with 25 mL of 80 % methanol solution. Thereafter, samples were filtered and centrifuged for 20 min and the supernatant was reserved. More supernatant was extracted from the precipitant with methanol 80 % and added to the first. The supernatant was diluted in water at 1:7 and then Folin-Ciocalteu reagent 1 N was used to conduct the reaction. After 3 min 1 ml of saturated Na_2CO_3 was added, and the absorbance of the mixture at 725 nm was measured by colorimetry after stand for 1 h.

2.2. Colour analysis

At harvest this fruit presents a uniform colour that ranges from bright to dark orange depending on the maturity being the colour a good indicative of this property (Salvador et al., 2007). The external colour the fruit under study was characterised

using two techniques. On the one hand, a colorimeter (CR-300, Konica Minolta Inc, Tokyo, Japan) was used to obtain the colour at three points of the equatorial part of the fruit. Hunter Lab colour coordinates were obtained by the average of three measures. On the other hand, the colour was also evaluated trough images of the two sides of each fruit. The image acquisition system consisted on a digital camera (EOS 550D, Canon Inc, Japan) arranged into a squared inspection chamber that included a calibrated and uniform illumination system composed of eight fluorescent tubes (BIOLUX 18 W/965, 6500 K, OsramGmbH, Germany). The angle between the axis of the lens and the sources of illuminationwas approximately 45° to avoid direct reflections to the camera (Diago et al., 2015), but due to the spherical shape of the samples these reflections could not be totally avoided this way and hence cross-polarization was also used (ElMasry et al., 2012).

A total of 180 images were obtained with a size of 2592 x 1944 pixels and a resolution of 0.11 mm/pixel. Figure 2 shows examples of images of the fruits in the three maturity stages. For each image, the mean RGB colour values of the pixels of the skin were obtained using the application Food_ColorInspector (free download at http://www.cofilab.com). RGB values were later converted to Hunter Lab colour space for analysis using the equations described in Mendoza et al. (2006) and HunterLab (1996) for illuminant D65 and standard observer 10^{o}. The HunterLab coordinates were finally transformed to the colour attributes Hunter L, Hunter h and Hunter C (Hutchings, 1999). In addition, RGB values were transformed into HSI (hue, saturation, intensity) values and other indices were estimated such as the ratios a/b and a/L and the CI (CI = 1000a/Lb) (Salvador et al., 2006).

2.3. Hyperspectral imaging

Hyperspectral images of the intact persimmons in the spectral range 450-1020 nm were acquired using a camera (CoolSNAP ES, Photometrics, USA) coupled to two LCTF (Varispec VIS-07 and NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). The illumination system consisted of 12 halogen lights arranged equally into a domo inspection chamber where whole fruits were manually introduced.

Hyperspectral images with a spatial resolution of 0.14 mm/pixel and a spectral resolution of 10 nm were captured in both sides of each fruit, which lead to a tagged database of 180 hyperspectral images. In each image, a ROI of 225 x 225 pixels in the central part of the fruit was selected and analysed as the average of spectrum of all pixels for ripeness and F analysis since these properties are quite uniformly distributed in the fruit. However, for the case of the astringency, the individual spectrum of each pixel in the ROI was included in the models due the uneven distribution in the fruit of ST responsible of the astringency. To obtain the relative reflectance of a pixel in the position (x, y) of the monochromatic wavelenght, the original reflectance was corrected using a dark and white reference (Spectralon 99 %, Labsphere, Inc, NH, USA) following the procedure described in Gat (2000).

2.4. Data analysis

ANOVA and Tukey multiple range test (Statgraphics Centurion XVI - Statpoint Technologies Inc., Virginia, USA) were used to show the effects of ripeness on colour parameters obtained with both, colorimeter and computer vision system. In this analysis, the three maturity stages were the observed values (Y) and the Hunter Lab colour coordinates captured by both the colorimeter and the vision systems were the predictive variables.

Hyperspectral images consisted of 67 wavelengths and therefore the spectra obtained from these images were distributed in a matrix with 67 columns each corresponding to the reflectance value of each band where the rows represented the fruits. In addition, the pixels were labelled as belonging to any of the ripening stages (M1, M2 and M3) and treatments (HA, LA, NA) to carry out the analysis for firmness and astringency prediction. First step was a pre-processing of data using SNV to remove scatter effects from original spectral data (The Unscrambler X 10.1, CAMO Software, Oslo, Norway). Classification models to sort the fruit by ripeness stage and treatment duration (astringency level) were developed using LDA, QDA and SVM (Dutta et al., 2016). The difference between LDA and QDA classifier is that LDA uses pooled covariance to assign an unknown sample to one of the pre-defined groups

while QDA uses the covariance of each group instead of pooling them (Naes et al., 2002). On the other hand, the SVM algorithm was developed based on the concept of hyperplane and support vectors, using a linear function kernel with C value set to 1. In addition, F prediction was conducted by PLS-R (Cheng et al., 2015b) using the RPD that was defined by Williams (1987) as the ratio of standard deviation of reference values in training set to the $RMSE_P$.

Hyperspectral systems capture a huge amount of information that is redundant and correlated, especially between contiguous wavelengths (Lorente et al., 2012). Therefore, PCA was used to know if it was possible to obtain good prediction using a reduced subset of bands. Four different PCA models were built, one of using the spectral data of the ripeness assessment and the other three PCA with data of the astringency assessment for each harvest. The variables wavelengths were chosen on the basis of the size of coefficients or loadings in the eigenvectors of the principal components.

3. Results and discussion

3.1. Maturity assessment

Several differences can be observed among the spectra of the fruit in the three ripening stages shown in Figure 3. Fruits of M1 gave higher reflection values than the others in the visible region, which is in agreement with the colour analysis. An absorption peak was found around the bands 670-680 nm only for fruits in M1 stage which could be due the presence of chlorophyll in the more unripe fruit (Lleó et al., 2011). However, the fruits in M2 stage are those which gave a higher reflection in the NIR region that can be due to the chemical differences among fruit at different ripeness. The absorption peak observed around 900-1050 nm could be assigned to water absorption band. This peak was higher in M3, which may be related to water content increases in the flesh during the onset of ripening, which in other fruits has been related to cell breakage and osmotic movement of water from the flesh to the skin.

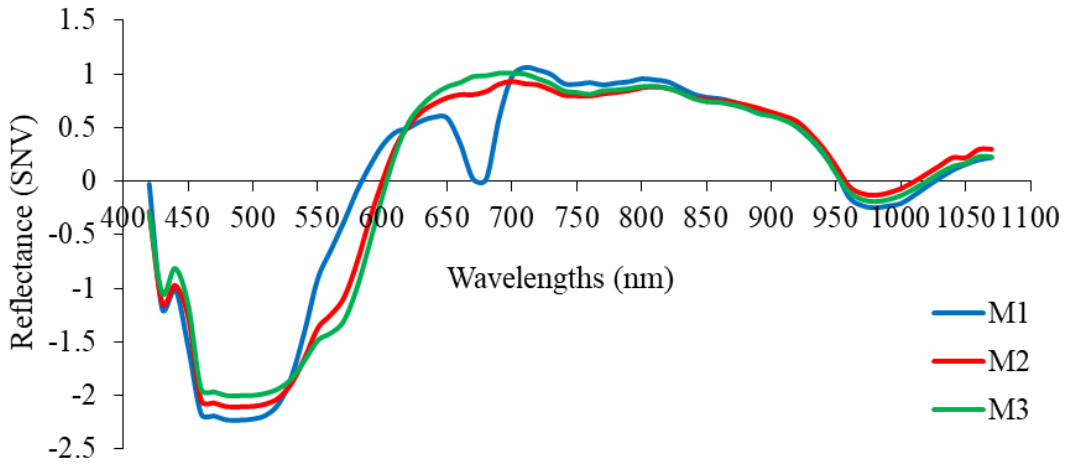


Figure 3. Average spectra of control fruits in three ripeness stages

The PCA model generated with the 67 wavelengths was analysed to identify the variables with the highest factor loadings since they reflected the importance of each wavelength in discriminating differences in the fruit (Wang et al., 2012). The loadings of the first two principal components were used for wavelength selection because these were responsible for 96% of the variance in the spectral data. The wavelengths corresponding to higher module values (peaks and valleys) at these particular principal components were selected as candidates for optimum wavelengths (Rodríguez-Pulido et al., 2013) (Figure 4). Four optimum wavelengths (450, 580, 680, and 1050 nm) were thus identified for discrimination purposes of different maturity stages. Wavelengths 450 nm and 680 nm are related with the presence of beta-carotene and chlorophyll a respectively. On the other hand, the importance of the wavelength 580 nm can be due to the colour changes during ripeness since it corresponds to the yellow colour. This would be in accordance with the ranges of h and C values shown in Table 3. The band 1050 nm could be related with an absorption region of water content although the peak is situated below 1000 nm (Lu & Peng, 2006).

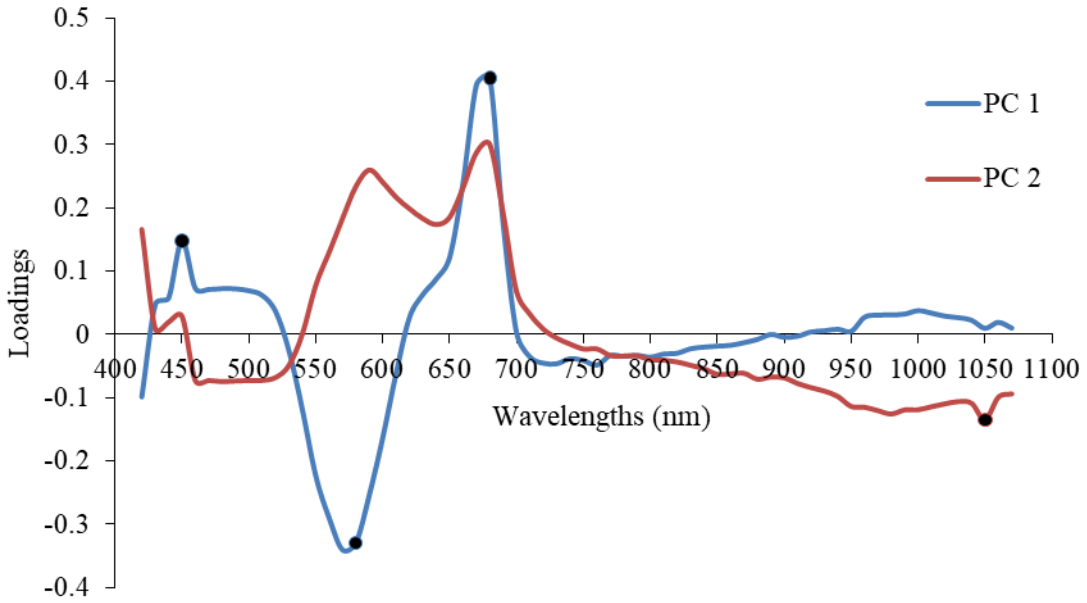


Figure 4. PC Loadings of the PC1 and PC2 showing the selected wavelengths for ripeness classification of 'Rojo Brillante' persimmon fruits.

Statistical models to classify the fruit into maturity stages were developed using the spectra of the full spectra and only the selected wavelengths. In order to build and validate the model, a 3-fold CV procedure was used (Simon, 2007). The data set of pixels was randomly partitioned into three disjoining subsets. The classifier development process was repeated three times using each two different subsets and the resulting classifiers used to classify the remaining test set. Finally the results of the three iterations were averaged. The four selected bands were used to build the models but also the possible combinations of three bands resulting that using only 580 nm, 680 nm, and 1050 nm, the results were similar to those achieved using the four bands. Using only these three selected wavelengths the success rate of correct classification was slightly lower (mean value of 94.8 %) than using the full spectrum (mean value of 98.5 %) as shown in Table 1.

Table 1. Ripeness classification of testing set by LDA, QDA and SVM using all and selected wavelengths.

Class	All wavelengths			Selected wavelengths		
	LDA	QDA	SVM	LDA	QDA	SVM
M1	99.5	99.8	99.1	98.6	99.3	98.4
M2	96.8	96.2	96.0	95.5	94.1	94.7
M3	99.0	100	99.8	83.7	93.9	94.9
Total	98.5	98.8	98.3	92.6	95.8	96.0

Different superscript letters in the same row indicate significant differences between groups (p-value<0.05), according to Tukey's test.

Comparing the three classification methods, all of them achieved a good classification above 98 % using the all wavelengths. Moreover, using only the three selected wavelengths only LDA showed an important reduction in the success rate while the other two classifiers still remain above 95 % which is considered as a good result for a non-destructive technique.

3.2. Firmness prediction

Table 2 shows the F evolution with the harvesting time (ripeness).

Table 2. Flesh firmness of 'Rojo Brillante' persimmon fruits before and after treatments in the three ripeness stages.

Group	M1	M2	M3
High astringent	$47.0^{a} \pm 4.3_{a}$	$29.0^{a} \pm 2.6_{b}$	$25.1^{a} \pm 3.4_{c}$
Low astringent	$44.7^{ab} \pm 2.6_{a}$	$30.9^{a} \pm 3.0_{b}$	$25.0^{a} \pm 4.7_{c}$
Deastringed	$40.6^{b} \pm 2.8_{a}$	$31.9^{a} \pm 2.1_{b}$	$21.1^{a} \pm 4.8_{c}$

Values are the flesh F (N) ± standard deviation. Different superscript letters in the same column (astringency) and different subscript letters in the same row (ripening) indicate significant differences between groups (p-value<0.05), according to Tukey's test.

A model based on PLS-R was built to know if it was possible to predict this property in this cultivar using the wavelengths selected in the previous study for ripeness assessment. For each fruit there was obtained only one global value of the flesh

firmness so the prediction model was built using the average values of the pixels selected for each fruit at the determined wavelengths of the hyperspectral images. CV leaving 5 % of samples for test was chosen to validate this study. This method splits randomly the calibration set into the training (95 %) and test (5 %), repeating the process 20 times. Results were achieved as the mean of the 20 repetitions.

The R^2 for the prediction of F was 0.80 and the RPD was 1.9. Viscarra-Rossel et al. (2006) suggested that calibration models will suffice for good quantitative application if RPD is larger than 1.8. The prediction results obtained was something higher than the minimum proposed but not as good as the prediction results of Wei et al. (2014) for 'Fangshi' persimmon who achieved a R^2 value of 0.91. However, in their work the F of the fruit ranged from 25 N to 1 N with large differences among the studied classes. In addition, during the ripening process of this cultivar not only drastic changes in F happened but also the skin begins to wrinkle and lose shine clearly affecting the reflectance. On the contrary, in the present work, theF gave values from 47 N to 21 N which means that these fruits are apparently firm in all maturity stages, which is logical since it is treated to be consumed as firm and crispy fruit. Figure 1 highlights the visual differences between a soft persimmon naturally deastringed and another deastringed using CO_2 treatment. Hence, for this fruit this prediction capability is considered as a good achievement taking into account the little differences between classes, especially between M2 and M3 classes.

A study was also carried out to analyse the possible correlation between the colour analysis and the F of the samples. The characterisation of the external colour was carried out using the colorimeter and the camera only for the control samples of the three stages to avoid the influence of the treatment in colour changes (Table 3). In general, the *L* and *b*, Hunter Lab coordinates, decreased but there were not statistical differences for M2 and M3. On the contrary, the value of a increased along the three stages. As a consequence of the changes observed in a and b, the h decreased and the C slightly increased along the three stages. These differences were observed in the measures given by both, the colorimeter and the camera, and reflect the loss of L of the fruit caused by the ripening process and the changes in the fruit from yellowish-orange to reddish-orange.

Table 3. Colour coordinates and attributes of the samples in the three harvests.

Stage	Colorimeter					Imaging				
	L	a	b	h	C	L	a	b	h	C
M1	58.9 ± 1.8 ^{a}	21.7 ± 3.3 ^{c}	34.8 ± 1.8 ^{a}	60.3 ± 4.4 ^{a}	40.7 ± 1.6 ^{c}	43.8 ± 1.0 ^{a}	27.8 ± 3.5 ^{c}	26.0 ± 0.5 ^{a}	49.5 ± 4.4 ^{a}	36.3 ± 2.0 ^{c}
M2	53.5 ± 1.9 ^{b}	34.5 ± 1.8 ^{b}	31.2 ± 1.4 ^{b}	46.5 ± 4.6 ^{b}	45.4 ± 1.6 ^{b}	33.9 ± 2.4 ^{b}	38.3 ± 2.5 ^{b}	20.5 ± 1.4 ^{b}	34.5 ± 3.8 ^{b}	42.1 ± 1.6 ^{b}
M3	52.6 ± 1.4 ^{b}	38.4 ± 1.7 ^{a}	30.8 ± 1.1 ^{b}	39.2 ± 1.9 ^{c}	48.3 ± 0.5 ^{a}	34.7 ± 2.0 ^{b}	41.2 ± 1.0 ^{a}	21.0 ± 1.1 ^{b}	28.6 ± 1.8 ^{c}	43.6 ± 1.3 ^{a}

Values are the mean of control samples in each harvest± standard deviation. Different superscript letters in the same column indicate significant differences between groups (p-value<0.05), according to Tukey's test.

The values of the colour attributes (L, h and C) of the colorimeter were higher than the ones obtained from the images. The higher differences were observed for the L values since the glossiness leads to a specular reflectance that reduces the contribution to the components a and b. In fact, colorimeter is very dependent on the scattering properties of the sample while the diffuse illumination of the vision system gives less dependency on the lightness of the sample than the simple illumination and filtering employed by the colorimeter (Trinderup et al., 2015). Despite the differences observed, good correlations where found between the values obtained by both methods (R^2 of 0.87, 0.80, and 0.96 for the L, C, and h respectively).

Linear regressions were performed between the different colour values, obtained with both the colorimeter and the camera, and the F. Table 4 summarises the results achieved for the coefficient of determination R^2 for each colour component using the imaging system and the colorimeter, respectively.

In general, better results are achieved with the imaging system which on the other hand makes sense since they integrate the colour of the whole surface of the fruit while colorimeter only measures in a small spot and thus increasing the variability. Good correlations are found in H (R^2 = 0.83), G (R^2 = 0.82) and h (R^2 = 0.81) or using simple ratios like a/b (R^2 = 0.83), G/R (R^2 = 0.83) or a/L (R^2 = 0.83). It is worthy of

interest that using the simple ratios measured with the imaging system could be obtained better correlations (R^2= 0.83) than using the CI that was the index used by Salvador et al. (2006) to estimate the F trough a colorimeter achieving a R^2 = 0.81.

Table 4. Results of F prediction using the different colour components measured with the imaging system and colorimeter.

	R	*G*	*B*	*H*	*S*	*I*	*G/R*	
R^2 Imaging	0.49	0.82	0.46	0.83	0.48	0.17	0.83	
	L	*a*	*b*	*CI*	*a/b*	*a/L*	*h*	*C*
R^2 Imaging	0.79	0.78	0.78	0.80	0.83	0.83	0.81	0.69
R^2 Colorimeter	0.66	0.78	0.63	0.77	0.78	0.78	0.77	0.76

R= red, G = green, B = blue, H = hue, S = saturation, I = intensity, CI = colour index, h = Hunter hue, C = chroma

3.3. Astringency prediction

The results of the measurements of ST of fresh weight for all maturity stages are shown in Table 5. It can be observed that the tannin content decreased in a similar way for the three maturity stages along with the duration of the treatment. The soluble tannins content decreased to values close to 0.4 % in the fruits treated for 12 h to 0.03 % in the fruits exposed for 24 h to CO_2.

Table 5. Soluble tannins content in 'Rojo Brillante' persimmon fruits before and after treatments in the three ripeness stages.

Group	M1	M2	M3
High astringent	0.61^{a} ± 0.09	0.65^{a} ± 0.06	0.63^{a} ± 0.07
Low astringent	0.45^{b} ± 0.04	0.43^{b} ± 0.10	0.39^{b} ± 0.06
Non astringent	0.03^{c} ± 0.00	0.03^{c} ± 0.00	0.03^{c} ± 0.00

Mean value ± standard deviation. Different superscript letters in the same column indicate significant differences between groups (p-value<0.05), according to Tukey's test.

Accordingly, Besada et al. (2010) reported that the CO_2-treatment applied for 12 h to fruit with firmness around 40 N led to a reduction of soluble tannins to values close 0.3%. Besides, it has been widely reported that a content of soluble tannins of 0.03 % after the CO_2-treatment is associated with a complete effectiveness of the deastringency process in 'Rojo Brillante' cultivar (Salvador et al., 2007, 2008).

Like in ripeness classification, three PCA models were analysed to identify the highest factor loadings in each ripeness stage. However, no wavelength selection could contribute to the astringency classification. This may be because tannins are mainly detected in the UV in the range 190-400 nm (Boulet et al., 2016), or in the NIR (2200-2300 nm) (Cozzolino et al., 2004). For this reason, the whole spectrum in the studied range (450-1020 nm) was necessary to discriminate the astringency.

Table 6 shows the results of astringency classification using the three classifiers. In general, QDA obtained the best overall classification but a reduction of the classification rate along with the maturity was observed, especially for the astringent fruits (HA and LA) in the M3 stage.

As it was shown in Table 2, a decrease of F in M1 between control and deastringed fruits was observed. However, in M2 and M3 there was no difference. This could be due because the effect of high CO_2 concentrations on the cell structure could be the cause of the important loss of firmness observed after deastringency treatment. But when the more ripe samples are treated with CO_2, no effect happens on flesh firmness because the loss of intercellular adhesion is already generalised due to the ripeness process (Salvador et al., 2007). Therefore, those changes detected by hyperspectral imaging are assigned to changes in the soluble tannins content and not to changes in texture.

Table 6. Astringency classification of test set by LDA, QDA and SVM.

	Class	Correct classification (%)		
		LDA	QDA	SVM
M1	HA	95.9	99.3	97.1
	LA	92.4	94.5	94.9
	NA	93.9	97.3	93.2
	Avg	94.1	97.0	95.1
M2	HA	93.2	96.2	92.7
	LA	93.0	95.3	91.1
	NA	93.9	95.6	95.9
	Avg	93.4	95.7	93.2
M3	HA	83.7	94.3	90.0
	LA	72.0	86.0	64.5
	NA	93.4	97.3	93.4
	Avg	83.0	92.5	82.7
Overall classification (%)		90.2 ^{b}	95.1 ^{a}	90.3 ^{b}

Different superscript letters in the same row indicate significant differences between groups (p-value<0.05), according to Tukey's test.

4. Conclusions

In this study, VIS-NIR hyperspectral imaging were evaluated as potential non-destructive methods to determine the flesh F, ripeness stage and the astringency level of 'Rojo Brillante' persimmon.

The characterisation of the colour showed that the L and b Hunter Lab coordinates decreased while the value of a increased along with the maturity. As a consequence the h decreased and the C slightly increased along the three stages using both colorimeter and image methods. Good correlations were found in some colour parameters like H (R^2 = 0.83), G (R^2 = 0.82) and h (R^2 = 0.81), but also using ratios like a/b (R^2 = 0.83), G/R (R^2 = 0.83) and a/L (R^2 = 0.83) with the data obtained by the imaging system improving previous results. Moreover, better correlations were obtained using these ratios than using the previously proposed CI (R^2 = 0.80) which

indicates the feasibility of images to assess the colour as a valid alternative to traditional and expensive colorimeters.

Using the hyperspectral system, three wavelengths (580, 680 and 1050 nm) were proposed as the optimum wavelengths for the classification of the fruits into three ripeness stages with high accuracy, more than 94 % of all samples were well classified for all of the used classifiers (LDA, QDA and SVM). Moreover, these wavelengths were used for flesh firmness prediction and the RPD value indicated that the obtained model is useful for good quantitative application. Regarding the astringency, the whole spectrum of the fruits needed to be used to classify the fruits into three levels of astringency: astringent fruit, fruit with a low-medium level of astringency and non-astringent fruit. The overall classification for the three ripeness stages was higher than 90 % for the three classifiers and higher than 95 % for QDA. These results indicate the potential proposed methodology based on hyperspectral imaging as a promising non-destructive tool to assess the internal quality of persimmon fruits destined to be de-astringed and rapidly marketed as fresh sweet fruit. However, more research is needed, involving more fruits from different regions and collected in different seasons to ascertain the discrimination power of the proposed

Acknowledgements

This work has been partially funded by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria de España (INIA) through research projects RTA2012-00062-C04-01 and RTA2012-00062-C04-03 with the support of European FEDER funds. Sandra Munera thanks to INIA the grant FPI-INIA #43 (CPR2014-0082) partially supported by the European Union FSE funds.

References

Arnal, L., & del Río, M. A. (2004). Quality of persimmon fruit cv. 'Rojo Brillante' during storage at different temperatures. *Spanish Journal of Agricultural Research 2*, 243-247.

Besada, C., Salvador, A., Arnal, L., & Martínez-Jávega, J. M. (2010). Optimization of the duration of deastringency treatment depending on persimmon maturity. *Acta Horticulturae 858*, 69-74.

Cheng, J. H., & Sun, D.-W. (2015). Rapid and non-invasive detection of fish microbial spoilage by visible and near infrared hyperspectral imaging and multivariate analysis. *LWT - Food Science and Technology 62*, 1060-1068.

Cheng, W., Sun, D.-W., & Cheng, J. H. (2016). Pork biogenic amine index (BAI) determination based on chemometric analysis of hyperspectral imaging data. *LWT - Food Science and Technology 73*, 13-19.

Cheng, W., Sun, D.-W., Pu, H., & Liu, Y. (2016). Integration of spectral and textural data for enhancing hyperspectral prediction of K value in pork meat. *LWT – Food Science and Technology 72*, 322-329.

Cozzolino, D., Kwiatkowski, M. J., Parker, M., Cynkar, W. U., Dambergs, R. G., Gishen, M., et al. (2004). Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta 513*, 73-80.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., & Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology 4*, 487-504.

Diago, M. P., Tardaguila, J., Aleixos, N., Millan, B., Prats-Montalban, J. M., Cubero, S., et al. (2015). Assessment of cluster yield components by image analysis. *Journal of the Science of Food and Agriculture 95*, 1274-1282.

Dutta, M. K., Sengar, N., Minhas, N., Sarkar, B., Goon, A., & Banerjee, K. (2016). Image processing based classification of grapes after pesticide exposure. LWT – *Food Science and Technology 72*, 368-376.

