JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

Use of near-infrared spectroscopy and chemometrics for the non-destructive identification of concealed damage in raw almonds (Prunus dulcis)

Cristian Rogel-Castillo, Roger B. Boulton, Arunwong Opastpongkarn, Guangwei Huang, and Alyson Elayne Mitchell

J. Agric. Food Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jafc.6b01828 • Publication Date (Web): 16 Jun 2016

ACS Publications

Journal of Agricultural and Food Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Use of near-infrared spectroscopy and chemometrics for the non-destructive identification of concealed damage in raw almonds (*Prunus dulcis*)

Cristian Rogel-Castillo[1], Roger Boulton[2], Arunwong Opastpongkarn[1], Guangwei Huang[3] and Alyson E. Mitchell[1]*

[1]Department of Food Science and Technology, University of California Davis, One Shields Avenue, Davis, CA 95616, United States; [2] Department of Viticulture and Enology, University of California Davis, One Shields Avenue, Davis, CA 95616, United States; [3]Almond Board of California, 1150 9th Street, Suite 1500, Modesto, CA 95354 United States;

*Address correspondence to this author at Department of Food Science and Technology, University of California, Davis, CA 95616 [telephone (530) 304-6618; fax (530) 752-4759; e-mail aemitchell@ucdavis.edu]

Abstract

Concealed damage (CD) is defined as a brown discoloration of the kernel interior (nutmeat) that appears only after moderate to high heat treatment (e.g. blanching, drying, roasting, etc.). Raw almonds with CD have no visible defects before heat treatment. Currently there are no screening methods available for detecting CD in raw almonds. Herein, the feasibility of using Near-Infrared Spectroscopy (NIR) between 1125–2153 nm for the detection of CD in almonds is demonstrated. Almond kernels with CD have less NIR absorbance in the region related with oil, protein and carbohydrates. Using partial least square regression and discriminant analysis (PLS-DA) and by selecting specific wavelengths, three classification models were developed. The calibration models have false positive and false negative error rates ranging between 12.4-16.1% and 10.6-17.2%, respectively. The percent error rates ranged between 8.2-9.2%. Second derivative pre-processing of selected wavelength resulted in the most robust predictive model.

Keywords

Prunus dulcis; Almond, Concealed Damage; Near Infrared Spectroscopy; Partial Least Square Regression; Discriminant Analysis.

Introduction

Concealed damage (CD) in raw almonds (*Prunus dulcis* (Mill.) D.A. Webb) is defined by the industry as a brown discoloration of the kernel interior (nutmeat) that appears only after moderate to high heat treatment (e.g. blanching, drying, roasting, etc.,) as shown in Figure 1. CD may develop anytime during harvest when rain occurs, or after harvest when kernels are in windrows or stockpiles and exposed to warm and moist environments.[1,2] Raw almond kernels with CD, have no visible defects on the interior or exterior surface of the kernel. Additionally, there are no visible signs of CD on the surface of whole roasted kernels.[3] CD is frequently associated with a strong bitter flavor(s) that can result in immediate consumer rejection.[1] Currently there are no screening methods available for detecting CD in raw almonds, or other nuts affected by CD, and processors often do not realize nuts are damaged until after they have been roasted.[1] Under current production practices, the most common methods for detecting CD involve visual inspection of roasted almonds after they are split open. Kernels with a "dark brown" color over ~50% of the interior of the kernel are considered to have CD.[4] A similar approach is used for hazelnuts.[5] Visual inspection and manual sorting is time-consuming, subjective, labor intensive and cannot be used to identify nuts with CD before heat treatments. This can result in significant product loss.

The current hypothesis is that the browning associated with CD is related to the Maillard reaction. Moisture can induced the hydrolysis of carbohydrates and potential availability of reducing sugars for Maillard browning reactions. For example, in macadamia nuts exposed to moisture during harvesting, increased levels of reducing sugars were observed in nuts with internal browning.[6] Similar observations were made in hazelnuts[7] and in almonds exposed to simulated rainfall.[2] In more recent studies, elevated levels of volatiles related to lipid oxidation and amino acid degradation were observed in almonds with CD.[8] Both lipid oxidation products and protein degradation products can serve as reactants in the Maillard browning reaction.

