

Feasibility of FT-NIR spectroscopy and Vis/NIR hyperspectral imaging for sorting unsound chestnuts

Giacomo Bedini¹, Swathi Sirisha Nallan Chakravartula¹, Giorgia Bastianelli¹, Romina Caccia¹, Mario Contarini², Carmen Morales-Rodríguez¹, Luca Rossini², Stefano Speranza², Andrea Vannini¹, Roberto Moschetti^{1*}, Riccardo Massantini^{1*}

¹ Department for Innovation in Biological, Agro-food and Forest System (DIBAF), University of Tuscia, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy

² Department of Agricultural and Forestry Sciences (DAFNE), University of Tuscia, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy

* Corresponding author: rmoschetti@unitus.it; massanti@unitus.it; Tel.: +39-0761-357496

Received: 15 February 2020; Accepted: 20 April 2020; Published: 30 April 2020

Abstract: Authors explored the potential use of Vis/NIR hyperspectral imaging (HSI) and Fourier-transform Near-Infrared (FT-NIR) spectroscopy to be used as in-line tools for the detection of unsound chestnut fruits (i.e. infected and/or infested) in comparison with the traditional sorting technique. For the intended purpose, a total of 720 raw fruits were collected from a local company. Chestnut fruits were preliminarily classified into sound (360 fruits) and unsound (360 fruits) batches using a proprietary floating system at the facility along with manual selection performed by expert workers. The two batches were stored at 4 ± 1 °C until use. Samples were left at ambient temperature for at least 12 h before measurements. Subsequently, fruits were subjected to non-destructive measurements (i.e. spectral analysis) immediately followed by destructive analyses (i.e. microbiological and entomological assays). Classification models were trained using the Partial Least Squares Discriminant Analysis (PLS-DA) by pairing the spectrum of each fruit with the categorical information obtained from its destructive assay (i.e., sound, $Y = 0$; unsound, $Y = 1$). Categorical data were also used to evaluate the classification performance of the traditional sorting method. The performance of each PLS-DA model was evaluated in terms of false positive error (FP), false negative error (FN) and total error (TE) rates. The best result (8% FP, 14% FN, 11% TE) was obtained using Savitzky-Golay first derivative with a 5-points window of smoothing on the dataset of raw reflectance spectra scanned from the hilum side of fruit using the Vis/NIR HSI setup. This model showed similarity in terms of False Negative error rate with the best one computed using data from the FT-NIR setup (i.e. 15% FN), which, however, had the lowest global performance (17% TE) due to the highest False Positive error rate (19%). Finally, considering that the total error rate committed by the traditional sorting system was about 14.5% with a tendency of misclassifying unsound fruits, the results indicate the feasibility of a rapid, in-line detection system based on spectroscopic measurements.

Keywords: *Castanea sativa* Mill.; healthy; unhealthy; hyperspectral imaging; discriminant analysis; NIR spectroscopy.

1. Introduction

Although chestnut supply chain is a niche sector at national level, Italy is the fourth largest producer of chestnuts in the world with around 53,000 Mg (FAOSTAT, 2020) and the largest exporter in terms of product value, followed by China. At present, the chestnut supply chain shows a modest technological level which, historically, has never been a problem to ensure a product marketable in terms of food safety and quality. However, in the last 15-20 years, supply chain's technology has failed to counteract the effects of climate change and the proliferation of native and non-native pathogens, which have significantly reduced the production and increased the incidence of hidden defects in the product. The consequences were catastrophic with the chestnut sector losing important market share overtime with product losses up to 90% in some production seasons (*e.g.* 2016). Among the issues concerning the sustainability of chestnut farming in Italy, the most significant are the health and marketability of the fruit which are compromised by impact damage and diseases caused by insects and fungal pathogens. These are responsible for (i) hidden damages and (ii) mycotoxin contamination due to which chestnut fruit is considered as an at-risk product (Rodrigues et al., 2012). The fungal species generally responsible for developing rot or mold are *Penicillium* spp., *Aspergillus* spp., *Fusarium* spp., *Gnomoniopsis castanea* Tamietti, *Phomopsis castanea* Sacc. Petr., *Acrospeira mirabilis* Berk. & Broome e *Sclerotinia pseudotuberosa* Rehm. (González et al., 2010). Fungal infection usually develops from flicker holes, larval tunnels and splits, as well as during the storage of fruits that are not properly dried after the washing and/or flotation process. Among the fungal parasites, *Gnomoniopsis castanea* Tamietti predominantly raises concerns considering the extent of the damage it causes. Most fungal contaminations are related to primary infestation of entomological nature. The major chestnut parasitic insects are tortrices (*Cydia splendana* Hb., *Cydia fagiglandana* Zel. e *Pammene fasciana* L.) and weevils (*Curculio elephas* Gyll.) (Paparatti and Speranza, 2005). The larva develops in the fruit and feeds on the amylose substratum of flesh. Upon completion of the larval stage (30–45 days) the insect exits from fruit and drops to the ground, leaving a hole of approximately 1-mm diameter. Although damage increases with the development of the larvae, even a very small attack may compromise the quality of the fruit. The exit hole provides external evidence of a previous infestation and therefore suggests internal damage.

Under existing production practices, the most widespread method for detecting unsound chestnuts involves floating, frequently combined with thermo-hydrotherapy (Sieber et al., 2007), followed by manual sorting. However, the traditional method is unreliable because it tends to discard excessive amount of sound product. Moreover, manual sorting is time-consuming, subjective, labour intensive and does not identify chestnuts with hidden damage. Consequently, the development of rapid, automatic and non-destructive techniques for chestnut sorting would be beneficial to food industry by reducing incidence of defective products, which are responsible for serious damage to brand reputation.

A variety of techniques have been reported for non-destructive selection of fruit and vegetables, including those based on electrical properties: near-infrared (NIR) spectroscopy (Saranwong et al., 2011), sound/noise/vibration (Liljedahl and Abbott, 1994), ultrasound, nuclear magnetic resonance (Zhang and McCarthy, 2013), X-ray (Haff and Toyofuku, 2008)