ElMasry, G., Cubero, S., Moltó, E., & Blasco, J. (2012). In-line sorting of irregular potatoes by using automated computer-based machine vision system. *Journal of Food Engineering 112*, 60-68.

ElMasry, G., Wang, N., Vigneault, C., Qiao, J., & ElSayed, A. (2008). Early detection of apple bruises on different background colors using hyperspectral imaging. *LWT - Food Science and Technology 41*, 337-345.

Folch-Fortuny, A., Prats-Montalban, J. M., Cubero, S., Blasco, J., & Ferrer, A. (2016). NIR hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus fruits. *Chemometrics and Intelligent Laboratory Systems 156*, 241-248.

Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report. Opto-Knowledge Systems Inc. OKSI.

Gómez-Sanchis, J., Blasco, J., Soria-Olivas, E., Lorente, D., Escandell-Montero, P., Martínez-Martínez, J. M., et al. (2013). Hyperspectral LCTF-based system for classification of decay in mandarins caused by Penicillium digitatum and Penicillium italicum using the most relevant bands and non-linear classifiers. *Postharvest Biology and Technology 82*, 76-86.

HunterLab. (1996). Applications note, Hunter Lab color scale. https://support.hunterlab.com/hc/en-us/article_attachments/201440625/an08_96a2.pdf *Accessed November 2016.*

Hutchings, J. B. (1999). Food color appearance. In Food Color Mechanisms (pp. 453-592). Gaithersburg, Maryland (USA): Aspen Publishers, Inc.

Lleó, L., Roger, J. M., Herrero-Langreo, A., Diezma-Iglesias, B., & Barreiro, P. (2011). Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening. *Journal of Food Engineering, 104*, 612-620.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology 5*, 1121-1142.

Lu, R., & Peng, Y. (2006). Hyperspectral Imaging for assessing peach fruit firmness. *Biosystems Engineering 93*, 161-171.

Matsuo, T., & Ito, S. (1982). A model experiment for de-astringency of persimmon fruit with high carbon dioxide: In vitro gelation of kaki-tannin by reacting with acetaldehyde. *Journal of Agricultural Food Chemistry 46,* 683-689.

Matsuo, T., Ito, S., & Ben-Arie, R. (1991). A model experiment for elucidating the mechanism of astringency removal in persimmon fruit using respiration inhibitors. *Journal of the Japanese Society for Horticultural Science 60*, 437-442.

Mendoza, F., Dejmek, P., & Aguilera, J. M. (2006). Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biology and Technology 41*, 285-295.

Mohammadi, V., Kheiralipour, K., & Ghasemi-Varnamkhasti, M. (2015). Detecting maturity of persimmon fruit based on image processing technique. *Scientia Horticulturae 184*, 123-128.

Munera, S., Besada, C., Blasco, J., Cubero, S., Salvador, A., Talens, P., et al. (2017). Astringency assessment of persimmon by hyperspectral imaging. *Postharvest Biology and Technology 125*, 35-41.

Naes, T., Isaksson, T., Fearn, T., & Davies, T. (2002). A user-friendly guide to multivariate calibration and classification. Chichester: NIR Publications.

Plaza, L., Colina, C., de Ancos, B., Sanchez-Moreno, C., & Cano, M. P. (2012). Influence of ripening and astringency on carotenoid content of high-pressure treated persimmon fruit (*Diospyros kaki* L.). *Food Chemistry 130*, 591-597.

Rodríguez-Pulido, F. J., Barbin, D. F., Sun, D.-W., Gordillo, B., González-Miret, M. L., & Heredia, F. J. (2013). Grape seed characterization by NIR hyperspectral imaging. *Postharvest Biology and Technology 76*, 74-82.

Salvador, A., Arnal, L., Besada, C., Larrea, V., Hernando, I., & Pérez-Munuera, I. (2008). Reduced effectiveness of the treatment for removing astringency in persimmon fruit when stored at 15 ºC: Physiological and microstructural study. *Postharvest Biology and Technology 49*, 340-347.

Salvador, A., Arnal, L., Besada, C., Larrea, V., Quiles, A., & Pérez-Munuera, I. (2007). Physiological and structural changes during ripening and deastringency treatment of persimmon cv. 'Rojo Brillante'. *Postharvest Biology and Technology 46*, 181-188.

Salvador, A., Arnal, L., Carot, J. M., Carvalho, C., & Jabaloyes, J. M. (2006). Influence of different factors on firmness and color evolution during the storage of persimmon cv. 'Rojo Brillante'. *Journal of Food Science 71*, 169-175.

Schmilovitch, Z., Ignat, T., Alchanatis, V., Gatker, J., Ostrovsky, V., & Felf€oldi, J. (2014). Hyperspectral imaging of intact bell peppers. *Biosystems Engineering 117*, 83-93.

Simon, R. (2007). Resampling strategies for model assessment and selection chapter 8, pp.173e186 in fundamentals of data mining in genomics and proteomics Dubitzky, Granzow and Berra reds. NY USA: Springer.

Taira, S. (1995). Astringency in persimmon. In H. F. Linskens, & J. F. Jackson (Eds.), Fruit analysis (pp. 97-110). Hannover, Germany: Springer.

Trinderup, C. H., Dahl, A., Jensen, K., Carstensen, J. M., & Conradsen, K. (2015). Comparison of a multispectral vision system and a colorimeter for the assessment of meat color. *Meat Science 102*, 1-7.

Vélez-Rivera, N., Gómez-Sanchis, J., Chanona-Pérez, J. J., Carrasco, J. J., Millan- Giraldo, M., Lorente, D., et al. (2014). Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. *Biosystems Engineering 122*, 91-98.

Viscarra-Rossel, A., McGlynn, R. N., & McBratney, A. B. (2006). Determining the composition of mineral-organic mixes using UV-VIS-NIR diffuse reflectance spectroscopy. *Geoderma 137*, 70-82.

Wang, W., Li, C., Tollner, E. W., Gitaitis, R. D., & Rains, G. C. (2012). Shortwave infrared hyperspectral imaging for detecting sour skin (Burkholderia cepacia)- infected onions. *Journal of Food Engineering 109*, 38-48.

Wei, X., Liu, F., Qiu, Z., Shao, Y., & He, Y. (2014). Ripeness classification of astringent persimmon using hyperspectral imaging technique. *Food and Bioprocess Technology 7*, 1371-1380.

Williams, P. C. (1987). Variables affecting near-infrared reflectance spectroscopic analysis. In P. Williams, & K. Norris (Eds.), Near-infrared technology in the agricultural and food industries (pp. 143-166). St. Paul, MN: American Association of Cereal Chemistry.

CHAPTER V

Discrimination of astringent and deastringed hard 'Rojo Brillante' persimmon fruit using a sensory threshold by means of hyperspectral imaging

Sandra Munera[a], Nuria Aleixos[b] Cristina Besada[c], Juan Gómez-Sanchís[d], Alejandra Salvador[c], Sergio Cubero[a], Pau Talens[e] and José Blasco[a]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[c] Departamento de Ingeniería Electrónica. Universitat de València. Av. Universitat, s/n, 46100 Burjassot, Valencia, Spain

[d] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[e] Departamento de Tecnología de Alimentos, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

Journal of Food Engineering (Under review)

Abstract

Persimmon fruit cv. 'Rojo Brillante' is an astringent cultivar due to its content of soluble tannins, which are insolubilised during the ripening of the fruit. Traditionally, the consumption of this cultivar has only been possible when the fruit is overripe and the texture is soft. Postharvest treatments based on exposing fruits to high CO_2 concentrations allow astringency removal while preserving high flesh firmness. However, the effectiveness of this treatment is controlled by means of slow destructive methods. The aim of this work is to study the application of hyperspectral imaging in the spectral range 450-1040 nm and to discriminate A and DA fruits non-destructively. The spectral information from three different areas of each fruit (calyx, middle and apex) was used to build models to predict the ST content using partial least squares regression. The results indicated that the model using the spectrum of the apex area was the most accurate. However, it was not possible to accurately predict fruits with very low levels of ST, especially in the case of DA fruits (42.2 %). Thus, classification models using partial least squares discriminant analysis were performed including other properties in order to discriminate between A and DA using an ST threshold. The most accurate models using all and the optimal wavelengths selected were those which focused on the middle and apex areas of the fruit, a correct classification rate of 87.0 % being achieved for A fruits and above 84.4 % for DA fruits. To date, there are only subjective and destructive analytical methods to monitor the effectiveness of the astringency removal treatments in persimmon. The results obtained in this study indicate that hyperspectral images can therefore be considered as an objective and non-destructive alternative in the control of this process.

1. Introduction

Spain is the number one producer of persimmon fruit (*Diospyros kakI* L.) in Europe and the third largest producer in the world, after China and South Korea (FAOSTAT, 2016). In the last twenty years, the land area devoted to cultivating this crop has risen from 2,000 to 14,000 ha, and production has increased from 33 to 310 thousand tons (FAOSTAT, 2016). Part of this growth is due to the increase in the production of the 'Rojo Brillante' cultivar. This cultivar, like other persimmon cultivars, is astringent at harvest and must be subjected to post-harvest treatments to remove astringency. The development of the de-astringency methods based on high CO_2 concentrations allowed removal of the astringency while preserving high flesh F (Arnal and Del Río, 2003), which has facilitated a rapid commercial expansion of this crop. Nowadays 'Rojo Brillante' persimmon is one of the most appreciated persimmon cultivars worldwide.

The conditions considered as standard for the complete elimination of astringency in this cultivar are 95 % CO_2 for 18-24 h at 20 ^{o}C. Under these conditions, the ST, responsible for astringency, are polymerised by the acetaldehyde produced to form insoluble compounds, which are non-astringent (Matsuo & Itoo, 1982; Taira et al., 1997; Salvador et al., 2008). However, the treatment may not be completely effective when the conditions of the process are not well controlled (Arnal & Del Río, 2003). In addition, the effectiveness of the treatment is also severely affected by the physiological state of the fruit. Small changes in the cellular structure can make the diffusion of CO_2 through the spaces difficult, the result being a low rate of anaerobic respiration and consequently less accumulation of acetaldehyde. This, in turn, leads to a lesser reduction of the ST (Salvador et al., 2007).

To commercialise this fruit, it is necessary to guarantee the complete removal of the astringency, since the presence of any astringency in the fruit can cause rejection by the consumer that will in turn affect future sales. The control of residual astringency in the fruits after the treatments can be performed destructively by measuring ST in the fruit pulp using the Folin–Denis method (Arnal and Del Río, 2004). However, in addition to being destructive, this method is slow and requires specialised equipment and personnel. An alternative is based on the reaction of the ST with $FeCl_3$. Tannic acid

complexes with ferric iron may consist of large highly coloured molecules that behave as colloids. Mixing them gives rise to a ferric complex that causes an intense black colour. The intensity of the black stain observed after impregnating a slice of the flesh with $FeCl_3$ reveals the presence of ST in the fruit and its intensity depends on their level (Gorini & Testoni, 1988; Munera et al., 2017b). Although this method is faster and easier than the analytical determination of ST, it is still destructive and subjective. Therefore, it is necessary to search for new rapid, reliable and non-destructive techniques. An alternative is based on the use of optical methods. Hyperspectral imaging is a promising optical technique for quality inspection of agricultural and food products that incorporates the main advantages of spectroscopy and imaging (Lorente et al., 2012). Thus, hyperspectral imaging can simultaneously acquire spectral and spatial information. In addition, the equipment used can be sensitive to different regions of the electromagnetic spectrum, such as the UV or IR (Gomez-Sanchís et al., 2014; Cortés et al., 2018). Their use has been widely studied to control the quality of fruit and vegetables during postharvest, for example to discriminate similar cultivars of nectarines with different properties (Munera et al., 2018), to discriminate gluten-free oats from cereals with gluten (Erkinbaev, Henderson and Paliwal, 2017), to detect decay lesions in citrus fruits (Folch-Fortuny et al., 2016) or mechanical damage in potatoes (López-Maestresalas et al., 2016). In recent years, several studies have been conducted to predict the content of ST or to assess the astringency in different varieties of persimmon fruit using spectroscopy (Zhang et al., 2013; Noypitak et al., 2014; Altieri et al., 2017; Cortés et al., 2017) and hyperspectral imaging (Munera et al., 2017). These works included the study of the best area of the fruit to measure the astringency, since the internal distribution may vary from the calyx area to the bottom. Most of the studies report successful prediction or classification models but they are not useful for precise prediction in fruits with low ST content, since they achieved limits of detection much higher than the minimum content of ST (0.10 %) that causes a sensation of astringency for most cultivars (Vidrih et al., 1994; Antoniolli et al., 2000; Antoniolli et al., 2003).

In the case of 'Rojo Brillante', ST values above 0.06 % can produce sensory astringency (Besada et al., 2010). Throughout the season, fruits of this cultivar exhibit

high astringency at harvest time with an ST content of between 0.80 % and 0.40 % (Salvador et al., 2007). Only when the fruit is over-ripe (which causes the total loss of firmness) does the loss of sensorial astringency occur. In that moment, the ST is around 0.04 % (Tessmer et al., 2017). In other studies in which the de-astringency treatment with high CO_2 concentration has been applied, an effective treatment has been associated with ST values of 0.01-0.03 % (Salvador et al., 2007; Salvador et al., 2008; Besada et al., 2008).

Hence, the main objective of this work was to study the application of hyperspectral imaging to predict the ST content in persimmon fruits and to discriminate A from DA persimmons using 0.04 % of ST as the threshold. Moreover, in order to establish a practical tool for use in industry, another goal is to determine which part of the fruit is the most appropriate to measure, as well as to reduce the amount of spectral information generated and speed up this process.

2. Material and methods

2.1 Fruit samples and experimental design

In this study, 300 persimmon fruits cv. 'Rojo Brillante' with similar size and no signs of external defects were analysed. In order to obtain fruit with different degrees of ripeness, 100 fruits were harvested every week over three consecutive weeks. The fruits were collected from an orchard in L'Alcúdia (Valencia, Spain) at commercial maturity. The maturity index used for harvesting was the external CI of the fruit. The CI commonly employed for 'Rojo Brillante' is CI = (1000a)/(Lb), where L, a and b are the colour coordinates in HunterLab colour space (Salvador et al., 2007). The average CI of the fruit at each harvest was 2.5, 3.9 and 9.3, respectively.

In each harvest, three homogeneous lots of fruit were submitted to different treatments to obtain fruit with different levels of ST, as follows: i) de-astringency treatment for 24 hours (40 fruits); ii) de-astringency treatment for 12 hours (30 fruits); and iii) no de-astringency treatment (30 fruits). In all cases, the de-astringency treatment was applied under standard conditions (95 % CO_2, at 20 °C, 90 % RH).

Hyperspectral images and the reference analyses were performed within 8 h after the treatment.

2.2 Hyperspectral imaging acquisition

The hyperspectral imaging system consisted of an industrial camera (CoolSNAP ES, Photometrics, AZ, USA) coupled to two LCTF (Varispec VIS-07 and NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). The camera was configured to acquire images with a size of 1392 × 1040 pixels and a spatial resolution of 0.14 mm/pixel. The working spectral range was defined between 450 nm and 1040 nm, capturing images every 10 nm. Thus, hypercubes with 60 images were captured. In order to avoid problems of unfocused images due to the refraction of light across this wide spectral range, the focus was adjusted on the central band of the acquisition interval (740 nm) and the images were captured using lenses capable of covering the whole spectral range without going out of focus (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA). To optimise the dynamic range of the camera, prevent the images from saturated regions and correct the spectral sensitivity of the different elements of the system, the maximum integration time of each band was calibrated by capturing the averaged grey level of a white reference standard (Spectralon 99%, Labsphere, Inc, NH, USA), corresponding to 90 % of the dynamic range of the camera.

The scene was illuminated using diffuse light from twelve halogen spotlights (37 W) (Eurostar IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current (12 V) and arranged equidistant from each other inside a hemispherical aluminium diffuser. The inner surface of the aluminium diffuser was painted white with a rough texture to maximise its reflectivity and minimise directional reflections, which could cause bright spots, the result being highly homogeneous light.

The fruits were introduced manually into a fruit holder in three different positions so as to obtain images from the top part of the fruit, the side, and the bottom. In this study, we have referred these areas as calyx, middle and apex areas respectively (Figure 1). Thus, three hyperspectral images were acquired for each fruit using ustomised software developed at IVIA, a total of 900 images being obtained.

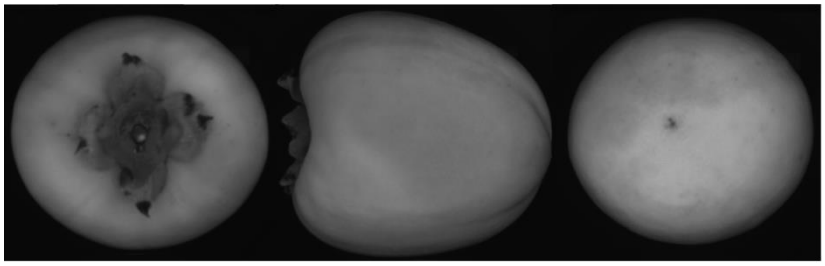


Figure 1. Hyperspectral images of the three areas of persimmon fruit acquired at 710 nm.

2.3 Reference analysis

The skin colour of each fruit was measured using a colorimeter (CR-300, Konica Minolta Inc., Tokyo, Japan). The mean value of the L, a and b colour coordinates (HunterLab colour space) was obtained as the average of three measurements in different parts of the fruit. The total colour difference (ΔE) between A and DA fruits was calculated by Equation (1):

$$\Delta E = \sqrt{(L_A - L_{DA})^2 + (a_A - a_{DA})^2 + (b_A - b_{DA})^2} \quad (1)$$

The F of the flesh was determined by means of a universal testing machine (4301, Instron Engineering Corp., MA, USA) equipped with an 8-mm puncture probe. The crosshead speed during testing was 1 mm s^{-1}. During the test, the force increased smoothly until it decreased drastically when the flesh was broken, and then the maximum peak force was registered. The results were expressed as the load (in N) required to break the flesh of the fruit on both sides after peel removal.

In order to assess the astringency of the fruits, each fruit was divided into two halves: one half was pressed against a 10 × 10 cm filter paper soaked in a 5 % $FeCl_3$ solution, which produced a dark print whose distribution and intensity gave information about the ST content in the pulp (Figure 2). The other half was used to obtain the ST content by the Folin-Denis method (Taira, 1995) based on the reduction of the Folin-Ciocalteu reagent by ST in alkaline solution (Arnal and Del Río, 2004). Taking into account the heterogeneous distribution of the tannins in the flesh (Figure

2), the samples for destructive analysis were taken from the lower part and near the apex, since the tannins take longer to be removed in this part.

The ANOVA, followed by Tukey's HSD test, was conducted using the software Statgraphics (Manugistics Corp., Rockville, USA) to find significant differences in the results of the physicochemical analysis related to the length of the de-astringency treatment. The groups of samples met the following three requirements: i) the observations being tested are independent within and among the groups; ii) the groups associated with each mean in the test are normally distributed; and iii) there is equal within-group variance across the groups associated to each mean in the test (homogeneity of variance).

Figure 2. Example of external and internal appearance of the fruit before and after de-astringency treatment. Visualisation of the distribution of ST using the alternative method of foils soaked in FeCl3. *A = astringent; DA = deastringed*

2.4 Image pre-processing

The reflectance captured by the camera is influenced by the intensity of the incoming light, the sensitivity of the sensor of the camera and the sensitivity of the LCFT, at the different wavelengths (Geladi, 2007). Thus, there is a need to correct these

effects to obtain the true reflectance of the sample. This is done using a reflectance standard (Spectralon 99%, Labsphere, Inc, NH, USA) through Equation (2) (Gat, 2000):

$$\rho_{xy}(x,y,\lambda) = \frac{\mathrm{R^{abs}}}{\mathrm{R^{abs}_{white}}} = \rho^{\mathrm{Ref}}(\lambda)\frac{\mathrm{R(x,y,\lambda)-R_{black}(x,y,\lambda)}}{\mathrm{R_{white}(x,y,\lambda)-R_{black}(x,y,\lambda)}} \quad (2)$$

Where ρ_{xy} is the reflectance of the fruit, $\rho^{Ref}(\lambda)$ is the standard reflectance of the white reference target (99 % in this work), $R(x,y,\lambda)$ is the radiance of the fruit captured by the CCD sensor of the camera, $R_{white}(x,y,\lambda)$ is the radiance captured by the CCD of the white reference target, and $R_{black}(x,y,\lambda)$ is the radiance captured by the CCD while avoiding any light source in order to quantify the electronic noise of the CCD.

The average reflectance spectrum of each area of each fruit was determined by averaging the relative reflectance spectra of all pixels included in the area using a binary mask which correctly removed the background and the leaves in the case of the area of the calyx.

The average reflectance spectrum of each area of each fruit was determined by averaging the relative reflectance spectra of all pixels included in the fruit area. This process was performed using a binary mask. For this, a thresholding between the background and the fruit was made at the wavelength of greater contrast between both regions (700 nm). In this way, it was possible to remove the background of the image from the fruit, easily. In the case of the calyx view, this allowed to remove also the leaves from the analyses. Being the contrast so high, the segmentation was quite accurate.

These operations were performed using HYPER-Tools (Mobaraki and Amigo, 2018) working under MATLAB R2017b (The MathWorks, Inc., MA, USA).

2.5 Multivariate data analysis

After the analysis of the ST content and knowing which fruit was A and DA, the spectra were randomly partitioned into two sets. For each area of the persimmon fruit, 201 fruit spectra (107 A and 94 DA) were used to calibrate the models and 99 fruit spectra (54 A and 45 DA) were used for independent validation or test set.

PLS-R was used to quantify the ST content and PLS-DA was used to classify the fruits as A and DA according to the threshold value of 0.04 % (Tesmeer et al., 2016). A model using the spectral information of each area (calyx, middle and apex) was performed.

Previously, the mean spectrum of each area of the persimmon fruit was filtered using the Savitzky-Golay smoothing filter (3-point smoothing window, second-order polynomial) to remove both additive and multiplicative effects, and pre-treated using standard normal variate to remove the scatter (Rinnan et al., 2009). Later, each resulting spectrum was normalised by mean centre. A 10-fold CV was used to obtain the optimal number of LVs as well as an estimation of the error rate of the models. The PLS-R models were evaluated by the R^2 and the RMSE between the predicted and the measured value of the reference parameter for calibration, CV and prediction. Furthermore, the RPD, defined as the ratio between the standard deviation of the reference data and RMSEP, was used (Williams, 1987). The results of the PLS-DA models were expressed as the percentage of correct classification (percentage of A or DA fruits correctly classified) and total accuracy (percentage of all fruits correctly classified) for calibration, CV and prediction.

In order to reduce the dimensionality of the hyperspectral images, the vector of regression coefficients was used. This method measures the association between each wavelength and the response (i.e. A and DA class) obtained by the PLS-DA model (Mehmood et al., 2012). The wavelengths with a high absolute value are selected, since they make the highest contribution to the classification, and those with a smaller absolute value are ignored.

The spectral pre-processing was carried out using HYPER-Tools (Mobaraki and Amigo, 2018) and the PLS models were performed using MATLAB R2017b (The MathWorks, Inc., MA, USA).

3. Results and discussion

3.1 Reference analysis

In general, the ST content in the fruits ranged from 1.18 % (non-treated fruits) to 0.01 % (fruit treated for 24 hours), while those fruits that were non-treated presented ST values from 0.37 % to 1.18 %, depending on the time of harvesting (Table 1). Thus, the mean value of the fruits collected in different moments was statistically different. The CO2 treatment for 12 hours resulted in fruits with a wide range of ST values between 0.66 % and 0.01 %. This meant that part of the fruits could already be consumed while others still needed more hours of treatment. In this case, the mean values of the three harvests were also statistically different. When the treatment was applied for 24 hours, all fruits reached an eatable stage and no statistical differences were found among the three times of harvesting. Using the threshold of 0.04 % for the ST value, a total of 161 fruits were considered as A and 139 as DA (Table 1).

Table 1. ST content and quantification of A and DA fruits.

Harvest	Treatment duration	Soluble tannins (%)			#A	#DA
		Min	Mean	Max		
1	0h	0.37	0.69^{b}	0.98	48	52
	12h	0.02	0.09^{e}	0.33		
	24h	0.01	0.02^{f}	0.03		
2	0h	0.45	0.61^{c}	0.77	53	47
	12h	0.01	0.11^{e}	0.31		
	24h	0.02	0.03^{f}	0.04		
3	0h	0.66	0.91^{a}	1.18	60	40
	12h	0.10	0.37^{d}	0.66		
	24h	0.02	0.03^{f}	0.04		
Total		0.01	0.32	1.18	161	139

Different letters indicate significant differences between groups (p-value<0.05), according to Tukey's (HSD) test. Min = minimum; Max = maximum; #A = number of astringent fruits; #DA = number of de-astringent fruits

The application of a de-astringency treatment with CO_2 does not usually have any effect on the colour in the early stages of ripeness. Only slight differences could be observed between A and DA fruit coordinates (Table 2). Although significant differences were found between L and b, they are barely perceptible to the human eye. According to the International Commission on Illumination (CIE), the value of ΔE obtained (1.9) indicates that, in general, the colour difference between the two classes of fruits is minimally perceptible (Mokrzycki and Tatol, 2011).

Table 2. Skin colour of A and DA fruits.

	L (*)			a			b (*)			ΔE	CI		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max		Min	Mean	Max
A	56.0	62.8	67.9	1.9	11.7	25.6	32.6	37.1	41.0	1.9	-0.03	5.1	13.5
DA	56.8	62.2	67.5	-2.25	13.4	30.5	32.3	36.7	41.7		-1.08	5.9	14.7

(*) indicate significant differences between groups (p-value<0.05). Min = minimum; Max = maximum; A = astringent; DA = deastringed; CI = colour index

As in the case of the colour, CO2 treatment does not usually affect the F of the fruit in the early stages of ripeness. However, it does give rise to a significant degree of softening in the following stages. These changes in firmness are related to the changes that take place in the cell structure (Salvador et al., 2007). Here, the mean value of the F was reduced from 47.3 N in A fruits to 43.7 N in DA fruits (Table 3).

Table 3. Flesh F of A and DA fruits.