Near-infrared spectroscopy (NIR) is a rapid and effective method for screening foods for specific chemical and physical characteristics.[9] NIR is advantageous as a screening method as it is non-destructive, can be used on whole foods, and produces no waste. The NIR spectral region (720 to 2500 nm) is ideally suited for foods as it contains absorbance bands that result primarily from three chemical bonds: C-H (fats, oil, hydrocarbons), O-H (water, alcohol) and N-H (protein). NIR spectroscopy is increasingly considered one of the more promising in-line detection methods for rapidly measuring specific chemical properties of food.[6] It has been successfully applied in detecting quality defects in macadamia kernels[7], walnuts[10], chestnut[11 - 12], hazelnuts[13 - 14], and soybean seed.[15] It has also been employed for food composition analysis including oleic and linoleic acid content in peanut seed[16], acidity and water content in hazelnuts[17], and characterization of shea tree nut fat profiles.[18]

Pearson (1999) was the first to recognize the use of NIR spectroscopy for the identification of CD in raw almonds[4, 19] and evaluated the transmission spectrum from 700–1400 nm in almonds soaked in water, and dried but not roasted. In these studies, almonds were either soaked for 30 minutes and exposed to 95% relative humidity for 30 hours (short moisture), or soaked for 60 minutes and exposed to 95% relative humidity for 60 hours (long moisture). Almonds were then dried at either 55 or 110°C. The higher temperature and shorter soak times produced the greatest amount of CD. Almonds with CD had enhanced absorption at 930 nm (oil absorption band). Raw almonds with CD could be distinguished from normal almonds at an error of 12.4% by using principal components of the absorbance, first derivative and second derivative spectra between 1000-1300 nm. Pearson (1999) recognized that collecting the NIR spectra over the full transmission range would be too slow to achieve desired inspection rates of 40 nuts/s and therefore tested the feasibility of using just 6 light emitting diodes at 660, 830, 880, 890, 940, and 950 nm[19]. These data resulted in a classification error rate of 14.3% for the validation set. More recently, Nakariyakul[20 - 21] achieve a higher classification rate using hyperspectral transmission and focusing on a sub-set of absorbing bands (760, 920, 935, and 970 nm) with a false negative error rate of 14.81%. Almonds used in this study were generated by Pearson (1999) as described above.

Herein, we present the development of a prediction model for the classification of almonds with CD using reflectance NIR in the extended range of the NIR spectrum (1125 – 2153 nm) and by employing data pre-processing and partial least square discriminant analysis (PLS-DA). Almonds evaluated in this study were exposed to controlled humidity environments that produced an internal nut moisture content of ~5% (control), 8% (mild CD) and 11% (100% CD). The percent CD in the raw almonds was validated using colorimetry as described previously[8].

Developing a rapid in-line screening method for detecting CD in raw almonds is a critical step towards improving quality control measures in almond processing and offers the advantage of sorting almonds with CD into product lines that do not require roasting or other heat treatments.

Materials and Methods

Sample preparation

Dehulled raw kernels (100 lbs., var. Nonpareil) were supplied by the Nickels Soil lab (Arbuckle, CA) in September 2013. Individual vessels containing ~100 gm were exposed to conditions that produced an internal kernel moisture of 5% (actual 5.4 ± 0.2), 8% (actual 8.6 ± 0.7) or 11% (actual 10.4 ± 1.5) in a controlled atmosphere (Thermo Scientific, Marietta, OH) at 45 ± 2°C. Under these conditions, CD is observed after 24 hours. The moisture content of the almonds was validated gravimetrically by drying samples (~1 g) at 95-105°C under vacuum to a constant weight. Moisture was determined in triplicate, and the results were averaged.

Near-infrared reflectance spectra measurement

NIR diffuse reflectance spectra were measured on single whole raw almond kernels using an extended MicroNIR 2200 spectrometer (JDSU, USA). The spectral range was collected from 1125 – 2153 nm using sampling intervals of 8 nm per pixel. The detector used was a 128 pixel uncooled element InGaAs (JDSU, USA). Reflectance spectra data (R) were converted to absorbance using the log (1/R) transformation. A Spectralon® SRM-99 Diffuse Reflectance Standard (Labsphere®, New Hampshire, USA) was used as white calibration reference. For each spectrum, 1000 scans with an integration time of 550 µs were averaged.

