Comparative study on the use of three different near infrared spectroscopy recording methodologies for varietal discrimination of walnuts

Short title: Varietal discrimination of walnuts using near infrared spectroscopy.

Julio Nogales-Bueno[1,2]; Luis Feliz[1]; Berta Baca-Bocanegra[2]; José Miguel Hernández-Hierro[2]; Francisco José Heredia[2]; Barroso, João Manuel[1]; Ana Elisa Rato[1]*;

[1] *Instituto de Ciências Agrárias e Ambientais Mediterrânicas, Universidade de Évora, Departamento de Fitotecnia, Apartado 94 7002 – 554 Évora, Portugal.*

[2] *Food Colour and Quality Laboratory, Área de Nutrición y Bromatología, Facultad de Farmacia, Universidad de Sevilla, 41012 Sevilla, Spain.*

*** Corresponding author:** Ana Elisa Rato

Phone: 266760800

E-mail: aerato@uevora.pt

Abstract

Walnut fruit (*Juglans regia* L.) is an internationally well-known product with an important tradition of consumption. Its health benefits and economic importance in the food industry make this nut an interesting research topic.

In this feasibility study, 200 walnut samples of 5 different varieties were collected and their NIR spectra were recorded with 3 different devices: a benchtop FT-NIR spectrograph, a hyperspectral imaging camera and a portable NIR device. Discriminant analyses were applied and different methods for the varietal discrimination of walnuts were obtained and compared.

Up to 96 and 84% of correct identification in internal and external validation were obtained. Better results were obtained covering the entire shell surface than collecting a unique random spectrum per sample. Moreover, FT-NIR and hyperspectral produce better models than the portable NIR one.

Keywords: near infrared, walnuts, hyperspectral imaging, chemometrics, discriminant analysis, *Juglans regia* L.

1. Introduction

The consumption of walnut fruit (*Juglans regia* L.) has a relevant importance in both health and economic fields. Health benefits of walnuts are due to their chemical composition. These nuts are rich in polyunsaturated fatty acids and tocopherols, being linoleic acid the most abundant fatty acid (Amaral, Casal, Pereira, Seabra, & Oliveira, 2003; Amaral, Cunha, Alves, Pereira, Seabra, & Oliveira, 2004; Pereira, Oliveira, Sousa, Ferreira, Bento, & Estevinho, 2008). Moreover, bioactive compounds with potential health benefits, such as dietary fibre, folic acid, polyphenolic compounds and other antioxidants, are present in walnuts (Kris-Etherton, Yu-Poth, Sabaté, Ratcliffe, Zhao, & Etherton, 1999; Larrosa, García-Conesa, Espín, & Tomás-Barberán, 2010). Therefore, it has been demonstrated that the regular consumption of walnuts is linked with a decrease of the risk of coronary heart disease, metabolic syndrome and other chronic diseases (Davis, Stonehouse, Loots, Mukuddem-Petersen, van der Westhuizen, Hanekom, et al., 2007; Kris-Etherton, Yu-Poth, Sabaté, Ratcliffe, Zhao, & Etherton, 1999).

The demonstrated health benefits together with the wide tradition of this nut in the human diet (since the pre-agricultural times) and their tasty sensory attributes (Sinesio & Moneta, 1997), have created and consolidated an important international walnut market. In 2017/2018, global walnut production was estimated at 870000 metric tons (kernel basis), consolidating the growing trend observed over the last 10 years (International Nut and Dried Fruit Council (INC), 2018). This positive trend has also been observed in Portugal, where walnut production reached 4600 metric tons in 2017. It is in Alentejo, the south-central region of Portugal, where is located the most important production area of walnuts in Portugal, yielding approximately 2000 metric tons. However, this production region is characterized by young orchards which did not reach yet the full production potential. According to the Instituto Nacional de Estatística (INE), in 2017, walnut production ranks fourth among other fruit nuts being an important agricultural commodity from Alentejo (Instituto Nacional de Estatística, 2017) In this region, the commercial walnut production occupies an area of about 4000 acres and the trend is still upwards. Most orchards are family owned with an average dimension between 7,5-50 acres distributed over different soil types. This orchard fragmentation leads to a heterogeneous quality of fruit production.