	Flesh firmness (N) (*)		
	Min	Mean	Max
A	37.4	47.3	58.9
DA	29.9	43.7	54.5

(*) indicate significant differences between groups (p-value<0.05). Min = minimum; Max = maximum; A = astringent; DA = deastringed;

3.2 Spectral analysis

The average spectra obtained for each measured area of A and DA fruits are illustrated in Figure 3. The spectra of all fruits followed a similar pattern in each area. Slight differences were present in the VIS region around 460 nm, 550-600 nm and 650-710 nm, where carotenoids, anthocyanins, chlorophylls and other pigments are responsible for fruit colour (Rajkumar et al., 2012). In the NIR region, some differences were found, especially in the apex area, around 750 nm, where a water absorption peak (OH second overtone) is observed (Siedliska et al., 2018; Williams and Norris, 1987). Noypitak et al. (2015) indicated that phenolic compounds are located between 940-1000 nm and the absorption peak of tannic acid is seen at 996 nm. In this case, slight differences were found close to these wavelengths between the A and DA spectra. However, it is not clear whether this corresponded to the ST content because a water absorption peak (third overtone of OH stretching vibration) dominates this part of the spectrum (Nicolaï et al., 2007).

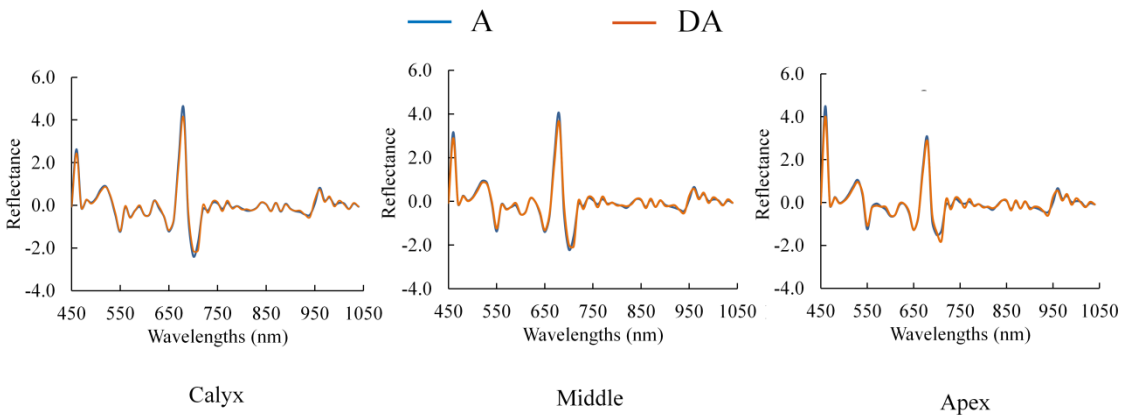


Figure 3. Mean pre-treated spectra of each area of astringent (A) and deastringed (DA) fruits.

3.3 Prediction of soluble tannins content

PLS-R models were performed to quantify the content of ST in each fruit using the spectral range of 450-1050 nm. Table 4 shows the results of the prediction of ST content obtained for the three areas of the fruit that were measured.

Table 4. Results of calibration and validation of the models to predict the ST content using hyperspectral imaging and the different areas of the fruit.

Area	#LV	Calibration		Cross validation		Prediction		
		R^2	RMSE	R^2	RMSE	R^2	RMSE	RPD
Calyx	15	0.71	0.18	0.54	0.23	0.49	0.25	1.4
Middle	12	0.71	0.17	0.60	0.21	0.69	0.19	1.8
Apex	13	0.76	0.16	0.64	0.20	0.73	0.18	1.9

#LV = number of latent variables

The optimal model was chosen when the number of LV yields the lowest RMSE for calibration and CV. Therefore, 15 LVs were determined for the calibration of the model of the calyx area, 12 for the middle area model and 13 for the apex area model. The model using the spectra obtained from the calyx area achieved the lowest R^2 (0.49) while the highest RMSE (0.25 %) was obtained in the test set. In contrast, the model built for the middle area offered better results, with an R^2 of 0.69 and an RMSE of 0.19 %, while the model obtained for the apex area achieved the highest R2 and the lowest RMSE of 0.73 and 0.18 %, respectively. The RPD values indicate that only the models that used measurements obtained in the middle and apex areas could discriminate between low and high ST values (RPD values between 1.5 and 2) (Nicolai et al., 2007). However, in all cases, the RMSE value was higher than the 0.04 % threshold, which means that the models were not altogether useful for accurate prediction in fruits with extremely low ST content.

The scientific literature contains other studies that achieve findings similar to ours but using mostly spectroscopy instead of hyperspectral imaging. For example, Noypitak et al. (2015) used interactance mode in the evaluation of ST using different areas of persimmons cv. 'Xichu', achieving, as best result, an R^2 of 0.93 and a high RMSE of 0.22 % but using the calyx area. However, the higher R^2 was probably achieved because most of the persimmons used had very low (0.02 %) or very high (1.6 %) tannin contents and only a few samples had intermediate values. In the case of 'Rojo Brillante', Cortés et al. (2017) developed models using spectra pre-treated with different techniques, achieving better results in terms of R^2 (0.91) and RMSE, above 0.08 %, using six measurement points distributed throughout the fruit.. In this case,

the ST content ranged from 0.023 to 0.75 but DA fruit were 20 %, while in our case they represent 46 % of the fruit in the models. Moreover, most of the error is introduced by fruits with very low ST values. Alitieri et al. (2017) also achieved a good prediction result with an R2 higher than 0.98 but in the cross validation set and using fruits with ST content values between 0.1 % and 1.7 %, which should be considered as astringent in all cases from a commercial point of view.

Figure 4 shows the prediction performance of the model using the test set and the data captured in the apex area. Taking into account the threshold of 0.04 %, only 77.8 % of A fruits and 42.2 % of DA fruits were correctly predicted. These results are clearly low and below those expected. Thus, the direct prediction of very low values of ST content (such as 0.04 %) does not seem to be possible with the procedure followed. This is probably because the concentration of ST is correlated with other major biochemical constituents such as pigments, water or other soluble solids like sugars that can mask the detection of constituents when the content is very low (Nicolaï et al., 2007). For this reason, a different approach to measuring astringency other than the direct estimation of ST was required. PLS-DA models were then developed to maximise the separation between classes A and DA, not only with respect to the differences in ST content, but also to capture the information contained in the spectra related to other properties that can contribute to make each class different.

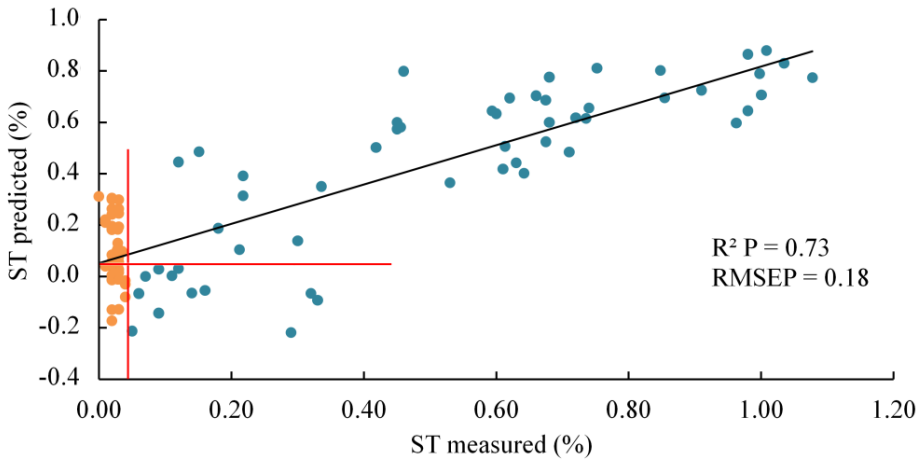


Figure 4. Prediction of the soluble tannins content in test set fruit using the apex area. The red lines indicate the threshold value of 0.04 %

3.4 Detection of astringent and de-astringent fruits

As in the prediction of ST content, the calyx, middle and apex areas were tested to distinguish A and DA fruits using PLS-DA models. The results of the classification models using hyperspectral imaging are presented in Table 5a.

Table 5a. Results of the classifications using the calyx, middle and apex areas and all wavelengths.

Area	#LV	Class	Calibration				Cross Validation				Prediction			
			#A	#DA	CC (%)	Acc (%)	#A	#DA	CC (%)	Acc (%)	#A	#DA	CC (%)	Acc (%)
Calyx	18	A	95	12	88.8	89.1	84	23	78.5	78.1	45	9	83.3	80.8
		DA	10	84	89.4		21	73	77.7		10	35	77.8	
Middle	13	A	101	7	94.4	91.5	96	11	89.7	86.6	47	7	87.0	88.9
		DA	11	83	88.3		16	78	83.0		4	41	91.1	
Apex	18	A	100	6	93.5	91.0	88	19	82.2	82.6	47	7	87.0	87.9
		DA	11	83	88.3		16	78	83.0		5	40	88.9	

#LV = number of latent variables; #A = number of astringent fruits; #DA = number of de-astringent fruits; CC = correct classification; Acc = accuracy

The calibration of the calyx and apex area models was performed using 18 LVs, while for the middle area model only 13 LVs were necessary. Furthermore, the internal CV of the middle area model presented the highest precision (86.6 %), then the apex area (82.6 %) and the calyx area model presented the lowest results, 78.1 %. This fact agrees with the previous results of the quantification of ST content, where the calyx area was the least precise part for this purpose.

The middle and calyx area models correctly classified more A fruits, 89.7 % and 78.5 % than DA fruits, 83.0 % and 77.7 %, respectively. In the case of the apex area, more DA fruits were correctly classified: 83.0 % versus 82.2 %.

The mean spectrum of each fruit of the test set was classified using the previously calibrated models. As in the calibration and CV, the model using the calyx area presented less precision, correctly classifying 83.3 % of A fruits and 77.8 % of DA fruits, and showing a total accuracy of 80.8 %. In the case of the middle and apex areas, their

prediction showed similar results with 87.0 % of A fruits and 91.1 % and 88.9 % DA fruits being classified correctly. Therefore, the total accuracy of the middle and apex area models, 88.9 % and 87.9 %, was higher than that of the calyx area. This fact agrees with the results obtained in the quantification of ST, where the calyx area was the least accurate area for this purpose (Table 4).

Previous studies have been conducted to classify the fruits according to their astringency using spectral information. It is noteworthy that the best results in terms of ST prediction have been reported when values of ST content are high (Zhang et al., 2013; Altieri et al., 2017; Cortés et al., 2017, Munera et al., 2017). In this line, Noypitak et al. (2015) reported on a model in which a classification accuracy of 97.1 % was achieved, assuming that the persimmon with a ST content lower than 0.8 % is non-astringent. However, as mentioned in the introduction section, the threshold of ST to detect astringency is not established and is highly dependent not only on the cultivar but also on the consumer's country of origin (Antoniolli et al., 2000; Antoniolli et al., 2002; Yamada et al., 2002; Tessmer et al 2016). In 'Rojo Brillante' persimmon it has been widely reported that sensorial astringency loss occurs when tannin content is lower than 0.04 % (Salvador et al 2007; Tessmer et al 2016). This means that the predictive models previously reported would not be valid for this cultivar. In the present study the threshold applied was 0.04% to guarantee the complete non-astringency of the fruits. Although the result of the ST predictive model might seem a priori unsatisfactory (42.2 % of DA fruits correctly classified), this is the first work in which such a low ST threshold has been established to guarantee the non-astringency of the fruits. The results reveal that the higher the established ST threshold is, the better the results provided by the predictive models are. This fact leads us to think that other attributes, besides the ST content, may influence the spectral response of persimmon.

Salvador et al. (2007) evaluated the physiological and structural changes that occur after the deastringency treatment with high CO_2 concentrations in persimmon 'Rojo Brillante' at different maturity stages. Some of the reported changes may affect the spectral information. In this way, a decline in the TSS, measured as ºBrix, occurs after deastringency treatment concomitant to the drop in ST as a response to the

deastringency process. On the other hand, after the CO_2 treatment a significant increase in pH is observed. This rise in the pH value is also related to the process by which soluble tannins, the acid components, become insoluble during the application of the treatment. It is noteworthy that the measurements of soluble solids in persimmon are related to ST, but also to sugars and acids, are located between 720 nm and 820 nm, phenolic compounds are between 940 nm and 1000 nm, and the tannic acid peaks at 996 nm.

In addition, it must be taken into account that the cellular microstructure can have an important effect on the spectral response. Hence, it has been reported that the deastringency process causes important changes in the cell microstructure. The insolubilisation of tannins occurs inside the vacuoles of tannin cells, which appeared to be filled with an insoluble material (like a compact mass) (Salvador et al., 2007). Thus, depending on the level of insolubilisation during the deastringency treatment, the number of cells in the parenchyma containing insoluble material will be greater or lower. Moreover, the CO_2 applied, in addition to triggering the insolubilisation of tannins, also brings about a progressive degradation of the parenchyma structure, affecting the cell walls and integrity of the cell membranes. The adhesion bonds between some cells are lost in certain areas and the intercellular spaces are filled with a soluble material. This effect becomes greater as the treatment time increases (Salvador et al 2007; Novillo et al., 2014). It should be noted that the declining firmness that occurs during the maturity process of persimmon fruit has been associated with a gradual loss of parenchyma structure due to degradation of the cell wall and membrane (Salvador et al., 2007; Tessmer et al., 2016). In the same way, the effect of high concentrations of CO_2 on the cellular structure is related to a loss of firmness.

These structural changes associated with both the maturity stage and the CO_2 treatment may have an important effect on the spectra, since firmness is related to the water content in the cells (water absorption peaks at 750 nm and 970 nm) and the structural status of the parenchyma. This may have an influence on the way the light interacts with the cells and is transmitted through the fruit and hence the spectral response received by the spectrometer, which allowed A fruits to be separated from DA fruits.

Regarding the colour, the treatment with CO_2 did not cause any great changes in fruit skin for the earlier stages of fruit maturity, although small differences were observed in the last stage due to changes in carotenoids, anthocyanins and chlorophylls related to wavelengths 450–720 nm, 460 nm, 550–600 nm and 650–710 nm (Rajkumar et al., 2012). However, since the colour has not previously been evaluated to detect the astringency of the 'Rojo Brillante' persimmon, a PLS-DA model was calibrated using the HunterLab colour coordinates L, a, b. As a result, 66.7 % of A and 33.6 % of DA fruits were correctly classified, showing a total precision of 52.5 % (Table 5b). This result indicates that traditional colour measures are not useful for the discrimination of A and DA fruits. However, from the results obtained using hyperspectral images, it is possible to present an alternative to those methods that are destructive, need chemical analysis, are subjective and only allow the inspection of a few samples per batch.

Table 5b. Results of the classification of A and DA fruits using the colour information.

#LV	Class	Calibration				Cross Validation				Prediction			
		#A	#DA	CC (%)	Acc (%)	#A	#DA	CC (%)	Acc (%)	#A	#DA	CC (%)	Acc (%)
2	A	78	29	72.9	57.2	77	30	72.0	58.7	36	18	66.7	52.5
	DA	57	37	39.4		53	41	43.6		29	16	35.6	

#LV = number of latent variables; #A = number of astringent fruits; #DA = number of deastringed fruits; CC = correct classification; Acc = accuracy

3.4.1 Selection of optimal wavelengths

In order to reduce the complexity of the system, the number of wavelengths used should be reduced because a large number of wavelengths increase the acquisition time while it reduces the performance of classifiers (Friedman, 1994). Numerous techniques have been employed to deal with this issue, such as restricting the information to just a few bands which reveal the most variability and therefore the most significant information in the hyperspectral image (Du and Sun, 2006). In this study, the vector of the regression coefficients was used. A total of 23 optimal

wavelengths were selected in the vector of each area, all of them being located across the VIS and NIR region (Figure 5).

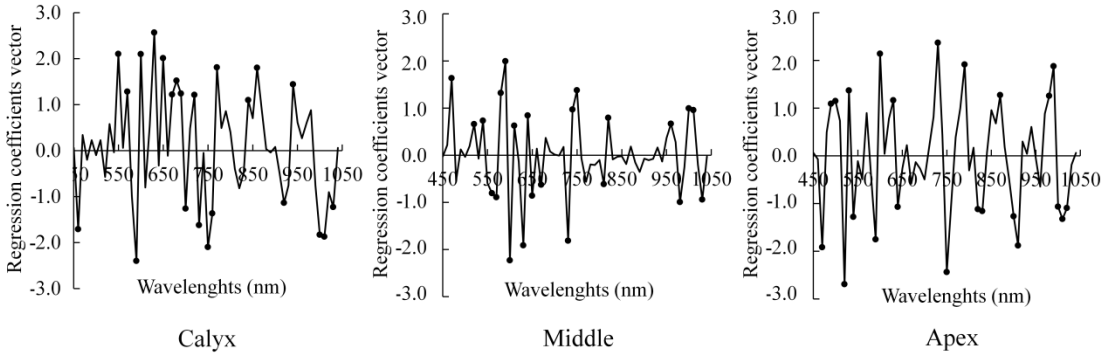


Figure 5. Regression coefficients vector of the PLS-DA model of each area with the optimal wavelengths selected.

The high number of wavelengths selected indicated that there are no specific ones that can be specifically linked to the tannins or other particular constituents related to astringency. Hyperspectral images show a high degree of collinearity and redundant information and this selection is probably a reduction of this information. More than half of the selected wavelengths for the three areas are located in the VIS region, which is related to the carotenoids, anthocyanins, chlorophylls and other pigments responsible for fruit colour, as previously commented. Several wavelengths were selected near the water absorption peaks, around 750 nm (first overtone of OH) and 970 nm (third overtone of OH) (Siedliska et al., 2018; Nicolaï et al., 2007; Williams and Norris, 1987). Other selected wavelengths are located around 850 nm, which is assigned to the absorption of acids and sugars (Yang et al., 2015). As commented earlier, phenolic compounds are located between 940–1000 nm and the absorption peak of tannic acid is seen at 996 nm (Noypitak et al., 2015). Several selected wavelengths are located in this region but it is not clear whether this corresponded to the ST content or to the water absorption peak.

The optimised classification models were built using the selected wavelengths as input. Results of the calibration are presented in Table 6. The models for the calyx and apex areas were performed using 15 and 13 LVs, while only eight LVs were necessary

to build the model for the middle area. In this case, the increased accuracy in CV made the results more similar in the calibration of the models. As in the case of the models built using the full spectra, the internal CV of the middle area presented the highest accuracy (88.1%), then the apex area with 83.1 % and the calyx area model presented the lowest results with only 78.6 % of total accuracy. In all cases, A fruit were detected better than DA fruits, which is in line with the principal aim of detecting astringent fruits among those that have been submitted to a CO_2 treatment.

Table 6. Results of the classification using calyx, middle, and apex area and optimal wavelengths selected.

Area	#V	#LV	Class	Calibration				Cross Validation				Prediction			
				#A	#DA	CC (%)	Acc (%)	A	DA	CC (%)	Acc (%)	#A	#DA	CC (%)	Acc (%)
Calyx	23	15	A	90	17	84.1	82.6	89	18	83.1	78.6	44	10	81.5	81.8
			DA	18	76	80.9		25	69	73.4		8	37	82.2	
Middle	23	8	A	97	10	90.7	90.0	96	11	89.7	88.1	47	7	87.0	86.9
			DA	10	84	89.4		13	81	86.2		6	39	86.7	
Apex	23	13	A	94	13	87.9	86.6	91	16	85.0	83.1	47	7	87.0	85.9
			DA	14	80	85.1		18	76	80.9		7	38	84.4	

#LV = number of latent variables; #A = number of astringent fruits; #DA = number of de-astringent fruits; CC = correct classification; Acc = accuracy

As in the classification performed using all wavelengths, the class of each fruit in the test set was predicted by introducing the mean spectrum of the fruit into the previously optimised models.

In the case of the middle and apex areas, their prediction of A fruit showed similar results between areas and using all and the optimal wavelengths, resulting in a correct classification of 87.0 % of A fruit. However, precision was lower for both areas in DA fruit, with respect to the previous models, i.e. 86.7 % and 84.4 % of DA fruit. Therefore, the total accuracy of the middle and apex area models was 86.9 % and 85.9 %. Despite the reduction in precision in the classification of DA fruit using fewer wavelengths, this is more desirable than the contrary. If a DA fruit is classified as A, it can be treated again with CO_2, but if an A fruit is classified as DA, this fruit goes directly to the

consumer. In the case of the calyx area model, 81.5 % of A and 82.2 % of DA fruit were classified correctly. The total accuracy was a little higher than when using all the wavelengths (81.8 %), although it was again the least accurate of the three areas.

4. Conclusions

The capability of VIS-NIR hyperspectral imaging to discriminate A and DA hard 'Rojo Brillante' persimmon fruits was investigated. Furthermore, as ST are heterogeneously distributed in the flesh of persimmon fruit, an individual study of three different areas of the fruit was carried out in order to find the most suitable to maximise the accuracy of the models.

The prediction of ST content in the fruits was performed using PLS-R models. The results obtained indicated that the model using the spectra of the apex area was the most accurate, R^2 of 0.71 with an RMSE of 0.18 and RPD 1.9. However, only 77.8% of A fruit and 42.2% of DA fruit were correctly classified when the threshold of 0.04 % was applied. Therefore, PLS-DA models were performed in order to maximise the separation between A and DA classes, which led to an improvement in the results. The most accurate models were those performed using middle and apex area spectra (88.9 % and 87.9 %), with a correct classification of 87.0 % of A fruit and 91.1 % and 88.9 % of DA fruit, respectively. When the discrimination of the fruit was performed using colour information, the accuracy in the classification was only 66.7 % for A and 33.6 % for DA fruit.

To reduce the huge amount of data captured by the hyperspectral systems, the vector of the regression coefficients of the PLS-DA model of each area was used to identify the optimal wavelengths. As when using all wavelengths, the most accurate models were those involving the middle and apex areas and 23 optimal wavelengths (86.9 % and 85.9 %), also with a correct classification of 87.0 % of A fruit and 86.7 % and 84.4 % of DA fruit, respectively.

According to these results, hyperspectral imaging combined with multivariate analysis has a great potential as a tool for rapid and non-destructive control of effectiveness of the astringency removal treatment applied to persimmon 'Rojo

Brillante'. Nevertheless, the results of this study need further experimentation on a larger set of fruits grown in different areas and harvested at different stages of ripeness before this could be effectively implemented in an in-line system.

Acknowledgements

This work was partially funded by INIA and FEDER funds through project RTA2015-00078-00-00. Sandra Munera thanks INIA for the FPI-INIA grant num. 43 (CPR2014-0082), partially supported by European Union FSE funds.

References

Altieri, G., Genovese, F., Tauriello, A. & Di Renzo, G.C. (2017). Models to improve the non-destructive analysis of persimmon fruit properties by VIS/NIR spectrometry. *Journal of the Science of Food and Agriculture 97*, 5302-5310.

Antoniolli, L.R., Castro, P.R.C., Kluge, R.A. & Scarpare Filho, J.A. (2000). Remoção da adstringência de frutos de caqui 'Giombo' sob diferentes períodos de exposição ao vapor de álcool etílico. *Pesquisa Agropecuária Brasileira*, 35, 2083-2091.

Antoniolli, L.R., Castro, P.R.C., Kluge, R.A. & Scarpare Filho, J.A. (2002). A remoção da adstringência de frutos de caquizeiro 'Giombo' sob diferentes temperaturas. *Pesquisa Agropecuária Brasileira*, 37, 687-691.

Arnal, L. & Del Río, M.A. (2004). Effect of cold storage and removal astringency on quality of persimmon fruit (*Diospyros kaki*, L.) cv. Rojo brillante. *Food Science and Technology International 10*, 179-185

Arnal, L. & Del Río, M.A. (2003). Removing astringency by carbon dioxide and nitrogen-enriched atmospheres in persimmon fruit cv. 'Rojo brillante'. *Journal of Food Science 68*, 1516-1518.

Besada, C., Salvador, A., Arnal, L. & Martínez-Jávega, J.M. (2010). Optimization of the duration of deastringency treatment depending on persimmon maturity. *Acta Horticulturae 858*, 69-74.

Besada, C., Arnal, L. & Salvador, A. (2008). Improving storability of persimmon cv. Rojo Brillante by combined use of preharvest and postharvest treatments. Postharvest Biology and Technology 50, 169-175

Cortés, V., Rodríguez, A. Blasco, J., Rey, B., Besada, C., Cubero, S., Salvador, A., Talens, P. & Aleixos, N. (2017). Prediction of the level of astringency in persimmon using visible and near-infrared spectroscopy. *Journal of Food Engineering 204*, 27-37.

Du, C.J. & Sun, D.W. (2006). Learning techniques used in computer vision for food quality evaluation: a review. *Journal of Food Engineering, 72*, 39–55.

Erkinbaev, C., Henderson, K., & Paliwal, J. (2017). Discrimination of gluten-free oats from contaminants using near infrared hyperspectral imaging technique. *Food Control 80*, 197-203.

FAOSTAT. http://www.fao.org/faostat/en/#data/QC - Accessed 02.11.18

Folch-Fortuny, A., Prats-Montalbán, J.M., Cubero, S., Blasco, J. & Ferrer, A. (2016). VIS/NIR hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus fruits. *Chemometrics and Intelligent Laboratory Systems 156*, 241-248.

Friedman J.H. (1994) An Overview of Predictive Learning and Function Approximation. In: From Statistics to Neural Networks. Cherkassky V., Friedman J.H., Wechsler H. (eds) NATO ASI Series, 136. Springer, Berlin, Heidelberg.

Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report, Opto- Knowledge Systems Inc. OKSI.

Geladi, P. L. M. (2007). Calibration standards and image calibration. In H. F. Grahn & P. Geladi (Eds.), Techniques and applications of hyperspectral image analysis (pp. 203–220). Chichester: Wiley.

Gómez-Sanchis J, Lorente D, Soria-Olivas E, Aleixos N, Cubero S, Blasco J (2014) Development of a hyperspectral computer vision system based on two liquid crystal tuneable filters for fruit inspection. Application to detect citrus fruits decay. *Food and Bioprocess Technology, 7*, 1047-1056.

Gorini, F. L. & Testoni, A. (1988). Maturazione, raccolta, conservazione e trasformazione dei frutti di kaki. *Agricoltura e Ricerca 95*, 81–88.