Data Pre-processing

NIR spectra are complex with broad overlapping NIR absorption bands making it often difficult to identify unique spectral features related to individual chemical components within a given sample. Therefore, a mathematical treatment (pre-processing) of NIR spectra is often used to correct for unwanted systematic sample-to-sample variation (e.g., kernel shape, roughness of kernel surface); help remove spectral baseline shift and scattering caused by particle size differences; reduce band overlapping; and enhance spectral differences.[22] Data pre-processing results in relevant NIR spectral data extraction without losing information while removing unwanted information (e.g., interferences or noise).[22]

After the acquisition, NIR spectra were converted to absorbance and pre-processed using either standard normal variate (SNV) or a 9- point second order Savitzky-Golay filter (second derivative preprocessing). These two techniques alone and a combination were compared to determine their effectiveness at removing baseline offsets. Data pre-processing was performed using R and R-Studio (version 0.98.1102). The following packages were used for preprocessing and PLS-DA: ChemometricsWithR[23], signal[24], plyr[25], dplyr[26] and caret[27].

Determination of CD by colorimetry

After NIR spectra were acquired, the almonds were roasted at 120°C for 90 min in a convection oven (Thermo Scientific, USA). Almonds were then split in half along the natural seam and the color of the internal kernel was measured using a ColorFlex colorimeter (HunterLab, USA) according to methods established previously[8]. The color values L* (lightness), C (Chroma), and h (hue), according to the CIE LCh color scale were recorded using a portsize of 0.4 inches with D65 optical sensor, 0° geometry and 10° angle of vision. Almonds with CD ($L^* \leq 71$) were identified and grouped separately from those with no concealed damage (NCD; $L^* > 71$).[8]

Prediction of almonds with concealed damage

Pre-processed spectra (SNV, second derivative and a combination of both) were analyzed using partial least squares discriminant analysis (PLS-DA). Almonds were separated into two groups (NCD and CD) using colorimetry. The data set (855 almonds) was then randomly divided into a calibration (655 almonds) and a validation (200 almonds) sample set. NCD and CD almonds were assigned constant values of 0 and 1 for a two-class model, respectively.

For the calibration model, repeated cross-validation was used to find the best model. Calibration models were evaluated based on the percentage false positive (% fp), percentage false negative (% fn) and percentage error rate (% ER). A fn was defined as the percentage of NCD almonds classified as those with CD, while fp was defined as CD almonds classified as NCD. The % ER represents the percentage of total almonds incorrectly classified by the predictive method.[4, 12]

Results and Discussion

A representative NIR spectrum (1125-2153 nm) of the NCD almonds, after SNV pre-processing, is shown in Fig. 2. The spectrum is characterized by broad and unresolved absorption bands and is similar to spectra for peanut[28], walnut[10], macadamiat[7], and shea nut.[18] To enhance spectral features and compensate for baseline offsets, a second-derivative of the absorbance data, with respect to wavelength, was calculated. In the second-derivative data, absorbance maxima are converted to minima (Fig. 3). The NIR spectra obtained after applying the second-derivative were characterized by 10 absorption bands. These bands correlate with the major constituents of raw almonds: lipid (50%), carbohydrates (~22%), and protein (~21%).[29] The absorption bands between 1165-1238 nm, 1692-1740 nm and 2064-2104 nm are associated with lipids. These include the C-H (-CH) second overtone stretching band (1200-1214 nm)[30], the C-H (-CH_2) first overtone stretching band (1700-1724 nm)[6, 31] and C-H combination band (~2098 nm).[32] The absorption bands in between 1408-1462 nm and 1902-1959 nm are associated with the H-OH second overtone of water[18] as well with protein. The absorption bands between 1692-1740 nm and 2064-2104 nm are associated with the absorption of protein (~1700–1850 nm) and amino acids (~2080 nm) respectively[32] and the region between 1902-1959 nm correlate to water and amides (~1910–1920 nm).[32] Additionally, the absorption band between 2064-2104 nm can be associated with the O-H and the carboxylic group (C=O-O) band of carbohydrates.[31]

An overlay of the averaged second derivative spectra for almonds classified as NCD and CD is also given in Fig. 3. The main differences between the NCD and CD spectra occur at 1432, 1457, 1505, 1513, 1708, 1918 and 2080, 2096 nm. The absorption bands at 1432, 1457, 1505, 1513, and 1918 nm correspond to protein[33], the band at 1708 nm corresponds to free fatty acids and oil[12, 33], and absorption at 2080 and 2096 nm correspond to carbohydrates. Almonds with CD present less absorbance in these regions indicating that kernels display decreased levels of lipids, protein and carbohydrates as compared to controls. These results corresponds to observations of King et al.[34] (1983) who reported that almonds with CD have lower crude fat (oil) and total carbohydrates as compared to almonds with no concealed damage. Additionally, we recently demonstrated higher levels of volatiles related to lipid oxidation and amino acid degradation in almonds with CD as compared to almonds with no concealed damage.[8] Taken together these results indicate the metabolic processes that activate the degradation of proteins, carbohydrates and lipids are involved in the development of CD. The free amino acids, sugars and products from the oxidation of lipids would be substrates for the Maillard reaction and supports the hypothesis that the Maillard reaction is involved in the formation of CD in almonds.