Contrary to other regions in the world where approximately two-thirds of walnut production is traded shelled, in Europe most of the walnuts production are traded in-shell and, particularly, Portuguese market for walnuts is half divided into walnuts with and without shell (Instituto Nacional de Estatística, 2017). Europeans value in-shell walnuts for considering them more natural and less processed. Moreover, in-shell walnuts can be better preserved than walnut kernels.

Walnuts can also be destined to the production of other products that can be consumed directly or used by food industry to improve food characteristics. As an example, the high lipid content of walnut kernel, between 50 to 75% (w/w) depending on the cultivar, encourages the use of walnuts for oil production (Amaral, Casal, Pereira, Seabra, & Oliveira, 2003; Rabadán, Pardo, Gómez, & Álvarez-Ortí, 2018). Producers have to decide how they are selling their walnuts. This decision is depending on a number of factors. Among them, walnut variety is one of the most important. Different walnut varieties have different organoleptic attributes and ratios of shell-kernel weight (Guerrero, Romero, Gou, Aleta, & Arnau, 2000; Sinesio & Moneta, 1997). For instance, varieties with a low shell-kernel ratio are usually destined to kernel sale and vice versa. Besides the difference among varieties in lipid content, the frequency of the apparition of defects, such as abnormal coloration of the skin and/or kernel, insect damage, detrimental disorders and fungal growth, is usually linked to walnut variety. Therefore, in the future the price of the walnut may be dependent largely on its variety. As a consequence, the processing and packing industries need reliable and accurate methods for identification and classification of walnut varieties.

Traditionally, discrimination of walnut varieties is carried out by the sensorial analysis of the kernel (Guerrero, Romero, Gou, Aleta, & Arnau, 2000; Sinesio & Moneta, 1997). This organoleptic evaluation is subjective, laborious, and time-consuming. In order to achieve accurate and automatic discrimination methods, other technologies have been gradually tested. Varietal discrimination has been tested by using some physical and chemical parameters, such as walnut oil viscosity, colour data, fatty acids profile, etc. (Bou Abdallah, Baatour, Mechrgui, Herchi, Albouchi, Chalghoum, et al., 2016; Martínez, Mattea, & Maestri, 2006; Rabadán, Pardo, Gómez, & Álvarez-Ortí, 2018). Moreover, in other studies, genetic analysis has been also applied to the classification of walnuts according to their varieties (Ciarmiello, Piccirillo, Pontecorvo, De Luca, Kafantaris, & Woodrow, 2011; Ma, Zhang, & Pei, 2011; Pop, Vicol, Botu, Raica, Vahdati, & Pamfil, 2013). All the aforementioned studies used destructive methods which also need polluting chemical reagents. Conversely, Ercisli, Sayinci, Kara, Yildiz, and Ozturk (2012) and Peng, Liu, Kong, Zhang, Yu, and He (2017) used visible image analysis and laser-induced breakdown spectroscopy (LIBS), respectively in walnut variety discrimination, achieving good results. Near infrared spectroscopy (NIRS) has been applied for the varietal discrimination of some nuts, such as almonds, although no studies has been developed so far for varietal classification of walnuts (Teixeira & Sousa, 2019). However, NIRS has been applied to walnuts for regional identification (Gu, Zhang, Li, Ma, Tu, Song, et al., 2018) or for the control of different quantitative parameters such as moisture, protein and fats (Yi, Sun, Zhu, Liu, & Lu, 2017). In addition, it is well known the possibilities of NIRS for the varietal classification of agricultural products by the use of imaging, benchtop and/or portable devices (Lacar, Lewis, & Grierson, 2001; Nogales-Bueno, Rodríguez-Pulido, Heredia, & Hernández-Hierro, 2015; Perez, Sanchez, Cano, & Garrido, 2001).

