



Determination of pectin content in orange peels by near infrared hyperspectral imaging

Amanda Teixeira Badaró^a, Juan Francisco Garcia-Martin^b, María del Carmen López-Barrera^b, Douglas Fernandes Barbin^{a,*}, Paloma Alvarez-Mateos^b

^a Department of Food Engineering, University of Campinas (UNICAMP), Rua Monteiro Lobato, 80, Cidade Universitária, Campinas-SP 13083-862, Brazil

^b Departamento de Ingeniería Química, Facultad de Química, Universidad de Sevilla, Sevilla 41012, Spain

ARTICLE INFO

Keywords:

Principal component analysis
Linear discriminant analysis
Partial least squares regression, near infrared spectra
Agriculture

ABSTRACT

Pectin has several purposes in the food and pharmaceutical industry making its quantification important for further extraction. Current techniques for pectin quantification require its extraction using chemicals and producing residues. Determination of pectin content in orange peels was investigated using near infrared hyperspectral imaging (NIR-HSI). Hyperspectral images from orange peel (140 samples) with different amounts of pectin were acquired in the range of 900–2500 nm, and the spectra was used for calibration models using multivariate statistical analyses. Principal component analysis (PCA) and linear discriminant analysis (LDA) showed better results considering three groups: low (0–5%), intermediate (10–40%) and high (50–100%) pectin content. Partial least squares regression (PLSR) models based on full spectra showed higher precision ($R^2 > 0.93$) than those based on few selected wavelengths (R^2 between 0.92 and 0.94). The results demonstrate the potential of NIR-HSI to quantify pectin content in orange peels, providing a valuable technique for orange producers and processing industries.

1. Introduction

Citrus are among the most cultivated crops around the world since they are beneficial for human consumption due to their nutritional and antioxidant properties. As a result, the citrus industry is responsible for millions of jobs (García-Martín, Olmo, & García, 2018; García, Olmo, & García, 2016). Spain is the first producing country of oranges and juices in the European Union and the fifth in the world. This agricultural activity is mainly located in Valencia, producing 3 million tons of citrus per year, of which, according to the Spanish Ministry of Agriculture, Fishing and Food (2017) (Ministerio de Agricultura, 2019), 1.5 million are oranges. The orange juice sector generates about 1.2 million tons waste per year in Spain. This waste is mainly used for animal feed (which is energy costly) or is ploughed into landfills. However, this organic waste has great potential for the food industry as source of many important and valuable compounds (Lessa, Gularte, Garcia, & Fajardo, 2017).

Orange waste can include peel, pulp, seeds, leaves and oranges without the quality requirements. This waste contains, as average, 16.9% (wt.) soluble sugars, 9.21% (wt.) cellulose, 10.5% (wt.)

hemicelluloses, and 42.5% (wt.) pectin, which is its most important component. The soluble sugars present in orange peel are glucose, fructose, and sucrose. The insoluble polysaccharides of the cellular wall of the orange peel are composed of pectin, cellulose, and hemicelluloses. Pectin and hemicelluloses are rich in galacturonic acid, arabinose, and galactose, and contain small amounts of xylose, rhamnose and glucose (Grohmann, Cameron, & Buslig, 1995; Rezzadori, Benedetti, & Amante, 2012). Pectin is composed of a main chain, which has a linear structure of α -D-galacturonic acids linked by α -(1,4)-glycosidic bonds, and a side chain that mainly contains neutral sugars (Liu, Fishman, Kost, & Hicks, 2003).

Pectin has several purposes in the food and pharmaceutical industry, as it is used as thickener, stabilizer, gelling agent, emulsifier and drug delivery vehicle (Wicker, Kim, Thirkield, Lin & Jung, 2014). Hence, the extraction of pectin from orange waste is very important and requires an economically feasible alternative. Among the different pectin extraction methods (dialysis, metal precipitation, membrane separation, etc.), alcohol precipitation is the most widely used methods despite its high energy and ethanol requirements (Ren et al., 2019). However, it is also important to quantify the amount of pectin in orange

* Corresponding author.

E-mail addresses: amandatbadaro@gmail.com (A.T. Badaró), jfgarmar@us.es (J.F. Garcia-Martin), mlopez91@us.es (M.d.C. López-Barrera), dffbarbin@unicamp.br (D.F. Barbin), palvarez@us.es (P. Alvarez-Mateos).

<https://doi.org/10.1016/j.foodchem.2020.126861>

Received 20 December 2019; Received in revised form 3 April 2020; Accepted 17 April 2020

Available online 18 April 2020

0308-8146/ © 2020 Elsevier Ltd. All rights reserved.

