Cortes-Lopez, V.; Talens Oliag, P.; Barat Baviera, JM.; Lerma-García, MJ. (2019). Discrimination of intact almonds according to their bitterness and prediction of amygdalin concentration by Fourier transform infrared spectroscopy. Postharvest Biology and Technology. 148:236-241. https://doi.org/10.1016/j.postharvbio.2018.05.006

Discrimination of intact almonds according to their bitterness and prediction of amygdalin concentration by Fourier Transform infrared spectroscopy

Victoria Cortés, Pau Talens, José Manuel Barat, María Jesús Lerma-García*

Departamento de Tecnología de los Alimentos, Universitat Politècnica de València, Camino de Vera s/n, 46022, Spain

*Corresponding author:

M.J. Lerma-García, e-mail: malerga1@tal.upv.es

ABSTRACT

Intact almond kernels (*N*=360, half sweet and half bitter) were analyzed using attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) for the prediction of amygdalin concentration and to classify them according to their bitterness. Amygdalin concentrations for sweet and bitter almonds, determined by high performance liquid chromatography, were between 0.7-350 and 15000-50000 mg kg^{-1}, respectively. Concentrations were successfully predicted by applying partial least squares (PLS) to the pre-treated spectral data with R^2_p of 0.951 and RMSEP of 0.398. Additionally, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and PLS-DA models were constructed to classify samples according to their bitterness. All three models provided a satisfactory discrimination of almonds into sweet and bitter categories, providing overall accuracy values of 83.3 %, 86.1 % and 98.6 %, respectively. The results indicate the potential of ATR-FTIR spectroscopy for the reliable, easy and fast prediction of amygdalin concentration, and for almond classification according to their bitterness.

Keywords: ATR-FTIR, amygdalin concentration, bitterness, intact almonds, PLS, almond discrimination

1. Introduction

Almonds are a very much valued nut due to their high nutritional and sensory attributes (Grane-Teruel *et al.*, 2001). The almond has been positioned in the market as a healthy and versatile product in its different uses, which had led to a doubling of demand around the world (Velasco & Aznar, 2016). Spain has the biggest area of almond cultivation in the world with more than 600,000 hectares in 2015 (Velasco & Aznar, 2016). Since 2010, Spanish production has remained relatively stable, reaching 48,000 tons in 2014 (Velasco & Aznar, 2016). This production represents 4.5 % of the world total and it places Spain in third place behind the USA and Australia.

Sweet almond is the most commonly consumed form, however, bitter almonds are also valued, primarily for the extraction of flavour extracts, which are processed before consumption to remove the poisonous substances (Borrás *et al.*, 2014). Almond bitter flavour is due to the presence of cyanogenic glucosides, like amygdalin and prunasin (Sánchez-Pérez *et al.*, 2008). In mature almonds, the only cyanogenic glucoside found is amygdalin, since prunasin (normally found in roots, leaves and kernel of immature almonds) is converted into amygdalin during maturation. Almond bitter flavour is a consequence of the enzymatic hydrolysis produced by β-glucosidase, which produces benzaldehydes, sugars and hydrogen cyanide. Therefore, because of their toxicity and because chemical components are different among sweet and bitter almonds, it is important to distinguish them for two main reasons: 1) to guarantee homogeneous lots of almonds in food industry and 2) to avoid possible health problems related to the consumption of bitter almonds.

Some interesting analytical techniques have been applied to almond quality control, for example, vibrational spectroscopy methods, such as Fourier transform

infrared (FTIR) spectroscopy (Ellis *et al.*, 2005). FTIR is fast, easy, non-destructive and a relatively inexpensive technique (Vahur *et al.*, 2009). Samples do not need any pre-treatment to register spectra in a few seconds (Ellis *et al.*, 2002). In addition, other advantage of this spectroscopy technique is its application in foodstuffs that can be fresh, dried, liquid or solid (Dogan *et al.*, 2007). The quantification and qualification of almond quality by FTIR consists of acquiring a fingerprint characteristic of any point of the almond, providing thus information about their grade of bitterness, the type of variety or the rate of spoilage, among others. The integration of FTIR and chemometrics in tandem could provide an excellent methodology that could be able to qualitatively and/or quantitatively classify sweet and bitter almonds based on extracted spectra features. Some applications of the FTIR spectroscopy previously published in the almond field included the quality control of medicinal almonds (Chun-Song *et al.*, 2017), the classification of almond cultivars by measuring the spectra in almond oil (Beltrán *et al.*, 2009) and in grounded almonds (Valdés *et al.*, 2013), among others. Only one study has discriminated almonds according to their bitterness, but using NIR and Raman spectroscopy (Borrás *et al.*, 2014). Micklander *et al.* (2002) predicted the amygdalin concentration in bitter almonds using Raman spectroscopy. However, calibration was performed by spiking sweet almonds with different concentrations of amygdalin standard, and not with real values of samples which could be established with other techniques such as high performance liquid chromagraphy (HPLC).

The aim of this work is the application of ATR-FTIR spectroscopy followed by multivariate analysis of the spectral data as a non-destructive methodology for the prediction of amygdalin concentration (measured by HPLC) of intact almonds and for the classification of almonds according to their bitterness.

2. Materials and methods

2.1. Reagents and almond samples

To determine amygdalin concentration, an amygdalin standard for HPLC from Sigma-Aldrich (St. Louis, Missouri, USA) of ≥ 97.0 % of purity was used. Different solvents, such as deionized water (obtained using an Aquinity deionizer from Membrapure GmbH (Berlin, Germany), acetonitrile (ACN, HPLC Far UV/Gradient Grade, J.T. Baker, The Netherlands), acetone (Panreac, Barcelona, Spain) and methanol (MeOH, AGR ACS, ISO, Ph.Eur. Assay ≥ 99.8 %, Labkem, Barcelona, Spain) were also used.

A batch of 360 almonds, kindly provided by Agricoop (Alicante, Spain), were employed in this work. This batch is composed by 180 sweet and 180 bitter almonds. Although the genetic variety of the bitter almonds is unknown, the sweet almonds were selected from six different genetic varieties (30 almonds each), which are 'Comuna', 'Guara', 'Largueta', 'Marcona', 'Planeta' and 'Rumbeta'. All samples included in the work were selected for absence of damage, and of similar colour and size.