López-Maestresalas, A., Keresztes, J.C., Goodarzi, M., Arazuri, S. & Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control 70*, 229-241.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology 5*, 1121–1142.

Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering 93*, 161-171.

Matsuo, T. & Itoo, S. (1982). A model experiment for de-astringency of persimmon fruit with high carbon dioxide treatment: In vitro gelation of kaki-tannin by reacting with acetaldehyde. *Agricultural and Biological Chemistry 46*, 683-689

Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems 118*, 62–69.

Mokrzycki, W. & Tatol, M. (2011). Color difference Delta E - A survey. *Machine Graphics and Vision 20*, 383-411.

Munera, S., Amigo, J.M., Aleixos, N., Talens, P., Cubero, S. & Blasco, J. (2018). Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine. *Food Control 86*, 1-10.

Munera, S., Besada, C., Aleixos, N., Talens, P., Salvador, A., Da-Wen, S., Cubero, S. & Blasco, J (2017). Non-destructive assessment of the internal quality of intact persimmon using colour and VIS/NIR hyperspectral imaging. *LWT - Food Science and Technology 77*, 241-248.

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys,W., Theron, K. I., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology 46*, 99-118.

Novillo, P., Salvador, A., LLorca, E., Hernando, I. & Besada, C. (2014). Effect of CO2 deastringency treatment on flesh disorders induced by mechanical damage in persimmon. Biochemical and microstuctural studies. *Food Chemistry 145*, 454-463.

Noypitak, S., Terdwongworakul, A., Krisanapook, K., Kasemsumran, S., (2015). Evaluation of astringency and tannin content in 'Xichu' persimmons using near infrared spectroscopy. *International Journal of Food Properties 18*, 1014–1028.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G.S.V. & Gariepy, Y., (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194–200.

Rinnan, Å., van den Berg, F. & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry 28*, 1201-1222.

Salvador, A., Arnal, L., Besada, C., Larrea, V., Quiles, A. & Pérez-Munuera, I., 2007. Physiological and structural changes during ripening and deastringency treatment of persimmon cv. 'Rojo Brillante'. *Postharvest Biology and Technology 46*, 181–188.

Salvador, A., Arnal, L., Besada, C., Larrea, V., Quiles, A. & Pérez-Munuera, I. (2008). Reduced effectiveness of the treatment for removing astringency in persimmon fruit when stored at 15 ºC. Physiological microstructural study. *Postharvest Biology and Technology 49*, 340–347.

Siedliska, A., Baranowski, P., Zubik, M., Mazurek, W. & Sosnowska, B. (2018). Detection of fungal infections in strawberry fruit by VNIR/SWIR hyperspectral imaging. *Postharvest Biology and Technology 139*, 115-126.

Taira, S., Ono, M. & Matsumoto, N. (1997). Reduction of persimmon astringency by complex formation between pectin and tannins. *Postharvest Biology and Technology 12*, 265-271.

Taira, S. (1995). Astringency in persimmon. In: Linskens, H.F., Jackson, J.F. (Eds.), Fruit analysis (pp 97–110). Hannover, Germany: Springer.

Tessmer, M.A., Besada, C., Hernando, I., Appezzato-da-Glória, B., Quiles, A. & Salvador, A. (2016). Microstructural changes while persimmon fruits mature and ripen. Comparison between astringent and non-astringent cultivars. *Postharvest Biology and Technology 120*, 52-60.

Vidal, M., & Amigo, J.M. (2018). HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems 172*, 174-187.

Vinzi, V., Chin, W. W., Henseler, J., & Wang, H. (2010). Handbook of partial least squares. Wang, H. (Ed.). Berlin: Springer.

Williams, P., & Norris, K. (1987). Near-infrared Technology in the Agricultural and Food Industries. St. Paul, USA: American Association of Cereal Chemists.

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

Yamada, M., Taira, S., Ohtsuki, M., Sato, A., Iwanami, H., Yakushiji, H., Wang, R., Yang, Y. & Li, G. (2002). Varietal differences in the ease of astringency removal by carbon dioxide gas and ethanol vapor treatments among Oriental astringent persimmons of Japanese and Chinese origin. *Scientia Horticulturae 94*, 63-72.

Yang, C. H., Sun, D. W., Pu, H., Wang, N. N. & Zhu, Z. (2015). Rapid detection of anthocyanin content in lychee pericarp during storage using hyperspectral imaging coupled with model fusion. *Postharvest Biology and Technology 103*, 55-65.

Zhang, P., Xue, Y., Li, J., Feng, X. & Wang, B. (2013). Research on non-destructive measurement of firmness and soluble tannin content of 'mopanshi' persimmon using Vis/NIR diffuse reflection spectroscopy. *Acta Horticulturae 996*, 447-452.

III. POMEGRANATE

CHAPTER VI

Machine vision and chemometrics for quality monitoring of intact 'Mollar de Elche' pomegranate fruit and arils during maturity

Sandra Munera[a], Francisca Hernández[b], Nuria Aleixos[c], Sergio Cubero[a] and José Blasco[a]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Grupo de Fruticultura y Técnicas de Producción. Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández de Elche, Carretera de Beniel, Km 3.2, 03312 Orihuela, Spain

[c] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

Postharvest Biology and Tecnology (Under review)

Abstract

Pomegranate fruit cv. 'Mollar de Elche' were collected at seven different harvest times. Colour and hyperspectral images of the intact fruit and arils were acquired at each harvest. Physicochemical properties such as total soluble solids, titratable acidity, maturity index, BrimA, internal colour, total phenolic compounds content and antioxidant activity were measured in the juice of each fruit. In order to relate the colour (L*, a*, b*) and spectral (720-1050 nm) data obtained from the images of the intact fruit and arils to their physicochemical properties, partial least square regression models were calibrated. The discrimination of the different maturity stages was carried out using partial least square discriminant analysis models. Similar results were obtained in the prediction of the physicochemical properties using the colour and hyperspectral images of the intact fruit. However, the predictions achieved for the information about the arils were higher using hyperspectral imaging. In the discrimination of maturity stage, the highest accuracies were obtained using hyperspectral imaging, where 95 % of intact fruits and 100 % of arils where correctly classified. These results indicate the great potential of machine vision techniques, especially hyperspectral imaging, for monitoring the quality of intact 'Mollar de Elche' pomegranate fruit and arils.

1. Introduction

In recent times, pomegranate (Punica granatum L.) fruit has gained great importance because it is a source of sugars, organic acids and bioactive compounds (Opara et al., 2009), thus reporting positive health benefits (Viuda-Martos et al., 2010).

In Europe, Spain is the largest producer of this fruit. About 60 % of the total production (65,165 t) (MAPA, 2018) is exported, 'Wonderful', 'Mollar Valenciana' and 'Mollar de Elche' being the most important cultivars (Melgarejo et al., 2010). This last cultivar has a sweet taste and soft seeds that make it very much appreciated by consumers and it has its own Protected Designation of Origin.

The commercial quality of this fruit is based on external attributes such as size, shape and colour (Boussa et al., 2019). However, the colour of the skin does not always indicate its suitability for consumption, and internal attributes such as total soluble solids and acidity also have to be considered to meet market requirements (Fawole & Opara, 2013; Boussa et al., 2019). Since pomegranate is a non-climacteric fruit, it is very important that fruits are harvested at their proper ripening stage to obtain their highest potential with respect to nutritional, functional and sensory properties (Nuncio-Jáuregui et al., 2014).

The acceptability of the pomegranate fruit by consumers relies on its health benefits and organoleptic properties. Nevertheless, the difficult and time-consuming separation of arils from the rind and membranes limits the consumption of this fruit. Therefore, consumers are increasingly expressing a preference for packaged ready-to-eat arils and numerous studies have been carried out to evaluate the quality and to develop different strategies to extend the shelf life of this product (Esteve-Peña et al., 2016; Maghoumi et al., 2013; Martínez-Romero, 2013; Özdemir and Gökmen, 2017; Belay et al., 2017).

Nowadays, the quality control of pomegranate fruit is still performed by traditional methods. This is mainly due to the fact that both the rind and the arils are delicate and can be damaged by the mechanical operations of machines, but also because its production is relatively low and there are no machines that have been properly adapted to such tasks. However, scientific efforts are beginning to be made in order to create new non-destructive techniques for this purpose. For instance, spectroscopy has been investigated to assess the microbial (Adiani et al., 2018) and physicochemical (Arendse et al., 2017; Arendse et al., 2018c) quality in order to detect *Ectomyelois ceratoniae* infestation (Khodabakhshian et al., 2016; Jamshidi et al., 2019), and also to predict rind scald (Arendse et al., 2018b). Dielectric spectroscopy has been applied to study the ripeness of this fruit (Castro-Giráldez et al., 2013).

A luster sensor has also been employed to measure the glossiness of the rind in order to determine the quality of intact pomegranate fruits during postharvest storage (Czieczor et al., 2018). Other techniques based on imaging, such as X-rays, have been investigated to quantify the volume of the different parts of this fruit (Salmanizadeh et al., 2014; Arendse et al., 2016a), and to detect blackheart disease and false codling moth (Arendse et al., 2016b). NMR was used to determine the effect of physiological changes induced by *Alternaria* and *Aspergillus*, and to detect blackheart (Zhang & McCarthy, 2012) and internal decay (Khoshroo et al., 2009). A computer vision system was developed to sort the arils of automated peeled fruits into different categories depending on the colour (good, immature or rotten arils) in real time (Blasco et al., 2009). This system was also able to separate raw material from the arils in the commercial line such as pieces of skin or internal membranes coming from the automatic peeling process.

Hyperspectral imaging, which integrates both spectral and spatial information (Lorente et al., 2012), has been applied as a powerful analytical processing tool for rapid, non-destructive inspection of the internal and external quality attributes in fruits with thick rind (Arendse et al., 2018a) such as orange (Liu et al., 2008), lime (Teerachaichayut & Ho, 2017) or banana (Rajkumar et al., 2012). However, only Khodabakhshian et al. (2017) have investigated the use of multispectral imaging to determine texture and TSS of intact pomegranate fruit.

Accordingly, the aim of the present study was to evaluate the capability of both machine vision techniques – colour and hyperspectral imaging – to predict the physicochemical properties and the maturity stage of pomegranate fruits cv. 'Mollar de Elche' using the information about both intact fruit and arils.

2. Material and methods

2.1. Fruit samples

A total of 210 pomegranate fruit cv. 'Mollar de Elche' samples were collected from a commercial orchard located in San Isidro (Alicante, Southern Spain). Seven harvests were carried out during the 2018 season, from the end of July (90 days after blooming) to the end of October (180 days after blooming), when the fruits were at the full ripeness stage. In each harvest, 30 fruits with no external damage were randomly collected.

In each harvest, all intact fruits were cleaned and weighed and the equatorial D was measured. Then, images of all fruits were captured as explained in the next section. After the analysis of the intact fruits, arils were carefully extracted by hand. A total of 20 arils per fruit were randomly selected and the surface moisture was removed by using paper towels. Then, images of the arils were acquired. The rest of the arils were squeezed and the juice thus obtained was used to analyse the chemical properties of each fruit.

From the seven harvesting times, three different maturity stages (immature, half-ripe and ripe) were identified according to the visual features of the intact fruit and the arils (Figure 1).

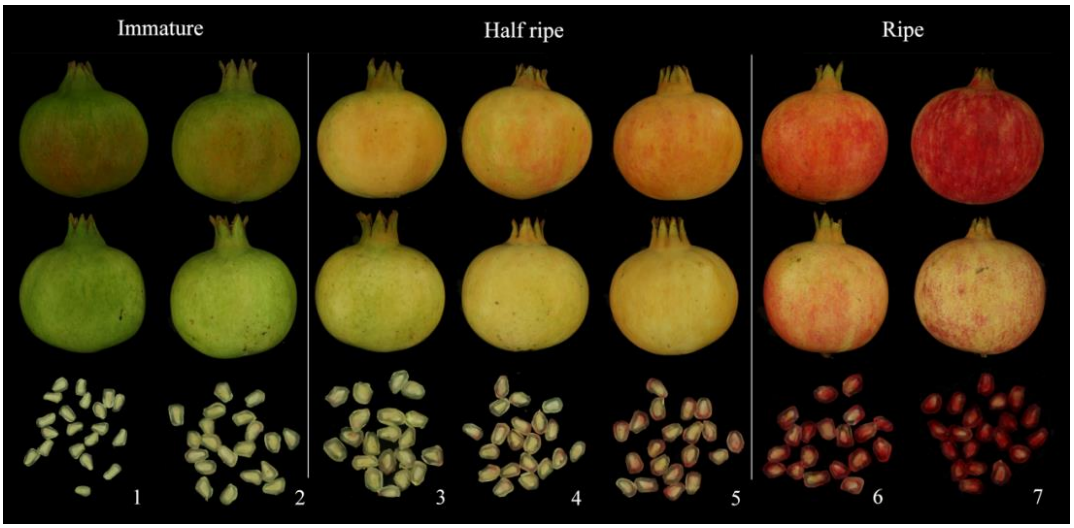


Figure 1. Example of the appearance of the two opposite sides of intact fruits and arils at each harvest and maturity stage.

2.2 Colour image acquisition and processing

The images were captured using a colour imaging system arranged inside a square inspection chamber consisting of a digital camera (EOS 550D, Canon Inc, Japan) and eight BIOLUX 18W/965 fluorescent tubes (Osram GmbH, Germany) with a colour temperature of 6500 K. Polarising filters were placed in front of the lamps and on the camera lenses to eliminate specular bright spots that could alter the true colour.

The fruit samples were introduced manually upon a holder, and two images were acquired of opposite sides of each intact fruit and one image of the 20 arils on a black background. The images were processed using customised software developed at IVIA (FoodImage-Inspector v4.0, freely available at http//www.cofilab.com) to analyse the colour. First, segmentation was performed. Because of the high contrast between the dark background and the samples, initially a threshold seemed to be sufficient to separate the fruits from the background. However, in the case of the darkest fruits or arils, some confusion appeared at the borders and therefore a more sophisticated method was used. Using this application, regions of interest belonging to a particular class (in this case background and different colours of the samples such as white, reddish or greenish) were selected. Then, using the RGB coordinates of the selected pixels and the class they were assigned to, the software built a discriminant analysis model based on the Bayes theorem. This model allowed any pixel in the image to be classified into one of the predefined classes (background or fruit). This process is explained in detail in Blasco et al. (2009).

The RGB colour coordinates obtained were converted to L*, a*, b* coordinates (CIELAB colour space), which offer a perception of colour closer to that of the human eye (Blasco et al., 2017). In addition, images of a colour reference target (ColorChecker Digital SG, X-Rite, MI, USA) were also captured as a colour reference. A total of 420 mean L*, a*, b* values of the intact fruits (two opposite sides) and 210 mean L*, a*, b* values of arils were extracted.

2.3 Hyperspectral image acquisition and processing

The hyperspectral imaging system was composed of an industrial camera (CoolSNAP ES, Photometrics, AZ, USA), coupled to a liquid crystal tuneable filter (Varispec NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). The camera was configured to acquire images with a size of 1392 × 1040 pixels and a spatial resolution of 0.14 mm per pixel at 34 different wavelengths every 10 nm, in the working spectral range of 720–1050 nm. In order to avoid problems of unfocused images due to the refraction of light across this wide spectral range, the focus was

adjusted on a central band of the acquisition interval and the images were captured using lenses capable of covering the entire spectral range without going out of focus (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA). To optimise the dynamic range of the camera, prevent saturated images and correct the spectral sensitivity of the different elements of the system, a calibration of the integration time of each band was performed by capturing the average grey level of a white reference target (Spectralon 99 %, Labsphere, Inc, NH, USA) corresponding to 90 % of the dynamic range of the camera.

The scene was illuminated by indirect light from twelve halogen spotlights (37 W) (Eurostar IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current (12 V) and arranged equidistant from each other inside a hemispherical aluminium diffuser. The fruit samples were introduced manually into a holder. The inner surface of the aluminium diffuser was painted white to maximise its reflectivity and given a rough texture in order to minimise directional reflections, which could cause bright spots, the result being a highly homogeneous light.

For the acquisition of the colour, three images per fruit (two of the opposite sides of the intact fruit and one of 20 random arils) were acquired using the hyperspectral imaging system.

The image processing started with the correction of the relative reflectance by using equation (1) (Gat, 2000):

$$\rho_{xy}(x,y,\lambda) = \frac{\mathrm{R^{abs}}}{\mathrm{R^{abs}_{white}}} = \rho^{\mathrm{Ref}}(\lambda)\frac{\mathrm{R(x,y,\lambda)-R_{black}(x,y,\lambda)}}{\mathrm{R_{white}(x,y,\lambda)-R_{black}(x,y,\lambda)}} \quad (1)$$

where $\rho^{Ref}(\lambda)$ is the standard reflectance of the white reference target (99 % in this work), $R(x,y,\lambda)$ is the reflectance of the fruit captured by the charge-coupled device (CCD) sensor of the camera, $R_{white}(x,y,\lambda)$ is the reflectance captured by the CCD of the white reference target, and $R_{black(}x,y,\lambda)$ is the reflectance captured by the CCD while avoiding any light source in order to quantify the electronic noise of the CCD.

Later, the background was removed using the clustering based method k-means, carried out using the toolbox HYPER-Tools (Mobaraki & Amigo, 2018) for MATLAB R2017b (The MathWorks, Inc. MA, USA). This method assigns each pixel of the image

to the k cluster whose centre is nearest, by minimising the sum of the squared distances of each pixel to its corresponding centre (Amigo et al., 2008).

The relative reflectance spectrum of all pixels of each fruit sample was finally extracted, resulting in a total of 420 mean spectra of the intact fruit (two opposite sides) and 210 mean spectra of arils.

2.4 Chemical properties

2.4.1 Total soluble solids, titratable acidity and maturity indexes

After obtaining the juice of each fruit, the total soluble solids (TSS) were determined using a digital refractometer (Atago N-20, Atago, Bellevue, Wash., USA) at 20 ºC and results were expressed as % of TSS. The TA was determined using an 877 Titrino plus acid-base potentiometer (Metrohm AG, Herisau, Switzerland). The TA results were obtained using 0.1 mol L^{-1} NaOH and expressed as g of citric acid per L. The ratio TSS/TA, or MI, was calculated for each sample. Furthermore, because MI does not always correlate well with the perception of sweetness or sourness in the fruit (Jordan et al., 2001), the BrimA index was also calculated using Eq. (2):

$$BrimA = TSS - k * TA \quad (2)$$

where *k* is the tongue's sensitivity index ranging from 2 to 10. According to Arendse et al., (2017), *k* value of 2 was used to avoid a negative BrimA value.

2.4.2 Total polyphenolic compounds

The content of the TPC was determined using the Folin-Ciocalteu method described by Singleton et al. (1999). Absorption was measured at 760 nm using a UV–VIS Spectrophotometer (Helios Gamma model, UVG 1002E, Mercers Row, Cambridge, UK). Calibration curves, with a concentration range between 0 and 0.25 g gallic acid per L, were used for the quantification of TPC, showing a good correlation ($R^2 \geq 0.996$). TPC was expressed as g of gallic acid equivalents L^{-1}.

2.4.3 Antioxidant activity

Following Nuncio-Jáuregui et al. (2015), a methanol extract from each sample was prepared to analyse the antioxidant activity (AA) by mixing 1 mL juice with 10 mL of MeOH/water (80:20, v/v) + 1 % HCl, before sonicating at 20 °C for 15 min and then leaving them for 24 h at 4 °C. The extract was then sonicated again for 15 min, and centrifuged at 25,058 × g for 10 min. The free radical scavenging capacities were determined by three methods, ABTS (Re et al., 1999), DPPH radical (Brand-Williams et al., 1995), and FRAP (Benzie and Strain, 1996). Calibration curves, in the range 0.5 and 5 mmol Trolox L^{-1}, were used to quantify AA, showing a good correlation ($R^2 \geq 0.998$). The absorbance of the three methods was measured using a UV–VIS Spectrophotometer (Helios Gamma model, UVG 1002E, Mercers Row, Cambridge, UK), and the results were expressed as mmol Trolox equivalents L^{-1}.

2.4.4 Data analysis

Finally, one-way ANOVA and multiple-range tests were used to compare the physicochemical parameters in each harvest and maturity stage. The method used to discriminate among the means (Multiple Range Test) was Tukey's procedure. Significance was defined at $p \leq 0.05$. The groups of samples met the following three requirements: i) the observations being tested are independent within and among the groups; ii) the groups associated with each mean in the test are normally distributed; and iii) there is equal within-group variance across the groups associated with each mean in the test (homogeneity of variance).

These analyses were performed using StatGraphics software (Manugistics, Inc., Rockville, USA).

2.5. Chemometric methods for quantitative and qualitative analysis

The multivariate data analysis started by partitioning the colour and spectral data of the intact fruit and arils randomly into two sets: two thirds of the samples (training set) were used to calibrate the models and for cross-validation (CV), while the

remaining third was used for independent test prediction (test set). In the case of the intact fruit, a total of 280 mean spectra were included in the training set and the other 140 were used in the test set. For the arils, 140 mean spectra were part of the training set and the other 70 were in the test set.

PCA was carried out on the colour and spectral data of the calibration set to obtain an overview of the main source of variance between the samples of different maturity stages. PCA transforms the variables of a dataset into a new set, called the principal components, using linear combinations of the original variables (Basha et al., 2018). This method is usually used for pattern recognition, classification and feature extraction without previous knowledge of the data. The aim of applying PCA was to reduce the dimensionality of the dataset, while retaining most of the variability.

For the quantification of the properties of the samples and discrimination of their maturity, PLS-R and PLS-DA were used. The PLS-R method searches for a linear regression model of latent variables by projecting prediction variables X (colour or spectral data) and response variables Y (reference properties such as W, D, TSS, TA, MI, BrimA, AA or TPC) into a new latent space where the covariance between these latent variables is maximised. The goal is to find the latent multidimensional direction in the data space that explains the direction of the maximum multidimensional covariance in the reference properties space (Lorente et al., 2012). On the other hand, when the Y variable is categorical, PLS-DA is performed in order to sharpen the separation between groups of observations by maximising the covariance between the colour or spectral data and the classes, such that a maximum separation among these classes is obtained (Lorente et al., 2012).

In order to reduce the variability among samples due to light scatter (Rinnan, van den Berg & Engelsen, 2009), the mean spectra of the intact fruit and arils were pre-processed using pre-treatments such as SNV and Savitzky Golay derivative (3-point smoothing window, second-order polynomial). Mean centring was applied to normalise the full spectrum.

The selection of the optimal number of LVs, as well as the estimation of the error rate of the PLS-R and PLS-DA models, were performed using a 10-fold CV on the

training set. In this process, the data were split evenly into 10 sets, leaving one of the sets out in each iteration of the validation procedure.

The predictive capability of the PLS-R models was evaluated by the RMSE and the R^2 between the predicted and the measured values of the quality indices for calibration, CV and prediction. Furthermore, RPD, defined as the ratio between the standard deviation of the reference data and RMSEP was also used (Williams, 1987). An RPD between 1.5 and 2 means that the model can discriminate between low and high values of the response variable. A value between 2 and 2.5 indicates that coarse quantitative predictions are possible, and a value between 2.5 and 3 or above corresponds to good and excellent prediction accuracy, respectively (Nicolaï et al., 2007).

The results of the PLS-DA models were expressed as a percentage of correct classification and total accuracy for calibration, CV and prediction.

3. Results and discussion

3.1 Physicochemical changes during maturity

Table 1 summarises the mean and standard deviation of the physicochemical properties measured in the intact fruit and juice samples at each harvest and maturity stage.

The physical properties of the intact fruit such as W or D are important from a commercial viewpoint because these attributes influence consumer preference (Fawole & Opara, 2013). In this case, the W increased from 237 g in the first harvest to 456 g in the last one, while the D increased from 78.9 mm to 95.6 mm. This increase was more noticeable when the fruit passed from the immature to the half-ripe stage. The fruit usually continues growing even after the optimum harvesting time, due to cell expansion from uptake of water and other nutrients (Shwartz et al., 2009). According to the specifications of the 'Granada Mollar de Elche' Protected Designation of Origin (http://www.agroambient.gva.es/en/pc_granadamollarelche), fruits weighing less than 125 grams must be excluded.

The TSS increased from 12.0 % in the first harvest to 16.6 % in the last harvest, the values of the last three harvests being non-statistically different. As regards the physical parameters, the increase was more evident when the fruit passed from the immature to the half-ripe stage. This agrees with Nuncio-Jáuregui et al. (2014), who studied three different stages of maturity of 'Mollar de Elche' and obtained TSS values of 14.6, 15.4 and 15.9 %. According to the specifications (http://www.agroambient.gva.es/en/pc_granadamollarelche), this indicates that the minimum value of TSS in ripe fruit is 14 %.

Table 1. Physicochemical properties of the pomegranate fruit at each harvest and maturity stage.