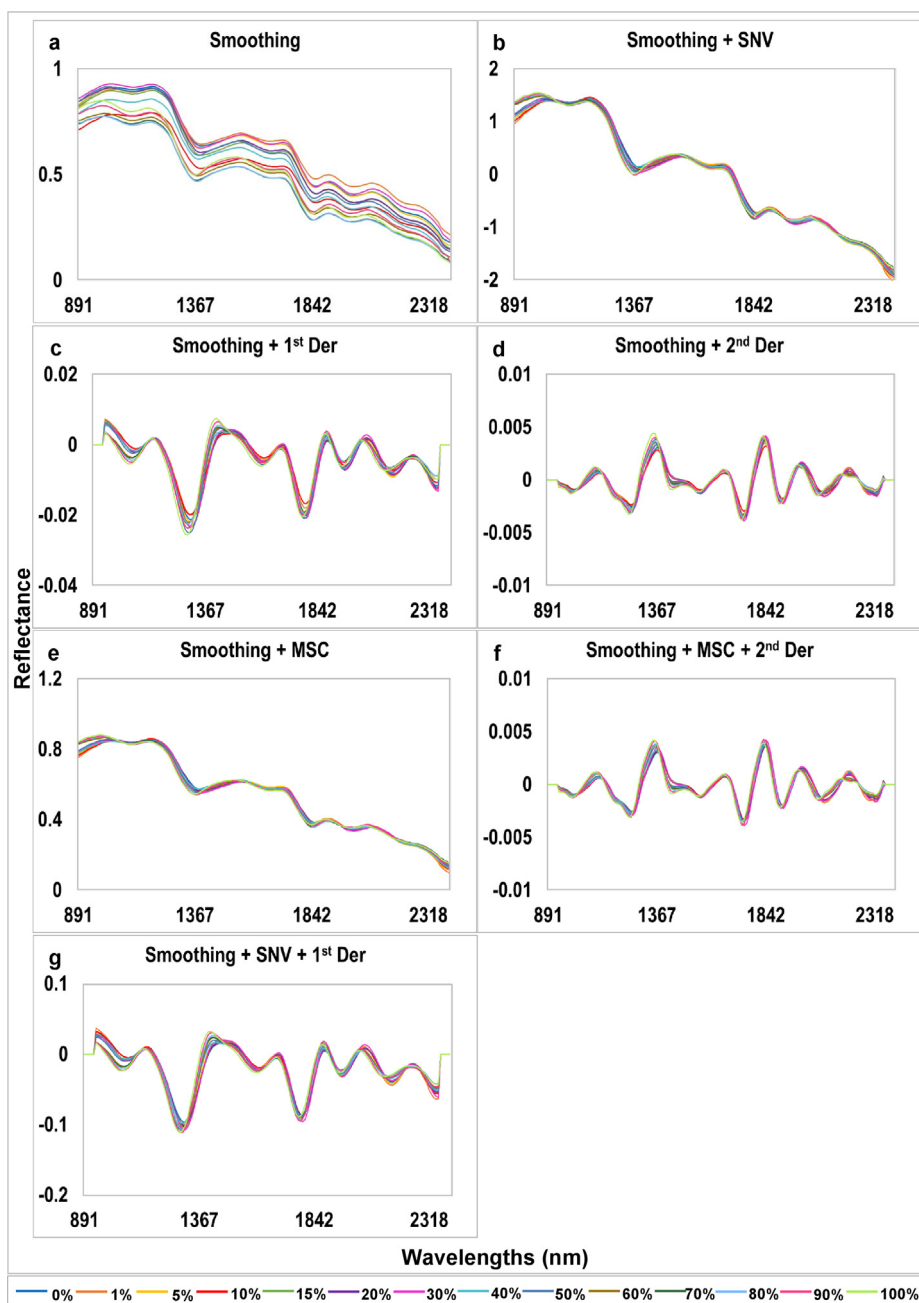


Fig. 1. Spectra of orange peels and pectin: a) Smoothing, b) Smoothing + SNV, c) Smoothing + 1st derivative, d) Smoothing + 2nd derivative, e) Smoothing + MSC, f) Smoothing + MSC + 2nd derivative, g) Smoothing + SNV + 1st derivative. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

residue, in order to justify investments in the waste processing and extraction methods. The current techniques for pectin quantification requires the previous pectin extraction. In this sense, the commonly-used sulphuric acid-carbazole colorimetric method for pectin content measuring in citrus involves pectin extraction, pectin depolymerisation, reaction of pectin with carbazole and absorbance measurement at approximately 525 nm (Wang, Chuang, & Hsu, 2008).

Some works have been reported using vibrational techniques, more specifically near infrared spectroscopy and Fourier transform near infrared spectroscopy, to determine some parameters in orange, including orange firmness, peel thickness and total pectin content (Bizzani, Flores, Colnago, & Ferreira, 2017; Chen et al., 2018). Although this techniques have achieved good results for many parameters, the fact that it only detects a single point of the sample at a time can be a barrier

for complex samples (Feng & Sun, 2012).

In this context, near infrared hyperspectral imaging (NIR-HSI), as an emerging technology, offers numerous advantages over conventional analysis methods and can overcome some problems faced by near infrared spectroscopy. Combined to chemometrics, NIR-HSI has achieved very precise results for many compounds and properties in food, as colour and pH in meat (Elmasry, Sun, & Allen, 2012), fibre in flour (Badaró, Morimitsu, Ferreira, Clerici, & Barbin, 2019) and melamine in milk powder (Forchetti & Poppi, 2017). As the hypercube data are massive, chemometrics is used to reduce the high dimensionality to the most meaningful dimension (data simplification) without compromising the information contained into the original image (Feng, Makino, Oshita, & Martín, 2018). Hence, HSI could provide a distribution map of pectin in the orange waste, as other authors have