2.2. ATR-FTIR

FTIR spectra were registered using a Tensor 27 spectrometer from Bruker Optics (Milan, Italy) dotted with a deuterated triglycine sulphate (DTGS) detector which is coupled to an ATR accessory (Specac Inc., Woodstock, Georgia, USA) composed of a zinc selenide (ZnSe) crystal. Spectra were registered in the absorbance mode as the mean of 32 scans in the 4000–600 cm^{-1} spectral range and using a resolution of 4 cm^{-1}. The

FTIR spectrometer was controlled using the OPUS software version 5.0 (Bruker Optics). The acquisition spectral time was 20 s. Each almond kernel (with skin) was placed on the ZnSe crystal and the spectra was measured. Two points were collected for each sample on each side of the almond, and the mean of both spectra were employed for statistical analysis. After each measurement, the crystal was cleaned using acetone and dried with a cellulose tissue.

2.3. Amygdalin extraction and quantification by HPLC

After spectral measurement, each almond was immersed in hot water to eliminate almond skin. After keeping almonds at room temperature to allow them to be dried, they were grinded with a porcelain mortar, and mixed with 20 mL MeOH. This suspension was maintained under constant stirring during 24h, being the supernatant finally passed through a syringe filter of polytetrafluoroethylene (PTFE) (0.22 μm, Scharlab, Barcelona, Spain). The filtered supernatant was directly injected into the HPLC system when obtained from sweet almonds, and it was diluted with MeOH in a 1:10 (v/v) proportion for bitter almonds in order to have amygdalin concentrations within the linear range of the calibration curve.

Supernatants were then analyzed by HPLC (LaChrom Elite" liquid chromatograph from Hitachi Ltd.,Tokyo, Japan), fitted with an auto-sampler (model L-2200) and an ultraviolet (UV) detector (model L-2400). Amygdalin was separated at a flow rate of 1.0 mL min^{-1} with a mixture of ACN and water (20:80, v/v) using an isocratic elution and a Liquid Purple C18 analytical column (250 x 4.6 mm i.d., 5 μm) (Análisis Vínicos, Tomelloso, Spain). Detection was monitored at 218 nm. Sample injection volume was 20 μL.

128

129 *2.4. Spectra pre-processing and chemometric data treatment*

130

131 Spectral pre-processing and statistical analysis were performed using the
132 statistical software program 'The Unscrambler X' (version 10.3, Camo Process SA,
133 Trondheim, Norway).

134 Various pre-treatment techniques were simultaneously applied to the ATR-FTIR
135 data, such as standard normal variate (SNV) to correct multiplicative interferences,
136 baseline shift variations and curvilinearity (Barnes *et al.*, 1989), the second derivate (with
137 2.3-gap-segment) and Savitzky-Golay smoothing using 15 points to extract useful
138 information (Cortés *et al.*, 2016, Rodriguez-Saona *et al.*, 2001) and to improve the signal-
139 to-noise ratio (Gorry, 1990, Savitzky & Golay, 1964).

140 After spectra pre-treatment and before multivariate analysis, data spectral
141 variation was analyzed by principal component analysis (PCA) and the defective spectra
142 due to a problem of acquisition were eliminated based on Hotelling's T^2 and squared
143 residual statistics (Beghi *et al.*, 2018).

144 To proceed with chemometric analysis, a spectral data matrix was constructed.
145 The 360 almond samples (180 sweet and 180 bitter) were introduced in rows, while both
146 X- and Y-variables were introduced in columns. The X-variables (also called predictors),
147 were the spectral data, while the Y-variables (or responses) were the amygdalin
148 concentrations established by HPLC or, in the case of discriminant models, a dummy
149 variable.

150 To construct the chemometric models for both, amygdalin prediction and
151 discrimination of almonds according to their bitterness, the total number of samples (N =
152 360) was divided in two sets: training and evaluation. The training set contained 80 % of

153 the almonds, which were randomly selected. Once a model is constructed, it is internally
154 validated using full cross-validation (CV, leave-one-out method) (Huang *et al.*, 2008).
155 The evaluation set contained the remaining 20 % of the samples (Soares *et al.*, 2013).
156
157 *2.4.1. PLS model construction to predict amygdalin content*
158 A PLS model was constructed to predict amygdalin concentrations. For PLS, the
159 covariance between the linear functions of the spectral variations (X-variables) and the
160 corresponding defined value of amygdalin concentration (Y-variable) was maximized.
161 The performance of the model was evaluated by the number of PLS factors (latent
162 variables, LV), the root mean square error of calibration, cross-validation and prediction
163 (RMSEC, RMSECV and RMSEP, respectively), and by the determination coefficient for
164 calibration, cross-validation and prediction (R^2_C, R^2_{CV}, R^2_P, respectively). Satisfactory
165 models are characterized by low RMSE, high R^2 and small differences between RMSEC
166 and RMSEP. Large differences could indicate the introduction of too many LVs in the
167 model (Bureau *et al.*, 2009).
168
169 *2.4.2. Discrimination of sweet and bitter almonds*
170 LDA, QDA and PLS-DA models were constructed to discriminate almonds according to
171 their bitterness. These models are supervised algorithms based on the relationship
172 between spectral intensity and sample characteristics; in this study, spectral variations
173 were the X-variables and sweet and bitter categories were Y-variables. In the case of PLS-
174 DA, a reference value (sweet almonds = 0 and bitter almonds = 1) was assigned to each
175 sample. In this study, a 0.5 threshold value was selected for the construction of the PLS-
176 DA models (Cortés *et al.*, 2016, Camilo *et al.*, 2012). Predicted values higher than 0.5
177 indicated that the sample belongs to the bitter class. For LDA and QDA, the Y-variable

was a categorical value created by assigning different letters to sweet and bitter almonds. Both, LDA and QDA require a number of variables lower than the number of subjects for model construction (Wu *et al.*, 2003, Sádecká *et al.*, 2016). Therefore, it is necessary to perform a variable reduction before a statistical procedure can be applied. The reduction can be done by PCA scores as input data, since the principal components (PCs) are found as linear transformations that are uncorrelated (Rodriguez-Campos *et al.*, 2011).

Finally, the performance of the three discriminant models was estimated and compared by accuracy, which is defined as the ratio of evaluation set samples correctly assigned into their respective categories.