Initially, multiple PLS-DA models were evaluated using the full wavelength region from 1125-2153 nm after data pre-processing (SNV, second derivative, and SNV and the second derivative). In general, the best predictive models give low percentage error rates (i.e. the highest percentage of correct classification). Herein we found that using the full wavelength region resulted in models with high percentage error rates and therefore PLS-DA models were developed using only relevant portions of the NIR spectra. Table 1 summarize the prediction performance of the calibration models and validation models, which were selected as they had the lowest percent of error (% ER), false positive (% fp) and false negative (% fn) rates. A large data set (200 samples) was analyzed to optimize the prediction models. The lowest % ER (8.2%) was obtained using only second derivative pre-processing as compared to 9.2% when using SNV pre-processing, and 8.2% when using SNV and second derivative pre-processing. Although the % fp rate was higher for this model (17.2%) as compared to the SNV (12.6%) and SNV and second derivative pre-processing (10.6%), the % fn was significantly lower (12.4%) as compared to these models (16%).

Previous studies employing IR to build models to discriminate differences between CD and NCD focused on the absorbance range between 700-1400 nm[5] and on selected wavelengths within the 700-1400 nm absorbance range.[20 - 21] Comparisons of these results with results obtained herein are summarized in Table 2. Although the % fp were lower across these studies (0.7-5.4%), as compared with our results (12.4 – 16.1%), the rate of % fn were significantly higher (11.1-62.96%) than those obtained using our predictive models (10.6 – 17.2%). Additionally, the % ER ranged from 5.8–27.5% whereas our predictive models gave a much narrower range of 8.2 – 9.2%. The three PLS-DA models presented herein offer significant improvements in the prediction capabilities and are able to identify almonds with CD with 90.8 – 91.8% certainty based upon calibration models. Although any of the three models presented could be considered for further development of a rapid in-line screening method for detecting CD in raw almonds, the PLS-DA model based on the second derivative spectra and utilizing four wavelength ranges (i.e.1408–1462, 1692–1740, 1902–1959 and 2064-2104 nm) gives the lowest rate of false negatives and may be the best choice for further method development.

Our results indicate that these PLS-DA predictive models offer advantages over previously reported models and that CD is related to the degradation of lipids, carbohydrates and proteins in almonds.

REFERENCE

(1) Reil, W, J.M. Labavitch, D. Holmberg. 1996. *Harvesting. In Almond Production Manual.* W.C. Micke, editor. University of California, Division of Agriculture and Natural Resources (publication 3364). Oakland, CA. pp. 260-264.

(2) Kader, A.A. and J.F. Thompson. 1992. *Postharvest handling systems: Tree nuts in Postharvest Technology of Horticultural Crops*. A.A. Kader, editor. University of California, Division of Agriculture and Natural Resources, publication 3311. pp. 254.

(3) Halbrook, W. U; Fuller, G.; Whitehand, L. C. Almond Nutmeat Moisture and Water Activity and its Influence on Fungal Flora and Seed Composition. *J. Food Sci.* **1983**, *48*, 615-617.

(4) Pearson, T. C. Spectral Properties and Effect of Drying Temperature on Almonds with Concealed Damage. *LWT--Food Sci. Technol*. **1999**, *32*, 67-72.

(5) Pannico, A.; Schouten, R. E.; Basile, B.; Romano, R.; Woltering, E. J.; Cirillo, C. Non-destructive detection of flawed hazelnut kernels and lipid oxidation assessment using NIR spectroscopy. *J. Food Eng.* **2015**, *160*, 42-48.

(6) Huang, H.; Yu, H.; Xu, H.; Ying, Y. Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. *J. Food Eng.* **2008**, *87*, 303-313.

(7) Guthrie, J.; Greensill, C.; Bowden, R.; Walsh, K. Assessment of quality defects in macadamia kernels using NIR spectroscopy. *Aust. J. Agric. Res.* **2004**, *55*, 471-476.