Nevertheless, most of the aforementioned studies are carried out in walnut kernels, the final product whose properties make the walnut a really interesting nut. Considering the characteristics of the southern Europe walnut market and its quality requirements, it might be interesting to develop non-destructive methods to varietal discrimination of in-shell walnuts. Moreover, these methods could be also developed and applied in other parts of the world. Varietal discrimination of in-shell walnuts would be interesting for producers, regardless of the final destination of their product. The aim of this study is to develop and optimize methodologies for the discrimination of five of the main walnuts varieties present in the Portuguese market by the use of three different NIR devices: a benchtop FT-NIR spectrograph, a hyperspectral imaging camera and a portable NIR device. Moreover, different orientations in spectra acquisition, spectral pre-treatments and chemometric analyses will be tested in order to obtain a robust and reliable methodology for the varietal discrimination of walnut samples. For the best of our knowledge, this is the first time that the aforementioned objectives have been jointly carried out.

2. Materials and methods

2.1. Walnut samples

Walnut samples of ‘Chandler’, ‘Franquette’, ‘Howard’, ‘Lara’ and ‘Tulare’ varieties were collected in 2018-2019 season, from Fruteco, a fruit producer’s association with a walnut production area of around 1500 acres in Alentejo. After sample collection, faulty samples were identified and removed by mechanical methods. Then, a total of 200 in-shell walnuts were collected, 40 samples for each variety. Samples were individually identified and stored until the different spectroscopic analyses were carried out.

2.2. Spectroscopic data acquisition

Three different near infrared (NIR) spectral matrixes were obtained from 3 different devices: a benchtop FT-NIR spectrograph, a hyperspectral imaging camera and a portable NIR device. As Figure 1 describes, 6 different geometrical replicates were acquired for each walnut and device. Each individual spectrum was acquired following a longitudinal or transversal axe of the nut: two spectra were acquired in the longitudinal axe (at the top and the bottom of the walnut), whereas 4 spectra were acquired in transversal axes of the nut, by rotating it 90 degrees between each acquisition. The different spectra acquired were labelled as Top (T), Bottom (B), Lateral 1 (L1), Lateral 2 (L2), Face 1 (F1) and Face 2 (F2). Therefore, a total of 1200 walnut spectra were acquired in each device (40 walnuts/variety × 5 varieties × 6 spectra/walnut).

For each device, spectral matrixes were presented in three different formats:

(1) All-spectra format: one matrix with 6 spectra for each walnut, i.e., all the spectra acquired in each device. A total of 1200 spectral samples per device.

(2) Average format: one matrix with the average spectrum of each walnut. A total of 200 spectral samples per device.

(3) Random format: one matrix with a randomly selected spectrum of each walnut. This format tries to imitate the random spectra acquisition in a separation line, where the orientation of the walnut can not be easily controlled. A total of 200 spectral samples per device.

2.2.1. Benchtop FT-NIR device.

An FT-NIR spectrometer (MPA, Opus Bruker, Germany) was used for the acquisition of the FT-NIR spectra of the walnuts in the range of 10499.34-3594.93 cm^{-1} (952.44-2781.70 nm). Absorbance spectra were obtained after a total of 32 scans with an average resolution of 16 cm^{-1} following a modification of the method described in Milinovic, Garcia, Rato, and Cabrita (2019). The background signal was corrected before each walnut spectra collection. The Opus v.7.5 software (Bruker Optik GmbH, Germany) was employed for spectral data collection.