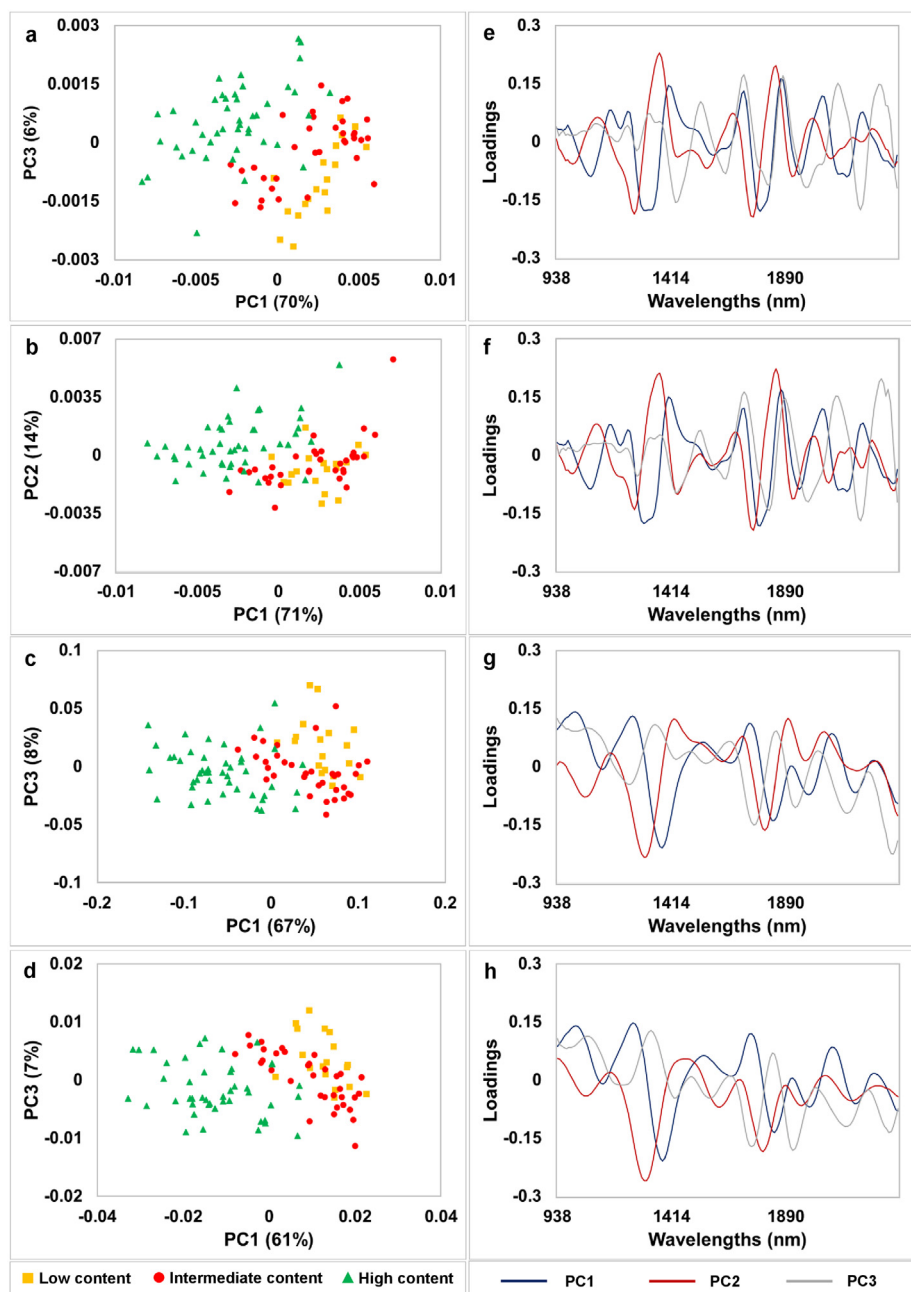


Fig. 2. PCA scores and loadings of orange peels and pectin: (a and e) Smoothing + 2nd derivative, (b and f) Smoothing + MSC + 2nd derivative, (c and g) Smoothing + SNV + 1st derivative, (d and h) Smoothing + 1st derivative. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

developed in other fruit such as peach (Zhu, Huang, Wu, Chen, & He, 2017). For quantitative assessment, partial least squares regression (PLSR), a type of linear analysis, is widely used to establish the correlation between spectral data of the hypercubes and real quantities or concentrations of the compound measured by ordinary laboratory assessments (Feng et al., 2018). NIR-HSI coupled with PLSR could be used in a preliminary analysis of orange waste to determine its constituents (such as pectin) in a quick way, to further decide whether to extract them or not based on the concentrations determined, being an alternative to improve extraction yield.

Thus, the hypothesis of this work is that NIR-HSI has a potential application for identification of pectin content in orange peels. This hypothesis will be verified by identification of the wavelengths that most contribute in the development of models with good precision and accuracy (defined as optimum wavelengths) and development of a

prediction model to quantify pectin content in orange peels.

2. Material and methods

2.1. Pectin extraction from raw material

Lanelate oranges harvested from January to July of 2019 were supplied by Export Orange S.L. The oranges were ground (1 cm × 1 cm × 1 cm) and the juice was separated from the peel using a strainer. The peel was washed with water at 100 °C, and washed again with water at 50 °C. This step was carried out to remove most of sugars. Then, the peel was strained and oven dried at 45 °C for 2 days. After, the peel was milled using a milling machine and sieved with an ASTM N°50 mesh (300 μm diameter). A sample of 5 g was dispersed in water at a solid-liquid ratio of 1:40 (w/v) and the pH of the mixture was

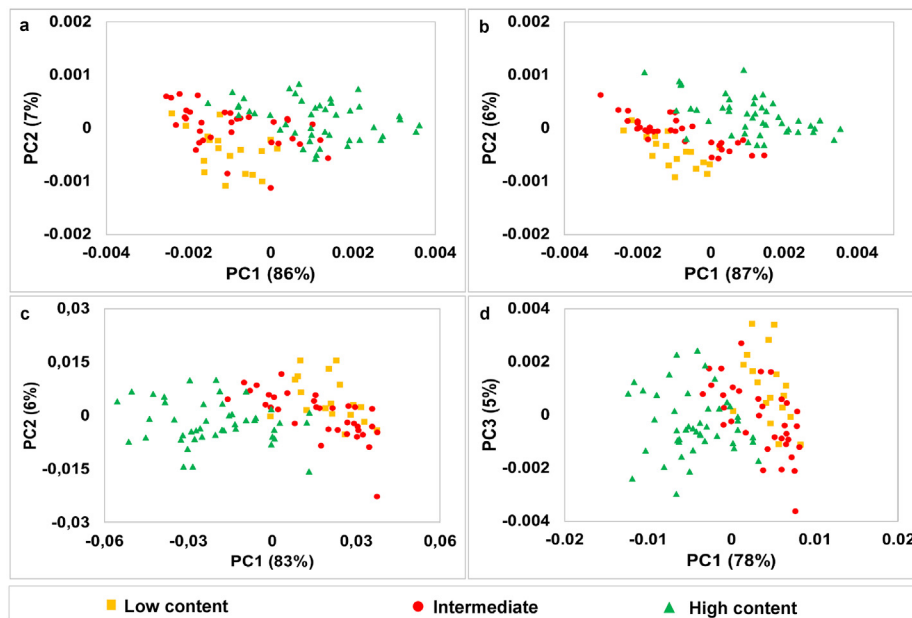


Fig. 3. PCA scores of orange peels and pectin after variable selection: a) Smoothing + 2nd derivative, b) Smoothing + MSC + 2nd derivative, c) Smoothing + SNV + 1st derivative, d) Smoothing + 1st derivative. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
LDA models performance for qualitative analysis of pectin in orange peels.

Pre-treatment	Wavelengths (nm)	Sensitivity (Validation)			Specificity (Validation)			Accuracy – Calibration Model (%)
		C1	C2	C3	C1	C2	C3	
Smoothing	1100, 1290, 1442, 1775, 1918, 2032, 2118, 2289	0.58	0.67	1.00	0.83	0.81	1.00	83.51
Smoothing + SNV	1318, 1442, 1775, 1871, 2032, 2213	0.67	0.40	0.93	0.70	0.81	1.00	84.54
Smoothing + 1st Derivative	1014, 1252, 1376, 1547, 1747, 1842, 1918, 1985, 2080, 2185, 2251	0.67	0.67	0.93	0.83	0.81	1.00	92.78
Smoothing + 2nd Derivative	1071, 1166, 1242, 1318, 1404, 1595, 1718, 1785, 1871, 1947, 2042, 2147, 2213	0.58	0.73	1.00	0.87	0.81	1.00	90.72
Smoothing + MSC	1318, 1452, 1775, 1871, 2051, 2213	0.67	0.47	0.93	0.73	0.81	1.00	85.57
Smoothing + MSC + 2nd Derivative	1081, 1166, 1233, 1309, 1404, 1595, 1718, 1785, 1871, 1947, 2042, 2137, 2213	0.58	0.67	1.00	0.83	0.81	1.00	88.66
Smoothing + SNV + 1st Derivative	1014, 1252, 1376, 1547, 1747, 1842, 1918, 1985, 2080, 2185, 2251	0.67	0.60	0.93	0.80	0.81	1.00	91.75

C1: low pectin content (0–5%); C2: intermediate pectin content (10–40%); C3: high pectin content (50–100%).