	Intact fruit		Juice			
Harvest / *maturity stage*	W (g)	D (mm)	TSS (ºBrix)	TA (g/L)	MI	BrimA
1	$237^{f} \pm 25$	$78.9^{d} \pm 3.1$	$12.0^{e} \pm 1.4$	$2.6^{a} \pm 0.3$	$4.6^{f} \pm 0.3$	$6.7^{e} \pm 1.0$
2	$274^{e} \pm 33$	$82.9^{c} \pm 3.8$	$13.7^{d} \pm 0.9$	$2.5^{ab} \pm 0.1$	$5.5^{e} \pm 0.4$	$8.7^{d} \pm 0.8$
Immature	*$256^{C} \pm 34$*	*$80.9^{C} \pm 4.0$*	*$12.8^{C} \pm 1.4$*	*$2.6^{A} \pm 0.2$*	*$5.1^{C} \pm 0.6$*	*$7.7^{C} \pm 1.3$*
3	$305^{d} \pm 45$	$84.5^{c} \pm 3.8$	$14.8^{c} \pm 0.6$	$2.4^{b} \pm 0.2$	$6.2^{d} \pm 0.5$	$10.0^{c} \pm 0.6$
4	$372^{c} \pm 40$	$90.8^{b} \pm 4.1$	$15.7^{b} \pm 0.8$	$2.2^{c} \pm 0.3$	$7.3^{c} \pm 0.8$	$11.4^{b} \pm 0.8$
5	$423^{b} \pm 64$	$94.8^{a} \pm 3.9$	$16.4^{a} \pm 0.8$	$2.1^{cd} \pm 0.2$	$7.9^{b} \pm 0.5$	$12.2^{a} \pm 0.6$
Half ripe	*$367^{B} \pm 70$*	*$90.0^{B} \pm 5.8$*	*$15.6^{B} \pm 1.0$*	*$2.2^{B} \pm 0.3$*	*$7.1^{B} \pm 0.9$*	*$11.2^{B} \pm 0.7$*
6	$414^{b} \pm 59$	$93.3^{ab} \pm 4.1$	$16.2^{a} \pm 0.8$	$2.0^{d} \pm 0.3$	$8.4^{a} \pm 1.0$	$12.3^{a} \pm 0.7$
7	$456^{a} \pm 66$	$95.6^{a} \pm 4.1$	$16.6^{a} \pm 0.8$	$2.0^{d} \pm 0.2$	$8.3^{a} \pm 0.6$	$12.6^{a} \pm 0.6$
Ripe	*$435^{A} \pm 66$*	*$94.5^{A} \pm 4.2$*	*$16.4^{A} \pm 0.8$*	*$2.0^{C} \pm 0.2$*	*$8.4^{A} \pm 0.8$*	*$12.5^{A} \pm 0.7$*

Mean value ± standard deviation. Different lowercase letters in the same column indicates statistical difference between harvests (p-value < 0.05); different capital letters in the same column indicates statistical difference between maturity stages (p-value < 0.05). W = weight; D = diameter; TA = titratable acidity; TSS = total soluble solids; MI = maturity index.

In contrast, the tendency of TA was to decrease during maturity. In general, the values of TA in the maturity stages were statistically different, from 2.6 g L^{-1} in immature fruit to 2.0 g L^{-1} in ripe fruit. This trend agrees with Nuncio-Jáuregui et al. (2014), who obtained TA values of 2.5, 2.4 and 2.3 g L^{-1} during maturity. The

established minimum value of TA in the ripe fruit of this cultivar is 1.8 g L^{-1} and the maximum is 2.4 g L^{-1} (http://www.agroambient.gva.es/en/pc_granadamollarelche).

The MI is commonly used to define the 'taste' of pomegranate fruit during development (Shwartz et al., 2009). The value of MI increased from 4.6 to 8.4 with no statistical difference between the last two harvests. As in the other parameters, the main changes were observed when the fruit passed from the immature to the half-ripe stage. The minimum value of MI in the ripe fruit of this cultivar has to be 6.0 and the maximum 9.0 (http://www.agroambient.gva.es/en/pc_granadamollarelche). In this case the value of ripe fruit was 8.4.

As mentioned, although MI is commonly used, this ratio does not always correlate well with the perception of maturity. Jordan et al. (2001) proposed the BrimA index due to the fact that sugars and acids have opposite effects on flavour, and the tongue is more sensitive to acidity. In this work, BrimA was found to be more related to flavour than MI. The values of BrimA increased from 6.7 to 12.5, with no statistical differences among the three harvests, as in the case of MI.

With regard to the bioactive compounds, Table 2 summarises the mean and standard deviation of the TPC content and the results of the AA measured in the juice samples at each harvest and maturity stage using the DPPH, ABTS and FRAP methods.

Polyphenols have high antioxidant capacity and are also responsible for major organoleptic characteristics, especially colour and taste properties. For this reason, TPC content is an important quality parameter of pomegranate fruit. The TPC content measured in the juice of the fruit in this study decreased from 2.7 to 1.2 g L^{-1}. This decrease was more pronounced in the fourth harvest, the immature fruit being richer in these compounds. Kulkarni and Aradhya (2005) also reported a 54.5 % reduction in TPC during the initial stage of fruit development and the decrease continued until the fruit was considered fully ripened. Something similar happened with AA, which is related to the level of TPC. The AA decreased from 8.0 to 2.5 mmol Trolox L^{-1} using the DPPH method, from 4.2 to 1.9 mmol Trolox L-1 using ABTS and from 6.4 to 3.1 mmol Trolox L^{-1} using FRAP; this decrease was more pronounced in the fourth harvest. However, these results are different to those obtained by Nuncio-Jáuregui et al. (2014), who presented TPC values of 3.7, 3.3 and 2.7 g L^{-1} and an AA of 7.0, 6.5 and 6.6

mmol Trolox L^{-1} using the DPPH method. As Mena et al. (2011) pointed out, the variation in the concentration of these parameters can vary to a large degree depending on factors such as growing area, weather conditions or the influence of processing techniques such as thinning.

Table 2. Antioxidant activity and total polyphenol content of the pomegranate juice at each harvest and maturity stage.

Harvest / *maturity stage*	TPC (g GAE/L)	Antioxidant activity (mM Trolox/100 ml)		
		DPPH	ABTS	FRAP
1	$2.7^{a} \pm 0.3$	$8.0^{a} \pm 0.2$	$4.2^{a} \pm 0.7$	$6.4^{a} \pm 0.9$
2	$2.4^{b} \pm 0.2$	$6.8^{b} \pm 0.4$	$3.8^{b} \pm 0.4$	$5.8^{b} \pm 0.6$
Immature	*$2.5^{A} \pm 0.3$*	*$7.4^{A} \pm 0.7$*	*$4.0^{A} \pm 0.6$*	*$6.1^{A} \pm 0.8$*
3	$2.4^{b} \pm 0.2$	$5.1^{c} \pm 0.7$	$3.9^{b} \pm 0.4$	$6.0^{b} \pm 0.9$
4	$1.2^{c} \pm 0.1$	$3.3^{d} \pm 0.1$	$1.7^{c} \pm 0.2$	$3.0^{cd} \pm 0.5$
5	$1.2^{c} \pm 0.1$	$3.1^{e} \pm 0.2$	$1.6^{e} \pm 0.1$	$2.5^{e} \pm 0.3$
Half ripe	*$1.6^{B} \pm 0.6$*	*$3.8^{B} \pm 1.0$*	*$2.4^{B} \pm 1.1$*	*$3.8^{B} \pm 1.7$*
6	$1.1^{c} \pm 0.1$	$2.8^{f} \pm 0.2$	$1.8^{c} \pm 0.2$	$2.8^{de} \pm 0.3$
7	$1.2^{c} \pm 0.1$	$2.5^{g} \pm 0.2$	$1.9^{c} \pm 0.2$	$3.1^{c} \pm 0.3$
Ripe	*$1.2^{C} \pm 0.1$*	*$2.7^{C} \pm 0.2$*	*$1.8^{C} \pm 0.2$*	*$3.0^{C} \pm 0.4$*

Mean value ± standard deviation. Different lowercase letters in the same column indicates statistical difference between harvests (p-value < 0.05); different capital letters in the same column indicates statistical difference between maturity stages (p-value < 0.05). TPC = total polyphenolic compounds

3.2 Colour and spectral data analysis

Colour images of intact fruit and arils were acquired in each harvest and the L*, a* and b* colour coordinates were calculated. The average mean value of each colour parameter of the intact fruit and the arils at each maturity stage are presented in Figure 4. In the rind, only a* (which goes from green to red) increased during maturity and all values of the three maturity stages were statistically different. Therefore, the colour of the rind evolved from green to reddish (Figure 1), mainly due to the decrease in chlorophylls and carotenoids content and the increase in the synthesis of pigments

such as anthocyanins (Zhao et al., 2015). In the case of arils, the three coordinates evolved during maturity. L* and b* (which goes from blue to yellow) decreased, especially from the second to the third maturity stage, while a* increased greatly. Therefore, the colour of the arils evolved from the immature white arils to reddish (Figure 1) due to the synthesis of anthocyanins (Gil et al., 1995). These changes observed in the colour of the intact frits and arils coincide with those found in the study by Melgarejo et al. (1997).

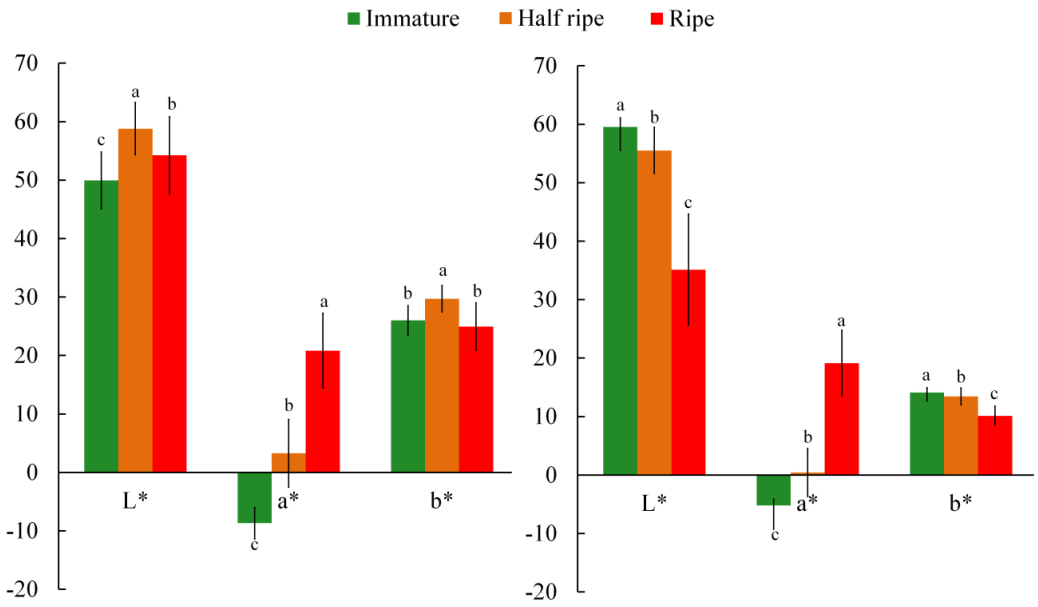


Figure 4. Colour coordinates of the intact fruit (left) and arils (right) at each maturity stage. *The points in the graph are the mean value and bars are standard deviation. Different letters in the same parameter indicates significant differences between harvests (p-value<0.05), according to Tukey's (HSD) test.*

The average mean reflectance spectra obtained from the hyperspectral images for intact fruit and arils are presented in Figure 5. The spectra showed similar trends for intact fruit and arils but the reflectance intensity is different in each maturity stage. This means that they have similar constituents but in different concentrations. In the case of the intact fruit, these differences in reflectance intensity were located around 720–750 nm, close to the chlorophyll absorption peak of 680 nm, and the valley present in the region 960–990 nm, primarily assigned to water absorption bands. This valley was more pronounced in the most mature fruit because the water content

increases in the tissues during the onset of ripening, due to cell breakage and osmotic movement of water (Rajkumar et al., 2012).

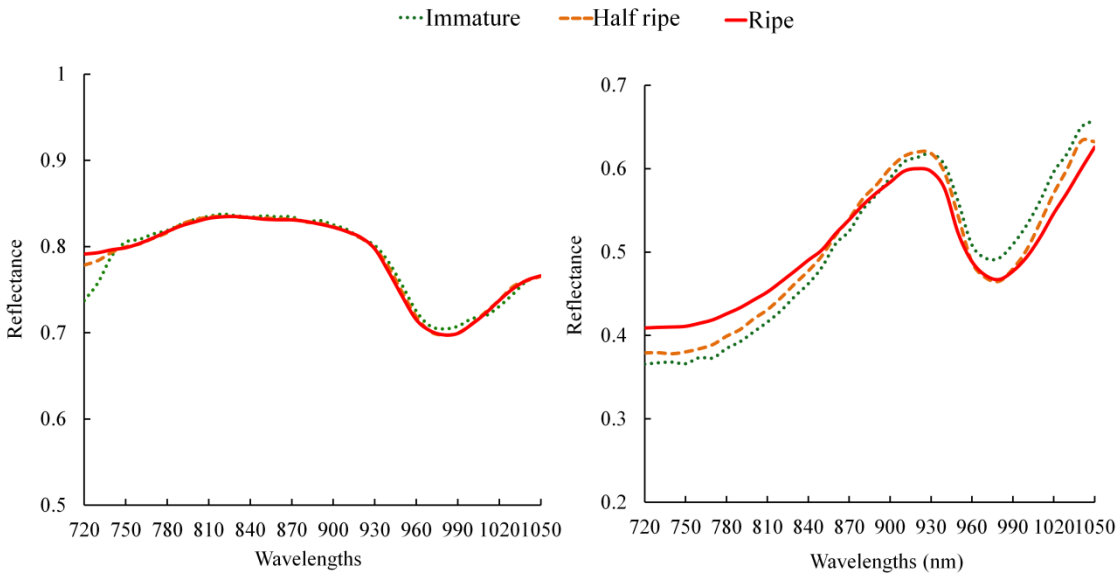


Figure 5. Mean spectra of intact fruit (left) and arils (right) at each maturity stage.

In the mean spectra of the arils, the differences were more pronounced than in the intact fruit. These were mostly visualised between 720–800 nm, with a peak around 920 nm and a valley between 960 and 990 nm, assigned to acids, sugars and water absorption (Yang et al., 2015). However, as Nicolaï (2007) pointed out, the spectra are dominated by the water spectrum with overtone bands of the OH-bonds at 760 nm and 970 nm, and sophisticated multivariate statistical techniques are needed to extract the useful information from these spectra.

3.3 Principal component analysis

A PCA was performed to explore the colour and spectral data of the intact fruit and arils and to obtain an overview of the distribution of the samples in a non-supervised way. Figures 6a and 6b show the score plot and loadings of the PCA using colour and spectral data of the intact fruit and arils.

Using colour data of the intact fruit, the first two PCs of the model explained 98.5 % of the total variance (75.7 % and 22.8 %) and using the data from arils, the first and third PCs explained 98.6 % of the total variance (98.1 % and 0.5 %).

In the case of the spectral data of intact fruit, the first two PCs explained 93.6 % of the total variance (84.7 % and 8.9 %) and the first and third PCs explained 85.4 % using the spectral data from arils (84.7 % and 0.7 %).

The scores plot shows the grouping of the fruit in the three maturity stages in all cases. The colour data of the intact fruit shows more separation of the three maturity stages than the spectral data in which the half-ripe and ripe stages were overlapped. This was probably due to the fact that most of the changes that occur during maturity involve pigments that are related to visible wavelengths such as chlorophyll or anthocyanin, among others. The loadings suggest that a* in the first PC and L* in the second PC could be the most important colour parameters for the monitoring of maturity in the intact fruit. In the case of the spectral data, although the loadings obtained for PC1 and PC2 might offer information on the most important wavelengths to distinguish the maturity stages, this was not useful since separation was not evident in the score plot.

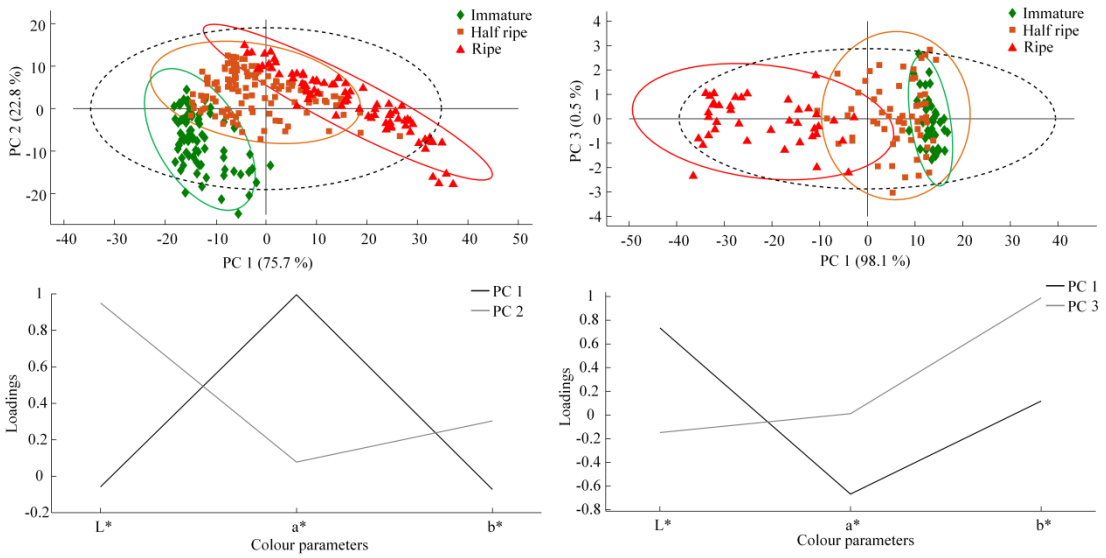


Figure 6a. Principal components analysis using colour data. Score plot and loadings of intact fruit (left) and arils (right).

In the case of arils, the best separation of maturity stages was achieved by the spectral data because, in the model based on colour data, immature and half-ripe were overlapped. The corresponding loadings show that although the three colour parameters gave information, this was not sufficient to obtain a clear separation of maturity stages in the score plot. The loadings of the spectral model show that in PC1 the region 960–1050 nm and in PC3 the wavelengths 730, 770, 850, 960 and 980 nm could be the most important wavelengths for the monitoring of maturity in arils.

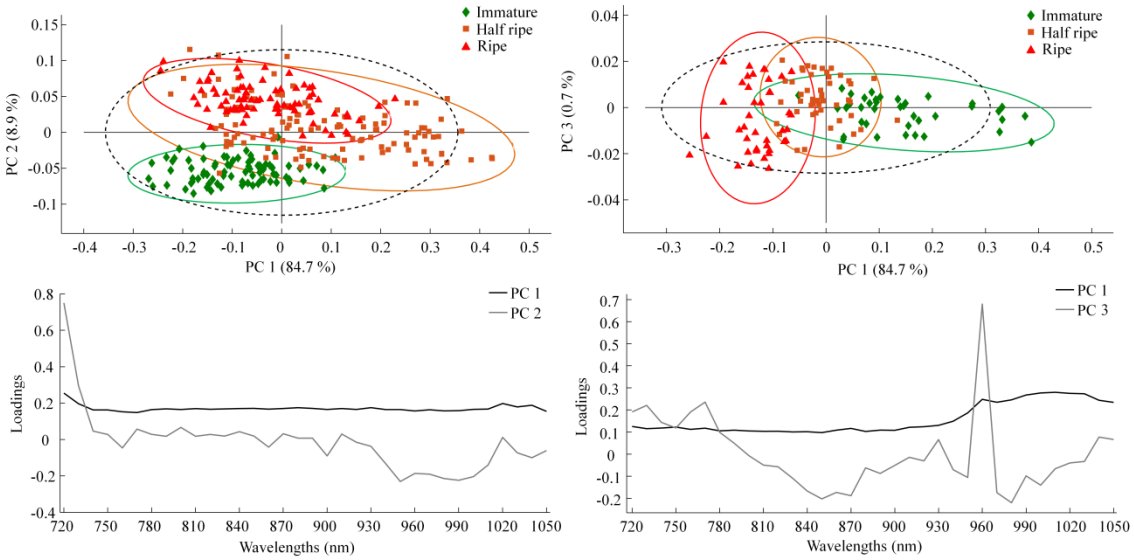


Figure 6b. Principal components analysis using spectral data. Score plot and loadings of intact fruit (left) and arils (right).

Although the PCA maximises the variance in the first components, this does not guarantee the separability of data by classes due to its unsupervised nature (Jolliffe, 2002). For this reason, supervised models need to be investigated for use in quantitative and qualitative analyses that are capable of identifying the maturity of the fruit and predicting their physicochemical properties.

3.4. Prediction of the physicochemical properties

Table 3 and 4 show the results of the calibration and validation of the models to predict the physicochemical properties using colour and spectral data of the intact fruit and arils.

Table 3. Results of the prediction of physicochemical properties using the colour data of the intact fruit and arils.

	Property		#LV	Calibration		Cross Validation		Prediction		RPD
				R^2	RMSE	R^2	RMSE	R^2	RMSE	
Intact fruit	W		3	0.65	53.3	0.64	54.0	0.62	58.1	1.6
	D		3	0.67	4.09	0.66	4.14	0.64	4.49	1.6
	TSS		3	0.75	0.94	0.72	0.99	0.68	0.94	1.7
	TA		3	0.53	0.23	0.45	0.24	0.47	0.24	1.4
	MI		3	0.81	0.66	0.81	0.67	0.78	0.70	2.1
	BrimA		2	0.85	0.85	0.84	0.90	0.81	0.92	2.2
	Arils colour	L*	3	0.78	5.46	0.69	6.51	0.77	5.47	2.1
		a*	3	0.86	3.95	0.83	4.42	0.85	4.05	2.6
		b*	3	0.54	1.49	0.47	1.59	0.49	1.44	1.4
	TPC		2	0.82	0.29	0.78	0.32	0.81	0.30	2.3
	DPPH		2	0.89	0.68	0.87	0.74	0.84	0.83	2.5
	ABTS		2	0.77	0.57	0.72	0.62	0.72	0.61	1.9
	FRAP		2	0.71	0.90	0.66	0.98	0.75	0.94	1.9
Arils	TSS		3	0.51	1.32	0.37	1.52	0.44	1.23	1.3
	TA		2	0.41	0.25	0.37	0.26	0.37	0.26	1.3
	TSS/TA		3	0.68	0.87	0.62	0.95	0.55	1.00	1.4
	BrimA		3	0.62	1.37	0.50	1.59	0.53	1.41	1.4
	TPC		2	0.62	0.42	0.51	0.48	0.57	0.45	1.4
	DPPH		2	0.63	1.22	0.48	1.47	0.61	1.28	1.6
	ABTS		2	0.55	0.79	0.43	0.90	0.51	0.80	1.4
	FRAP		2	0.56	1.11	0.45	1.26	0.49	1.30	1.4

#LV = numbers of latent variables; W = weight; D = equatorial diameter; TSS = total soluble solids; TA = titratable acidity; MI = TSS/TA; TPC = total phenolic compounds

Table 4. Results of the prediction of physicochemical properties using the spectral data of the intact fruit and arils.

	Property		#LV	Calibration		Cross Validation		Prediction		RPD
				R^2	RMSE	R^2	RMSE	R^2	RMSE	
Intact fruit	W		9	0.75	45.0	0.71	48.7	0.67	54.2	1.7
	D		5	0.73	3.69	0.70	3.87	0.71	4.10	1.8
	TSS		6	0.80	0.84	0.77	0.90	0.71	0.89	1.8
	TA		4	0.51	0.23	0.48	0.23	0.46	0.24	1.4
	MI		8	0.82	0.64	0.80	0.68	0.71	0.81	1.8
	BrimA		6	0.88	0.78	0.86	0.84	0.85	0.79	2.6
	Arils colour	L*	7	0.79	5.37	0.70	6.40	0.68	6.45	1.8
		a*	7	0.83	4.41	0.76	5.22	0.75	5.15	2.0
		b*	7	0.54	1.49	0.43	1.65	0.45	1.50	1.3
	TPC		9	0.88	0.24	0.84	0.27	0.86	0.25	2.7
	DPPH		5	0.93	0.55	0.91	0.61	0.91	0.62	3.4
	ABTS		9	0.85	0.46	0.81	0.51	0.83	0.47	1.9
	FRAP		9	0.82	0.72	0.78	0.79	0.85	0.74	2.4
Arils	TSS		10	0.82	0.80	0.70	1.07	0.77	0.82	2.0
	TA		5	0.51	0.22	0.34	0.26	0.46	0.24	1.4
	TSS/TA		10	0.83	0.62	0.74	0.78	0.78	0.72	2.1
	BrimA		10	0.89	0.73	0.81	0.97	0.88	0.72	2.7
	TPC		12	0.92	0.19	0.83	0.29	0.87	0.25	2.7
	DPPH		11	0.94	0.48	0.88	0.70	0.92	0.57	3.6
	ABTS		11	0.90	0.37	0.79	0.56	0.81	0.52	2.2
	FRAP		10	0.84	0.66	0.67	0.98	0.76	0.89	2.1

#LV = numbers of latent variables; W = weight; D = equatorial diameter; TSS = total soluble solids; TA = titratable acidity; MI = TSS/TA; TPC = total phenolic compounds

3.4.1 Weight and equatorial diameter

The physical parameters W and D were predicted using only the information of the intact fruit. The models using colour data were calibrated using 3 LVs (Table 3). The R^2 of prediction (R^2_P) was 0.62 and 0.64 with an $RSME_P$ of 58.1 g and 4.49 mm for W and D, respectively. When W and D were correlated with the spectral data, the models were calibrated using 9 and 5 LVs (Table 4). The R^2_P was 0.64 and 0.71, respectively, and the $RSME_P$ was 54.2 g and 4.10 mm. The results of W prediction using colour and spectral data are in accordance with Arendse et al. (2018c), who obtained an R^2 of 0.62 in the 'Wonderful' cultivar using Fourier-transform near infrared diffuse reflectance spectroscopy.

Although hyperspectral imaging obtained more accurate results in W and D, the RPD value obtained using colour data was 1.6 for both parameters, and 1.7 and 1.8 using spectral data, respectively. These values indicated that the models can only discriminate the lower from the higher values of the response variable, but this is not sufficient, and hence they are not recommended for a quality control application.

3.4.2 Total soluble solids, titratable acidity and maturity indexes

The models to predict the organoleptic properties such as TSS and TA using the colour data of the intact fruit were calibrated using 3 LVs (Table 3). The R^2_P obtained was 0.68 and 0.47 with an $RSME_P$ of 0.94 % and 0.24 g L^{-1}. In the case of the arils, the models were calibrated using 3 and 2 LVs, respectively (Table 4). The R^2_P obtained was 0.44 and 0.37 with an $RSME_P$ of 1.23 % and 0.26 g L^{-1}. When TSS and TA were correlated with the spectral data, the models using intact fruit information were calibrated using 6 and 4 LVs. The R^2_P was 0.71 and 0.46 and the $RSME_P$ was 0.89 % and 0.24 g L^{-1}. The models using the arils information were calibrated using 10 and 5 LVs, obtaining an R^2_P of 0.77 and 0.46 and the $RSME_P$ was 0.82 % and 0.24 g L^{-1}, respectively.