and X-ray computed tomography imaging (Donis-González et al., 2014) and others (Nicolai et al., 2007; Singh et al., 2010; Wang et al., 2010). Among these non-destructive techniques, NIR spectroscopy presents several advantages including minimal need for sample preparation and a wide range of applications in the food sector. In addition, it is fast, environmental friendly, easy to apply and highly suited for rapid on-line inspection (Pasquini, 2003; Wang et al., 2010). Its efficacy in detecting and classifying fungal infections has been demonstrated on food products such as soybeans (Wang et al., 2004) and dates (Teena et al., 2014). Moreover, NIR spectroscopy was proved to be effective for the detection of insects or insect damage in various food commodities: e.g., blueberry (Peshlov et al., 2009), cherry (Xing and Guyer, 2008; Xing et al., 2008), fig (Burks et al., 2000), flour (Wilkin et al., 1986), green soybean (Sirisomboon et al., 2009), jujube (Wang et al., 2011), wheat (Baker et al., 1999), seeds of *Picea abies* L. (Tigabu et al., 2004) and seeds of *Cordia africana* Lam. (Tigabu and Odén, 2002). Insects and larvae can be detected directly due to their haemolymph, lipids and/or chitin content (Rajendran, 2005; Moschetti et al., 2014a) or indirectly due to subsequent damages such as internal browning or darkening, dehydration or fungal contamination (Wang et al., 2011). The efficiency of NIR spectroscopy to identify insect damage or infestation depends on the type of spectroscopic measurement (transmittance, reflectance and interactance) (Wang et al., 2010; Wang et al., 2011). Most NIR spectroscopy applications described in the literature are based on spot measurements. Moreover, NIR spectrophotometers are conveniently classified according to the type of monochromator which may affect the speed and quality of measurements. Specifically, Fourier Transform (FT) spectrophotometers use an interferometer in order to generate modulated light in which the time domain signal of the light reflected or transmitted by the sample onto the detector can be converted into a spectrum via a fast Fourier transform (Nicolai et al., 2007). Various authors have observed radial and circumferential quantitative variations of certain quality parameters in different fruits (Peiris et al., 1999) such as kiwi (Martinsen and Schaare, 1998) and melons (Long and Walsh, 2006). This has an impact on the application of single point spectroscopy for fruit and vegetables analysis. In fact, depending on the heterogeneity of the analysed chemical parameter, different values can be obtained depending on where the acquisition was made. Therefore, it might be necessary to perform the spectral acquisition at several positions on the same fruit or vegetable, which may result inefficient when, for example, skin disorders are to be detected. In such cases hyperspectral imaging (HSI) is suggested; in fact, through the concurrent acquisition of the spatial and spectral domains in a 3-D matrix (or hypercube), HSI provides spectral information at pixel level from the whole surface of the sample. In both NIR and HSI, data handling is almost always required and represents a crucial step before analysis (Wu and Sun, 2013). Nicolai et al. (2007), Burger and Gowen (2011) give exhaustive overviews of the most common chemometric methods that are useful in dealing with issues related to data handling of conventional NIR dataset and hypercube, which are usually affected by a “curse of dimensionality”. The hypercube is a complex data structure characterized by both spatial and spectral domains as well as signal noise and high-correlated features. Thus, HSI data needs to be carefully handled in order to enhance and discard important and useless information

respectively, and consequently reduce requirements in computational load. Finally, proper data handling in HSI is essential for successful implementation in on-line or in-field systems.

The objective of the present study was to demonstrate the feasibility of using both conventional and imaging NIR spectroscopy for the detection of unsound chestnuts, exploring their potential to be fitted, or retrofitted, into wide range of sorting systems.

2. Materials and Methods

2.1. Samples preparation

Two batches of sound (S) and unsound (U) chestnut fruits of approximately 5 kg each (*Castanea sativa* Mill. Cv. Marrone Fiorentino), harvested in October 2018, were sampled from a local storage facility (Mastrogregori s.r.l., Viterbo, Italia) and sorted through the traditional method (i.e. floating followed by manual sorting). The batch S (sound) corresponded to the chestnuts sorted as healthy fruits, while the batch U (unsound) to chestnuts sorted as infested/infected fruits. The batches were transferred to the laboratory, and further selected to remove impact-damaged and/or germinated chestnuts and stored at 4 ± 1 °C until further analyses. A total of 720 fruits were selected (360 per batch, S and U) for the analyses and each fruit was labelled with a unique ID. Prior to the spectrophotometric analysis, samples were left at room temperature for at least 12 h.

2.2. Spectral measurements (i.e., non-destructive assays)

2.2.1. Vis/NIR hyperspectral imaging

The hyperspectral scans were performed using the hyperspectral camera (mod. PFD4K-CL-65-V10E, 400-1000 nm, ~0.78 nm resolution) equipped with a lens mod. OLE23 with focal length 23 cm and number F 2.4 (Specim Spectral Imaging Ltd., Oulu, Finland). The distance between the camera and the samples was 30 cm while, the lighting source consisted of 6 halogen lamps of 35W (mod. DECOSTAR 51 ALU 41866 WFL, Osram, Munich, Germany). The Lumo Scanner 1.2 software (Specim Spectral Imaging Ltd., Oulu, Finland) was used for the control of the camera and transition stage (mod. Labscan 40×20, Specim Spectral Imaging Ltd., Oulu, Finland). The scanning speed was set to 3.24 mm s^{-1} with an exposure time of 35 ms. The hypercubes were acquired in reflectance at a resolution of 1788×1568 pixels (H×W). Before each scan, dark and white references were acquired for the calibration. The dark reference (0% of reflectance) was acquired with the camera shutter closed, while the white reference consisted of the scan of a 70%-reflective Teflon standard. The relative hyperspectral image (I) was obtained by applying the following equation (Eq. 1) (ElMasry, Wang and Vigneault, 2009):

$$(1) \quad I = \frac{I_0 - D}{W - D} \times 100$$

where I_0 is the hyperspectral raw image, W is the white reference and D is the dark reference. The calibration is used to transform the reflectance value of each pixel of the raw hyperspectral image (I_0) into relative reflectance values (I, unitless measurement). This

allows to compare hyperspectral images resulting from different samples and the spectra with those contained in the spectral libraries.

Each scan was performed on eight samples at a time, which were scanned on four sides: (i) convex side, C; (ii) flat side, P; (iii) hilum side, B and (iv) torch side, T (Fig. 1). Subsequently, the raw hypercube of each fruit was (i) clipped from the raw scan (Fig. 2), (ii) calibrated using the eq. 1, and (iii) segmented for removing the background and the edges in each HSI image. The resulting Region Of Interest (ROI, Fig. 3) was used to measure the mean reflectance spectrum of each chestnut fruit. A matrix of 720 samples (i.e. 360 sound and 360 unsound fruits) \times 776 wavebands (i.e. 400-1000-nm sensitivity, \sim 0.78-nm resolution) was acquired for each chestnut side. The flow chart of the procedure described above is shown in Figure 4. Matlab R2017b software was used to remove background from hypercubes (Mathworks Inc., Natick, USA).