Table 2
PLSR model performance for quantitative analysis of pectin in orange peels.

Pre-treatment	Wavelengths (nm)	LV	R_{CV}^2	RMSECV	R_p^2	RMSEP	SEP	RPD	RER
Smoothing	Full spectra	6	0.93	8.54	0.93	9.16	8.27	4.11	12.09
	1271, 1366, 1452, 1690, 1766, 1861, 1947, 2261, 2365	6	0.93	8.61	0.93	8.84	8.04	4.22	12.44
Smoothing + SNV	Full spectra	6	0.95	7.43	0.94	7.99	7.15	4.75	13.98
	1271, 1328, 1452, 1690, 1766, 1871, 1947, 2261, 2346	5	0.95	7.57	0.94	8.42	7.59	4.47	13.17
Smoothing + 1st Derivative	Full spectra	3	0.93	8.71	0.95	7.94	7.66	4.43	13.06
	1223, 1395, 1652, 1728, 1813, 1899, 2166, 2318	2	0.94	8.37	0.92	9.73	9.20	3.69	10.87
Smoothing + 2nd Derivative	Full spectra	4	0.95	7.42	0.96	6.50	6.32	5.37	15.81
	1062, 1214, 1318, 1433, 1690, 1861, 2051, 2118, 2204, 2270	2	0.94	8.15	0.94	8.03	7.33	4.63	13.64
Smoothing + MSC	Full spectra	4	0.96	6.68	0.94	8.23	7.88	4.30	12.68
	1328, 1452, 1690, 1766, 1871, 1937, 2194, 2237	4	0.95	7.60	0.94	8.03	7.70	4.41	12.99
Smoothing + MSC + 2nd Derivative	Full spectra	4	0.95	7.30	0.96	6.94	6.67	5.09	14.99
	1033, 1318, 1423, 1690, 1861, 2118, 2204, 2270	2	0.94	7.85	0.94	8.32	7.85	4.32	12.74
Smoothing + SNV + 1st Derivative	Full spectra	5	0.95	7.13	0.95	7.96	7.51	4.52	13.32
	1290, 1395, 1499, 1652, 1728, 1899, 2166, 2299	3	0.95	7.54	0.94	8.44	7.90	4.29	12.66

adjusted to 1.5 with 0.1 M HCl. Subsequently, the mixture was placed in a water bath between 70 and 85 °C at 300 rpm for 70 min. After, the mixture was centrifuged at 5000 rpm for 20 min. The supernatant (solution of pectin extract) was removed and the liquid residue was oven dried at 45 °C for 2 days. Then, the dried residue was milled and sieved with the ASTM N°50 mesh.

2.2. Sample preparation and image acquisition

Standard pectin (Apple pectin powder – 100% purity, Solgar, Inc. Leonia, USA) was added to the dried residue of orange peel without pectin in the following percentages: 0, 1, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90 and 100% (being 0% residue of orange peel without pectin addition, and 100% pure pectin without residues of orange peel). Ten