2.2.2. Hyperspectral imaging.

Hyperspectral images were acquired following a modification of the method described in Nogales-Bueno, Hernández-Hierro, Rodríguez-Pulido, and Heredia (2014). Briefly, walnuts were disposed in groups of 5 samples under the hyperspectral camera. This device consists of a Xenics® XEVA-USB InGaAs camera (320 × 256 pixels; Xenics Infrared Solutions, Inc., Leuven, Belgium), a spectrograph (Specim ImSpector N17E Enhanced; Spectral Imaging Ltd., Oulu, Finland) covering the spectral range between 900 and 1700 nm (spectral resolution of 3.25 nm), two 70W tungsten iodine halogen lamps (Prilux®, Barcelona, Spain), a mirror scanner (Spectral Imaging Ltd., Oulu, Finland) and a computer system. Hyperspectral images were recorded using a 50 Hz frame rate and an exposure time of 9 ms using the instrument acquisition software SpectralDAQ v. 3.62 (Spectral Imaging Ltd., Oulu, Finland).

Reflectance spectra were corrected by means of a two point calibration. For that, a white reference spectrum (Labsphere Inc., North Sutton, USA) and a dark current spectrum were acquired in each collection session. Then, a segmentation procedure was applied to the images in order to extract the spectral information of the samples and remove the information of the background. This procedure allowed extracting the average spectrum of each walnut in the image. Finally, a number of noisy wavelengths at both extremes of the spectral range were identified and only the subrange 950-1650 nm was transformed to absorbance values and saved. Segmentation was carried out in the software Matlab (R2018a; TheMathWorks, Inc., MA, USA) and SPSS 25.0 (SPSS,Inc.,Chicago,IL,USA).

2.2.3.Portable MicroNIR device

The third spectral matrix was obtained by means of the use of a portable MicroNIR device (VIAVI, Santa Rosa, CA, USA). This device allows the acquisition of spectral samples in situ. It can be controlled with a laptop or a tablet, which is also the energy source for the MicroNIR. Six spectra were acquired for each sample following a modification of the method described in Baca-Bocanegra, Nogales-Bueno, García-Estévez, Escribano-Bailón, Hernández-Hierro, and Heredia (2019). Spectra were recorded using 9.3 ms as integration time and 100 as scan count. Background and reference spectra were acquired for each walnut and sample spectra were automatically corrected by the instrument acquisition software, Micro-NIR™ Pro v.2.2 (VIAVI, Santa Rosa, CA, USA). Finally, spectral matrix were constructed and saved in absorbance values.

2.3. Chemometrics

2.3.1.Spectral pre-treatments

Several spectral treatments were tested in each spectral matrix: multiplicative scattering correction (MSC), standard normal variate (SNV), detrend and different derivatives. These pre-treatments were carried out by means of the statistical software The Unscrambler® X (CAMO Software AS., Oslo, Norway) with the aim of removing the undesirable effects that light scattering, sample texture or geometry have in the spectral data.

2.3.2.Principal component analysis (PCA)

The existence of patterns in the different NIR matrixes was tested with the application of PCA. This qualitative analysis allows obtaining information about the latent structure of the spectral matrix and it is an important source of knowledge to evaluate the suitability of posterior discriminant methods. PCA was carried out in The Unscrambler® X software.

Previously to the application of PCA to any spectral matrix, Mahalanobis distance (H) was evaluated for each spectra with the software Win ISI (v1.50) (Infrasoft International, LLC, Port. Matilda, PA, USA). Samples were ranked in order of their H distance from the mean spectrum of the entire sample set, and the $H > 3$ criterion was applied for spectral outliers detection and they were removed from the spectral matrix.

2.3.3. Linear discriminant analysis (LDA)

Linear discriminant analysis was applied in several occasions with two different purposes: (1) to discriminate between walnut and background in the segmentation of hyperspectral images and (2) to discriminate spectral samples between the different walnut varieties.