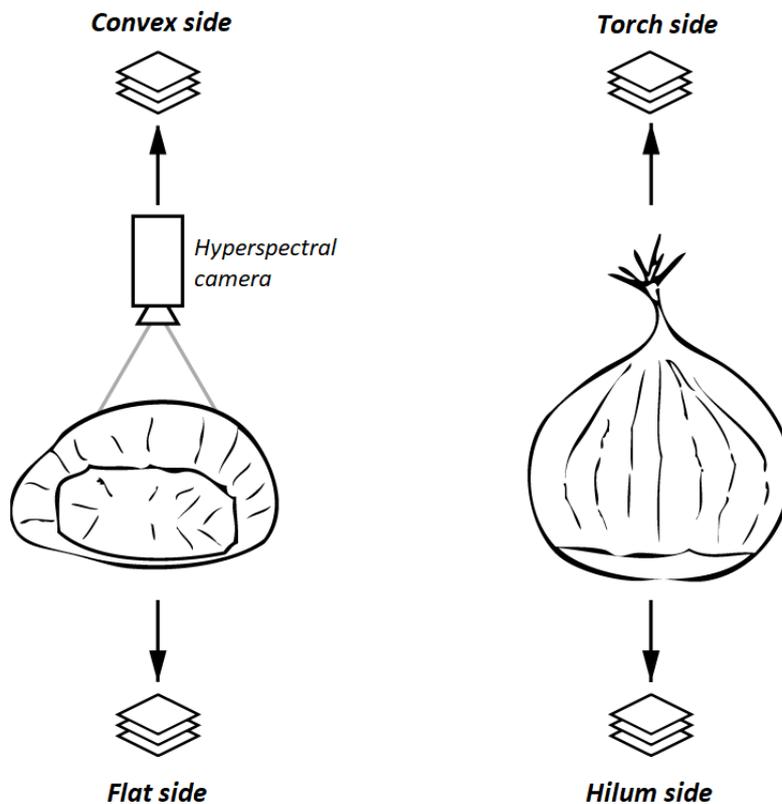


Figure 1. Overview of the four sides of hyperspectral image acquisition.

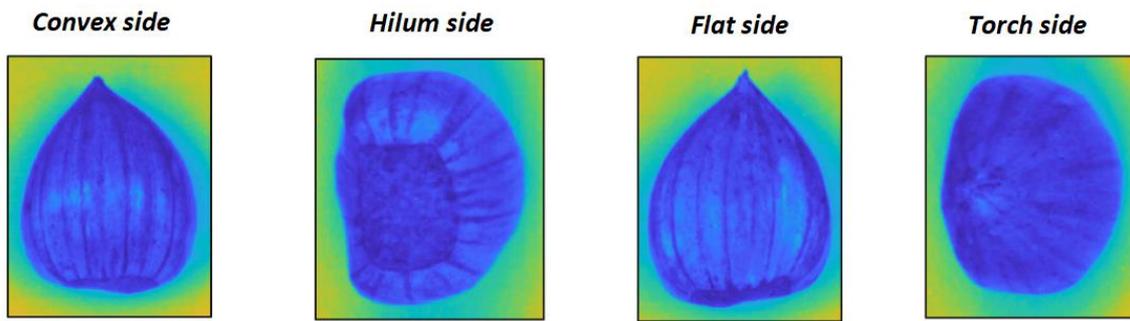


Figure 2: Clippings of the raw HSI scans of the four sides of sample.

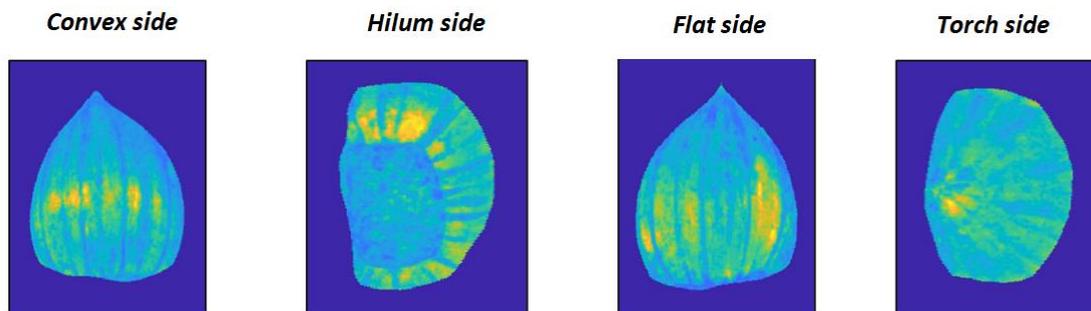


Figure 3. Region of interest (ROI) for the four scanned sides of sample; the background is in dark blue color.

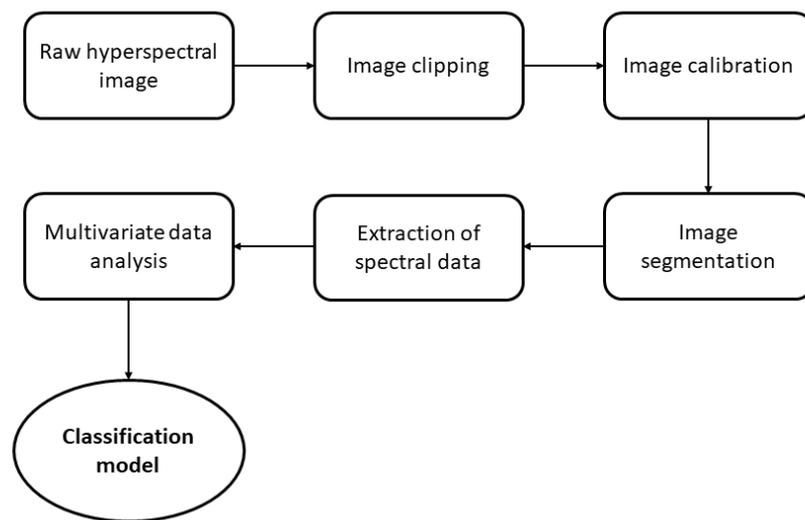


Figure 4. Flowchart of the main steps involved in the process of hyperspectral image analysis.

2.2.2. FT-NIR spectroscopy

The Antaris II spectrophotometer (Thermo Scientific, Madison, WI, USA) was used for acquisition of Fourier's transformed spectra in the 1000-2500 nm spectral region with a

resolution of 0.3 nm. The spectra were acquired in absorbance, using internal standard of the instrument as a reference. Each spectrum was an average of 30 scans and was acquired only on the flat side of the fruit (P). The dimensions of the resulting spectral matrix in terms of number of rows and number of columns (M×N) was equal to 720×5001.

2.3. Microbiological and entomological measurements (i.e., destructive assays)

Immediately after the non-destructive analysis, each fruit was longitudinally dissected into two halves using a sterilized scalpel. Sections were analysed for the identification of damage from fungal parasites and/or insects. As the first step, a pulp fragment was inoculated into a petri dish containing Potato-Destrose Agar (PDA) as growth medium, added with antibiotics (i.e. ampicillin and streptomycin). The plate was then incubated at 25 ± 1 °C for 7 days and the fungal pathogens were identified. As the second step, the dissected fruit was visually evaluated to identify and classify the species of insects responsible for any damage.