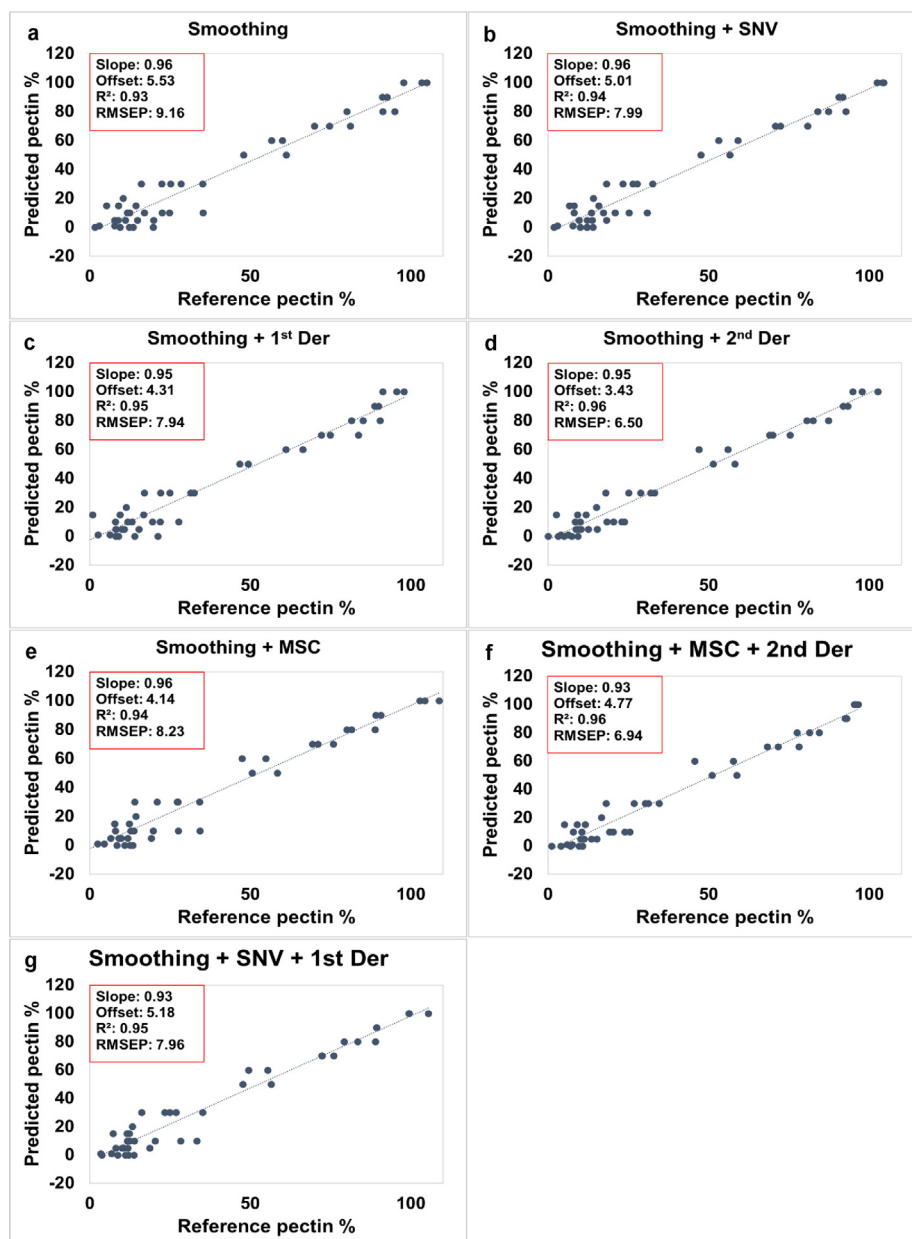


Fig. 4. Predicted vs. reference plot for the PLSR models obtained with full spectra: a) Smoothing, b) Smoothing + SNV, c) Smoothing + 1st derivative, d) Smoothing + 2nd derivative, e) Smoothing + MSC, f) Smoothing + MSC + 2nd derivative, g) Smoothing + SNV + 1st derivative.

samples of 5 g of each percentage were prepared, totalizing 140 samples. Samples were placed in Petri dishes of 3 cm of diameter for image acquisition.

NIR hyperspectral images were acquired from each sample in the reflectance mode using a SWIR camera (Headwall Photonics SWIR M series, Massachusetts, USA), in the range of 900–2500 nm, with an illuminator of 75 W and a scanning speed of 14.7 mm/s. The program automatically subtracted the white (~99% reflectance) and dark (0% reflectance) references from subsequently acquired images.

2.3. Spectra extraction and multivariate analysis

Image segmentation and spectrum extraction were performed using a code developed by the research group using the open software Python (version 3.7.0; Python Software Foundation License). The averaged reflectance spectra were smoothed (Savitzky-Golay) and mean centred prior to multivariate data analysis. Data were pre-treated with Standard

Normal Variate (SNV), Multiplicative Scatter Correction (MSC), first derivative (1st Der) (Savitsky-Golay smoothing, 11 points window, 2nd order polynomial), second derivative (2nd Der) (Savitsky-Golay smoothing, 11 points window, 2nd order polynomial) and a combination of MSC + 2nd der, and SNV + 1st der. After that, qualitative and quantitative analysis such as Principal Component Analysis, Linear Discriminant Analysis and Partial Least Square Regression were carried out. The multivariate data analysis was performed using The Unscrambler X 10.4 software.

2.4. Principal components analysis (PCA)

PCA was performed as exploratory data analysis, in order to obtain an overview of the variation among samples, identify clusters and outliers. This step is applied to reduce the spectral information into principal components (PCs), which are a linear combination of the variables in the spectra data and contain most of the relevant

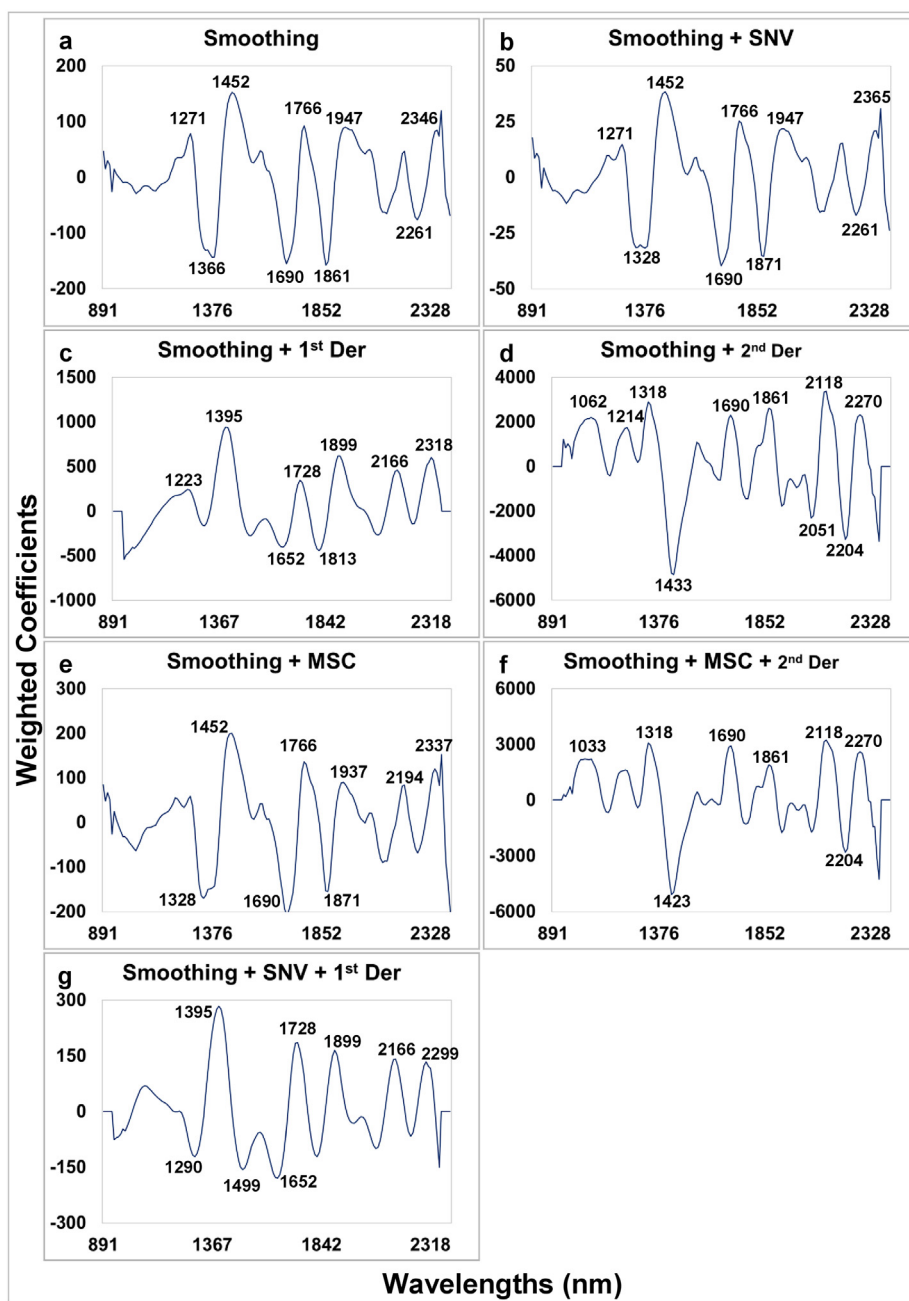


Fig. 5. Regression coefficients for the PLSR models obtained with full spectra: a) Smoothing, b) Smoothing + SNV, c) Smoothing + 1st derivative, d) Smoothing + 2nd derivative, e) Smoothing + MSC, f) Smoothing + MSC + 2nd derivative, g) Smoothing + SNV + 1st derivative.