(1) For hyperspectral image segmentation 60 walnut and 60 background spectra were manually extracted from the hypercubes with ENVI 4.7 (ITT Corporation, White Plains, N.Y., USA) and saved in a spectral matrix. These 120 spectra were exported to SPSS 25.0 and a LDA was carried out. NIR hyperspectral wavelengths were used as dependent variables and a categorical variable with the membership of the spectra to walnut or background was used as factor. In order to select the minimum number of wavelengths for the development of the discriminant method, LDA was constructed with the stepwise feature and setting the probability of entrance of a new variable in 0.001.

(2) The discriminations of walnut varieties were carried out with the spectral matrixes obtained from the three NIR devices. These analyses were developed in The Unscrambler® X. This software allows developing a PCA from the spectral data and then using the PCA scores as dependent variables of the LDA. In that way, a low number of variables are introduced in the LDA algorithm without a high loss of spectral information. Walnut variety was used as categoric variable. In order to obtain a quantitative measure of the predictive capability of the developed methods, the spectral matrixes were randomly divided into calibration and validation sets. For each variety, 75% of the samples were used for the development of

discrimination methods and internal validation, while the remaining 25% were allocated to external validation.

3. Results and discussion

3.1. Segmentation of hyperspectral images

From the 215 wavelengths introduced in the stepwise LDA, only 3 wavelengths were retained and used in the discrimination model generated (1007.91, 1213.17 and 1363.41 nm). Following a modification of the method described in Rodríguez-Pulido, Hernández-Hierro, Nogales-Bueno, Gordillo, González-Miret, and Heredia (2014), the Fisher discriminant functions were introduced in a Matlab script. This script evaluates each pixel in the hyperspectral images and identifies them as walnut or background, i.e., it creates a segmentation mask for each image. Then, background was discarded, the number of objects (walnuts) in each mask was automatically counted and numbered by the script and the average spectrum of each walnut was calculated and saved. Figure 2 shows the whole process, from image acquisition to the extraction of the average spectra of each walnut in the image.

3.2. Principal component analysis. Spectral structure of NIR matrixes

Different spectral pre-treatments were applied to the FT-NIR, hyperspectral and MicroNIR matrixes. SNV combined with a second derivative were the most efficient pre-treatments for all matrixes. In order to check if the different geometrical replicates can be part of a homogeneous spectral matrix, a PCA was applied. It could be seen that for each device, geometrical replicates are overlapped in the space defined by the first and second principal components (PCs) (Fig. S1). Spectral outliers were identified according the H criterion: 52, 38 and 47 outliers were respectively identified in the FT-NIR, hyperspectral and MicroNIR matrixes. In any case, spectral outliers represent less than 5% of the corresponding sample set.

In order to look for variety trends among the spectral data, two different approaches were carried out for the three devices: (1) for each walnut, a single spectrum was calculated by averaging its six geometrical replicates. (2) For each walnut, a single spectrum was randomly selected from its six geometrical replicates. In both cases, a 200-sprectrum matrix was obtained for each device.

(1) Spectral outliers were identified and removed following the H criterion. Three, 1 and 6 outliers were respectively found in the FT-NIR, hyperspectral and MicroNIR matrixes. Then PCA were applied and the scores of the PCs were plotted. Figure 3 (a, b, c) shows the representations where a higher varietal trend can be observed for each device. It can be observed that there are partial separations between some varieties. Among all NIR spectroscopes the hyperspectral device allows obtaining a better separation between varieties. As can be observed in Fig. 3b, ‘Franquette’ and ‘Howard’ varieties are almost completely separated from other samples.

(2) In the case of random matrixes, 9, 1 and 8 spectra were identified as spectral outliers and removed from their respective matrixes. New PCA were developed and score plots were created (Fig. 3 d, e and f). Similar trends were found for the three spectral matrixes, although the trends were worse defined in this case.