3. Results and discussion

3.1. Determination of amygdalin concentration in sweet and bitter almonds by HPLC

The amygdalin concentration of the 360 almonds (bitter and sweet) was measured using the extraction and chromatographic conditions previously reported. The mean amygdalin concentration and ranges obtained for each class are shown in Table 1. The amygdalin concentration of bitter almonds (comprised between 15000 and 50000 mg kg^{-1}) was more than 400 times higher than the medium value of sweet almonds. On the other hand, the concentration in sweet almonds varied among the different varieties: the lowest content was obtained for 'Planeta' (0.7 mg kg^{-1}), while the highest concentration was registered for 'Guara' (350 mg kg^{-1}). These concentrations are similar to those previously published by other authors (Lee *et al.*, 2013).

Table 1

Amygdalin concentrations, with minimum and maximum values in parentheses, of the sweet and bitter almonds measured by HPLC.

Almonds	Genetic variety	Amygdalin concentration (mg kg^{-1})
Bitter	Mix of varieties	31000 (15000-50000)
Sweet	'Planeta'	20.4 (0.7-211.3)
	'Comuna'	40.0 (1.0-174.7)
	'Largueta'	70 (20-180)
	'Rumbeta'	80 (3-19)
	'Marcona'	65 (6-173)
	'Guara'	150 (50-350)

3.2. ATR-FIR spectral analysis

The raw mean spectra of sweet and bitter intact almonds observed between 4000 and 600 cm^{-1} is shown in Fig. 1. A total of 16 regions, which corresponded to each peak or shoulder observed, are evidenced in the spectra. Each one of these regions represented structural of functional group information, as indicated in Table 2. The whole spectra are the combination of many almond constituents, which are mainly represented by the combination of O-H stretching, C-H bending and C-O stretching. The broad band (3140-2808 cm^{-1}) containing three different peaks was mainly assigned to the stretching

vibrations of CH_2 functional group (Hernández & Zacconi, 2009, Maqsood & Benjakul, 2010). In this sense, this band could be associated with the saturated fatty acids fraction present in almonds. The band appearing between 1880–1680 cm^{-1} was due to the stretching movement of the typical ester carbonyl functional group (C=O) of the triacylglyceride esters (Beltrán *et al.*, 2009, Vlachos *et al.*, 2006). The band located around and 1490–1406 cm^{-1} is associated with the presence of CH bending vibrations in CH_3 (Hernández & Zacconi, 2009).

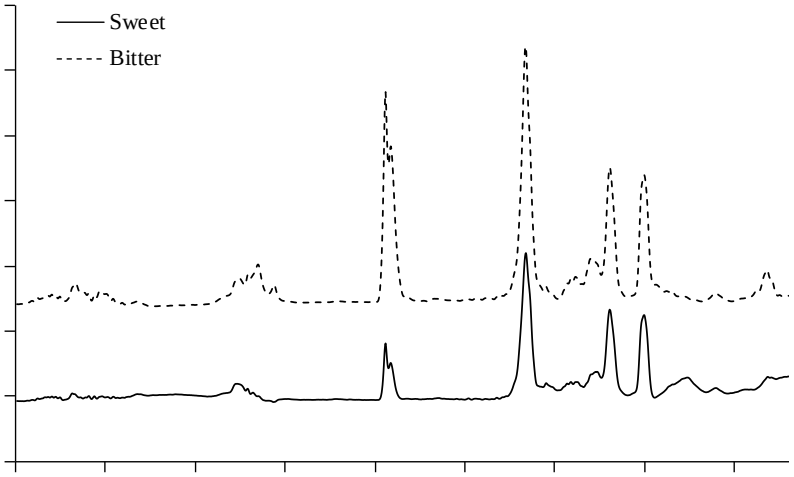


Fig. 1. Raw mean ATR-FTIR spectra for sweet and bitter almonds.

Table 2

Bands observed in the 4000 – 600 cm^{-1} region of the ATR-FTIR spectra of sweet and bitter almonds[a].

Identification no.	Wavenumber range (cm^{-1})	Functional group	Nominal frequency	Mode of vibration
1	3792-3660	O-H	3741	Stretching
2	3660-3535	O-H	3629	Stretching
3	3535-3400	O-H	3471	-
4	3140-2985	=C-H (trans)	3025	Stretching
5	2985-2882	-C-H (CH_2)	2930	Stretching (asym)
6	2882-2808	-C-H (CH_2)	2860	Stretching (sym)
7	2430-2280	alkane	2360	Stretching
		alkane	2341	Stretching
8	1880-1680	-C=O (ester)	1740	Stretching
9	1680-1604	-C=C- (cis)	1654	Stretching
10	1604-1490	N-H	1531	Bending
11	1490-1406	-C-H (CH_3)	1446	Bending (asym)
12	1406-1296	O-H	1365	Bending (in plane)
13	1296-1170	-C-O	1218	Stretching
14	1170-950	-C-O	1033	Stretching
15	950-830	-HC=CH- (cis)?	906	Bending (out of plane)
16	760-600	C-H	680	Bending (out of plane)

[a]According to references Rohman *et al.*, 2011, Chen *et al.*, 2010, Mbonyiryivuze *et al.*, 2015, Lingegowda *et al.*, 2012, Conrad *et al.*, 2014, Lerma-García *et al.*, 2010.

234 *3.3. PCA analysis*

235

236 The data from the pre-treated ATR-FTIR spectra were analyzed by PCA because
237 is a simply, rapidly and accurately way to identify the different almond groups. The two
238 first PCs summarized the 86 % of accumulative contribution of the original data, which
239 means that nearly all the variation of the variables were explained by these two PCs. The
240 first principal component (PC1) offers the main contribution (69 %), while the second
241 one (PC2) explained 17 %. Fig. 2 shows the score plot and the X-loading plot of the
242 training set samples based on the two first PCs. The score plot indicates that sweet and
243 bitter almonds can be grouped into two separate groups. Specifically, sweet almonds are
244 located on the positive axis of PC1, while the bitter almonds were located on the same
245 axis but on the opposite side (negative axis of PC1). The X-loading plot indicated the
246 absorbance peaks with greatest discriminatory effect on PCA. In particular, most of these
247 peaks (Fig. 2B) were found at the nominal frequencies included in Table 2, such as 1218,
248 1740, 2860 and 2930 cm^{-1}, while the other main wavelength (2350 cm^{-1}) is found within
249 the wavenumber range of region 7.