information (Rachmawati, Rohaeti, & Rafi, 2017). Moreover, through PCA scores, data from both sets of experiments were split into calibration and validation sets (independent sets), using Kennard-Stone algorithm (Kennard & Stone, 1969), in a ratio of 70 and 30%, respectively. In addition, the loadings plot were used to manually select the wavelengths that contained the relevant information about the sample, defined as optimum wavelengths. These optimum wavelengths were chosen based on the most prominent peaks and valleys for further qualification quantification analysis (LDA).

2.5. Linear discriminant analysis (LDA)

Linear discriminant analysis (LDA) is a classification method used to differentiate samples into different classes. As a particularity of LDA method, the number of variables must be smaller than the number of

samples (Manley & Baeten, 2018). Therefore, optimum wavelengths selected in the loadings plot of PCA were used as the data input to perform LDA analysis in all pre-treated data. LDA was performed with a leave-one-out cross-validation method and an external validation was carried out with the independent dataset. Models performance was measured in terms of sensitivity, specificity (of validation set) and accuracy of calibration model. Values of sensitivity and selectivity close or equal to 1.00 and accuracy of 100% show good discriminative power.

2.6. Partial least square regression (PLSR)

Partial Least Square Regressions (PLSR) was performed to test the ability of NIR-HSI data to quantify the different percentages of pectin added to the residues of orange peel. First, calibration models were developed using full spectra, and, then, the weighted regression

coefficients (B_w) resulting from these PLSR models were used for the development of reduced models. The wavelengths with the highest values (regardless of sign) were selected as optimum wavelengths. All the models were developed based on a full cross validation (leave-one out), and the models performance was evaluated in terms of coefficient of determination and root mean square error of cross-validation (R_{CV}^2 , RMSECV) and prediction (R_p^2 , RMSEP), the ratio performance to deviation (RPD) and the ratio of error range (RER), both based on the standard error of validation (SEP). It is expected coefficients of determination close to 1, errors close to 0, RPD and RER above 3 and 10, respectively (Malley, Yesmin, & Eilers, 2002).

3. Results and discussion

3.1. Spectral analysis

The average HSI reflectance spectra of orange peel and pectin in each percentage are presented in Fig. 1. The smoothed spectra (Fig. 1a) showed very similar shape for all samples, only differing in the intensity of reflectance across the spectral region. Overall, samples with lower amount of pectin had higher reflectance. This baseline difference was minimized after spectra pre-treatment with SNV and MSC (Fig. 1b and 1e, respectively). Peaks were observed around 1214 and 1728 nm, which correspond to stretching vibrations of C–H in the first and second overtone. A valley at 1306 nm indicated stretching vibration of C–H or a combination of deformation. The peak around 1918 nm is related to the stretching vibration of C=O in the second overtone of amides. At 2100 nm, there was a peak which denotes a combination of O–H deformation and C–O stretching vibration, associated to starch.

After 2nd derivatives and the combination of pre-treatments (MSC + 2nd der), other peaks and valleys were highlighted along the spectra (Fig. 1d and 1g). Most of these denote stretching vibrations of C–H in the first and second overtones or combination of deformation. In addition, the 2nd derivative highlighted in the spectra a peak around 2000 nm, which denote a combination of vibration of O–H and C–O deformation in starch, and a valley at 2080 nm corresponding to a combination of vibrations of O–H stretching and O–H deformation, associated to sucrose and starch (Osborn, Fearn, & Hindle, 1993).

3.2. Principal components analysis (PCA)

According to the literature, the performance of classification models based on multi-classes seems to be worse than classifications based on group of classes (Ziegler et al., 2016). Therefore, in order to achieve a better performance for models with different percentages of pectin, three classes were created and defined as “Low content” (samples with 0–5% of pectin), “Intermediate content” (samples with 10–40% of pectin) and “High content” (samples with 50–100% of pectin), based on the average pectin content in orange peels reported by other authors (El-Nawawi & Shehata, 1987; Rouse & Crandall, 1976), which is roughly 30% (wt.) on a dry matter basis. Then, principal component analysis was performed on reflectance spectra with all pre-treatments based on these classes. Those pre-treatments that better separated the groups of samples are displayed in Fig. 2. Fig. 2a, 2c and 2d show the scores plots of PC1 and PC3 for spectra pre-treated with 2nd derivative, a combination of SNV and 1st derivative, and 1st derivative only, respectively; whereas, Fig. 2b shows the scores plot of the first two principal components for spectra pre-treated with a combination of MSC and 2nd derivative. Overall, samples with high content of pectin are spread in the negative side of PC1, while samples with a low content of pectin are in the positive side. Accordingly, samples with an intermediate content of pectin are disposed between the other two groups.

The first three principal components explained over 75% of the variance among samples, and the loadings plot of these three PCs are represented in Fig. 2(e-h). Most of the peaks and valleys observed in these plots agree with the respective spectra, especially those associated

to cellulose, starch and sucrose. Consequently, these loadings plot were used in order to identify the wavelengths that most contributed to separate the groups of samples. Then, these optimum wavelengths were selected, and a new PCA was recalculated for each of the pre-treatments (Fig. 3 a-d). Fig. 3a and 3b showed that over 90% of the variance between samples were explained by the first two PCs, with samples of low pectin content displayed in the left side of PC1 and negative part of PC2, and samples of high pectin content in the right side of PC1 and positive part of PC2. Samples of intermediate pectin content remained in between the other two classes of samples. After variable selection, data pre-treated with SNV + 1st der and only 1st derivative showed similar behaviour on PCA plot as with full spectra, with PC1 and PC2 explaining 89% of the variance among samples.