Therefore, PCA can help to find some trends among walnut samples. The PCs represent the major part of the spectral variability of the original sample set and some of this variability is linked with the differences among different varieties. However, these trends do not allow visualizing a complete separation of all the varieties. PCA is an unsupervised dimensionality reduction method that is able to recognize underlying patterns. In order to look for underlying class structure how well and what causes this separation, supervised pattern recognition methods, such as LDA, are more adequate. In addition, it has been checked that spectral variability is not linked with the different geometrical replicates recorded, i.e., it does not matter the part of the walnut shell in which the spectra is recorded (Fig. S1). PCA performed on the data from random replicates produced similar trends than those performed on the data from the average walnut spectra (Fig. 3). These findings might have important consequences in the future development of spectral methods to control in-shell walnut in the field or industry.

Moreover, taking into account the results obtained in the different PCA developed for the 3 devices, it seems that hyperspectral imaging system can produce better results than FT-NIR and MicroNIR. Due to the rounded but irregular shape and hard texture of the walnut shell, it is difficult to perfectly adapt the measure port of these types of spot spectrometers to the shell. Therefore, FT-NIR and MicroNIR measures can be partially contaminated with external light. This problem is avoided in the image system by the segmentation process, which automatically selects only walnut pixels.

3.3. Linear discriminant analysis. Varietal discrimination

With the aim of confirm the findings of the section 3.2, LDA were applied to the walnut spectral data to obtain discrimination methods for an easy classification of the different varieties. Spectral matrixes were randomly divided into calibration and validation sets and the respective spectral outliers were removed from these sets. Previously to the development of the discriminant analysis, the statistical software carried out PCAs with the calibration data. Then, it automatically selected the scores of the principal components obtained and developed the discriminant analysis from these variables.

Table 1 shows the results obtained: the number and the percentage of samples correctly classified in internal and external validation of the methods developed from the average walnut spectra and from the randomly selected spectra. Quite good results were obtained in any case, although methods developed from average spectra were more accurate than those obtained from the random samples. Random selection tries to imitate the random spectra acquisition in a separation line, where the orientation of the walnut can not be easily controlled. Taking that into account, the results obtained in the case of randomly selected samples are really interesting, i.e., 88 and 70% for FT-NIR and 92 and 74% for hyperspectral imaging of correct classification in internal and external validation respectively. The MicroNIR device produced somewhat worse results that could perhaps be improved by using a different measurement port with better adaptation to the surface of the walnut. It is also necessary to take into account other premises such as the preliminary nature of this study and the possibility of some minor contaminations of each varietal sample set by walnuts of different varieties during the collection process. This last premise is due to the characteristics of walnut collection methodology. In general, walnuts are collected from the ground and it is possible that fruits of different varieties, coming from neighbour trees, are collected by mistake.

Similar deductions can be inferred when LDA results are expressed in form of sensitivity and specificity of the models (Fig. 4). In the receiver operating characteristic (ROC) curves, it can be appreciate that the major difference between the models developed from average and random spectra is a loss of sensitivity. There are sensitivity losses of 5 and 11% in internal and external validations whereas in the case of specificity, these losses only reach 1 and 3% respectively. Therefore, in random models the number of false negatives, samples that belong to a specific variety but that are mistakenly identified as a different variety, is quite bigger than in the average models. However, in random and average models, the number of false positives, samples that belong to different varieties but that are mistakenly identified as the variety studied, is similar.

Conclusions

Near infrared (NIR) spectroscopy is an interesting and suitable technique for the study of in-shell walnuts. Different spectroscopic devices have been tested for NIR spectra acquisition and it has been developed methods for walnut varietal discrimination. In order to achieve good results, it is recommendable to obtain most of the spectral information of the sample. However, based on the preliminary results of this study, it could be possible to obtain suitable methods for varietal discrimination starting from individual and random walnut spectra. Moreover, FT-NIR and hyperspectral devices seem to be the more useful tools for varietal discrimination than MicroNIR portable device. Nevertheless, more comprehensive studies should be developed in order to obtain more robust and accurate methods.