250

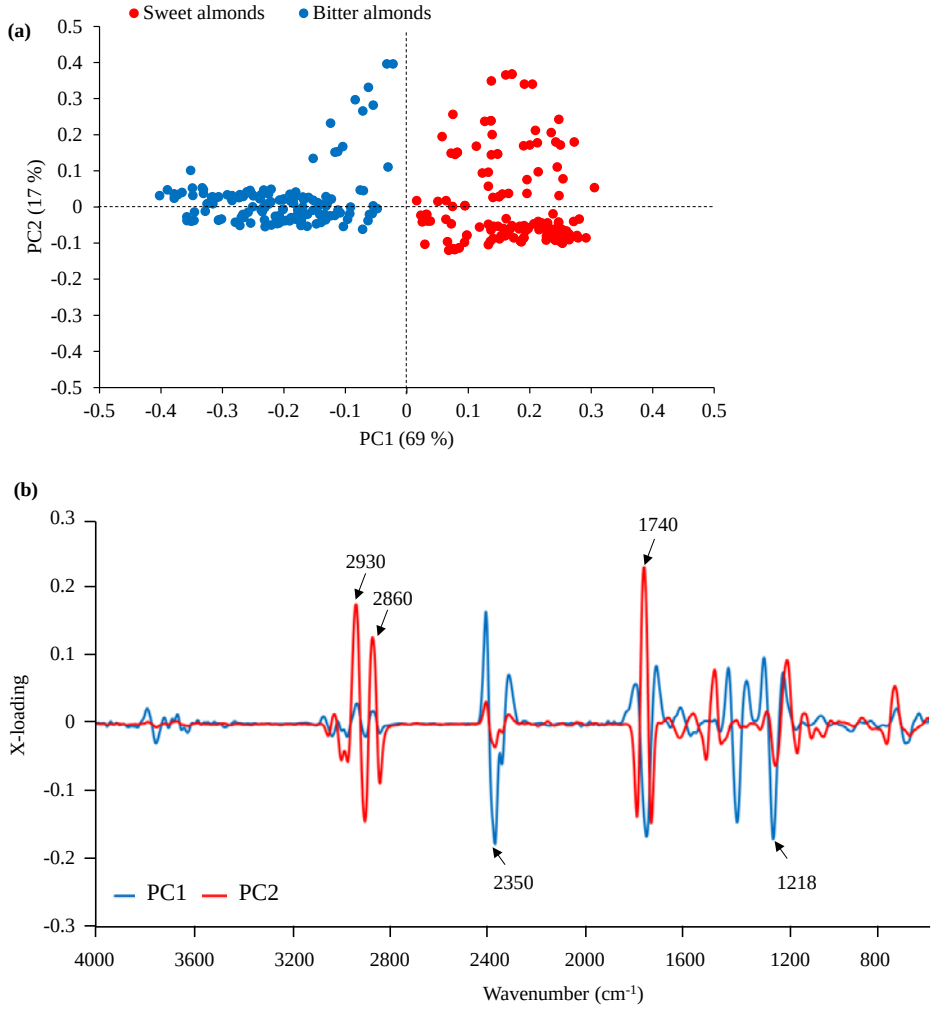
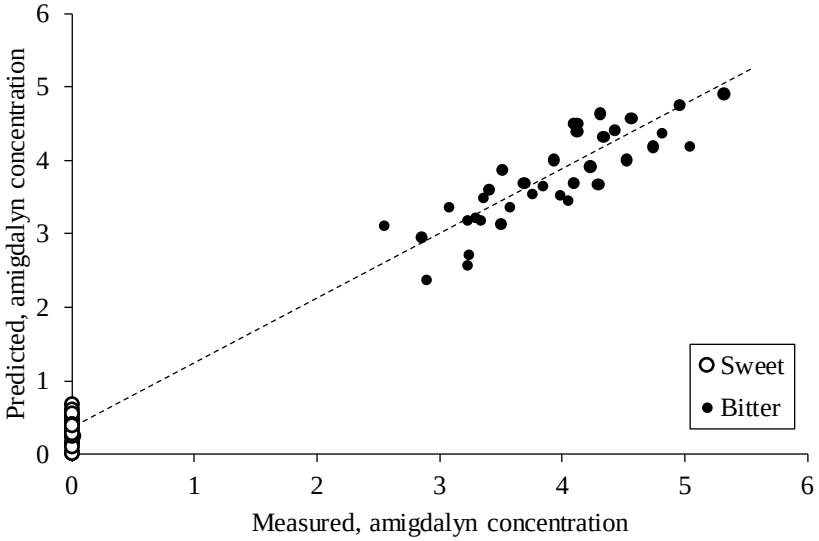


Fig. 2. PCA (A) score plot and (B) X-loading plot of pre-treated spectral data of almonds.

3.4. Prediction of amygdalin concentration

A PLS model was constructed in order to predict the amygdalin concentrations of almond samples. A total of 13 LVs were selected by the model. For the training set samples, R^2_C and RMSEC were, respectively, 0.949 and 0.348, whereas, after applying leave-one-out cross-validation technique, R^2_{CV} and RMSECV values were 0.899 and 0.492, respectively. The high R^2 value and the low values of RMSEC and RMSECV indicate good accuracy and precision of the PLS model (Saraiva *et al.*, 2017). Finally, when the performance of the PLS model was estimated by the evaluation set samples, the

263 R^2_P and RMSEP values were 0.951 and 0.398, respectively. Taking into consideration
264 that the RMSEP value was small, it could be concluded that the PLS model constructed
265 also provided a good performance in the prediction of amygdalin concentration in the
266 evaluation samples. The good results are also evidenced in Fig. 3.
267



268

269 **Fig. 3.** Measured versus predicted amygdalin concentration by PLS in the evaluation set.
270