3.3. Linear discriminant analysis (LDA)

The selected wavelengths from PCA loadings were used as predictors for LDA models, in order to evaluate the pre-treatments that better discriminated samples into their respective classes. Table 1 shows the models performance of LDA for qualitative analysis of pectin in orange peels. As in PCA, samples containing different percentages of pectin were grouped into three classes defined as “Low content” (samples with 0–5% of pectin), “Intermediate content” (samples with 10–40% of pectin) and “High content” (samples with 50–100% of pectin). These classes were used to develop the classification models of samples added of pectin. Data pre-treated with 1st derivative and a combination of SNV + 1st derivative provided the calibration models with the best accuracy (92.78 and 91.75%, respectively), and high values of sensitivity and specificity (0.60–1.00), thus demonstrating good capacity of the models to discriminate samples of residues of orange peel with different pectin content. Calibration models developed with pre-treated data with 2nd derivative and a combination of MSC + 2nd derivative had the second and third best accuracy (90.72 and 88.66%, respectively). However, the sensitivity and selectivity of these models were lower than those, ranging from 0.58 to 1.00. Data that was only smoothed, and SNV and MSC pre-treatments showed a suitable accuracy for calibration models (83.51, 84.54 and 85.57%, respectively). However, they did not improve the model discriminant power, since sensitivity and specificity of some classes were lower than 0.50. In addition, Class 3 had less misclassified samples when compared to Classes 1 and 2. This was also observed in the PCA score plots (Fig. 2 e-h), where, visibly, the distance between Classes 1 and 2 were closer than Class 3, so there was an overlap of samples from those classes.

3.4. Partial least Square regression (PLSR)

The PLSR models performances for pectin content in orange peels are presented on Table 2. The predicted values of each PLSR model based on full spectra are displayed in Fig. 4. Overall, these models showed good precision, with high values of coefficient of determination (over 0.93), and errors slightly lower than those obtained by models based on optimum wavelengths, for both cross-validation and external validation.

The development of representative and accurate models with few wavelengths is more functional and interesting, especially for industrial applications. The weighted regression coefficients (B_w) provide information about the model quality, and those with large absolute value can be used as optimum wavelengths to develop reduced models (Kamruzzaman, Makino, Oshita, & Liu, 2015). Therefore, the important wavelengths were selected based on these weighted regression coefficients obtained by the full PLSR models, and the reduced models were developed. The results obtained by the reduced models also showed good coefficient of determination, which ranged from 0.92 to 0.94. Additionally, the errors of most models were slightly higher than those with full spectra, but still very close. Moreover, all the models based on full spectra or reduced models had $RPD > 3$ and $RER > 10$, which

showed a good calibration performance of the models.

The weighted regression coefficients carry a lot of information regarding the variables in the dataset. While those of high absolute values have a large contribution in the model, those of small values have a little contribution in the model (Elmasry et al., 2012). The weighted regression coefficients of the PLSR models developed in this work (Fig. 5) showed strong peaks and valleys, which are positive and negatively related to the important variables for pectin quantification. Most of these peaks and valleys are the same to those observed in the pectin spectra, as show in Fig. 1. In addition, they had spectral characteristics and did not show a lot of noise, suggesting that the models can be considered reliable and this technique can be considered for further application.

4. Conclusion

Hyperspectral imaging technique showed a great potential to classify orange peel samples according to the pectin content. As PCA, LDA was able to separate the groups of samples containing different percentages of pectin into the same three classes (low, intermediate and high pectin content). Among the different combinations of data pre-treatments assayed, 1st derivative and a combination of SNV + 1st derivative achieved the best accuracy (92.78 and 91.75%, respectively), and the highest sensitivity and specificity (between 0.6 and 1.0). Additionally, PLSR models for pectin content quantification based on full spectra showed excellent precision, providing high coefficients of determination (over 0.93), whereas those of the PLSR models built with the most contributing wavelengths ranged between 0.92 and 0.94.

Hence, the study confirms the hypothesis that NIR-HSI can be used for quantification of pectin content in orange peels. The results showed that this technique holds potential as an alternative to the carbazole colorimetry method to quantify pectin in orange peels, and to categorize orange peel samples into groups of different pectin concentration, and can be used to justify investments in the waste processing and extraction methods.

CRedit authorship contribution statement

Amanda Teixeira Badaró: Writing - original draft, Methodology, Investigation, Validation, Software. **Juan Francisco Garcia Martin:** Conceptualization, Writing - original draft, Writing - review & editing, Investigation, Visualization, Funding acquisition. **María del Carmen López Barrera:** Formal analysis, Writing - review & editing, Software, Visualization. **Douglas Fernandes Barbin:** Conceptualization, Formal analysis, Supervision, Software, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition. **Paloma Alvarez Mateos:** Conceptualization, Resources, Supervision, Methodology, Writing - review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Dr. Douglas Fernandes Barbin would like to acknowledge the University of Seville for the mobility grant (VIPIT-2019-I.3) awarded under the "VI Plan Propio de Investigación y Transferencia" of the University of Seville, and São Paulo Research Foundation (FAPESP), Young Researchers Award (Grant no. 2015/24351-2). Amanda Teixeira Badaró would like to thank São Paulo Research Foundation (FAPESP) for the financial support (project numbers 2017/17628-3 and 2019/

06842-0). This work was in part financed by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