Other previously tested non-destructive techniques yielded different results in the prediction of these compounds in several cultivars of pomegranate. Arendse et al. (2017) and (2018c) used spectroscopy to predict TSS and TA of 'Wonderful'

pomegranates using the intact fruit information, obtaining an R^2 of 0.78 and 0.77, respectively. Using the arils information they obtained an R^2 of 0.88 and 0.87. Khodabakhshian et al. (2017) predicted the TSS of intact pomegranate fruit cv. Ashraf using a multispectral system and obtained an R^2 of 0.94. Zhang and McCarthy (2013) employed NMR to assess TSS and TA in 'Wonderful' pomegranates, obtaining an R^2 of 0.12 and 0.54.

In this study, the values of RPD obtained for TSS using colour data were 1.7 and 1.3, indicating a fair model performance using intact fruit information but a poor model performance using the information coming from the arils. When the spectral data were used, the values were 1.8 and 2.0, indicating a good model performance using both the intact fruit and the arils information. Thus, quantitative predictions are possible using hyperspectral imaging with intact fruit and arils information. For TA, the RPD value obtained using the intact fruit and arils information and both techniques was 1.4, which indicated a poor performance of the model.

The prediction of MI and BrimA using colour data of intact fruit was performed by means of two models calibrated using 3 and 2 LVs, respectively (Table 3). The R^2_P was 0.78 and 0.81 and the $RSME_P$ was 0.70 and 0.92. Using the colour data from arils, the two models were calibrated using 3 LVs (Table 3). The R^2_P was 0.55 and 0.53 and the RSMEP was 1.00 and 1.41. In the case of spectral data, the R^2_P of MI and BrimA using the intact fruit information was 0.71 and 0.85 and the $RSME_P$ was 0.81 and 0.79 when the models were calibrated using 8 and 6 LVs (Table 4). Using arils information, the R^2_P was 0.78 and 0.88 and the $RSME_P$ was 0.72 when the models were calibrated using 10 LVs (Table 4). Similar results were obtained by Arendse et al. (2018c) in intact 'Wonderful' pomegranates, R^2 of 0.78 and 0.79, and in arils, R^2 of 0.82 and 0.83 (Arendse et al., 2017). Zhang and McCarthy (2013) used nuclear magnetic resonance to assess MI in 'Wonderful' pomegranates, obtaining an R^2 of 0.63.

The values of RPD of MI and BrimA obtained using colour data were 2.1 and 2.2, which indicated a good model performance. However, using the information from the arils, the value of RPD of both indices was 1.4, which indicated a poor performance of the model. The values of RPD using intact fruit were 1.8 and 2.6, indicating a fair model performance for MI and excellent prediction accuracy for BrimA. The values obtained

using arils information were 2.1 for the prediction of MI, which means that it is possible to predict this index, and 2.7 for BrimA, indicating an excellent model performance.

The PLS-R models calibrated using the colour data showed a limited potential of TSS, MI and BrimA prediction when the arils information was used. These results demonstrated greater potential of hyperspectral imaging compared to conventional colour imaging for predicting these properties in pomegranate fruit cv. 'Mollar de Elche' pomegranate using both intact fruit and arils information.

3.4.3 Internal colour

Regarding the internal colour, the prediction models of the colour coordinates of the arils, L*, a* and b*, were predicted using only the information of the intact fruit. The three models using colour data were calibrated using 3 LVs (Table 3). The R^2_P obtained was 0.77, 0.85 and 0.49 and the $RSME_P$ was 5.47, 4.05 and 1.44, respectively. In the case of spectral data, the models were calibrated using 7 LVs (Table 4). The R^2_P was 0.68, 0.75 and 0.45 and the $RSME_P$ was 6.45, 5.15 and 1.50.

The RPD values obtained for L*, 2.6 and 2.0, indicated an excellent model performance using colour data and a fair model performance using spectral data. The values obtained for a*, 2.1 and 1.8, indicated a good model performance using colour data and a fair model performance using spectral data. In the case of b*, both techniques presented an RPD of 1.4 and 1.3, which means a poor model performance. Arendse et al. (2018c) predicted the a* coordinate of arils of the 'Wonderful' cultivar using spectroscopy and obtained a similar result to that of hyperspectral imaging: an R^2 of 0.71. These results showed that colour information of the rind had a better correlation with L* and a* of arils colour than the spectral data in the NIR region.

3.4.4 Total polyphenolic compounds

In the case of TPC, the models using the colour data of the intact fruit and arils were calibrated using 2 LVs (Table 3). The R^2_P was 0.81 and 0.57 and the $RSME_P$ was 0.30 and 0.45 g L^{-1}. Using spectral data, the model of the intact fruit and arils were

calibrated using 9 and 12 LVs (Table 4). The R^2_P was 0.86 and 0.87 and the $RSME_P$ was 0.25 and 0.25 g L^{-1}. Arendse et al. (2018c) and Arendse et al. (2017) also predicted TPC in intact fruit and arils using spectroscopy and obtained similar results to those achieved using hyperspectral imaging: an R^2 of 0.83 using the intact fruit information and 0.87 using the arils information.

The RPD values obtained for colour data were 2.3 and 1.4, indicating a good model performance using the intact fruit information but a poor model performance using arils information. In contrast, the spectral data models obtained a value of 2.7, indicating an excellent performance using both intact fruit and arils information. As in the case of TSS, MI and BrimA, these results demonstrated the greatest potential of hyperspectral imaging for predicting TPC in 'Mollar de Elche' pomegranate using the intact fruit and arils information.

3.4.5 Antioxidant activity

Antioxidant activity was predicted by correlating the values obtained by means of the DPPH, ABTS and FRAP methods, and colour and spectral data of the intact fruit and arils. The three models using the colour data of the intact fruit were calibrated using 2 LVs and the R^2_P was 0.84, 0.72 and 0.75 and the $RSME_P$ was 0.83, 0.61 and 0.94 mmol Trolox L^{-1} (Table 3). The three models using the colour data from arils were also calibrated using 2 LVs but in this case the R^2_P was 0.61, 0.51 and 0.49 and the $RSME_P$ was 1.28, 0.80 and 1.30 mmol Trolox L^{-1}.

When the spectral data of the intact fruit was used, the model for DPPH was calibrated using 5 LVs and the models for ABTS and FRAP were calibrated with 9 LVs (Table 4). The R^2_P was 0.91, 0.83 and 0.85 and the $RSME_P$ was 0.62, 0.47 and 0.74 mmol Trolox L^{-1}. Using the arils information, the models for DPPH and ABTS were calibrated using 11 LVs and the model for FRAP was calibrated using 10 LVs. The R^2_P was 0.92, 0.81 and 0.76 and the $RMSE_P$ was 0.57, 0.52 and 0.89 mmol Trolox L^{-1}.

The RPD values obtained for the DPPH method were 2.5 and 1.6 for colour data, indicating a very good model performance using the intact fruit information and fair model performance using the arils information. Using the spectral data, the RPD values

were 3.4 and 3.6, indicating an excellent model performance using spectral data of the intact fruit and arils. In the case of the ABTS method, the value of RPD using colour data was 1.9 and 1.4, which means that the model using intact fruit had a fair performance and a poor model performance using arils information. When spectral data was used, the RPD values were 1.9 and 2.2, which means that the model has a fair and good model performance using the intact fruit and arils information, respectively.

The values obtained for the FRAP method using colour data were the same as with the ABTS method, but 2.4 and 2.1 using spectral data, indicating a good model performance using the spectral data of the intact fruit and arils.

As in the case of TSS, MI, BrimA, and TPC, the PLS-R models calibrated using the colour data showed a limited potential of AA prediction when the arils information was used.

3.5 Classification according maturity stage

Tables 5 and 6 show the calibration and validation results of the models to discriminate the maturity stage using colour and spectral data of the intact fruit and arils.

The model using the colour data of the intact fruit was calibrated using 3 LVs, obtaining a total accuracy of 85.7 % in the CV. In the prediction of the test set, the total accuracy was 84.3 %, in which 80.0 % of immature fruit, 81.7 % of half-ripe fruit and 92.5 % of ripe fruit were correctly classified. In the case of arils, the model was calibrated using 2 LVs, a total accuracy of 85.7 % being obtained in the CV as when the data of the intact fruit were used. In the prediction of the test set, the total accuracy was 85.7 %, in which 85.0 % of immature fruit, 80.0 % of half-ripe fruit and 95.0 % of ripe fruit were correctly classified.

The model using the spectral data of intact fruit was calibrated using 11 LVs, obtaining a total accuracy of 95.0 % in the CV (Table 6). In the prediction of the test set, the total accuracy was also 95.0 %, in which 100 % of immature fruit, 95.0 % of half-ripe fruit and 90.0 % of ripe fruit were correctly classified. When the arils information was used, the model was calibrated using 9 LVs, obtaining a total accuracy

of 92.9 % in the CV. In the prediction of the test set, the total accuracy was 100 %, in which all fruit were correctly classified into their corresponding classes. These results are similar to those obtained using magnetic resonance imaging (Khoshroo et al., 2009), but for semi-ripe, ripe and over-ripe fruit (100 %, 98.5 % and 100 %, respectively).

Table 5. Results of classification by maturity stage using the colour data of the intact fruit and arils.

	#LV	Class	Calibration			Cross validation			Prediction		
			I	HR	R	I	HR	R	I	HR	R
Intact fruit	3	I	71	11	0	69	12	0	32	6	0
		HR	9	100	3	11	96	5	8	49	3
		R	0	9	77	0	12	75	0	5	37
		CC (%)	88.8	83.3	96.3	86.3	80.0	93.8	80	81.7	92.5
		A (%)		88.6			85.7			84.3	
Arils	2	I	38	10	0	38	11	0	17	5	0
		HR	2	46	3	2	45	3	2	24	1
		R	0	4	37	0	4	37	0	0	19
		CC (%)	95.0	75.0	92.5	95.0	75.0	92.5	85.0	80.0	95.0
		A (%)		86.4			85.7			85.7	

#LV = number of latent variables; I = immature; HR = half ripe; R = ripe; CC = correct classification; A = accuracy

Although the two machine vision techniques discriminated the maturity stages of the intact fruit and arils with good results, hyperspectral imaging was more accurate than colour imaging.

The results of this study have been confirmed only in fruit of the cultivar 'Mollar de Elche', which is a sweet cultivar of high economic importance in Spain. These experiments should therefore be confirmed in other cultivars and with different seasons.

Table 6. Results of classification by maturity stage using the spectral data of the intact fruit and arils.

	#LV	Class	Calibration			Cross validation			Prediction		
			I	HR	R	I	HR	R	I	HR	R
Intact fruit	11	I	80	0	0	80	0	0	40	0	0
		HR	0	116	1	0	110	4	0	57	4
		R	0	4	79	0	10	76	0	3	36
		CC (%)	100	96.7	98.8	100	91.7	95.0	100	95.0	90.0
		A (%)		98.2			95.0			95.0	
Arils	9	I	40	0	0	36	0	0	20	0	0
		HR	0	60	0	3	59	5	0	30	0
		R	0	0	40	0	1	35	0	0	20
		CC (%)	100	100	100	90.0	98.3	87.5	100	100	100
		A (%)		100			92.9			100	

#LV = number of latent variables; I = immature; HR = half ripe; R = ripe; CC = correct classification; A = accuracy

4. Conclusions

In this work, the potential of colour and hyperspectral imaging has been evaluated to monitor the quality of 'Mollar de Elche' intact pomegranate fruit and arils during maturity.

Different maturity stages could be observed in a non-supervised way by means of PCA using the colour and spectral data of the intact fruit and arils. Later, PLS-R models were performed to predict the physicochemical properties of intact fruit and the arils using the colour and the spectral data (750–1050 nm). The physicochemical parameters that were predicted better (RPD > 2) using colour imaging were the maturity MI and BrimA indices, the L* and a* colour coordinates, the AA using the DPPH method and TPC. All of them were performed using the intact fruit information. When hyperspectral imaging was used in the intact fruit, the physicochemical parameters that were predicted better (RPD < 2) were BrimA, a* colour coordinate, the AA using the DPPH and FRAP methods and TPC. For the arils, all physicochemical parameters studied were correctly predicted (RPD > 2) except TA.

PLS-DA models were performed to classify the fruit according to the maturity stage. The models using colour data achieved an accuracy of 84.3 % and 85.7 % for intact fruit and arils, respectively. However, when the spectral data were used, more accurate models were obtained, achieving an accuracy of 95.0 % and 100 %.

These findings demonstrate that colour imaging could be used as a potential tool to monitor some physicochemical properties and maturity of the intact fruit. However, hyperspectral imaging has demonstrated a better and greater potential in both intact fruit and arils.

Acknowledgements

This work was partially funded by INIA and FEDER funds through project RTA2015-00078-00-00. Sandra Munera thanks INIA for the FPI-INIA grant num. 43 (CPR2014-0082), partially supported by European Union FSE funds.

References

Adiani, V., Gupta, S., Ambolikar, R. & Variyar, P.S. (2018). Development of rapid method to assess microbial quality of minimally processed pomegranate arils using FTIR. *Sensors and Actuators B: Chemical, Volume 260*, 800-807.

Arendse, E., Fawole O. A., Magwaza. L.S. & Opara, U. L. (2018a). Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. *Journal of Food Engineering 207*, 11-23.

Arendse, E., Fawole O. A., Magwaza. L.S., Nieuwoudt, H. & Opara, U. L. (2018b). Evaluation of biochemical markers associated with the development of husk scald and the use of diffuse reflectance NIR spectroscopy to predict husk scald in pomegranate fruit. *Scientia Horticulturae 232*, 240-249.

Arendse, E., Fawole O. A., Magwaza. L.S., Nieuwoudt, H. & Opara, U. L. (2018c). Fourier-transform near infrared diffuse reflectance spectroscopy and two spectral acquisition modes for evaluation of external and internal quality of intact pomegranate fruit. *Postharvest Biology and Technology 138*, 91-98.

Arendse, E., Fawole, O.A., Magwaza, L.S., Nieuwoudt, H.H. & Opara, U.L. (2017). Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy combined with chemometrics. *Biosystems Engineering 159*, 22-32.

Arendse, E., Fawole O. A., Magwaza. L.S. & Opara, U.L. (2016a). Non-destructive characterization and volume estimation of pomegranate fruit external and internal morphological fractions using X-ray computed tomography. *Journal of Food Engineering 186*, 42-49.

Arendse, E., Fawole O. A., Magwaza. L.S. & Opara, U.L. (2016b). Estimation of the density of pomegranate fruit and their fractions using X-ray computed tomography calibrated with polymeric materials. *Biosystems Engineering 148*, 148-156

Basha, N., Nounou, M. & Nounou, H. (2018). Multivariate fault detection and classification using interval principal component analysis. *Journal of Computational Science 27*, 1-9.

Blasco, J., Munera, S., Aleixos, N., Cubero, S. & Moltó, E. (2017). Machine vision-based measurement systems for fruit and vegetable quality control in postharvest. In: Bernd Hitzmann (Ed.). Measurement, modeling and automation in advanced food processing (Pp 71-91). Advances in Biochemical Engineering/Biotechnology 161. USA: Springer.

Blasco, J., Cubero, S., Gómez-Sanchís, J., Mira, P. & Moltó, E. (2009). Development of a machine for the automatic sorting of pomegranate (*Punica granatum*) arils based on computer vision. *Journal of Food Engineering 90,* 27–34.

Boussa, F., Zaouay, F., Burlo-Carbonell, F., Nuncio-Jáuregui, N., Gmati, M., El Arbi, B., Melgarejo, P., Hernández, F., Mars, M. (2019). Combined effects of cropping system and harvest date determine quality and nutritional value of pomegranate fruits (*Punica granatum* L. cv. Gabsi). *Scientia Horticulturae 249*, 419-431.

Brand-Williams, W., Cuvelier, M.E. & Berset, C. (1995) .Use of free radical method to evaluate antioxidant activity. *LWT - Food Science and Technology 28*, 25–30.

Castro-Giráldez, M., Fito, P.J., Ortolá, M.D. & Balaguer, N. (2013). Study of pomegranate ripening by dielectric spectroscopy. *Postharvest Biology and Technology 86*, 346–353.

Centre for the Promotion of Imports (CPI). (2018). Exporting fresh pomegranates to Europe. https://www.cbi.eu/ - Accessed 23.01.19

Czieczor, L., Bentkamp, C., Damerow, L. & Blanke, M. (2018). Non-invasive determination of the quality of pomegranate fruit. *Postharvest Biology and Technology 136*, 74-79

Fawole, O.A. & Opara, U.L. (2013). Developmental changes in maturity indices of pomegranate fruit: A descriptive review. *Scientia Horticulturae 159*, 152-161

Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report, Opto- Knowledge Systems Inc. OKSI.

Generalitat Valenciana. Conselleria de Agricultura, medio ambiente, cambio climático y desarrollo rural. Pliego de condiciones de la denominación de origen protegida "Granada Mollar de Elche" / "Granada de Elche". http://www.agroambient.gva.es/ - Accesed 14.02.19

Gil, M.I., García-Viguera, C., Artés, F., Tomás-Barberán, F.A. (1995). Changes in pomegranate juice pigmentation during ripening. *Journal of the Science of Food and Agriculture 68*, 77-81.

Jamshidi, B., Mohajerani, E., Farazmand, H., Mahmoudi, A. & Hemmati, A. (2019). Pattern recognition-based optical technique for non-destructive detection of Ectomyelois ceratoniae infestation in pomegranates during hidden activity of the larvae. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 206*, 552-557

Jolliffe, I. T. (2002). Principal component analysis (2nd ed.). New York: Springer.

Jordan, R., Seelye, R. & McGlone, A. (2001). A sensory-based alternative to brix/acid ratio. *Food Technology 55*, 36–44.

Khodabakhshian, R., Emadi, B., Khojastehpour, M. & Golzarian, M.R. (2017). Determining quality and maturity of pomegranates using multispectral imaging. *Journal of the Saudi Society of Agricultural Sciences16*, 322–331.

Khodabakhshian, R., Emadi, B., Khojastehpour, M. & Golzarian, M.R. (2016). Carob moth, Ectomyelois ceratoniae, detection in pomegranate using visible/near infrared spectroscopy. *Computers and Electronics in Agriculture, Volume 129*, 9-14.

Khoshroo A., Keyhani A., Zoroofi R.A., Rafiee S., Zamani Z. & Alsharif M.R. (2009). Classification of pomegranate fruit using texture analysis of MR images. *Agricultural Engineering International: the CIGR Ejournal* 11, manuscript 1182.

Kulkarni, A.P. & Aradhya S.M. (2005). Chemical changes and antioxidant activity in pomegranate arils during fruit development. *Food Chemistry 93*, 319-324.

Liu, M., Hu, S., Lin, H., & Guo, E. (2008). Hyperspectral laser-induced fluorescence imaging for non-destructive assessing soluble solids content of orange. *IFIP 258*, 51-59.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology 5*, 1121–1142.

Manera, F.J., Legua, P., Melgarejo, P., Brotons, J.M., Hernández, F. & Martínez, J.J. (2013). Determination of a colour index for fruit of pomegranate varietal group "Mollar de Elche". Scientia Horticulturae 150, 360–364.

MAPA 2018, Anuario de estadistica, Ministerio de Agricultura, Pesca y Alimentación 2017. Available on-line at https://www.mapa.gob.es/estadistica/pags/anuario/2017/anuario/AE17.pdf Last visit March 2019.

Melgarejo, P., Hernández, F., Legua, P., 2010. El Granado. Proceedings of I Jornadas Nacionales sobre el Granado: Producción, Economía, Industrialización, Alimentación y Salud. Universidad Miguel Hernández de Elche, Departamento de Producción Vegetal y Microbiología, Alicante, Spain, pp. 8–37.

Melgarejo, P., Martínez-Valero, R., Guillamón, J.M., Miro, M. & Amorós, A. (1997). Phenological stages of the pomegranate tree (*Punica granatum* L.). *Annals of Applied Biology 130*, 135-140.

Mena, P., García-Viguera, C., Navarro-Rico, J., Moreno, D.A., Bartual, J. & Saura, D. (2011). Phytochemical characterisation for industrial use of pomegranate (*Punica granatum* L.) cultivars grown in Spain. *Journal of the Science of Food and Agriculture, 91*, 1893-1906

Mobaraki, N. & Amigo, J.M. (2018). HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems 172*, 174-187.

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys,W., Theron, K. I., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*, 99-118.

Nuncio-Jáuregui, N., Calín-Sánchez, A., Carbonell-Barrachina, A. & Hernández, F. (2014). Changes in quality parameters, proline, antioxidant activity and color of pomegranate (Punica granatum L.) as affected by fruit position within tree, cultivar and ripening stage. *Scientia Horticulturae 165*, 181-189.

Nuncio-Jáuregui, N., Munera-Picazo, S., Calín-Sánchez, A., Wojdyło, A., Hernández, F. & Carbonell-Barrachina, A.A. (2015). Bioactive compound composition of pomegranate fruits removed during thinning. *Journal of Food Composition and Analysis* 37, 11-19.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G.S.V. & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194-200.

Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry 28*, 1201-1222.

Salmanizadeh, F., Nassiri, S.M., Jafari, A. & Bagheri, M.H. (2014). Volume estimation of two local pomegranate fruit (Punica granatum L.) cultivars and their components using non-destructive X-ray computed tomography technique. *International Journal of Food Properties 18*, 439-455.

Singleton, V.L., Orthofer, R. & Lamuela-Raventos, R.M. (1999). Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. *Methods in Enzymology 299*, 152–178.

Shwartz, E., Glazer, I., Bar-Ya'akov, I., Matityahu, I., Bar-Ilan, I., Holland, D. & Amir, R. (2009). Changes in chemical constituents during the maturation and ripening of two commercially important pomegranate accessions. *Food Chemistry 115*, 965-973.

Teerachaichayut, S. & Ho, H.T. (2017). Non-destructive prediction of total soluble solids, titratable acidity and maturity index of limes by near infrared hyperspectral imaging. *Postharvest Biology and Technology 133*, 20-25.

Viuda-Martos, M., Fernández-López, J. & Pérez-Álvarez, J.A. (2010). Pomegranate and its many functional components as related to human health: a review. *Comprehensive Reviews in Food Science and Food Safety 9*, 635–654.

Williams, P.C. (1987). Variables affecting near-infrared reflectance spectroscopic analysis. In: Williams, P., Norris, K. (Eds). Near-infrared Technology in the Agricultural and Food Industries. (pp. 143–166). St. Paul, MN: American Association of Cereal Chemists.

Yang, C. H., Sun, D. W., Pu, H., Wang, N. N., & Zhu, Z. (2015). Rapid detection of anthocyanin content in lychee pericarp during storage using hyperspectral imaging coupled with model fusion. *Postharvest Biology and Technology 103*, 55-65.

Zhang., L. & McCarthy, M.J. (2012). Black heart characterization and detection in pomegranate using NMR relaxometry and MR imaging. *Postharvest Biology and Technology 67*, 96–101.

Zhao, X., Yuan, Z., Yin, Y. & Feng, L. (2015). Patterns of pigment changes in pomegranate (punica granatum l.) peel during fruit ripening. *Acta horticulturae 1089*, 83-89.

IV. LOQUAT

CHAPTER VII

Discrimination of common defects on 'Algerie' loquat fruit using hyperspectral imaging and machine learning techniques

Sandra Munera[a], Juan Gómez-Sanchís[b], Nuria Aleixos[c], Joan Vila-Francés[b], Sergio Cubero[a], Esteban Soler[d], Giancarlo Colelli[e] and José Blasco[a]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera Km 4.5, 46113, Moncada, Valencia, Spain

[b] Departamento de Ingeniería Electrónica. Universitat de València. Av. Universitat, s/n, 46100 Burjassot, Valencia, Spain

[c] Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain

[d] Cooperativa de Callosa d'en Sarrià. Partida Micleta, s/n. 03510 Callosa d'en Sarriá, Alicante, Spain

[e] Departament of Science of Agriculture, Food and Environment, University of Foggia, Via Napoli 25, 71122, Foggia, Italy

Pending submission

Abstract

Loquat fruit (*Eriobotrya japonica* L.) is susceptible to mechanical damages and physiological disorders. VIS-NIR hyperspectral imaging was used to discriminate internal and external common defects of loquat cv. 'Algerie'. Three classifiers, PLS, RF and XGBOOST, and different spectral pre-processing techniques were evaluated in order to discriminate the sound and defect features according to three approaches. In the first approach, the fruit pixels were classified into two classes, sound or defect; in the second approach, the defects were considered as internal or external defects; and in the third approach each type of defect, purple spot, bruise, scars and flesh browning, were considered separately. The results indicated that the hyperspectral imaging combined with XGBOOST classifier could discriminate the pixels between sound and defect with an accuracy of 97.5 % and between sound or internal or external defect with an accuracy of 96.7 %. It was also possible to distinguish between the different defects with an accuracy of 95.9 %.

1. Introduction

Loquat fruit (*Eriobotrya japonica* L.) is native from China, which is the first producing country in the world. Although loquat is a minor crop in Spain, this is the main loquat-producing country in the Mediterranean region and the main exporter in the world (Besada et al., 2017). The production is concentrated on the 'Algerie' cultivar, which accounts for more than 80 % of total production. Furthermore, its interest lies in the fact that loquat trees are harvested during a short period (from mid-April to the end of April), when there is low competition with other fruit on the market (Ballester et al., 2018).

Loquat is a very delicate fruit easily damaged by wind that allows the leaves to scratch the fruit favouring the appearance of rusetting. In most areas of Spain, screens protect the crop in order to avoid wind damage while humidity and temperature is controlled by irrigation (Soler et al., 2007). This fruit is also susceptible to bruising because of mechanical damage during harvest or postharvest handling (Cañete et al., 2015). Regarding to physiological disorders, loquat fruit is highly sensitive to purple spot which is characterised by an extensive area of slightly depressed surface of purple colour and irregular shape that affects up to 30 % of the exposed face of the fruit. This disorder only affects the epidermal fruit tissue and localised fruit calcium deficiency is the most accepted cause (Gariglio et al., 2002). Other disorder is the browning of the flesh due to high temperatures before or after harvest and long periods of storage (Kader, 1999).