271 *3.5. Discrimination of sweet and bitter almonds*
272

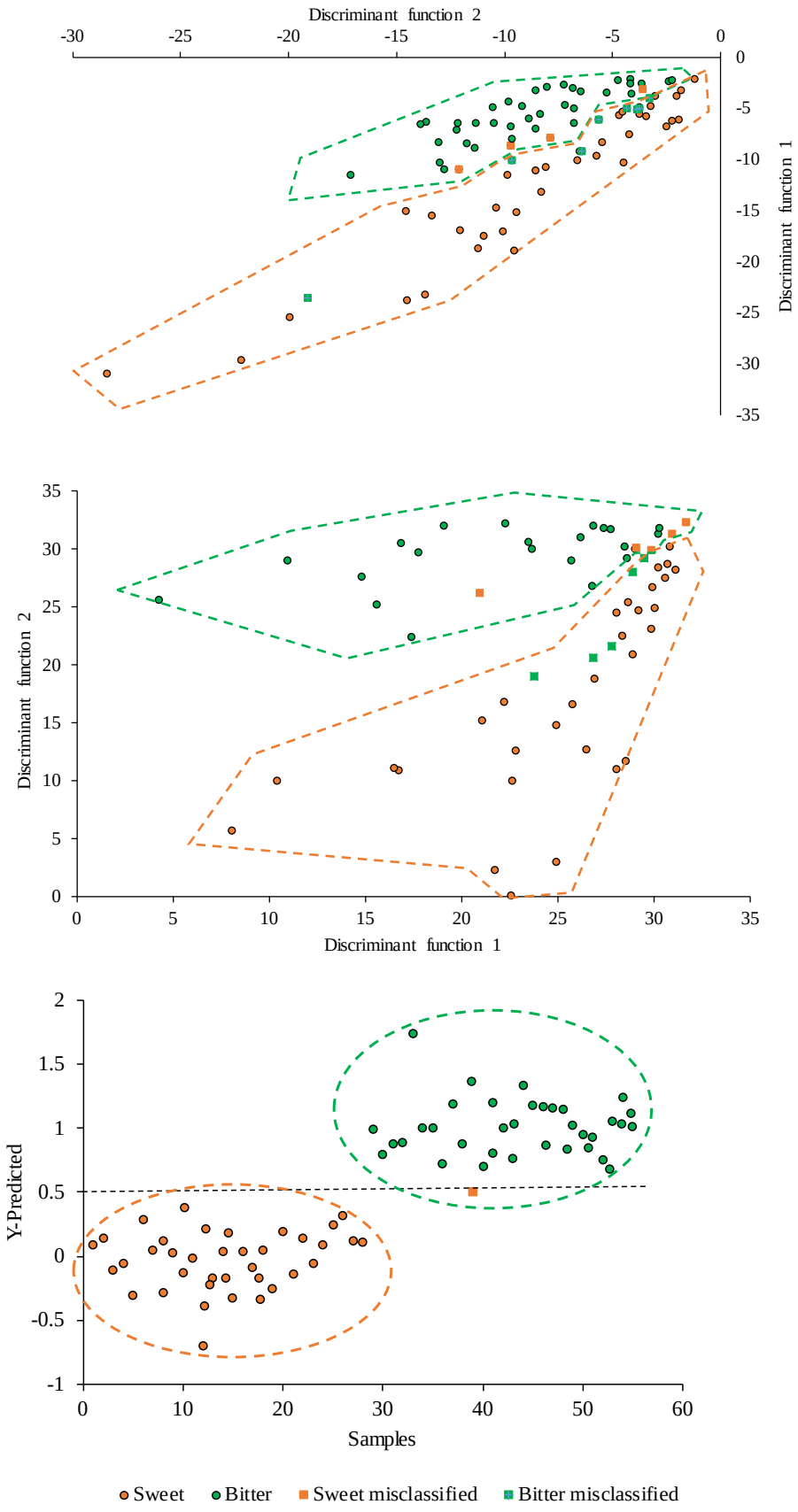
273 To discriminate almonds according to their bitterness, LDA, QDA and PLS-DA
274 were also applied to the pre-treated spectral data. For LDA and QDA, a PCA model was
275 first developed to reduce variables. The first 9 PCs explained 99.5 % of the variance of
276 the spectral data, and then they were used for model construction. Table 3 presents the
277 assignation of the evaluation set samples into the two studied classes (sweet and bitter)
278 and a summary of the classification accuracy using PLS-DA, QDA and LDA. Moreover,
279 as it can be observed in Fig. 4, only one sample of bitter category was not correctly
280 assigned for PLS-DA, while for LDA and QDA, four and five samples of sweet category,

and eight and five samples of bitter category, respectively, were not correctly assigned. Thus, ATR-FTIR spectroscopy gave good classification performances on both classes studied, especially with the PLS-DA model with an accuracy percentage of 98.6 %, although the results of the QDA and LDA models were also very satisfactory with 86.1 % and 83.3 % of accuracy, respectively.

Table 3

Assignation of the evaluation set samples into the two studied classes and overall accuracy using LDA, QDA and PLS-DA.

Method	**True classes**	**Predicted classes (%)**		**Overall accuracy (%)**
		Sweet	*Bitter*	
LDA	*Sweet*	88.89	11.11	83.3
	Bitter	22.22	77.78	
QDA	*Sweet*	86.11	13.89	86.1
	Bitter	13.89	86.11	
PLS-DA	*Sweet*	100	-	98.6
	Bitter	2.78	97.22	



292

293 **Fig. 4.** Discrimination plots of the (A) LDA, (B) QDA and (C) PLS-DA models

294 constructed to classify the evaluation set almonds into bitter and sweet categories.

4. Conclusions

The results obtained by the discriminant methods (LDA, QDA and PLS-DA) indicate that the proposed ATR-FTIR technique is a promising alternative to identify bitter and sweet almonds. Moreover, the appropriate use of the PLS model can provide useful information for the prediction of amygdalin concentration in almonds. All these advantages combined with the saving time and non-destructive analysis of a large number of samples allow operators to quickly monitor characterization of intact almonds with the purpose of a better management of homogeneous lots.

Acknowledgements

Victoria Cortés López thanks the Spanish Ministry of Education, Culture and Sports for the FPU grant (FPU13/04202). The authors wish to thank the cooperative Agricoop for kindly donating the almonds.

References

Barnes, R. J., Dhanoa, M. S., Lister, S. J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc., 43, 772–777. DOI: 10.1366/0003702894202201.

Beghi, R., Giovenzana, V., Tugnolo, A., Guidetti, R., 2018. Application of visible/near infrared spectroscopy to quality control of fresh fruits and vegetables in large-scale mass distribution channels: a preliminary test on carrots and tomatoes. J. Sci. Food Agric., 98, 2729–2734. DOI: 10.1002/jsfa.8768.