Acknowledgments

The authors acknowledge the collaboration of FRUTECO the producer’s association for supplying samples and the assistance of the technical staff. This study was supported by FEDER and by National funds through the Programa Operacional Regional ALENTEJO 2020 (ALT20-

03-0145- FEDER-000005 – Eficiencia da tecnologia NIR para avaliação da maturação e qualidade de frutos). This work was also funded by National Funds through FCT – Foundation for Science and Technology under the Project UID/AGR/00115/2013.

References

Amaral, J. S., Casal, S., Pereira, J. A., Seabra, R. M., & Oliveira, B. P. P. (2003). Determination of Sterol and Fatty Acid Compositions, Oxidative Stability, and Nutritional Value of Six Walnut (Juglans regia L.) Cultivars Grown in Portugal. *Journal of Agricultural and Food Chemistry, 51*(26), 7698-7702.

Amaral, J. S., Cunha, S. C., Alves, M. R., Pereira, J. A., Seabra, R. M., & Oliveira, B. P. P. (2004). Triacylglycerol Composition of Walnut (Juglans regia L.) Cultivars: Characterization by HPLC-ELSD and Chemometrics. *Journal of Agricultural and Food Chemistry, 52*(26), 7964-7969.

Baca-Bocanegra, B., Nogales-Bueno, J., García-Estévez, I., Escribano-Bailón, M. T., Hernández-Hierro, J. M., & Heredia, F. J. (2019). Screening of Wine Extractable Total Phenolic and Ellagitannin Contents in Revalorized Cooperage By-products: Evaluation by Micro-NIRS Technology. *Food and Bioprocess Technology, 12*(3), 477-485.

Bou Abdallah, I., Baatour, O., Mechrgui, K., Herchi, W., Albouchi, A., Chalghoum, A., & Boukhchina, S. (2016). Essential oil composition of walnut tree (Juglans regia L.)' leaves from Tunisia. *Journal of Essential Oil Research, 28*(6), 545-550.

Ciarmiello, L. F., Piccirillo, P., Pontecorvo, G., De Luca, A., Kafantaris, I., & Woodrow, P. (2011). A PCR based SNPs marker for specific characterization of English walnut (Juglans regia L.) cultivars. *Molecular Biology Reports, 38*(2), 1237-1249.

Davis, L., Stonehouse, W., Loots, D. T., Mukuddem-Petersen, J., van der Westhuizen, F. H., Hanekom, S. M., & Jerling, J. C. (2007). The effects of high walnut and cashew nut diets on the antioxidant status of subjects with metabolic syndrome. *European Journal of Nutrition, 46*(3), 155-164.

Ercisli, S., Sayinci, B., Kara, M., Yildiz, C., & Ozturk, I. (2012). Determination of size and shape features of walnut (Juglans regia L.) cultivars using image processing. *Scientia Horticulturae, 133*, 47-55.

Gu, X., Zhang, L., Li, L., Ma, N., Tu, K., Song, L., & Pan, L. (2018). Multisource fingerprinting for region identification of walnuts in Xinjiang combined with chemometrics. *Journal of Food Process Engineering, 41*(4), e12687.

Guerrero, L., Romero, A., Gou, P., Aleta, N., & Arnau, J. (2000). Perfil sensorial de diferentes muestras de nuez (Juglans regia L.)/Sensory profiles of different walnuts (Juglans regia L.). *Food Science and Technology International, 6*(3), 207-216.

Instituto Nacional de Estatística, I. P. (2017). Estatísticas Agrícolas 2017. In I. P. Instituto Nacional de Estatística (Ed.), 2018 ed.). Lisboa, Portugal: Instituto Nacional de Estatística, I.P.

International Nut and Dried Fruit Council (INC). (2018). Nut and dried fruit statistical yearbook. In). Reus (Spain).

Kris-Etherton, P. M., Yu-Poth, S., Sabaté, J., Ratcliffe, H. E., Zhao, G., & Etherton, T. D. (1999). Nuts and their bioactive constituents: effects on serum lipids and other factors that affect disease risk. *The American Journal of Clinical Nutrition, 70*(3), 504s-511s.