Traditionally, quality inspection in packinghouses has been carried out by operators who visually assessed external features of the skin related to the quality standards of the market. As the decisions made by operators are affected by psychological factors such as fatigue or acquired habits, there is a high risk of human error in classification processes, and this is one of the most important drawbacks that can be prevented by automated inspection systems based on computer vision (Cubero et al., 2011). Computer vision methods based on colour cameras have been designed with the intention of emulating the human eye; however, hyperspectral imaging offers the possibility of going far beyond the capabilities of the human eye. For instance, some

damage or defect can be often observed in particular regions outside the visible spectrum, or their detection may be enhanced at certain specific wavelengths (Blasco et al., 2017). Hyperspectral imaging has been previously used in numerous works to detect physical damages or defects in pears (Lee et al., 2014), peaches (Zhang et al., 2015), apples (Zhang et al., 2018), mangoes (Velez-Rivera et al., 2014), oranges (Li et al., 2011) or potatoes (López-Maestresalas et al., 2016). Regarding to loquat fruit, only Yu et al. (2014) have previously used hypespectral imaging to detect other type of defects in loquats cv. 'Luoyangqing'. They used PLS and the mean spectra of each defect to calibrate the models. In the present work, the main objective is to develop predictive models to discriminate the pixels corresponding to purple spot, russeting, bruises and flesh browning of loquat cv. 'Algerie' (Figure 1) by using hyperspectral imaging combined with two robust machine learning techniques, RF and XGBOOST. Furthermore, three approaches were proposed: in the first approach, the samples are classified into two classes, sound or defect; in the second approach the defects are considered as internal or external defects; and in the third approach each type of defect is considered separately.

2. Material and methods

2.1. Fruit samples

In this study, loquat fruit cv. 'Algerie' was obtained from the quality inspection line of the Callosa d'en Sarrià Agricultural Cooperative (Alicante, Spain). A batch of 130 samples composed of fruit without defects and with different defects as flesh browning, bruises, russeting and purple spots was selected (see representative samples in Figure 1).

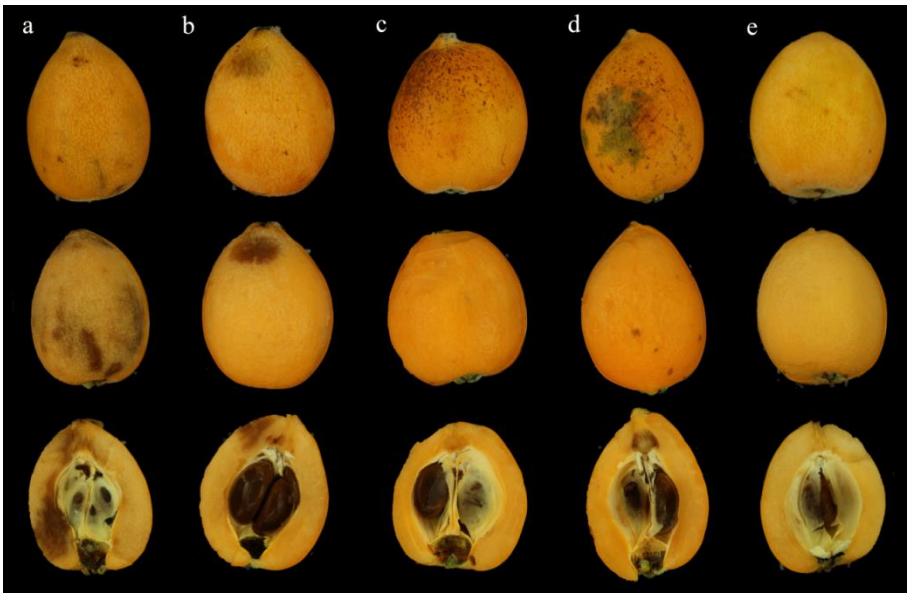


Figure 1. Example of defects in 'Algerie' loquat fruit. Internal defects: flesh browning (a) and bruise (b); external defects: russeting (c) and purple spot (d); sound (e).

2.2 Hyperspectral image acquisition and processing

The system was composed of an industrial camera (CoolSNAP ES, Photometrics, AZ, USA), coupled to two LCTF (Varispec VIS-07 and NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). The camera was configured to acquire images with a size of 1392 x 1040 pixels and a spatial resolution of 0.14 mm/pixel at 60 different wavelengths every 10 nm, in the working spectral range of 450 nm − 1040 nm. In order to avoid problems of unfocused images due to the refraction of light across this wide spectral range, the focus was adjusted on the central band of the acquisition interval (740 nm) and the images were captured using lenses capable of covering the whole spectral range without going out of focus (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA). To optimise the dynamic range of the camera, prevent saturated images and correct the spectral sensitivity of the different elements of the system, a calibration of the integration time of each band was performed capturing the averaged grey level of a white reference target (Spectralon 99%, Labsphere, Inc, NH, USA) corresponding to 90 % of the dynamic range of the camera.

The scene was illuminated by indirect light from twelve halogen spotlights (37 W) (Eurostar IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current

(12 V) and arranged equidistant from each other inside a hemispherical aluminium diffuser. The samples were introduced manually onto a fruit holder. The inner surface of the aluminium diffuser was painted white with a rough texture to maximise its reflectivity in order to minimise directional reflections, and providing highly homogeneous light.

The image processing started with the correction of the relative reflectance by using the equation (1) (Gat, 2000):

$$\rho_{xy}(x,y,\lambda) = \frac{\mathrm{R^{abs}}}{\mathrm{R^{abs}_{white}}} = \rho^{\mathrm{Ref}}(\lambda)\frac{\mathrm{R(x,y,\lambda)-R_{black}(x,y,\lambda)}}{\mathrm{R_{white}(x,y,\lambda)-R_{black}(x,y,\lambda)}} \quad (1)$$

Where $\rho^{Ref}(\lambda)$ is the standard reflectance of the white reference target (99 % in this work), $R(x,y,\lambda)$ is the reflectance of the fruit captured by the sensor of the camera, $R_{white}(x,y,\lambda)$ is the reflectance the white reference target, and $R_{black(}x,y,\lambda)$ is the reflectance captured while avoiding any light source to quantify the electronic noise of the sensor.

The spectral data was manually extracted from the fruit identifying the regions of interest (ROI) of sound and defective features. A total of 22140 pixels were selected (sound = 7733; purple spot = 1738; russeting = 478; bruise = 5871; browning = 6320) and considered as samples.

The correction of the images and the selection of the ROIs to extract the pixel spectra were performed using customised software developed at IVIA (Hyperspectral Image Inspector, freely available at http//www.cofilab.com, Spain).

2.3 Data analysis

The data obtained from sound and defective features was randomly partitioned into the calibration set (15497 pixel samples) and test set (6643 pixel samples). The method used to this end was the Kennard-Stone algorithm (Kennard and Stone, 1969) which allows selecting samples with a uniform distribution over the predictor space using a Euclidean distance.

Three different classifiers were tested, PLS considered as baseline method in this work and two robust machine learning methods, RF and XGBOOST.

All operations were implemented using the 'mlr Package' (Bischl et al., 2016) in R (R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org).

PLS searches for a linear multivariate model of latent variables by projecting prediction variables X and response variables Y into a new latent space, where the covariance between these latent variables is maximised. The goal is to find the latent multidimensional direction in the data space that explains the direction of the maximum multidimensional covariance in the Y space (Lorente et al., 2012). When the Y variable is quantitative, PLS regression is performed (Lorente et al., 2012). When the Y variable is categorical, PLS discriminant analysis is performed in order to sharpen the separation between groups of observations by maximising the covariance between the X and Y, such that a maximum separation among classes is obtained (Lorente et al., 2012).

RF is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalisation error for forests converges to a limit as the number of trees in the forest becomes large. The generalisation error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation among them (Breiman, 2001).

XGBOOST is based on Gradient Boosting (Friedman, 2001) which is an ensemble technique that attempts to create a strong learner from a given number of weak learners, i.e. models that only perform slightly better than random guessing (Schapire, 1999). XGBoost uses a tree ensemble model, which is a set of classification and regression trees (CART). This type of boosting, using trees as base learners is called Tree Boosting. Because of one tree might not be enough to obtain good results, multiple CARTs can be used together and the final prediction will be the sum of each CART's score (Nobre & Neves, 2019).

2.3.1 Optimisation of the classifiers parameters

The selection of the proper combination of the training parameters for each classifier was performed by means of 200 random models using Monte Carlo method. The training parameters tested for PLS were the number of latent variables (from 5 to 10), probability function of the model output (softmax and Bayesian) and method (kernelpls, widekernelpls, simpls and oscorepls). For RF the number of trees allowed in each model (from 500 to 1000) was tested, the maximum size allowed for the node of a tree (from 10 to 100) and the number of randomly chosen variables used to build the trees (from 2 to 20). And the training parameters tested for XGBOOST were the number of times the data are passed to the model during training (from 10 to 80), maximum depth of a tree allowed to control the overfitting (from 1 to 40), minimum number of samples of a node to be a terminal node in order to control the overfitting (from 1 to 20), minimum loss of information to divide a node (from 0.30 to 0.85), constant of pruning to avoid overfitting (from 0.001 to 0.8), degree of randomness in the division of the data set to build the model (0.1 to 0.9) and the ratio of variables chosen to build each tree (from 0.1 to 0.9).

The calibration of the classifiers was obtained by means of a 3-fold cross validation with 3 repetitions in order to evaluate the robustness of the model with the division of the calibration set.

2.3.2 Calibration of the models

After determining the best combination of parameters for each classifier, several pre-processing techniques were applied to the spectra to obtain the best results in the three classification approaches. The techniques used were Standard Normal Variate (SNV), Moving Average + SNV (movav+SNV), Savitzky Golay smoothing + SNV (SG+SNV), first derivative + SNV (1D+SNV), second derivative + SNV (2D+SNV), Gap segment derivative + SNV (GapD+SNV). The raw data (RAW) was also used to build the models.

The calibration of each model with the combination of spectral pre-processing was obtained by means of a 3-fold cross validation with 10 repetitions in order to know their robustness with the division of the calibration set.

At the end of this process, the combination of the best classifier and the best spectral pre-processing technique were available to predict the class of the validation set.

The results of the calibration set and test set were presented in a confusion matrix. Due to in Approach I the classification was binary, the performance of the models was expressed in terms of the area under the receiver operating characteristic curve (AUC) and accuracy. For the Approaches II and III, the performance of the models was expressed in terms of accuracy.

3. Results and discussion

3.1 Spectral analysis

The mean spectrum of the sound and the defective features are presented in Figure 2.

The reflectance of the sound spectrum was higher than the defects such as scar, bruise and purple spot in the region of 530-930 nm as several authors stated previously in peaches (Zhang et al., 2015), oranges (Li et al., 2011), jujube (Wu et al., 2016) and loquat (Yu et al., 2014). As Yu et al. (2014) pointed out, this difference might be attributed to the different tissue structure in pulp and skin cells. However, flesh browning presented a similar pattern to sound feature in this region. In VIS region, carotenoids (500 nm), chlorophylls (680 nm) and other pigments are present and are responsible of colour (Rajkumar et al., 2012). The differences between sound and defective features are mainly due to the degradation of these pigments. In NIR region, the reflection valley around 950-1030 nm, primarily assigned to water absorption (Lu and Peng, 2006) was more pronounced in bruise spectra because more water is free in the flesh due to the rupture of the cells (López-Maestresalas et al., 2016). On the contrary, purple spot, flesh browning and russeting presented more reflectance in this

region due to the difference in the tissue structure, which presented less water content.

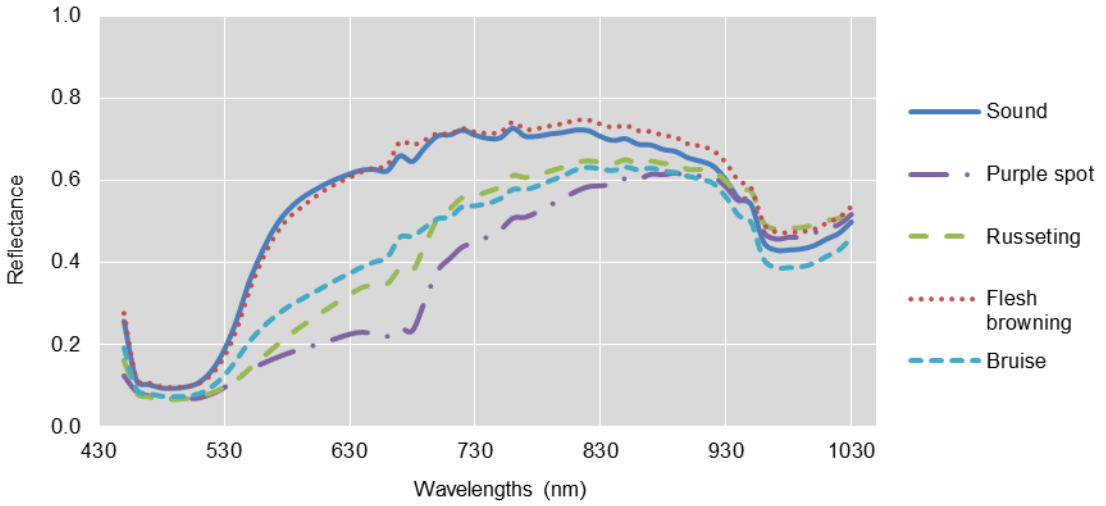


Figure 2. Mean spectra of sound and the defective features

3.2 Classification of sound and defective features

In this study, three approaches were proposed in order to detect sound and defective features. In the first approach the samples were classified into two classes, sound or defect, in the second approach, the samples of defective features were considered as internal or external defects, and in the third approach, each type of defect was considered separately.

3.2.1 Approach I

In this approach PLS, RF and XGBOOST were evaluated to calibrate the classification models using different spectral pre-processing techniques. Figure 3 shows the AUC for the three classifiers and all pre-processing techniques obtained using a 3-fold validation. The variability in the performance of the three classifiers was not very high, being AUC higher than 0.85 in all cases. PLS was the method that presented the lower results of AUC, less than 0.95 in all cases. However, this method

obtained the highest value of the three classifiers when 2D+SNV pre-processing was used.

The two machine learning methods presented similar results being the AUC higher than 0.95 except using D1 and D2 + SNV. However, XGBOOST obtained, in the main cases, the highest values of AUC using different spectral pre-processing techniques. The best result was obtained using the raw data. Thus, the best combination to classify the samples as sound or defect was using XGBOOST and the raw data.

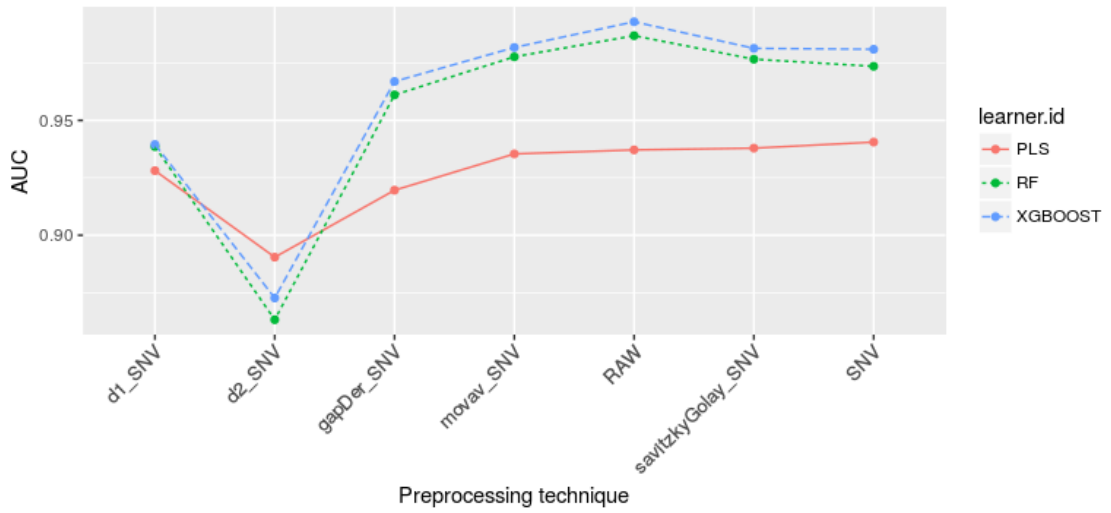


Figure 3. Evolution of AUC vs. pre-processing method for PLS, RF and XGBOOST classifiers in the model calibration of Approach I.

The results of the calibration and test of the model using the combination of XGBOOST and raw data are presented in Table 1. The results of calibration set showed a good performance, 99.9 % of sound samples and 100 % of defect samples were correctly classified. When the test set was introduced in the model, the performance was also good but the accuracy was reduced, 92.0 % of sound samples and 98.9 % of defect samples were correctly classified. Thus, the total accuracy was slightly reduced from 99.9 % to 97.5 %, being the discrimination performance of defect features higher than sound features.

Table 1. Results of calibration and test set for the Approach I using XGBOOST classifier and raw spectra.

Calibration set				
Class	Sound	Defect	Acc (%)	Total Acc (%)
Sound	6339	1	99.9	99.9
Defect	0	9157	100	
Test set				
Class	Sound	Defect	Acc (%)	Total Acc(%)
Sound	1178	102	92.0	97.5
Defect	65	5295	98.8	

Acc = accuracy

3.2.2 Approach II

In order to reduce the time in the optimisation of the parameters of each classifier and according to the previous results obtained by PLS, in this approach only the two machine learning techniques RF and XGBOOST were evaluated.

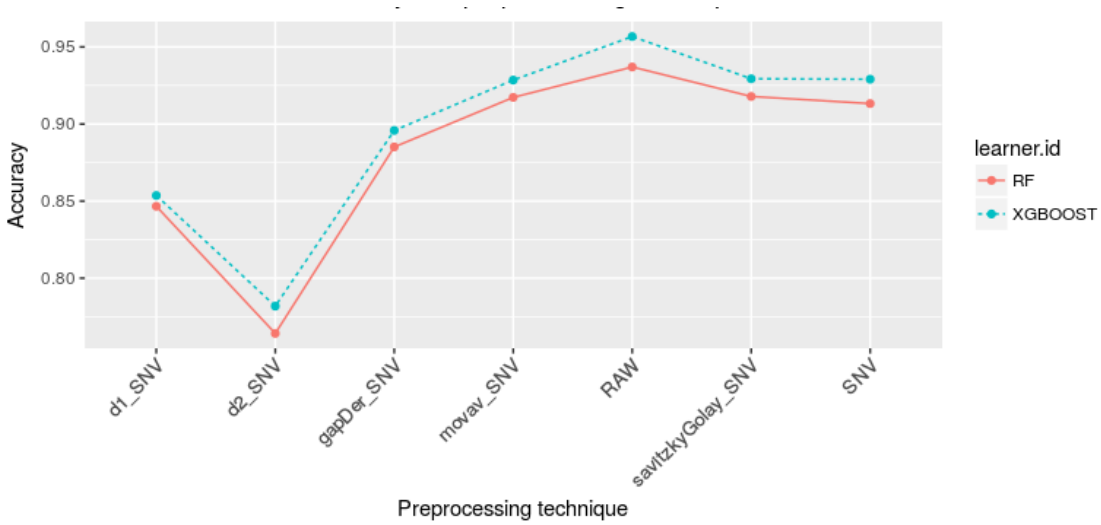


Figure 4. Evolution of mean accuracy vs. preprocessing technique for RF and XGBOOST clasiffiers in the model calibration of Approach II.

Figure 4 shows the accuracy for the two classifiers for each pre-processing technique. Both classifiers continued presenting similar results being the accuracy

higher than 90.0 % except using the derivative pre-processing techniques. However, XGBOOST obtained in all cases the highest accuracy, being the best result obtained using also the raw data.

Thus, the best combination to classify the samples as sound or external or internal defect was also using XGBOOST and the raw data.

Table 2 shows the results of the calibration and validation of the model using the combination of XGBOOST and raw data. The results of calibration set showed also a good performance as Accuracy I, 100 % of sound samples, 99.9 % of external defect samples and 100 % of internal defect samples were correctly classified. When the test set was introduced in the model, the performance was also good but the accuracy was reduced, 95.5 % of sound samples, 93.1 % of external defect samples and 98.0 % of internal defect samples were correctly classified. Thus, the total accuracy was reduced from 99.9 % to 96.7 %.

Table 2. Results of calibration and test set for the Approach II using XGBOOST classifier and raw spectra.

Calibration set						
Class	Sound	External defect	Internal defect	Acc (%)		Total Acc (%)
Sound	6381	0	0	100	100	99.9
External defect	0	1205	1	99.9	99.9	
Internal defect	0	0	7910	100		
Test set						
Class	Sound	External defect	Internal defect	Acc (%)		Total Acc (%)
Sound	1269	0	60	95.5	95.5	96.7
External defect	0	978	72	93.1	97.0	
Internal defect	76	11	4177	98.0		

Acc = accuracy

The discrimination performance of the internal defect pixels was higher than the external defect pixels. The misclassified pixels of these defective features were

considered as internal defect pixels. In the case of sound feature, the misclassified pixels were considered as internal defects.

3.2.3 Approach III

In this approach both machine learning techniques XGBOOST and RF were also tested to calibrate the classification models using different spectral pre-processing techniques.

As figure 5 shows, the two classifiers continued presenting similar results being the accuracy higher than 0.90 except using the derivative pre-processing. However, XGBOOST obtained, one more time, the highest accuracy in all cases, being the best result obtained using also the raw data.

Therefore, the best combination to classify the samples as sound, purple spot, bruise, russeting or flesh browning was also using XGBOOST and the raw data.

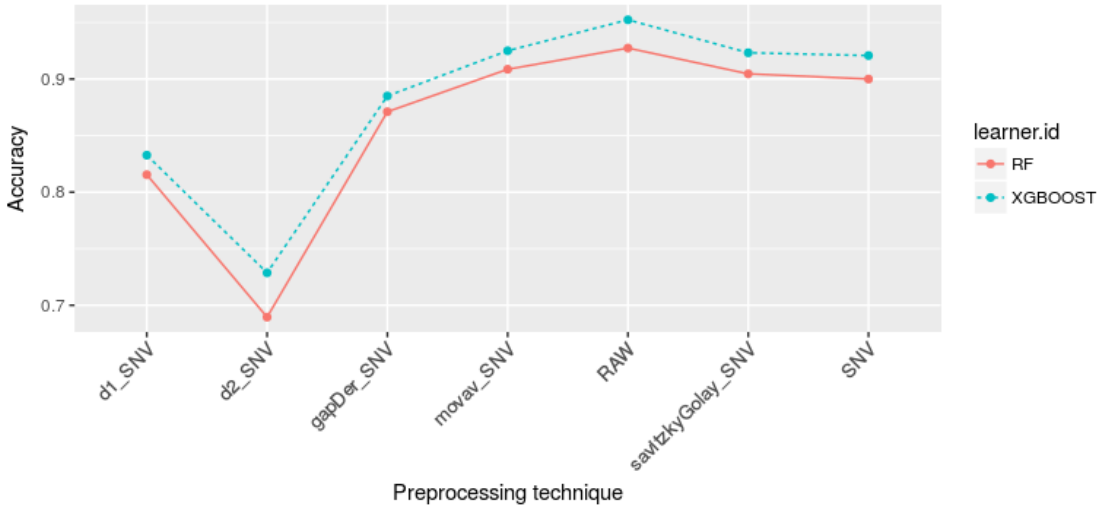


Figure 5. Evolution of mean accuracy vs. preprocessing technique for RF and XGBOOST clasiffiers in the model calibration of Approach III.

.

Table 3 shows the results of the calibration and validation of the model using the combination of XGBOOST and raw data. The results of calibration set showed also a good performance as Accuracy I and II, more than 99.7 % of sound and defective features were correctly classified. When the test set was introduced in the model the

performance was also good but the accuracy was reduced, 95.5 % of sound samples, 93.1 % of external defect samples and 98.0 % of internal defect samples were correctly classified. Thus, the total accuracy was reduced from 99.9 % to 95.9 %.

The discrimination performance of the bruise pixels was higher than the flesh browning pixels. These misclassified features were mainly considered as sound pixels by the model. In the case of the sound feature, the misclassified pixels were considered as internal defects, flesh browning or bruise.

Table 3. Results of calibration and test set for Approach III using XGBOOST classifier and raw spectra.

Calibration set								
Class	Sound	1	2	3	4	Acc (%)		Total Acc (%)
Sound	6433	0	0	1	0	99.9	99.9	
1	0	815	0	0	0	100		
2	0	1	363	0	0	99.7	99.9	99.9
3	0	0	0	2917	0	100		
4	0	0	0	0	4967	100		
Test set								
Class	Sound	1	2	3	4	Acc (%)		Total Acc (%)
Sound	1235	0	0	22	42	95.1	95.1	
1	0	890	4	29	0	96.4		
2	0	2	109	3	0	95.6	96.1	95.9
3	0	62	1	2883	8	97.6		
4	80	0	0	17	1256	92.8		

1 = Purple spot; 2 = russeting; 3 = Bruise; 4 = flesh browning; Acc = accuracy

3.2.4 Approach comparation

In the three approaches, XGBOOST and the raw spectra was the best combination for the classification of sound and defective features. This novel classifier is started to be used in very different fields such as in the detection of pesticides residues in grapes

(Mohite et al., 2017), in pre-diabetes diagnosis (Yang et al., 2019) or in the prediction of the direction of stock market prices (Basak et al., 2019).

The accuracy in the three approaches was high, being correctly discriminated more than 95.7 % of pixels.

The discrimination of defective features was more accurate when the pixels were classified as sound or defect in Approach I, discriminating 98.8 % of defective pixels correctly. On the contrary, when the defective features were classified separately (Approach III), 96.1 % of defective features were correctly discriminated.

In the case of sound features, the best discrimination was obtained in Approach II (95.5 %), when the defective features were discriminated between internal and external defects. Sound features were slightly worse discriminated than defective features in the three approaches. In the second and third approaches, it was possible to know that sound features were misclassified as internal defect and flesh browning specially and vice versa. This fact agrees with the similar pattern spectrum between both features (Figure 2).