Beltrán, A., Prats, M.S., Maestre, S. M., Grané, N., Martín, M.L., 2009. Classification of four almond cultivars using oil degradation parameters based on FTIR and GC data. J. Am. Oil Chem. Soc., 86, 51–58. DOI: 10.1007/s11746-008-1323-x.

Borrás, E., Amigo, J. M., van den Berg, F., Boqué, R., Busto, O., 2014. Fast and robust discrimination of almonds (*Prunus amygdalus*) with respect to their bitterness by using near infrared and partial least squares-discriminant analysis. Food Chem., 153, 15–19. DOI: 10.1016/j.foodchem.2013.12.032.

Bureau, S., Ruiz, D., Reich, M., Gouble, B., Bertrand, D., Audergon, J. M., Renard, C. M., 2009. Rapid and non-destructive analysis of apricot fruit quality using FT-near-infrared spectroscopy. Food Chem., 113, 1323–1328. DOI: 10.1016/j.foodchem.2008.08.066.

Camilo, A., Bezerra, L. F., Pimentel, M. F, Coelho, M. J., 2012. Detection of adulteration in hydrated ethyl alcohol fuel using infrared spectroscopy and supervised pattern recognition methods. Talanta, 93, 129–134. DOI: 10.1016/j.talanta.2012.01.060.

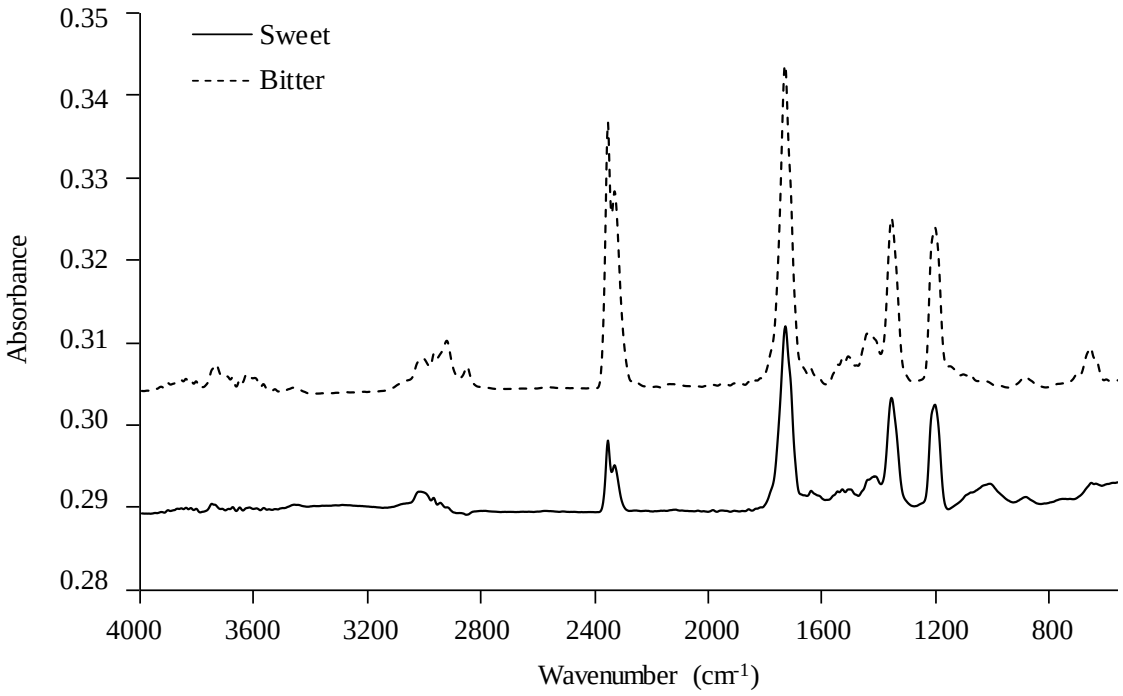
Chen, H., Ferrari, C., Angiuli, M., Yao, J., Raspi, C., Bramanti, E., 2010. Qualitative and quantitative analysis of wood samples by Fourier transform infrared spectroscopy