Each fruit was reclassified into sound ($Y = 0$) and unsound ($Y = 1$) as per the procedure outlined in the flow chart (Fig. 5). The Y vector was used as binary-class vector for the supervised learning of classification algorithm.

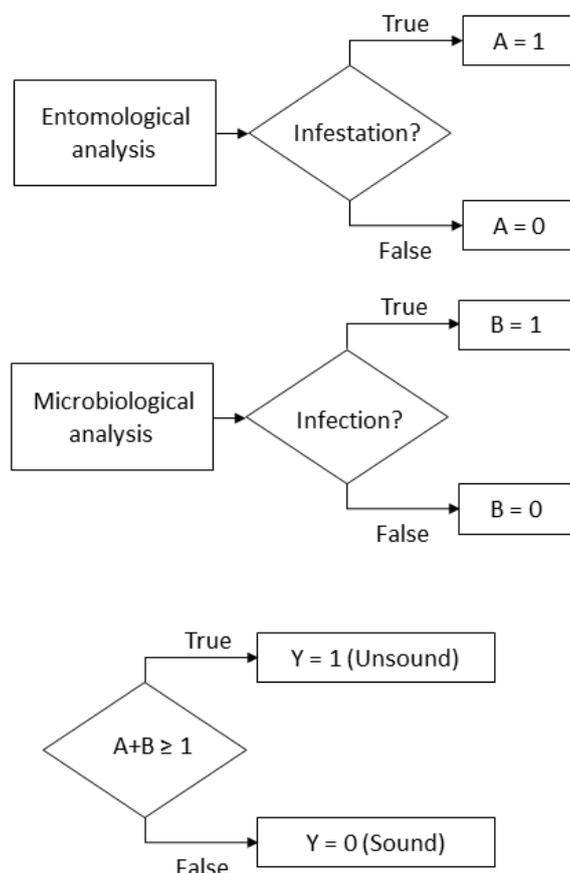


Figure 5. Flowchart outlining the reclassification of fruits based on entomological and microbiological analyses.

2.4. Chemometrics

2.4.1. Data handling and spectral pre-treatments

Matlab R2017b software with PLS_Toolbox v8.5.6 software (Eigenvector Research Inc., WA, USA) was used for data handling and model computing. Chemometrics was performed following a combination of spectral pre-treatments, including: Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), Savitzky-Golay first and second derivatives (*D1f* and *D2f*, respectively) with a second or third order polynomial fitted over a window of five (S5), seven (S7), nine (S9) or eleven (S11) features (Savitzky and Golay, 1964; Boysworth and Booksh, 2008) and Mean Centering (MC). Every possible combination of pre-processing was also tested and only best results, in terms of model performances, were retained.

2.4.2. Classification models

Classification models were computed using Partial Least Squares Discriminant Analysis (PLS-DA) using the SIMPLS algorithm (de Jong, 1993). PLS-DA, a variant of PLS, was used for research of useful linear combinations between independent variables, with the aim of obtaining a more stable regression, by discarding the most irrelevant and unstable information which allowed to resolve any problem of collinearity. This was chosen because the Y variable was categorical. In fact, the classes identified by destructive analysis (sound as S and unsound as U) were used as response variable (Y) in the development of PLS-DA models where predictive variables (X) were represented by the spectral profiles of each chestnut fruit. The optimal number of latent variables (LVs) was selected as described by Moschetti et al. (2017, 2018). The classification performance of PLS-DA models was determined in terms of False Positive rate or First Type Error (FP – Eq. 2), False Negative rate or second-type error (FN – Eq. 3) and Total Error rate (TE - Eq. 4) and reported for both calibration (C) and cross-validation (CV) procedures (Fawcett, 2006):

$$(2) \quad \text{False Positive rate (FP)} = \frac{\text{False positive}}{\text{False positive} + \text{True negative}} = 1 - \text{specificity}$$

$$(3) \quad \text{False Negative rate (FN)} = \frac{\text{False negative}}{\text{False negative} + \text{True positive}} = 1 - \text{sensitivity}$$

$$(4) \quad \text{Total Error rate (TE)} = \frac{\text{False positive} + \text{False negative}}{\text{Total Positive} + \text{Total Negative}} = 1 - \text{accuracy}$$

wherein, True Positive (TP) corresponds to the number of sound samples correctly classified as sound; True Negative (TN), the number of unsound samples correctly classified as unsound; False Negative (FN), the number of sound samples incorrectly classified as unsound; and False Positive (FP), the number of unsound samples incorrectly classified as sound. A test with a low False Negative rate rarely classifies a sample as unsound even though there is no damage. The sound class sensitivity takes values between 0 and 1 and describes the model ability to correctly recognize samples belonging to that class; conversely, a test with a low False Positive rate rarely classifies as positive an unsound sample. Also, the sound class specificity takes values between 0 and 1 and describes the model ability to reject

samples of all other classes. A perfect predictor would be described as 100% sensitive, which means that all sound samples are correctly identified as sound, and 100% specific, that no unsound sample is wrongly identified as sound. The Total Error rate (TE) of a binary model is calculated as the arithmetic mean of the FP and FN error rates. TE was used to select models in terms of predictability, while a venetian blinds cross-validation with 10 splits was used to select the proper number of LVs and to evaluate each model in terms of robustness (i.e. the model ability to withstand small variations in the operative conditions of test) (Ballabio and Consonni, 2013). The cross-validation method was preferred because of its capability of providing a better representation when the dataset contains many samples (Wise, 2009). The outliers' analysis was not carried out.

3. Results

3.1. Spectral measurements (i.e., non-destructive assays) – an overview

3.1.1. Vis/NIR hyperspectral imaging

HSI technology allowed to acquire spectra that can be divided into two ranges: the visible band (Vis) from 400 nm to 780 nm, and the near infrared band (NIR) from 780 nm to 1000 nm. Figure 6 shows the reflectance spectra acquired for the four sides of the fruit using the Vis/NIR hyperspectral camera. Bearing in mind that a “by eye” evaluation of spectral data may lead to misleading conclusions, the average spectra apparently showed the evidence of spectral differences between sound and unsound fruits using HSI.