(8) Rogel-Castillo, C.; Zuskov, D.; Chan, B. L.; Lee, J.; Hong, G.; Mitchell, A. E. Effect of Temperature and Moisture on the Development of Concealed Damage in Raw Almonds (Prunus dulcis). *J Agric Food Chem.* **2015**, *63*, 8234-8240.

(9) Büning-Pfaue, H. Analysis of water in food by near infrared spectroscopy. *Food Chemistry.* **2003**, 82, 107-115.

(10) Jensen, P. N.; Sørensen, G.; Engelsen, S. B.; Bertelsen, G. Evaluation of Quality Changes in Walnut Kernels (*Juglans regia L.*) by Vis/NIR Spectroscopy. *J. Agric. Food Chem.* **2001**, *49*, 5790-5796.

(11) Liu, J.; Li, X. Y.; Li, P. W.; Wang, W.; Zhang, J.; Zhang, R.; Liu, P. Nondestructive detection of moldy chestnut based on near infrared spectroscopy. *Afr. J. Agric. Res.* **2010**, *5*, 3213-3218.

(12) Moscetti, R.; Haff, R. P.; Saranwong, S.; Monarca, D.; Cecchini, M.; Massantini, R. Nondestructive detection of insect infested chestnuts based on NIR spectroscopy. *Postharvest Biol. Technol.* **2014**, *87*, 88-94.

(13) Moscetti, R.; Haff, R. P.; Aernouts, B.; Saeys, W.; Monarca, D.; Cecchini, M.; Massantini, R. Feasibility of Vis/NIR spectroscopy for detection of flaws in hazelnut kernels. *J. Food Eng.* **2013**, *118*, 1-7.

(14) Pannico, A.; Schouten, R. E.; Basile, B.; Romano, R.; Woltering, E. J.; Cirillo, C. Non-destructive detection of flawed hazelnut kernels and lipid oxidation assessment using NIR spectroscopy. *J. Food Eng.* **2015**, *160*, 42-48.

(15) Wang, D.; Dowell, F.; Ram, M.; Schapaugh, W., Classification of fungal-damaged soybean seeds using near-infrared spectroscopy. *Int. J. Food Prop.* **2004**, *7*, 75-82.

(16) Tillman, B. L.; Gorbet, D. W.; Person, G. Predicting Oleic and Linoleic Acid Content of Single Peanut Seeds using Near-Infrared Reflectance Spectroscopy. *Crop Sci.* **2006**, *46*, 2121-2126.

(17) Bellincontro, A.; Fracas, A.; DiNatale, C.; Esposito, G.; Anelli, G.; Mencarelli, F. Use of NIR Technique to Measure the Acidity and Water Content of Hazelnuts. *Acta Hortic.* **2005**, *686*, 499-503.

(18) Davrieux, F.; Allal, F.; Piombo, G.; Kelly, B.; Okulo, J. B.; Thiam, M.; Diallo, O. B.; Bouvet, J. M. Near infrared spectroscopy for high-throughput characterization of Shea tree (*Vitellaria paradoxa*) nut fat profiles. *J Agric Food Chem*. **2010**, *58*, 7811-7819.

(19) Pearson, T. C., Use of near infrared transmittance to automatically detect almonds with concealed damage. *LWT--Food Sci. Technol*. **1999**, 32, 73-78.

(20) Nakariyakul, S., Internal damage inspection of almond nuts using optimal near-infrared waveband selection technique. *J. Food Eng.* **2014**, *126*, 173-177.

(21) Nakariyakul, S.; Casasent, D. P. Classification of internally damaged almond nuts using hyperspectral imagery. *J. Food Eng.* **2011**, *103*, 62-67.

(22) Metrohm NIRSystem. Herisau, Switzerland. Monograph: NIR Spectroscopy A guide to near-infrared spectroscopic analysis of industrial manufacturing processes. Retrieved April 16, 2016 from: http://www.mep.net.au/wpmep/wp-content/uploads/2013/05/MEP_Monograph_NIRS_81085026EN.pdf

(23) Wehrens, R. Chemometrics With R: Multivariate Data Analysis in the Natural Sciences and Life Sciences. **2011**, Springer, Heidelberg.

(24) Signal developers. signal: Signal processing. **2015**. URL: http://r-forge.r-project.org/projects/signal/.

(25) Wickham, H. The Split-Apply-Combine Strategy for Data Analysis. *J. Stat. Softw.* **2011**. 40, 1-29. URL http://www.jstatsoft.org/v40/i01/.