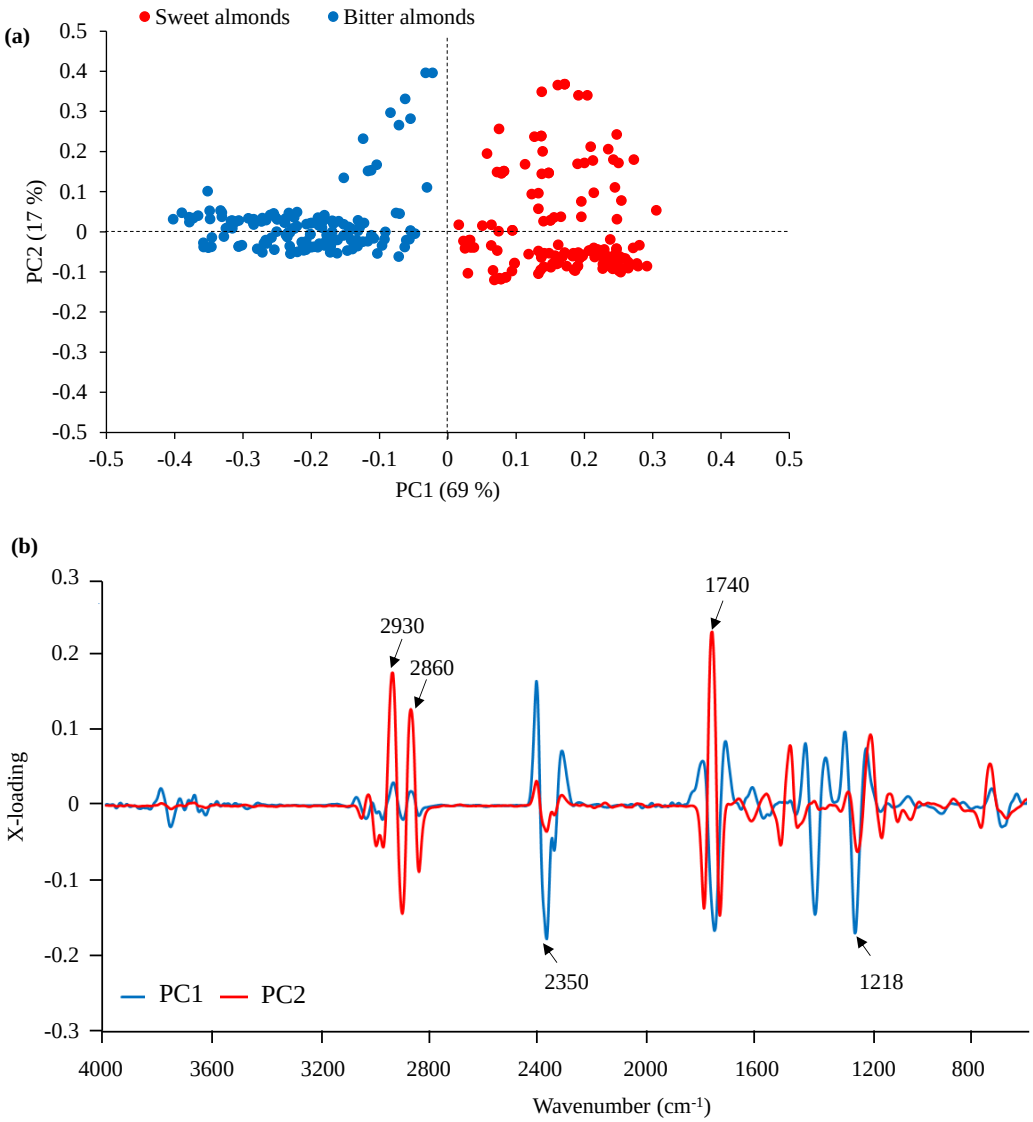
335 and multivariate analysis. Carbohydr. Polym., 82, 772–778. DOI:
336 10.1016/j.carbpol.2010.05.052.
337 Chun-Song, C., Can-Jian, W., Liang, J., Chi-Chou, L., Hua, Z., Zhin-Feng, Zhang, Z. F.,
338 2017. A new approach for identification of medicinal almonds by fourier transform
339 infrared spectroscopy and systematic clustering of characteristic peaks. Chin. J. Nat.
340 Med., 15, 703–709. DOI: 10.1016/S1875-5364(17)30100-0.
341 Conrad, A. O., Rodriguez-Saona, L. E., McPherson, B. A., Wood, D. L., Bonello, P.,
342 2014. Identification of Quercus agrifolia (coast live oak) resistant to the invasive
343 pathogen Phytophthora ramorum in native stands using Fourier-transform infrared
344 (FT-IR) spectroscopy. Front. Plant Sci., 5, 521. DOI: 10.3389/fpls.2014.00521.
345 Cortés, V., Ortiz, C., Aleixos, N., Blasco, J., Cubero, S., Talens, P., 2016. A new internal
346 quality index for mango and its prediction by external visible and near-infrared
347 reflection spectroscopy. Postharvest Biol. Technol., 118, 148–158. DOI:
348 10.1016/j.postharvbio.2016.04.011.
349 Dogan, A., Siyakus, G., Severcan, F., 2007. FTIR spectroscopic characterization of
350 irradiated hazelnut (*Corylus avellana L.*). Food Chem., 100, 1106–1114. DOI:
351 10.1016/j.foodchem.2005.11.017.
352 Ellis, D. I., Broadhurst, D., Clarke, S. J., Goodacre, R., 2005. Rapid identification of
353 closely related muscle foods by vibrational spectroscopy and machine learning.
354 Analyst, 130, 1648–1654. DOI: 10.1039/B511484E.
355 Ellis, D. I., Broadhurst, D., Kell, D. B., Rowland, J. J., Goodacre, R., 2002. Rapid and
356 quantitative detection of the microbial spoilage of meat by Fourier transform infrared
357 spectroscopy and machine learning. Appl. Environ. Microbiol., 68, 2822–2828. DOI:
358 10.1128/AEM.68.6.2822-2828.2002.

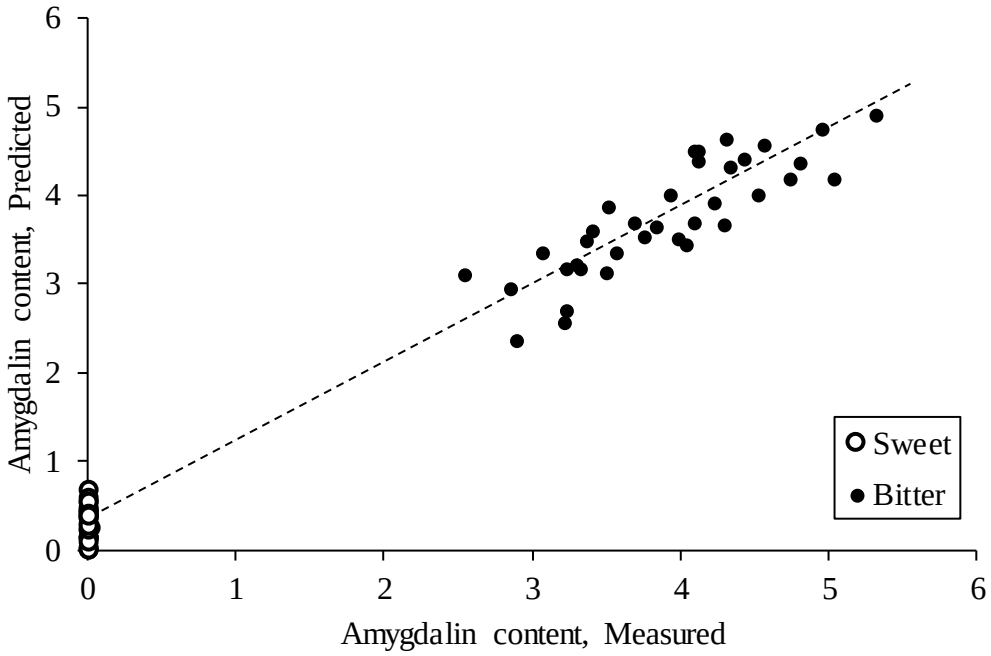
359 Gorry, P. A., 1990. General least-squares smoothing and differentiation by the
360 convolution (Savitzky-Golay) method. Anal. Chem., 62, 570–573. DOI:
361 10.1021/ac00205a007.
362 Grane-Teruel, N., Prats-Moya, M. S., Berenguer-Navarro, V., Martin-Carratala, M. L.,
363 2001. A possible way to predict the genetic relatedness of selected almond cultivars.
364 J. Am. Oil Chem. Soc., 78, 617–619. DOI: 10.1007/s11746-001-0314-z.
365 Hernández, S., Zacconi, F., 2009. Aceite de almendras dulces: Extracción, caracterización
366 y aplicación. Quim. Nova, 32, 1342–1345. DOI: 10.1590/S0100-
367 40422009000500044.
368 Huang, H., Yu, H., Xu, H., Ying, Y., 2008. Near infrared spectroscopy for on/in-line
369 monitoring of quality in foods and beverages: a review. J. Food Eng., 87, 303–313.
370 DOI: 10.1016/j.jfoodeng.2007.12.022.
371 Lee, J., Zhang, G., Wood, E., Rogel Castillo, C., Mitchell, A. E., 2013. Quantification of
372 amygdalin in nonbitter, semibitter, and bitter almonds (*Prunus dulcis*) by UHPLC-
373 (ESI) QqQ MS/MS. J. Agric. Food Chem., 61, 7754–7759. DOI: 10.1021/jf402295u.
374 Lerma-García, M. J., Ramis-Ramos, G., Herrero-Martínez, J. M., Simó-Alfonso, E. F.,
375 2010. Authentication of extra virgin olive oils by Fourier-transform infrared
376 spectroscopy. Food Chem., 118, 78–83. DOI: 10.1016/j.foodchem.2009.04.092.
377 Lingegowda, D. C., Kumar, J. K., Prasad, A. D., Zarei, M., Gopal, S., 2012. FTIR
378 spectroscopic studies on Cleome gynandra–comparative analysis of functional group
379 before and after extraction. Rom. J. Biopys, 22, 137–143.
380 Maqsood, S., Benjakul, S., 2010. Comparative studies of four different phenolic
381 compounds on in vitro antioxidative activity and the preventive effect on lipid
382 oxidation of fish oil emulsion and fish mince. Food Chem., 119, 123–132. DOI:
383 10.1016/j.foodchem.2009.06.004.