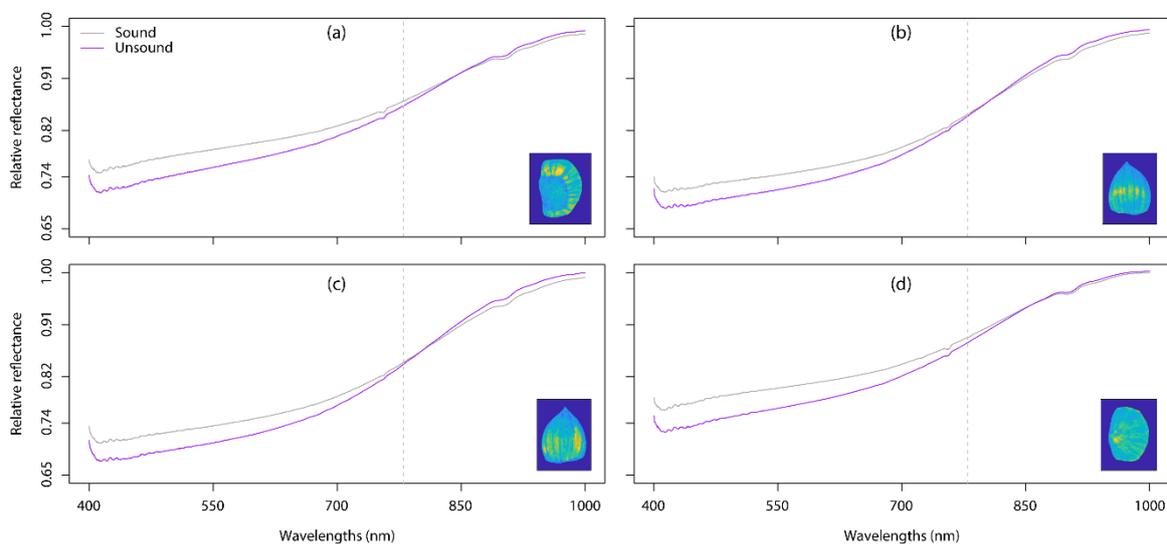


Figure 6. Vis/NIR averaged spectra (i.e. from 400 to 1000 nm) of sound and unsound chestnuts acquired with hyperspectral camera on (a) hilum, (b) convex, (c) flat and (d) torch side of fruit.

3.1.2. FT-NIR spectroscopy

Figure 7 shows both sound and unsound average raw spectra acquired with the FT-NIR Antaris II spectrophotometer. The spectra showed higher absorbance in infested fruits. This is confirmed by Moscetti et al. (2014a) and might be explained with a decrease in reflectance due to the majority of the light beam absorbed by the infested tissue, as pointed out by Wang et al. (2011).

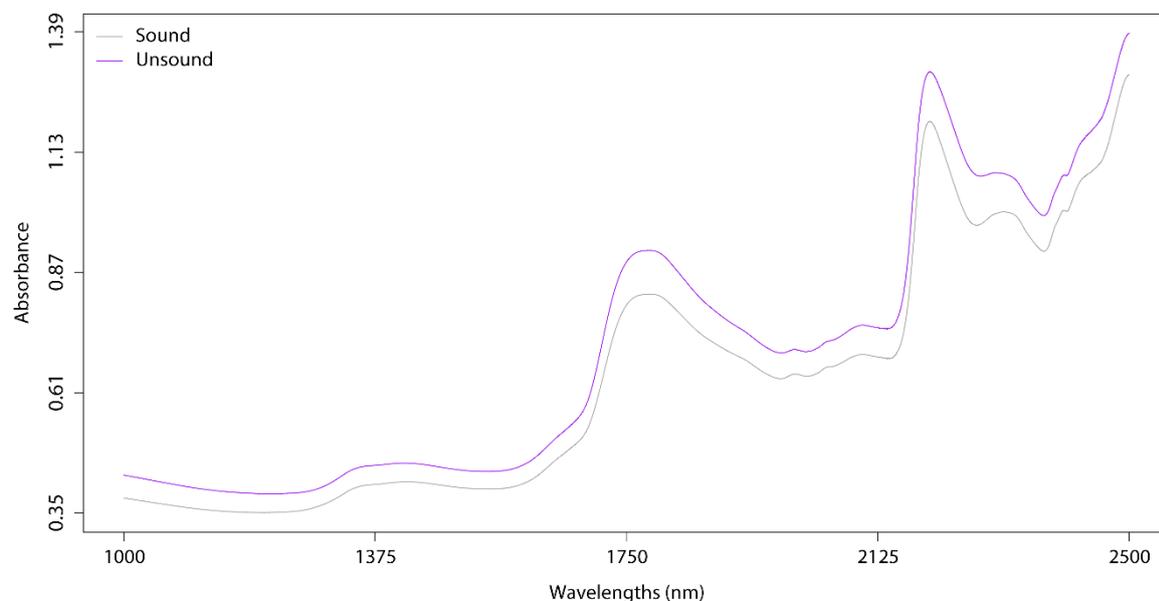


Figure 7. NIR averaged spectra (i.e., from 1000 to 2500 nm) of sound and unsound chestnuts acquired with FT-NIR spectrophotometer on the flat side (P) of fruit.

3.2. Microbiological and entomological measurements (i.e., destructive assays)

Based on the results obtained from various surveys performed in different production seasons in central Italy (Viterbo), the total error rate of the traditional sorting operations (flotation and/or manual sorting) may vary considerably from 15 to 20%. Data collected during the 2018 season showed a first type error (or False Positive, FP) equal to 16.1%, which outweighs over the second type error (or False Negative, FN) of 12.5%. Therefore, the total error of classification (TE) committed by the industry was about 14.5%, with a tendency of misclassifying unsound fruits. These results were in accordance with those observed by (Moscetti et al., 2014a; 2014b) in the 2012 production season.

3.3. Classification models

Performance metrics of each PLS-DA model were obtained using the class labels from destructive analysis as response variable (Y) and the average spectrum of each chestnut fruit as predictive variable (X) as shown in Tab. 1.

The best fitting models were all obtained without any pre-treatment of optical dispersion removal (or scattering). FT-NIR data was successfully pre-treated by applying a Savitzky-Golay filter with a 5-points smoothing window followed by the Mean Centering, while Vis/NIR HSI data required a first derivative Savitzky-Golay filter with a 5-points smoothing window and a subsequent Mean Centering.