(26) Wickham, H.; Francois, R. dplyr: A Grammar of Data Manipulation. R package version 0.4.3. **2015**. https://CRAN.R-project.org/package=dplyr

(27) Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; Benesty, M.; Lescarbeau, R.; Ziem, A.; Scrucca, L.; Tang, Y.; Candan, C. caret: Classification and Regression Training. R package version 6.0-68. **2016**. https://CRAN.R-project.org/package=caret

(28) Govindarajan, K. N.; Kandala, C. V. K.; Subbiah, J. NIR Reflectance Spectroscopy for Nondestructive Moisture Content Determination in Peanut Kernels. *Trans. ASABE.* **2009**, *52*, 1661-1665.

(29) (USDA) US Department of Agriculture, Agricultural Research Service, 2015. USDA National Nutrient Database for Standard Reference, Release 27. Retrieved August 26, 2015 from: Nutrient Data Laboratory Home Page: http://ndb.nal.usda.gov/ndb/search

(30) Sathe, S.; Seeram, N.; Kshirsagar, H.; Heber, D.; Lapsley, K. Fatty acid composition of California grown almonds. *J. Food Sci*. **2008**, *73*, C607-C614.

(31) Moscetti, R.; Monarca, D.; Cecchini, M.; Haff, R. P.; Contini, M.; Massantini, R., Detection of Mold-Damaged Chestnuts by Near-Infrared Spectroscopy. *Postharvest Biol. Technol.* **2014**, *93*, 83-90.

(32) Fassio, A.; Cozzolino, D. Non-destructive prediction of chemical composition in sunflower seeds by near infrared spectroscopy. *Ind. Crops Prod.* **2004**, *20*, 321-329.

(33) Workman Jr, J.; Weyer, L., Appendix D. Spectra - Structure Correlations for Near Infrared. In *Practical guide and spectral atlas for interpretive near-infrared spectroscopy*, Second edition; CRC Press: 2012; pp 229 - 267.

(34) Jr, A. D. K.; Halbrook, W. U.; Fuller, G.; Whitehand, L. C. Almond Nutmeat Moisture and Water Activity and its Influence on Fungal Flora and Seed Composition. *J. Food Sci.* **1983**, *48*, 615-617.

Abbreviations Used

Concealed damage, CD; No concealed damage, NCD; NIR, Near Infrared Spectroscopy; SNV, Standard Normal Variate; PLS, Partial Least Square; DA, Discriminant Analysis; fp, False Positive; fn, False Negative; % ER., percentage error rate.

Funding Sources

The Almond Board of California provided financial support for this study. We would also like to acknowledge the support of the John Kinsella Endowed Chair in Food, Nutrition and Health.

Acknowledgements

The authors thank Franz Niederholzer (UC Farm Advisor Colusa/Sutter/Yuba Counties) for many thoughtful conversations on almond flavor and breeding and for providing almond samples.

FIGURE LEGENDS:

Figure 1. Color development in raw and roasted almonds (120°C for 90 min) exposed to 5% moisture (control) and 11% moisture (concealed damage).

Figure 2. The mean standard normal variate (SNV) pre-processed absorbance spectra of almonds with no concealed damage (NCD).

Figure 3. A comparison of the mean second derivative (Savitzky-Golay, 9-smoothing points) pre-processed spectra of almonds with no concealed damage (NCD) and concealed damage (CD).

Figure 1


RAW (CONTROL, NCD)
RAW (CONCEALED DAMAGE, CD)
ROASTED (CONTROL, NCD)
ROASTED (CONCEALED DAMAGE, CD)