- Badaró, A. T., Morimitsu, F. L., Ferreira, A. R., Clerici, M. T. P. S., & Barbin, D. F. (2019). Identification of fiber added to semolina by near infrared (NIR) spectral techniques. *Food Chemistry*, 289, 195–203. <https://doi.org/10.1016/j.foodchem.2019.03.057>.
- Bizzani, M., Flores, D. W. M., Colnago, L. A., & Ferreira, M. D. (2017). Non-invasive spectroscopic methods to estimate orange firmness, peel thickness, and total pectin content. *Microchemical Journal*, 133, 168–174. <https://doi.org/10.1016/j.micro.2017.03.039>.
- Chen, H., Xu, L., Jia, Z., Cai, K., Shi, K., & Gu, J. (2018). Determination of parameter uncertainty for quantitative analysis of shaddock peel pectin using linear and non-linear near-infrared spectroscopic models. *Analytical Letters*, 51(10), 1564–1577. <https://doi.org/10.1080/00032719.2017.1384479>.
- El-Nawawi, S. A., & Shehata, F. R. (1987). Extraction of pectin from Egyptian orange peel. Factors affecting the extraction. *Biological Wastes*, 20(4), 281–290. [https://doi.org/10.1016/0269-7483\(87\)90005-X](https://doi.org/10.1016/0269-7483(87)90005-X).
- Elmasry, G., Sun, D. W., & Allen, P. (2012). Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef. *Journal of Food Engineering*, 110(1), 127–140. <https://doi.org/10.1016/j.jfoodeng.2011.11.028>.
- Feng, C. H., Makino, Y., Oshita, S., & Martín, J. F. G. (2018). 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*, 84, 165–176. <https://doi.org/10.1016/j.foodcont.2017.07.013>.
- Feng, Y.-Z., & Sun, D.-W. (2012). Application of hyperspectral imaging in food safety inspection and control: A review. *Critical Reviews in Food Science and Nutrition*, 52(11), 1039–1058. <https://doi.org/10.1080/10408398.2011.651542>.
- Forchetti, D. A. P., & Poppi, R. J. (2017). Use of NIR hyperspectral imaging and multivariate curve resolution (MCR) for detection and quantification of adulterants in milk powder. *LWT - Food Science and Technology*, 76, 337–343. <https://doi.org/10.1016/j.lwt.2016.06.046>.
- García-Martín, J. F., Olmo, M., & García, J. M. (2018). Effect of ozone treatment on postharvest disease and quality of different citrus varieties at laboratory and at industrial facility. *Postharvest Biology and Technology*, 137, 77–85. <https://doi.org/10.1016/j.postharvbio.2017.11.015>.
- García, J. F., Olmo, M., & García, J. M. (2016). Decay incidence and quality of different citrus varieties after postharvest heat treatment at laboratory and industrial scale. *Postharvest Biology and Technology*, 118, 96–102. <https://doi.org/10.1016/j.postharvbio.2016.03.019>.
- Grohmann, K., Cameron, R. G., & Buslig, B. S. (1995). Fractionation and pretreatment of orange peel by dilute acid hydrolysis. *Bioresource Technology*, 54(2), 129–141. [https://doi.org/10.1016/0960-8524\(95\)00121-2](https://doi.org/10.1016/0960-8524(95)00121-2).
- Kamruzzaman, M., Makino, Y., Oshita, S., & Liu, S. (2015). Assessment of visible near-infrared hyperspectral imaging as a tool for detection of horsemeat adulteration in minced beef. *Food and Bioprocess Technology*, 8(5), 1054–1062. <https://doi.org/10.1007/s11947-015-1470-7>.
- Kennard, R. W., & Stone, L. A. (1969). Computer-aided design of experiments. *Technometrics*, 11(1), 137–148. <https://doi.org/10.1080/00401706.1969.10490666>.
- Lessa, E. F., Gularte, M. S., Garcia, E. S., & Fajardo, A. R. (2017). Orange waste: A valuable carbohydrate source for the development of beads with enhanced adsorption properties for cationic dyes. *Carbohydrate Polymers*, 157, 660–668. <https://doi.org/10.1016/j.carbpol.2016.10.019>.
- Liu, L. S., Fishman, M. L., Kost, J., & Hicks, K. B. (2003). Pectin-based systems for colon-specific drug delivery via oral route. *Biomaterials*, 24(19), 3333–3343. [https://doi.org/10.1016/S0142-9612\(03\)00213-8](https://doi.org/10.1016/S0142-9612(03)00213-8).
- Malley, D. F., Yesmin, L., & Eilers, R. G. (2002). Rapid analysis of hog manure and manure-amended soils using near-infrared spectroscopy. *Soil Science Society of America Journal*, 66, 1677–1686. <https://doi.org/10.2136/sssaj2002.1677>.
- Manley, M., & Baeten, V. (2018). Spectroscopic technique: near infrared (NIR) spectroscopy. In D.-W. Sun (Ed.), *Modern techniques for food authentication* (pp. 51–102). (2nd ed.). Academic Press.
- Ministerio de Agricultura. Superficies y producciones anuales de cultivos. (2017). <https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/superficies-producciones-anuales-cultivos/> Accessed 28 November 2019.
- Osborn, B. G., Fearn, T., & Hindle, P. H. (1993). *Theory of near infrared spectroscopy. Practical NIR spectroscopy with applications in food and beverage analysis* (pp. 13–35). (2nd ed.). London: Longman Singapore Publishers (Pte) Ltd.
- Rachmawati, Rohaeti, E., & Rafi, M. (2017). Combination of near infrared spectroscopy and chemometrics for authentication of taro flour from wheat and sago flour. *Journal of Physics: Conference Series*, 835. <https://doi.org/10.1088/1742-6596/835/1/012011>.
- Ren, J. N., Hou, Y. Y., Fan, G., Zhang, L. L., Li, X., Yin, K., & Pan, S. Y. (2019). Extraction of orange pectin based on the interaction between sodium caseinate and pectin. *Food Chemistry*, 283, 265–274. <https://doi.org/10.1016/j.foodchem.2019.01.046>.
- Rezzadori, K., Benedetti, S., & Amante, E. R. (2012). Proposals for the residues recovery: Orange waste as raw material for new products. *Food and Bioprocess Technology*, 90(4), 606–614. <https://doi.org/10.1016/j.fbp.2012.06.002>.
- Rouse, A. H., & Crandall, P. G. (1976). Nitric Acid Extraction of Pectin from citrus peel. *Proceedings of the Florida State Horticultural Society*, 89, 166–168.
- Wang, Y. C., Chuang, Y. C., & Hsu, H. W. (2008). The flavonoid, carotenoid and pectin content in peels of citrus cultivated in Taiwan. *Food Chemistry*, 106(1), 277–284.

- <https://doi.org/10.1016/j.foodchem.2007.05.086>.
- Wicker, L., Kim, Y., Kim, M. J., Thirkield, B., Lin, Z., & Jung, J. (2014). Pectin as a bioactive polysaccharide - Extracting tailored function from less. *Food Hydrocolloids*, 42(P2), 251–259. <https://doi.org/10.1016/j.foodhyd.2014.01.002>.
- Zhu, N., Huang, W., Wu, D., Chen, K., & He, Y. (2017). Quantitative visualization of pectin distribution maps of peach fruits. *Scientific Reports*, 7(9275), <https://doi.org/10.1038/s41598-017-09817-7>.
- Ziegler, J. U., Leitenberger, M., Longin, C. F. H., Würschum, T., Carle, R., & Schweiggert, R. M. (2016). Near-infrared reflectance spectroscopy for the rapid discrimination of kernels and flours of different wheat species. *Journal of Food Composition and Analysis*, 51, 30–36. <https://doi.org/10.1016/j.jfca.2016.06.005>.