As commented before, several studies have been previously carried out using hyperspectral imaging and different approaches in order to detect common defects in fruits. Wu et al. (2016) discriminated common defects on jujube. They also evaluated different pre-processing techniques and the best results were those using also the raw data combined with Soft Independent Modelling of Class Analogy (SIMCA). The percentage of correct classification of intact, cracked, bruised, and insect-infested jujubes was above 95.0 %. Zhang et al. (2015) discriminated common defects on peaches and obtained an accuracy of 93.3 % when sound, artificial defects and non-artificial defects were separated using two characteristic wavelengths at 925 nm and 726 nm. To distinguish the stem from the non-artificial defect regions, other two characteristic wavelengths at 650 nm and 675 nm were used. In the case of loquat, Yu et al. (2014) classified the sound and seven defective features using a PLS model obtaining an accuracy of 95.5 % using 12 optimal wavelengths. Then, all different defects were considered as defective feature and 92.7 % of samples were correctly classified using five minimum noise fraction bands.

The results obtained in this study using the full spectrum were good, pointing the way to perform a proper selection of optimal wavelengths to discriminate the sound and defective features and identifying the different common defects of 'Algerie' loquat in the fruit surface.

4. Conclusions

In this work, hyperspectral imaging combined with machine learning techniques has been evaluated to discriminate common defects of 'Algerie' loquat fruit such as purple spot, bruises, russeting or flesh browning. Three classifiers, PLS, RF and XGBOOST, and different pre-processing techniques were evaluated to discriminate the sound and defective features according to three approaches.

In the three approaches, the best result was obtained using XGBOOST and the data without any pre-processing. In Approach I, the fruit pixels were classified as sound or defect with an accuracy of 97.5 %. In Approach II, the fruit pixels were classified as sound, internal or external defect with an accuracy of 96.7 %. And in Approach III, the fruit pixels were classified as sound or purple spot, scar, bruise or flesh browning with an accuracy of 95.9 %.

These results indicate the potential proposed methodology based on hyperspectral imaging is a promising tool to assess the quality of loquat fruits. However, a proper selection of optimal wavelengths and the identification of the defects in the image of fruits are needed. These are the next steps of this work to study the feasibility of this technique to be implemented in line.

Acknowledgements

This work was partially funded by INIA and FEDER funds through project RTA2015-00078-00-00. Sandra Munera thanks INIA for the FPI-INIA grant num. 43 (CPR2014-0082), partially supported by European Union FSE funds.

References

Ballester, C., Buesa, I., Soler, E., Besada, C., Salvador, A., Bonet, L. & Intrigliolo, D.S. (2018). Postharvest regulated deficit irrigation in early- and intermediate-maturing loquat trees. *Agricultural Water Management 205*, 1-8.

Basak, S., Kar, S., Saha, S., Khaidem, L. & Dey, S.R. (2019). Predicting the direction of stock market prices using tree-based classifiers. The North American Journal of Economics and Finance 47, 552-567.

Besada, C., Sanchez, G., Gil, R., Granell, A. & Salvadora, A. (2017). Volatile metabolite profiling reveals the changes in the volatile compounds of new spontaneously generated loquat cultivars. *Food Research International 100*, 234-243.

Bischl B, Lang M, Kotthoff L, Schiffner J, Richter J, Studerus E, Casalicchio G, Jones Z (2016). "mlr: Machine Learning in R." *Journal of Machine Learning Research 17*, 1-5.

Blasco, J., Munera, S., Aleixos, N., Cubero, S. & Moltó, E. (2017). Machine vision-based measurement systems for fruit and vegetable quality control in postharvest. In: Bernd Hitzmann (Ed.). Measurement, modeling and automation in advanced food processing (Pp 71-91). Advances in Biochemical Engineering/Biotechnology 161. USA: Springer.

Breiman, L. (2001). Random Forests. In: Machine Learning, 45. Pp 5-32. Robert E. Schapire (Ed). Springer US.

Cañete, M.L., Hueso, J., Pinillos V. & Cuevas, J. (2015) Ripening degree at harvest affects bruising susceptibility and fruit sensorial traits of loquat (Eriobotrya japonica Lindl.). *Scientia Horticulturae 187*, 102–107.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. & Blasco, J. (2011). Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables. *Food and Bioprocess Technology 4*, 487-504.

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics 29*, 1189-1232.

Gariglio, N., Juan, M., Castillo, A., Almela, V. & Agustí, M. (2002). Histological and physiological study of purple spot of loquat fruit. Scientia Horticulturae 92, 255-263.

Kader, A. (1999). Loquat. Recommendations for Maintaining Postharvest Quality. http://postharvest.ucdavis.edu/Commodity_Resources/Fact_Sheets/ (Accessed March 20, 2019).

Kennard, R.W. & Stone, L.A. (1969). Computer aided design of experiments. *Technometrics 11*, 137-148.

Lee, W.H., Kim, M.S. Lee, H., Delwiche, S.R. Bae, H., Kim, D.Y. & Cho, B.K. (2014). Hyperspectral near-infrared imaging for the detection of physical damages of pear. *Journal of Food Engineering 130*, 1–7.

Li, J., Rao, X. & Ying, Y. (2011). Detection of common defects on oranges using hyperspectral reflectance imaging. *Computers and Electronics in Agriculture 78*, 38-48.

López-Maestresalas, A., Keresztes, J.C., Goodarzi, M., Arazuri, S., Jarén, C. & Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control 70*, 229-241.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology 5*, 1121–1142.

Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering 93*, 161–171.

Mohite, J., Karale, Y., Pappula, S., Ahammed, S.T.P., Sawant, S.D. & Hingmire, S. (2017). Detection of pesticide (Cyantraniliprole) residue on grapes using hyperspectral sensing. Proceedings of SPIE - The International Society for Optical Engineering 10217, Article number 102170P.

Nobre, J. & Neves, R.F. (2019). Combining Principal Component Analysis, Discrete Wavelet Transform and XGBoost to trade in the financial markets. *Expert Systems with Applications 125*, 181-194.

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G.S.V. & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering 108*, 194–200.

Schapire, R.E. (1999). A brief introduction to boosting. *Proceedings of the sixteenth international joint conference on artificial intelligence 2*, 1401-1406.

Soler, E., Martínez-Calvo, J., Llácer, G. & Badenes, M.L. (2007). Loquat in spain: production and marketing. *Acta Horticulturae 750*, 45-47.

Vélez-Rivera, N., Gómez-Sanchís, J., Chanona-Pérez, J., Carrasco, J.J., Millán-Giraldo, M., Lorente, D., Cubero, S. & Blasco, J. (2014). Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. *Biosystems Engineering 122*, 91-98.

Yang, X., Fang, T., Li, Y., Guo, L., Li, F., Huang, F. & Li. L. (2019). Pre-diabetes diagnosis based on ATR-FTIR spectroscopy combined with CART and XGBoots. *Optik 180*, 189-198.

Yu, K.Q., Zhao, Y.R., Liu, Z.Y., Li, X.L., Liu, F. & He, Y. (2014). Application of Visible and Near-Infrared Hyperspectral Imaging for Detection of Defective Features in Loquat. *Food and Bioprocess Technology 7*, 3077–3087.

Zhang, B., Liu, L., Gu, B., Zhou, J., Huang, J. & Tian, G. (2018). From hyperspectral imaging to multispectral imaging: Portability and stability of HIS-MIS algorithms for common defect detection. *Postharvest Biology and Technology 137*, 95–105.

Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C. & Huang, D. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (*Prunus persica*). *Computers and Electronics in Agriculture 114*, 14–24.

GENERAL DISCUSSION

This doctoral thesis presents different off-line and laboratory scale studies with the aim to evaluate the capability of hyperspectral imaging for the non-destructive monitoring of fruit quality in postharvest.

The first part, which corresponds to the **Chapters I, II and III**, studied the application of hyperspectral imaging in the assessment of different cultivars of **nectarine**. Previous to this thesis, several studies were performed with the aim of applying hyperspectral imaging in the analysis of stone fruits, detecting different types of defects in skin (Zhang et al., 2015; Li et al., 2016; Huang et al., 2015) or chilling injury (Pan et al., 2016; Sun et al., 2017). Apart from defects, few works were carried out to estimate other properties. Lu and Peng (2006) presented one of the first works to detect firmness in peaches using hyperspectral scattering and later, Lleó et al. (2011) classified peaches by maturity using multispectral indices.

In **Chapter I** and **II**, the cultivars 'Big Top' and 'Magique' were studied to monitor their ripeness and quality using reflectance and transmittance mode. Two indices, RPI and IQI, obtained by measuring the physicochemical properties destructively were predicted using PLS-R models. In the case of the reflectance mode, the R^2 values were 0.87 and 0.89 for RPI and IQI for 'Big Top', while for the 'Magique' cultivar these values were 0.91 and 0.89, respectively. Using transmittance mode only IQI was predicted and the R^2 values were 0.89 and 0.86 for 'Big Top' and 'Magique'. In both modes a selection of optimal wavelengths was performed and the results were similar to those using all wavelengths. However, each cultivar needed a particular set of wavelengths.

In transmittance mode (**Chapter II**), fruit was also classified by an F threshold (35 N), which indicates changes during postharvest ripening and the susceptibility to damage by bruising (Crisosto et al., 2001). As a result, around 95.0% of fruit of the two cultivars were correctly classified. Furthermore, the use of this mode, gave the possibility to detect split pit defect, which can be a big problem in nectarines because it can affect 45.0% of the fruit, depending on the cultivar and the season (IRTA, 2016). The accuracy of the PLS-DA models using all and the optimal wavelengths was higher than 90.0 %.

Therefore, both modes had a great potential to obtain and estimate the stage of ripeness of nectarines. The selection of one or the other mode would therefore

depend on the application. Taking into account that split pit can only be detected by transmittance.

Other application of hyperspectral imaging investigated in this doctoral thesis was the discrimination of nectarine cultivars with similar appearance but different taste (**Chapter III**). Previous studies were conducted to differentiate among nectarine cultivars using colour images (Font et al., 2014), but with clearly different appearance. In this case 'Big Top' (sweet) and 'Diamond Ray' (acid) cultivars were used due to their similar skin and flesh appearance. Hyperspectral imaging was compared with colour imaging and a trained panel, which achieved an accuracy of only 56.9 % and 54.5 %. The classification of these two cultivars by hyperspectral imaging was performed using two approaches based o the use of the individual spectrum of each pixel and on the use of mean spectrum of each fruit. In both cases the results of the PLS-DA models were better than the colour imaging and the trained panel, being the accuracy of mean spectrum approach higher than the individual spectrum, 94.4 % and 84.4 %. Furthermore, the use of the vector of regression coefficients of the PLS-DA model let to select 14 optimal wavelengths and to obtain similar results.

The second part, which corresponds to the **Chapters IV and V**, studies the application of hyperspectral imaging in the assessment of **persimmon** cv. 'Rojo Brillante'. In the last twenty years, the production of this fruit in Spain has increased from 33 to 310 thousand tons (FAOSTAT, 2016), due to the development of the de-astringency methods based on high CO_2 concentrations. These methods allow removing the astringency while preserving high flesh firmness (Arnal and Del Río, 2003). Nowadays 'Rojo Brillante' persimmon is one of the most appreciated persimmon cultivars worldwide. First, in **Chapter IV**, three maturity stages of the fruit were evaluated. Models using LDA, QDA and SVM were performed to discriminate the three stages. All of them achieved a good classification above 98.0 % using all wavelengths. Using only three wavelengths selected by PCA the success rate of correct classification was slightly lower but QDA and SVM still remain above 95.0 %. At the same time, F was evaluated by hyperspectral imaging obtaining a R^2 of 0.80 using the previous three wavelengths. This result was similar to those obtained using different parameters of the skin colour like H (R^2 = 0.83), G (R^2 = 0.82) and h (R^2 = 0.81), but also

using ratios like a/b (R^2 = 0.83), G/R (R^2 = 0.83), a/L (R^2 = 0.83) and CI (R^2 = 0.80) which indicates the feasibility of colour imaging to assess the firmness of this fruit.

Regarding to the astringency, several studies have been conducted to predict the content of ST or to assess the astringency in different varieties of persimmon fruit using spectroscopy (Zhang et al., 2013; Noypitak et al., 2014; Altieri et al., 2017; Cortés et al., 2017). In this doctoral thesis, two studies were performed to detect A fruit using hyperspectral imaging since the presence of any astringency in the fruit can cause rejection by the consumer that will in turn affect future sales. In **Chapters IV** and **V**, the fruit used presented different ripeness and was treated in closed containers at 20º C with 90 % of RH and 95.0 % of CO_2 during 12h and 24h, and also not treated. In the first study the fruits were classified using LDA, QDA and SVM as high astringency, medium astringency and deastringed corresponding to the time of the treatment. As a result 95.0 % of fruits were correctly classified using QDA. In the second study, the classification of the fruit was performed to discriminate between A and DA using 0.04 % of ST as threshold. Furthermore the prediction of ST was performed and the most appropriate area of the fruit and the optimal wavelengths were determined. In this case, the results obtained indicated that the PLS-R model using the spectra of the apex area was the most accurate, R^2 of 0.71. However, only 68.7 % of fruit were correctly classified when the threshold of 0.04 % was applied. When the PLS-DA models were performed, the most accurate models were those using middle and apex area spectra using all wavelengths (88.9 % and 87.9 %) and also using 23 optimal wavelengths (86.9 % and 85.9 %).

In the third part, the application of hyperspectral imaging in the assessment of pomegranate cv. 'Mollar de Elche' was studied. This cultivar is appreciated due to its good source of bioactive compounds and its sweet taste and soft seed. Previously, others studies have used other techniques such as spectroscopy (Arendse et al., 2017; Arendse et al., 2018), X-rays (Salmanizadeh et al., 2014) or MNR (Zhang and McCarthy, 2012) to assess the quality of this fruit using the intact fruit or the arils information of other cultivars. In **Chapter VI**, the prediction of physicochemical properties and the maturity stage of this fruit using the intact fruit and arils information were performed by means of two machine vision techniques, colour and hyperspectral imaging. The

images of each intact fruit and their arils and the physicochemical analyses were performed at seven different harvests. PLS-R and PLS-DA models were calibrated to predict the physicochemical properties and maturity of the fruit using the colour and spectral data of intact fruit and arils. The physicochemical parameters better predicted using colour imaging and intact fruit information were the MI and BrimA indices, the L* and a* colour coordinates, the AA using DPPH method and the TPC. When hyperspectral imaging was used, the physicochemical parameters better predicted were BrimA, a* colour coordinate, the AA using DPPH and FRAP methods and the TPC. When the arils information was used, all physicochemical parameters studied were correctly predicted (RPD > 2) except TA. Then, PLS-DA models were carried out to classify the fruit according three maturity stages delimited by the external changes and the moment of the commercial harvest. The models using colour data achieved an accuracy of 84.3 % and 85.7 % for intact fruit and arils, respectively. However, when the spectral data was used, more accurately models were obtained, achieving an accuracy of 95.0 % and 100 %. These results demonstrate that colour imaging can be used as interesting tool to monitor some physicochemical properties and maturity of the intact fruits. However, hyperspectral imaging demonstrated a great potential in both intact fruit and arils.

In the last part, the application of hyperspectral imaging in the assessment of loquat cv. 'Algerie' is studied. Despite this is an important cultivar in Spain, it is very sensitive to mechanical damages and physiologycal dirorders such us purple spot or flesh browning. When those defects are unavoidable and reach the quality inspection lines, inspection systems based on computer vision are needed in order to avoid the human error in classification processes (Cubero et al., 2011). In **Chapter VII**, hyperspectral imaging combined with machine learning techniques was used to discriminate common defects of 'Algerie' loquat fruit such as purple spot, bruises, russeting or flesh browning. Three classifiers, PLS, RF and XGBOOST, and different pre-processing techniques were evaluated to discriminate the sound and defective features according to three approaches. In the first approach, the best result was obtained using XGBOOST and the data without any pre-processing. The fruit pixels were classified as sound or defect with an accuracy of 97.5 %. In the second approach,

the best result was also obtained using XGBOOST and the data without any pre-processing. The fruit pixels were classified as sound, internal or external defect with an accuracy of 96.7 %. In the third approach, the best result was also obtained using XGBOOST and the data without any pre-processing. The fruit pixels were classified as sound or purple spot, scar, bruise or flesh browning with an accuracy of 95.9 %.

References

Altieri, G., Genovese, F., Tauriello, A. & Di Renzo, G.C. (2017). Models to improve the non-destructive analysis of persimmon fruit properties by VIS/NIR spectrometry. *Journal of the Science of Food and Agriculture 97*, 5302-5310.

Arendse, E., Fawole O. A., Magwaza. L.S., Nieuwoudt, H. & Opara, U. L. (2018). Fourier-transform near infrared diffuse reflectance spectroscopy and two spectral acquisition modes for evaluation of external and internal quality of intact pomegranate fruit. Postharvest Biology and Technology 138, 91-98.

Arendse, E., Fawole, O.A., Magwaza, L.S., Nieuwoudt, H.H. & Opara, U.L. (2017). Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy combined with chemometrics. Biosystems Engineering 159, 22-32.

Arnal, L. & Del Río, M.A. (2003). Removing astringency by carbon dioxide and nitrogen-enriched atmospheres in persimmon fruit cv. 'Rojo brillante'. *Journal of Food Science 68*, 1516-1518.

Cortés, V., Rodríguez, A. Blasco, J., Rey, B., Besada, C., Cubero, S., Salvador, A., Talens, P. & Aleixos, N. (2017). Prediction of the level of astringency in persimmon using visible and near-infrared spectroscopy. *Journal of Food Engineering 204*, 27-37.

Crisosto, C.H., Slaughter, D., Garner, D., & Boyd J. (2001). Stone fruit critical bruising thresholds. *Journal of The American Pomological Society 55*, 76-81.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. & Blasco, J. (2011). Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables. *Food and Bioprocess Technology 4*, 487-504.

FAOSTAT (2016). http://www.fao.org/faostat/en/#data/QC - Accessed 02.11.18

Font, D., Tresanchez, M., Pallejà, T., Teixidó, M., Martinez, D., Moreno, J., Palacín, J. (2014). An image processing method for in-line nectarine variety verification based on the comparison of skin feature histogram vectors. *Compututers and Electronics in Agriculture 102*, 112-119.

Huang, F., Zhang, S., Yang, Y., Man, Z., Zhang, X. & Wu, Y. (2015). Application of hyperspectral imaging for detection of defective features in nectarine fruit. *Transactions of the Chinese Society for Agricultural Machinery 11,* 252-259.

Institut de Recerca i Tecnologia Agroalimentàries (IRTA). (2016). XX Exposición de variedades de melocotón y nectarina. http://kp.eufrin.eu/ Accessed 03/11/18

Li, J., Chen, L., Huang, W., Wang, O., Zhang, B., Tian, X., Fan, S. & Li, B. (2016). Multispectral detection of skin defects of bi-colored peaches based on VISeNIR hyperspectral imaging. *Postharvest Biology and Technology 112*, 121-133.

Lleó, L., Roger, J.M., Herrero-Langreo, A., Diezma-Iglesias, B. & Barreiro, P. (2011). Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening. *Journal of Food Engineering 104*, 612-620.

Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering 93*, 161-171.

Noypitak, S., Terdwongworakul, A., Krisanapook, K., Kasemsumran, S., (2015). Evaluation of astringency and tannin content in 'Xichu' persimmons using near infrared spectroscopy. *International Journal of Food Properties 18*, 1014–1028.

Pan, L., Zhang, Q., Zhang, W., Sun, Y., Hua, P. & Tu, K. (2016). Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chemistry 192*, 134-141.

Salmanizadeh, F., Nassiri, S.M., Jafari, A. & Bagheri, M.H. (2014). Volume estimation of two local pomegranate fruit (Punica granatum L.) cultivars and their components using non-destructive X-ray computed tomography technique. *International Journal of Food Properties 18*, 439-455.

Sun, Y., Gu, X., Sun, K., Hu, H., Xu, M., Wang, Z., Tu, K. & Pan, L. (2017). Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches. *LWT - Food Science and Technology 75*, 557-564.

Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C. & Huang, D. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (*Prunus persica*). *Computers and Electronics in Agriculture 114*, 14-24.

Zhang, L. & McCarthy, M.J. (2012). Black heart characterization and detection in pomegranate using NMR relaxometry and MR imaging. *Postharvest Biology and Technology 67*, 96–101.

Zhang, P., Xue, Y., Li, J., Feng, X. & Wang, B. (2013). Research on non-destructive measurement of firmness and soluble tannin content of 'mopanshi' persimmon using Vis/NIR diffuse reflection spectroscopy. *Acta Horticulturae 996*, 447-452.

CONCLUSIONS

Considering the planned objectives and the results obtained in this thesis, the following conclusions can be drawn:

1. Hyperspectral imaging combined with chemometrics is capable to monitor the ripeness of two cultivars of nectarine using ripening indices. PLS-R models produced optimal prediction for both cultivars of around R^2 0.90 of RPI and IQI indices. A total of eight wavelengths were selected for 'Big Top' and seven for 'Magique' using VIP scores. The simplified models also yielded good performance in prediction with R^2 values of around 0.90 for both indices and both cultivars. The ripeness distribution maps facilitated the visual observation of the state of fruit ripening.

2. Hyperspectral transmittance imaging may be a potential non-destructive method to detect split pit nectarines and to monitor their ripeness. The detection of split pit fruits of the 'Big Top' cultivar using PLS-DA was successful, achieving 100 % correct classification for split pit fruit and 91.3 % for normal pit using all the captured wavelengths. The ripeness of the 'Big Top' and 'Magique' cultivars was determined by two indicators: IQI and an F threshold (35 N). The prediction of the IQI was performed by means of PLS-R models, obtaining an R^2 of 0.89 and 0.86 and an RPD of 2.7 and 2.6 for the 'Big Top' and 'Magique' cultivars. The results achieved from estimating the IQI were similar to those obtained in a previous study using the reflectance mode. The classification of the fruits by F was performed by PLS-DA, which correctly classified 95.7 % of the 'Big Top' fruits and 94.5 % of the 'Magique' fruits. An optimal wavelength selection was performed by means of forward i-PLS and the simplified models obtained similar results to those models that used all the wavelengths. A hierarchical model was built to evaluate the total internal quality of the 'Big Top' cultivar and the visualized results indicated that 10.3 % of 'ready to buy' fruits were classified as split pit and 6.9 % as 'hard'.

3. Hyperspectral imaging may have potential as a tool for rapid and non-destructive cultivar discrimination, allowing the selection of fruit that is better suited to the consumer's preferences. The classification of two cultivars of nectarine by colour

imaging or by a trained panel was very poor. However, hyperspectral imaging supported by chemometrics and optimised by reduction of the spectral and spatial information enabled classification more accurately. The use of the mean spectrum of the fruit as input of the predictive models provided classification accuracy of 94.4 %. To cope with the huge amount of data captured by the hyperspectral systems, the vector of the regression coefficients of a PLS-DA model identified 14 wavelengths which were selected as optimal, producing the best classification model with a classification accuracy of 96.3 %.

4. The potential proposed methodology based on hyperspectral imaging is a promising non-destructive tool to assess the internal quality of persimmon fruits destined to be deastringed and rapidly marketed as fresh sweet fruit. The characterisation of the colour showed that good correlations (R^2 > 0.80) were found in some colour parameters like H, G and h, but also using ratios like a/b, G/R and a/L which indicates the feasibility of images to assess the colour as a valid alternative to traditional and expensive colorimeters. Using hyperspectral imaging, three wavelengths (580, 680 and 1050 nm) were proposed as the optimum wavelengths for the classification of the fruits into three ripeness stages with high accuracy, more than 94 % of all samples were well classified for all of the used classifiers (LDA, QDA and SVM). These wavelengths were used for flesh firmness prediction and the RPD value indicated that the obtained model is useful for good quantitative application. Regarding the astringency, the whole spectrum was needed to be used to classify the fruits into three levels of astringency or time of treatment: astringent fruit (0h), fruit with a low-medium level of astringency (12 h) and non-astringent fruit (24h). The overall classification for the three ripeness stages was higher than 95 % for QDA classifier.

5. Hyperspectral imaging combined with multivariate analysis has a great potential as a tool for rapid and non-destructive control of effectiveness of the astringency removal treatment applied to persimmon ‘Rojo Brillante’. The prediction of ST content in the fruits was performed using PLS-R models to determine the astringency. The results obtained indicated that the model using the spectra of the apex area was the most

accurate. However, few fruit were correctly classified when the threshold of 0.04 % was applied. This means that changes in colour, F and other properties have to be considered to discriminate them, and not only the difference in ST content. PLS-DA models were performed to maximise the separation between A and DA classes. The most accurate models were those performed using middle and apex area spectra (88.9 % and 87.9 %). To reduce the huge amount of data captured, the vector of the regression coefficients of the PLS-DA model of each area was used to identify the optimal wavelengths. As when using all wavelengths, the most accurate models were those involving the middle and apex areas and 23 optimal wavelengths (86.9 % and 85.9 %).

6. Hyperspectral imaging has demonstrated a great potential in the quality monitoring of the intact 'Mollar de Elche' pomegranate fruit and arils. The physicochemical parameters better predicted (RPD > 2) using PLS-R and colour imaging were the MI and BrimA indices, the L* and a* colour coordinates, the AA using DPPH method and the TPC. All of them were performed using the intact fruit information. When hyperspectral imaging was used in the intact fruit, the physicochemical parameters better predicted (RPD < 2) were BrimA, a*, the AA using DPPH and FRAP methods and the TPC. When the arils information was used, all physicochemical parameters studied were correctly predicted (RPD > 2) except TA. PLS-DA models were performed to classify the fruit according maturity stage. The models using colour data achieved an accuracy of 84.3 % and 85.7 % for intact fruit and arils, respectively. However, when the spectral data was used, more accurately models were obtained, achieving an accuracy of 95.0 % and 100 %.

7. The potential proposed methodology based on hyperspectral imaging and machine learning techniques is a promising tool to detect common defects in 'Algerie' loquat fruit such as purple spot, bruises, russeting or flesh browning. Three classifiers, PLS, RF and XGBOOST, and different pre-processing techniques were evaluated to discriminate the sound and defective features according to three approaches. In the three approaches, the best result was obtained using XGBOOST and the data without any

pre-processing. In Approach I, the fruit pixels were classified as sound or defect with an accuracy of 97.5 %. In Approach II, the fruit pixels were classified as sound, internal or external defect with an accuracy of 96.7 %. And in Approach III, the fruit pixels were classified as sound or purple spot, scar, bruise or flesh browning with an accuracy of 95.9 %. A proper selection of optimal wavelengths and the identification of the defects in the image of fruits are needed. These are the next steps of this work to study the feasibility of this technique to be implemented in real time.