Figure 2

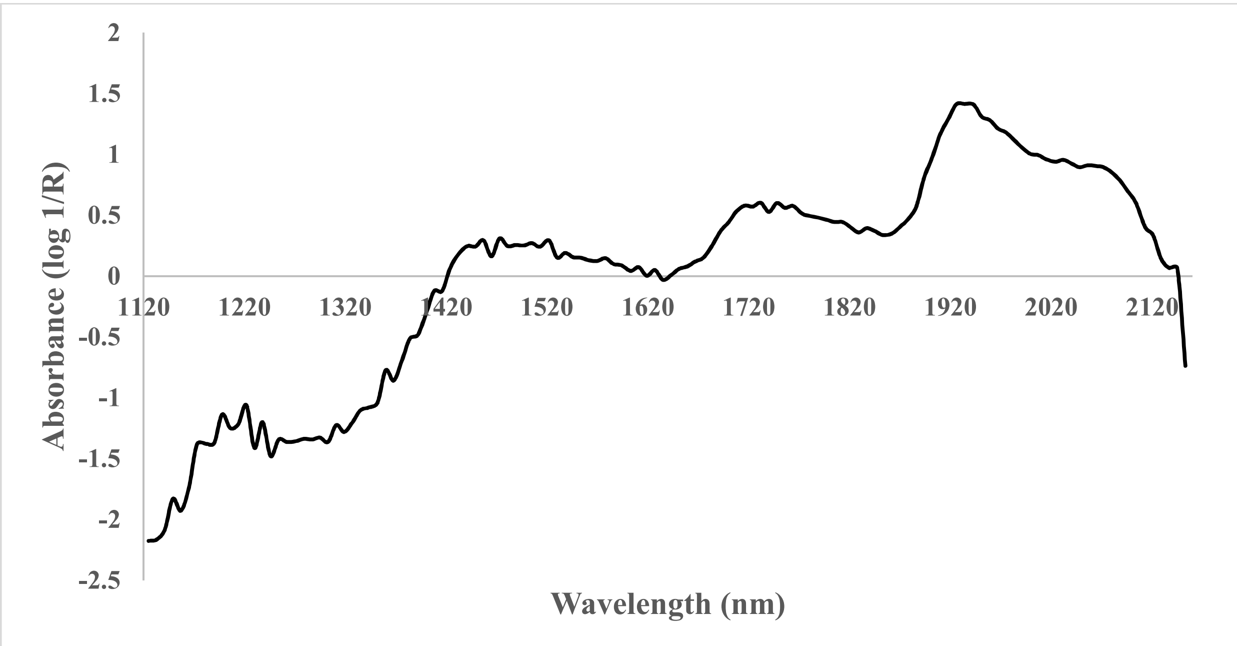

2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
Absorbance (log 1/R)
1120
1220
1320
1420
1520
1620
1720
1820
1920
2020
2120
Wavelength (nm)

Figure 3

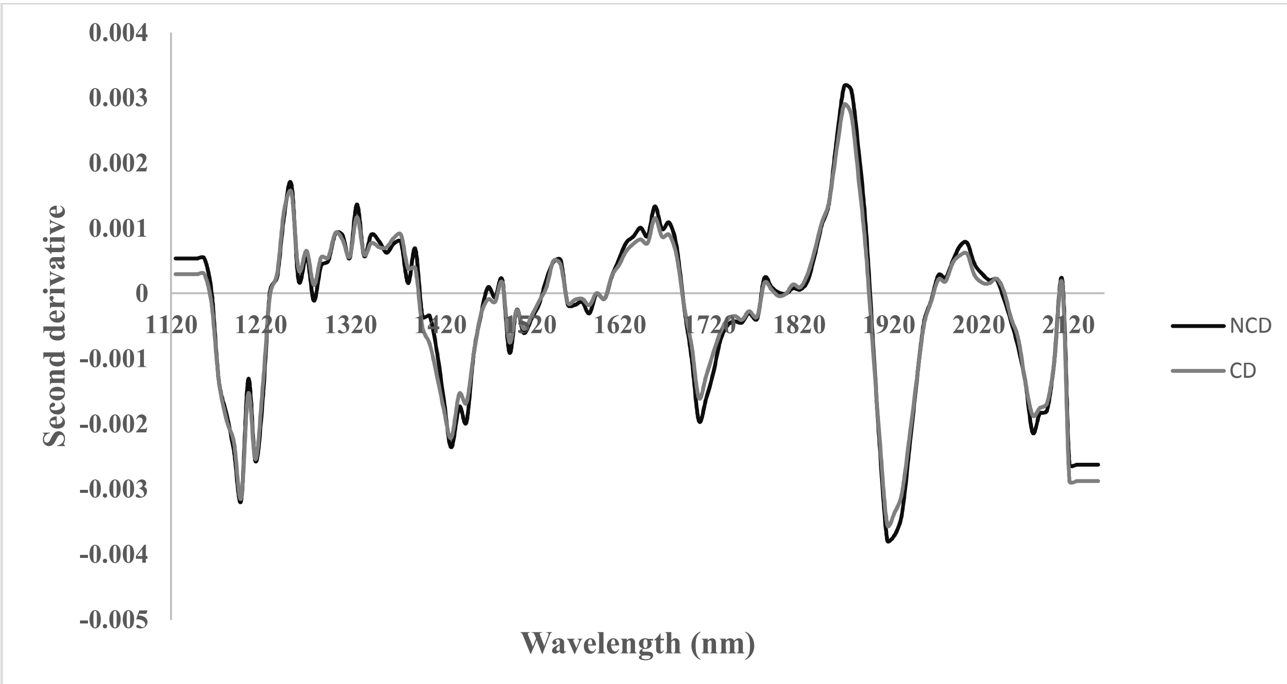

0.004
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
-0.004
-0.005
Second derivative
1120
1220
1320
1420
1520
1620
1720
1820
1920
2020
2120
Wavelength (nm)
NCD
CD