Table 1. Performance of PLS-DA models for chestnut sorting.

| Instrumentation | Scanning side | Spectral pre-treatment | | | | LVs | FP | | FN | | TE | |
|-----------------|---------------|------------------------|----|------|-------|-----|------|------|------|------|------|------|
| | | Scatt. C. | SG | Der. | Norm. | | C | CV | C | CV | C | CV |
| FT-NIR | P | - | 15 | - | MC | 13 | 0.19 | 0.19 | 0.14 | 0.15 | 0.16 | 0.17 |
| HSI (Vis/NIR) | B | - | 5 | 1st | MC | 14 | 0.07 | 0.08 | 0.14 | 0.14 | 0.10 | 0.11 |
| | C | - | 5 | 1st | MC | 13 | 0.07 | 0.08 | 0.15 | 0.15 | 0.11 | 0.12 |
| | P | - | 5 | 1st | MC | 13 | 0.07 | 0.09 | 0.14 | 0.14 | 0.11 | 0.11 |
| | T | - | 5 | 1st | MC | 11 | 0.07 | 0.08 | 0.14 | 0.15 | 0.10 | 0.11 |

Scatt. C., scatter correction; SG, Savitzky-Golay filter; Der., derivative; Norm, normalization (or mean centering); FP, false positive; FN, false negative; TE, total error; C, calibration; CV, cross-validation; P, flat side; B, hilum side; C, convex side; T, torch side

4. Discussion

The spectral analysis in the visible range (i.e., ~400-780 nm) is extremely useful for the quanti-qualitative assessment of pigments. In this spectral region, colour differences between samples are mainly described. Whereas, the ~780-2500-nm spectral band corresponds to the NIR region, which is particularly suitable for the analysis of chemical compounds (Polesello et al., 1981). In fact, NIR bands are mostly represented by overtones containing information about various molecular vibrations and functional groups, e.g. C-H, N-H, O-H, etc. (Polesello et al., 1981). Therefore, classification models based on the Vis/NIR spectral range are mainly colour oriented, while discriminant algorithms computed using NIR wavelengths are trained only for chemical pattern recognition. Bearing in mind that the FT-NIR and the Vis/NIR HSI setups used in the experimentation are part of different technical and technological paradigms, the previous assumption allows to speculate that one of the most significant reasons behind the difference in model performances between the two setups is the spectral range in which models were trained. Thus, NIR-based models were affected by the major constituents of sample. Raw chestnuts are mainly composed by water (~48.5%), carbohydrates (~45.5%), fibers (~8.0%), proteins (~2.5%), and lipids (~2.0%) (USDA, 2019). The water molecule has very significant NIR spectral bands, which show overtones at 760, 970, 1180, and 1450 nm and a combination at 1940 nm (Polesello et al., 1981). Considering that larvae metabolic activity and mould development may be responsible for changes in temperature and moisture content of chestnut (Rajendran, 2005), the performance of NIR-based classification models may be affected by the fruit health.

The spectral pre-treatments selection also highlighted that different approaches were needed from an instrument to another to improve spectral quality and then the final model performance. The only exception was the recourse to the scattering correction algorithm, which was not required for the treatment of spectra from both setups. In fact, light scattering is a physical characteristic of sample, and its correction is not frequently needed for the achievement of classification tasks, especially when light scattering significantly contributes to the between-variance classes. The baseline correction through the first derivative was fundamental only for HSI data, while, as expected, the Savitzky-Golay smoothing window size was larger for FT-NIR data, because of its highest spectral resolution. Finally, mean

centering was effective in improving classification in all cases. The effectiveness of spectral pre-treatments was in agreement with what was previously reported by Moscetti et al. (2014a).

In general, PLS-DA models developed from Vis/NIR HSI data had a higher effectiveness in product sorting than the FT-NIR-based model. Specifically, the best HSI model was obtained through the scan of the hilum-scar side of fruit, as previously observed by Moscetti et al. (2014a) and Moscetti et al. (2014b), with a total error rate equal to 10% and 11% in calibration and cross-validation, respectively. On the other hand, FT-NIR performed worse with a total error rate of 15% and 16% in calibration and cross-validation, respectively, due to the highest FP error rate (~19%). However, it should be specified that the technical characteristics of the Antaris II instrument allowed to scan the fruit on all four sides. Therefore, further investigation will be necessary before confirming a superior technology in terms of detection performance of chestnut fruit affected by hidden damage.

Finally, results represent an improvement over the conventional sorting system, which usually has a total error rate of ~15-20%. Nevertheless, all the models performed slightly worse than those reported by Moscetti et al. (2014a and 2014b) using spectra acquired from an Acousto-Optic-Tunable-Filter (AOTF) NIR spectrophotometer. For these reasons, further tests will be needed for the improvement of robustness and resilience of the models over the harvest seasons.

5. Conclusions

The feasibility of using FT-NIR spectroscopy and Vis/NIR hyperspectral imaging for chestnut sorting is confirmed by the investigation conducted during this experiment. Compared to traditional floating system, the performance of the techniques tested in this study is slightly better; in fact, the total error rate (%) obtained using HSI is approximately four points lower than traditional method. The best result obtained by analysing the hilar scar side suggests that an on-line sorting device should have a pivoting chestnut conveyor located up-line. However, the design of a sorting device based on the methods presented here should need to balance the need for simplicity of the system with cost and performance compared to other competitors. Further, the obtained results provide the basis for a superior detection system in terms of sensitivity, selectivity, non-destructive analysis and automation. However, in addition to further investigation of the possible use of FT-NIR spectroscopy, more research will be needed to test the effectiveness of new discriminating analysis algorithms, filters and image analysis methods, including the emerging and promising Deep Learning techniques: i.e., (i) Deep Chemometrics for spectra analysis and (ii) Convolutional Neural Networks (CNN) for the development of classification models.

Funding: This research was funded by the Latium region (Italy) through the SANCAST and FORECAST projects, and by ‘Departments of excellence 2018’ program (i.e. ‘Dipartimenti di eccellenza’) of the Italian Ministry of Education, University and Research for the financial support through the ‘Landscape 4.0 food, wellbeing and environment’ (DIBAF department of University of Tuscia).

Acknowledgments: The authors gratefully acknowledge Mastrogregori Aldo S.r.l. (Via del Ruscello, 16 - 01030 – Vallerano, Viterbo, Italy) for providing the samples; the master student Francesco Tonnetti for his valuable help and support during the experimentation; and Gianpaolo Moschetti for the English revision of the manuscript.

Conflicts of Interest: Authors have no conflicts of interest.