Lacar, F. M., Lewis, M. M., & Grierson, I. T. (2001). *Use of hyperspectral reflectance for discrimination between grape varieties*. New York: Ieee.

Larrosa, M., García-Conesa, M. T., Espín, J. C., & Tomás-Barberán, F. A. (2010). Ellagitannins, ellagic acid and vascular health. *Molecular Aspects of Medicine, 31*(6), 513-539.

Ma, Q., Zhang, J., & Pei, D. (2011). Genetic Analysis of Walnut Cultivars in China Using Fluorescent Amplified Fragment Length Polymorphism. *136*(6), 422.

Martínez, M. L., Mattea, M. A., & Maestri, D. M. (2006). Varietal and crop year effects on lipid composition of walnut (Juglans regia) genotypes. *Journal of the American Oil Chemists' Society, 83*(9), 791-796.

Milinovic, J., Garcia, R., Rato, A. E., & Cabrita, M. J. (2019). Rapid Assessment of Monovarietal Portuguese Extra Virgin Olive Oil's (EVOO's) Fatty Acids by Fourier-Transform Near-Infrared Spectroscopy (FT-NIRS). *European Journal of Lipid Science and Technology, 121*(3), 1800392.

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.

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.

Peng, J., Liu, F., Kong, W., Zhang, C., Yu, K., & He, Y. (2017). Rapid Identification of Varieties of Walnut Powder Based on Laser-Induced Breakdown Spectroscopy. *Transactions of the ASABE, 60*(1), 19-28.

Pereira, J. A., Oliveira, I., Sousa, A., Ferreira, I. C., Bento, A., & Estevinho, L. (2008). Bioactive properties and chemical composition of six walnut (Juglans regia L.) cultivars. *Food and Chemical Toxicology, 46*(6), 2103-2111.

Perez, D. P., Sanchez, M. T., Cano, G., & Garrido, A. (2001). Authentication of green asparagus varieties by near-infrared reflectance spectroscopy. *Journal of Food Science, 66*(2), 323-327.

Pop, I. F., Vicol, A. C., Botu, M., Raica, P. A., Vahdati, K., & Pamfil, D. (2013). Relationships of walnut cultivars in a germplasm collection: Comparative analysis of phenotypic and molecular data. *Scientia Horticulturae, 153*, 124-135.

Rabadán, A., Pardo, J. E., Gómez, R., & Álvarez-Ortí, M. (2018). Evaluation of physical parameters of walnut and walnut products obtained by cold pressing. *LWT, 91*, 308-314.

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*(0), 145-150.

Sinesio, F., & Moneta, E. (1997). Sensory evaluation of walnut fruit. *Food Quality and Preference, 8*(1), 35-43.

Teixeira, A. M., & Sousa, C. (2019). A review on the application of vibrational spectroscopy to the chemistry of nuts. *Food Chemistry, 277*, 713-724.

Yi, J., Sun, Y., Zhu, Z., Liu, N., & Lu, J. (2017). Near-infrared reflectance spectroscopy for the prediction of chemical composition in walnut kernel. *International Journal of Food Properties, 20*(7), 1633-1642.

Figure captions

Figure 1. Graphical description of the spectroscopic data acquisition.

Figure 2. Hyperspectral image segmentation. Description of the whole procedure and software used in each step: image acquisition, calibration, linear discriminant analysis, segmentation, numbering and spectra extraction.

Figure 3. Score plot of the principal components that show varietal differences in a better way. PCAs were performed from the following spectral data: a) average FT-NIR spectra, b) average hyperspectral image spectra, c) average MicroNIR spectra, d) random FT-NIR spectra, e) random hyperspectral image spectra and f) random MicroNIR spectra.

Figure 4. Receiver operating characteristic (ROC) curves of different LDA methods developed for average and random spectra.