Table 1. Results of the PLS-DA model using standard normal variate (SNV), second derivative, and SNV and second derivative pre-processing.

a) Calibration Model

	SNV	Second derivative	Second derivative + SNV
Wavelength selected (nm)	1408 – 1465 1902 – 1959	1408 – 1465 1692 – 1740 1902 – 1959 2064 – 2104	1408 – 1465 1692 – 1740 1902 – 1959
# Latent variables	4	7	4
ROC* / % Error rate	0.908/9.2	0.918/8.2	0.918/8.2
Specificity / % fn	0.839/16.1	0.876/12.4	0.840/16.0
Selectivity / % fp	0.874/12.6	0.828/17.2	0.894/10.6

* Area under the Receiver Operating Characteristics

b) Validation Model

	SNV	Second derivative	Second derivative + SNV
% Error rate	9	7	9
% false positive (% fp)	8	8	9
% false negative (% fn)	11	6	7

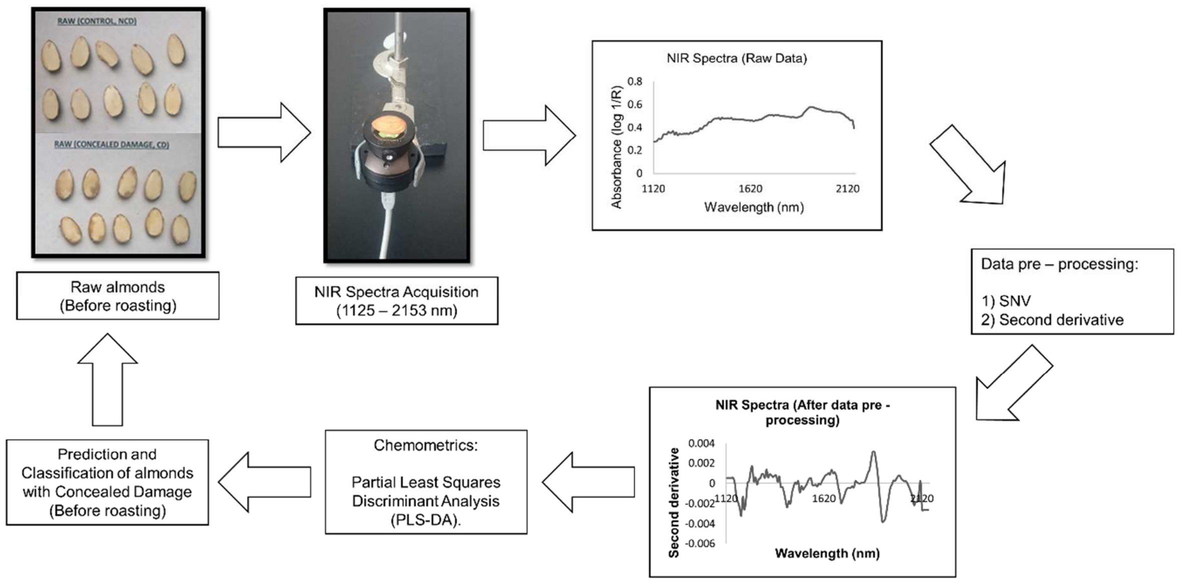
Table 2. Comparison of NIR validation results between methods used in the classification of almond with concealed damage (CD).

	Pearson, T.C (1998)	**Nakariyakul, S. (2014)**	**Nakariyakul et al. (2011)**	**Results Obtained Herein***
Range (nm)	700 – 1400	700 – 1400 (selected wavelength)	700 –1400 (selected wavelength)	1125–2153 (selected wavelength)
% Error rate	12.4 – 27.5	5.8	8.8	8.2 – 9.2
% False positive	0.7 – 5.4	2.91 – 3.41	0.58 – 1.74	12.4 – 16.1
% False negative	11.1 – 23.8	14.81 – 62.96	31.48 – 53.70	10.6 – 17.2

* Based on calibration models

411 **Table of Contents Graphics**

412



413