References

- Baker, J.E., Dowell, F.E. and Throne, J.E. (1999) 'Detection of parasitized rice weevils in wheat kernels with near-infrared spectroscopy', *Biological Control*, 16(1), pp. 88-90. doi: [10.1006/bcon.1999.0733](https://doi.org/10.1006/bcon.1999.0733)
- Ballabio, D. and Consonni, V. (2013) 'Classification tools in chemistry. Part 1: linear models. PLS-DA', *Analytical Methods*, 5(16), pp. 3790-3798. doi: [10.1039/c3ay40582f](https://doi.org/10.1039/c3ay40582f)
- Boysworth, M.K. and Booksh, K.S. (2008) 'Aspects of multivariate calibration applied to near-infrared spectroscopy', in Burns, D.A. and Ciurczak, E.W. (eds) *Handbook of Near-Infrared Analysis*. 3rd edn. New York: CRC Press, pp. 207-229.
- Burger, J. and Gowen, A. (2011) 'Data handling in hyperspectral image analysis', *Chemometrics and Intelligent Laboratory Systems*, 108(1), pp. 13-22. doi: [10.1016/j.chemolab.2011.04.001](https://doi.org/10.1016/j.chemolab.2011.04.001)
- Burks, C.S., Dowell, F.E. and Xie, F. (2000) 'Measuring fig quality using near-infrared spectroscopy', *Journal of Stored Products Research*, 36(3), pp. 289-296. doi: [10.1016/S0022-474X\(99\)00050-8](https://doi.org/10.1016/S0022-474X(99)00050-8)
- Donis-González, I.R., Guyer, D.E., Fulbright, D.W. and Pease, A. (2014) 'Postharvest noninvasive assessment of fresh chestnut (*Castanea* spp.) internal decay using computer tomography images', *Postharvest Biology and Technology*, 94, pp. 14-25. doi: [10.1016/j.postharvbio.2014.02.016](https://doi.org/10.1016/j.postharvbio.2014.02.016)
- ElMasry, G., Wang, N. and Vigneault, C. (2009) 'Detecting chilling injury in Red Delicious apple using hyperspectral imaging and neural networks', *Postharvest Biology and Technology*, 52(1), pp. 1-8. doi: [10.1016/j.postharvbio.2008.11.008](https://doi.org/10.1016/j.postharvbio.2008.11.008)
- FAOSTAT (2020) *Production quantities of Chestnut by country (2018)*.
- Fawcett, T. (2006) 'An introduction to ROC analysis', *Pattern Recognition Letters*, 27(8), pp. 861-874. doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010)
- Donis González, I.R., Fulbright, D.W., Ryser, E.T. and Guyer, D. (2010) 'Shell mold and kernel decay of fresh chestnuts in Michigan', *Acta Horticulturae*, 866, pp. 353-362. doi: [10.17660/actahortic.2010.866.45](https://doi.org/10.17660/actahortic.2010.866.45)
- Haff, R.P. and Toyofuku, N. (2008) 'X-ray detection of defects and contaminants in the food industry', *Sensing and Instrumentation for Food Quality and Safety*, 2(4), pp. 262-273. doi: [10.1007/s11694-008-9059-8](https://doi.org/10.1007/s11694-008-9059-8)
- de Jong, S. (1993) 'SIMPLS: An alternative approach to partial least squares regression', *Chemometrics and Intelligent Laboratory Systems*, 18(3), pp. 251-263. doi: [10.1016/0169-7439\(93\)85002-x](https://doi.org/10.1016/0169-7439(93)85002-x)

- Liljedahl, L.A. and Abbott, J.A. (1994) 'Changes in sonic resonance of "Delicious" and "Golden Delicious" apples undergoing accelerated ripening', *Transactions of the American Society of Agricultural Engineers*, 37(3), pp. 907-912. doi: [10.13031/2013.28158](https://doi.org/10.13031/2013.28158)
- Long, R.L. and Walsh, K.B. (2006) 'Limitations to the measurement of intact melon total soluble solids using near infrared spectroscopy', *Australian Journal of Agricultural Research*, 57(4), pp. 403-410. doi: [10.1071/AR05285](https://doi.org/10.1071/AR05285)
- Martinsen, P. and Schaare, P. (1998) 'Measuring soluble solids distribution in kiwifruit using near-infrared imaging spectroscopy', *Postharvest Biology and Technology*, 14(3), pp. 271-281. doi: [10.1016/S0925-5214\(98\)00051-9](https://doi.org/10.1016/S0925-5214(98)00051-9)
- Moscetti, R., Monarca, D., Cecchini, M., Haff, R.P., Contini, M. and Massantini, R. (2014b). 'Detection of mold-damaged chestnuts by Near-Infrared spectroscopy', *Postharvest Biology and Technology*, 93, pp. 83-90. doi: [10.1016/j.postharvbio.2014.02.009](https://doi.org/10.1016/j.postharvbio.2014.02.009)
- Moscetti, R., Haff, R.P., Saranwong, S., Monarca, D., Cecchini, M. and Massantini, R. (2014a). 'Nondestructive detection of insect infested chestnuts based on NIR spectroscopy'. *Postharvest Biology and Technology*, 87, 88-94. doi:[10.1016/j.postharvbio.2013.08.010](https://doi.org/10.1016/j.postharvbio.2013.08.010)
- Moscetti, R., Haff, R.P., Ferri, S., Raponi, F., Monarca, D., Liang, P. and Massantini, R. (2017). 'Real-time monitoring of organic carrot (var. Romance) during hot-air drying using Near-Infrared spectroscopy'. *Food and Bioprocess Technology*, 10(11), 2046-2059. doi: [10.1007/s11947-017-1975-3](https://doi.org/10.1007/s11947-017-1975-3)
- Moscetti, R., Raponi, F., Ferri, S., Colantoni, A., Monarca, D. and Massantini, R. (2018). 'Real-time monitoring of organic apple (var. Gala) during hot-air drying using near-infrared spectroscopy'. *Journal of Food Engineering*, 222, 139-150. doi: [10.1016/j.jfoodeng.2017.11.023](https://doi.org/10.1016/j.jfoodeng.2017.11.023)
- Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I. and Lammertyn, J. (2007) 'Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review', *Postharvest Biology and Technology*, 46(2), pp. 99-118. doi: [10.1016/j.postharvbio.2007.06.024](https://doi.org/10.1016/j.postharvbio.2007.06.024)
- Paparatti, B. and Speranza, S. (2005) 'Management of chestnut weevil (*Curculio* spp.)¹, insect key-pest in central Italy', *Acta Horticulturae*, 693, 551-556. doi: [10.17660/actahortic.2005.693.72](https://doi.org/10.17660/actahortic.2005.693.72)
- Pasquini, C. (2003) 'Near Infrared spectroscopy: fundamentals, practical aspects and analytical applications', *Journal of the Brazilian Chemical Society*. Brazilian Chemical Society, 14(2), pp. 198-219. doi: [10.1590/S0103-50532003000200006](https://doi.org/10.1590/S0103-50532003000200006)
- Peiris, K.H.S., Dull, G.G., Leffler, R.G. and Kays, S.J. (1999) 'Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: implications for the development and use of NIR spectrometric techniques', *HortScience*, 34(1), pp. 114-118. doi: [10.21273/hortsci.34.1.114](https://doi.org/10.21273/hortsci.34.1.114)
- Peshlov, B.N., Dowell, F.E., Drummond, F.A. and Donahue, D.W. (2009) 'Comparison of three near infrared spectrophotometers for infestation detection in wild blueberries using multivariate calibration models', *Journal of Near Infrared Spectroscopy*, 17(4), 203-212. doi: [10.1255/jnirs.842](https://doi.org/10.1255/jnirs.842)

- Polesello, A., Giangiaco, R. and Dull, G.G. (1981) 'Application of near infrared spectrophotometry to the nondestructive analysis of foods: a review of experimental results', *C R C Critical Reviews in Food Science and Nutrition*, 18(3), 203-230. doi: [10.1080/10408398309527363](https://doi.org/10.1080/10408398309527363)
- Rajendran, S. (2005) 'Detection of insect infestation in stored foods.', *Advances in food and nutrition research*, 49, pp. 163-232. doi: [10.1016/s1043-4526\(05\)49005-1](https://doi.org/10.1016/s1043-4526(05)49005-1)
- Rodrigues, P., Venâncio, A. and Lima, N. (2012) 'Mycobiota and mycotoxins of almonds and chestnuts with special reference to aflatoxins', *Food Research International*, 48(1), pp. 76-90. doi: [10.1016/j.foodres.2012.02.007](https://doi.org/10.1016/j.foodres.2012.02.007)
- Saranwong, S., Haff, R.P., Thanapase, W., Janhira, A., Kasemsumran, S. and Kawano, S. (2011) 'Short communication: a feasibility study using simplified near infrared imaging to detect fruit fly larvae in intact fruit', *Journal of Near Infrared Spectroscopy*, 19(1), pp. 55-60. doi: [10.1255/jnirs.915](https://doi.org/10.1255/jnirs.915)
- Savitzky, A. and Golay, M.J.E. (1964) 'Smoothing and differentiation of data by simplified least squares procedures', *Analytical Chemistry*, 36(8), pp. 1627-1639. doi: [10.1021/ac60214a047](https://doi.org/10.1021/ac60214a047)
- Sieber, T.N., Jermini, M. and Conedera, M. (2007) 'Effects of the harvest method on the infestation of chestnuts (*Castanea sativa*) by insects and moulds', *Journal of Phytopathology*, 155(7-8), pp. 497-504. doi: [10.1111/j.1439-0434.2007.01269.x](https://doi.org/10.1111/j.1439-0434.2007.01269.x)
- Singh, C.B., Jayas, D.S., Paliwal, J. and 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(2), pp. 118-125. doi: [10.1016/j.compag.2010.06.001](https://doi.org/10.1016/j.compag.2010.06.001)
- Sirisomboon, P., Hashimoto, Y. and Tanaka, M. (2009) 'Study on non-destructive evaluation methods for defect pods for green soybean processing by near-infrared spectroscopy', *Journal of Food Engineering*, 93(4), 502-512. doi: [10.1016/j.jfoodeng.2009.02.019](https://doi.org/10.1016/j.jfoodeng.2009.02.019)
- Teena, M.A., Manickavasagan, A., Ravikanth, L. and Jayas, D.S. (2014) 'Near infrared (NIR) hyperspectral imaging to classify fungal infected date fruits', *Journal of Stored Products Research*, 59, pp. 306-313. doi: [10.1016/j.jspr.2014.09.005](https://doi.org/10.1016/j.jspr.2014.09.005)
- Tigabu, M. and Odén, P. (2002) 'Multivariate classification of sound and insect-infested seeds of a tropical multipurpose tree, *Cordia africana*, with near infrared reflectance spectroscopy', *Journal of Near Infrared Spectroscopy*, 10(1), p. 45-51. doi: [10.1255/jnirs.320](https://doi.org/10.1255/jnirs.320)
- Tigabu, M., Odén, P.C. and Shen, T.Y. (2004) 'Application of near-infrared spectroscopy for the detection of internal insect infestation in *Picea abies* seed lots', *Canadian Journal of Forest Research*, 34, pp. 76-84. doi: [10.1139/X03-189](https://doi.org/10.1139/X03-189)
- Wang, D., Dowell, F.E., Ram, M.S. and Schapaugh, W.T. (2004) 'Classification of fungal-damaged soybean seeds using near-infrared spectroscopy', *International Journal of Food Properties*, 7(1), pp. 75-82. doi: [10.1081/JFP-120022981](https://doi.org/10.1081/JFP-120022981)
- Wang, J., Nakano, K., Ohashi, S., Takizawa, K. and He, J.G. (2010) 'Comparison of different modes of visible and near-infrared spectroscopy for detecting internal insect infestation in jujubes', *Journal of Food Engineering*, 101(1), pp. 78-84. doi: [10.1016/j.jfoodeng.2010.06.011](https://doi.org/10.1016/j.jfoodeng.2010.06.011)

- Wang, J., Nakano, K. and Ohashi, S. (2011) 'Nondestructive detection of internal insect infestation in jujubes using visible and near-infrared spectroscopy', *Postharvest Biology and Technology*, 59(3), pp. 272-279. doi: [10.1016/j.postharvbio.2010.09.017](https://doi.org/10.1016/j.postharvbio.2010.09.017)
- Wilkin, D.R., Cowe, I.A., Thind, B.B., McNicol, J.W. and Cuthbertson, D.C. (1986) 'The detection and measurement of mite infestation in animal feed using near infra-red reflectance', *The Journal of Agricultural Science*, 107, pp. 439-448. doi: [10.1017/s0021859600087244](https://doi.org/10.1017/s0021859600087244)
- Wise, B.M. (2009) 'Eigen vector research wiki - Using cross-validation. URL <http://wiki.eigenvector.com>. Accessed 11.09.2015'.
- Wu, D. and Sun, D.-W. (2013) 'Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review — Part II: Applications', *Innovative Food Science & Emerging Technologies*, 19, pp. 15-28. doi: [10.1016/j.ifset.2013.04.016](https://doi.org/10.1016/j.ifset.2013.04.016)
- Xing, J., Guyer, D., Ariana, D. and Lu, R. (2008) 'Determining optimal wavebands using genetic algorithm for detection of internal insect infestation in tart cherry', *Sensing and Instrumentation for Food Quality and Safety*, 2(3), pp. 161-167. doi: [10.1007/s11694-008-9047-z](https://doi.org/10.1007/s11694-008-9047-z)
- Xing, J. and Guyer, D. (2008) 'Comparison of transmittance and reflectance to detect insect infestation in Montmorency tart cherry', *Computers and Electronics in Agriculture*, 64(2), pp. 194-201. doi: [10.1016/j.compag.2008.04.012](https://doi.org/10.1016/j.compag.2008.04.012)
- Zhang, L. and McCarthy, M.J. (2013) 'Assessment of pomegranate postharvest quality using nuclear magnetic resonance', *Postharvest Biology and Technology*, 77, pp. 59-66. doi: [10.1016/j.postharvbio.2012.11.006](https://doi.org/10.1016/j.postharvbio.2012.11.006)



© 2020 by the authors. Licensee Italian Society for Horticultural Science (Società di Ortoflorofrutticoltura Italiana; SOI), Sesto Fiorentino (Firenze), Italy. This work is an open access article distributed under a Creative Commons Attribution-NonCommercial (CC BY NC) 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>).