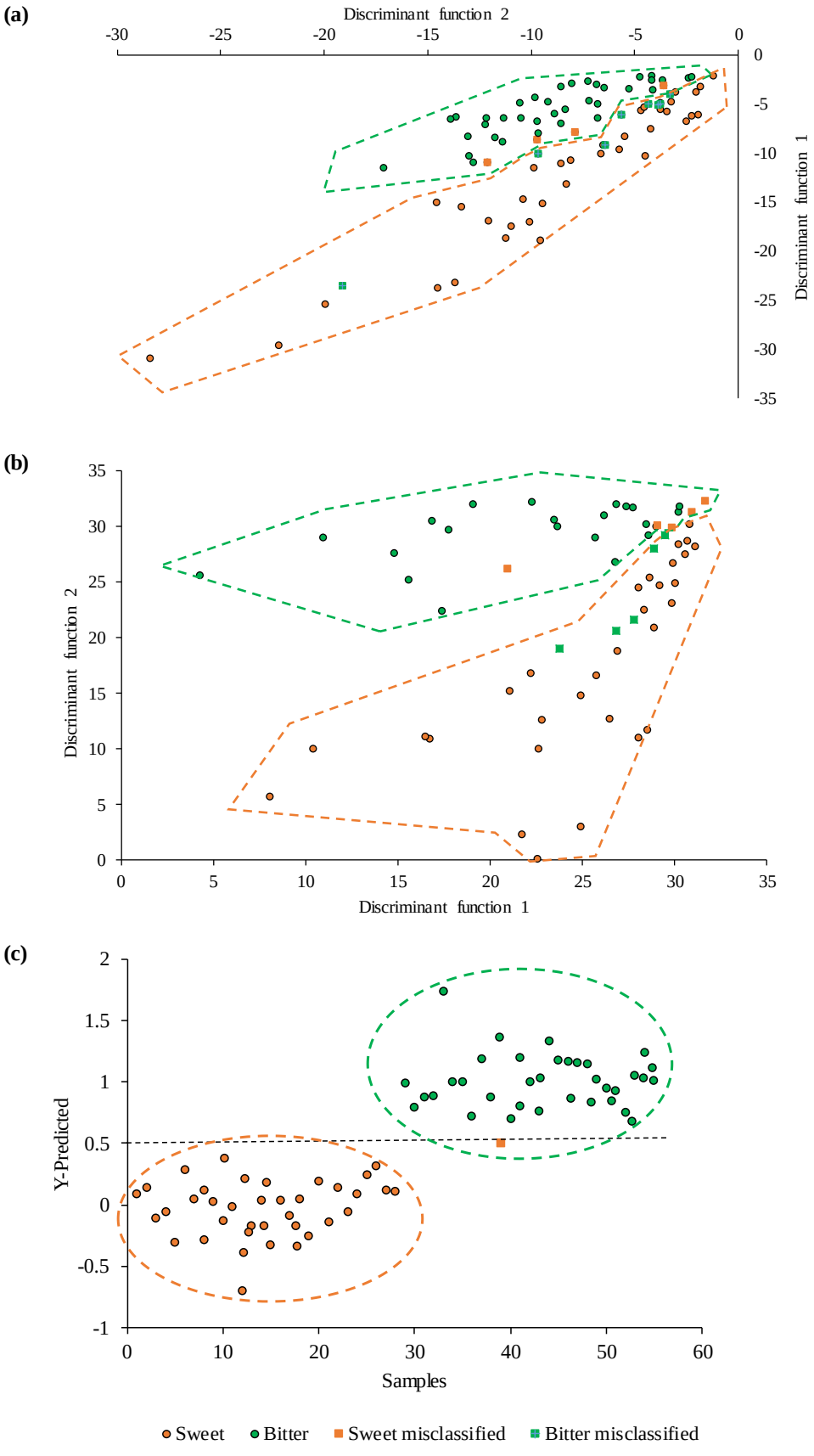
384 Mbonyiryivuze, A., Mwakikunga, B., Dhlamini, S. M., Maaza, M., 2015. Fourier
385 transform infrared spectroscopy for sepia melanin. Phys. Mat. Chem., 3, 25–29.
386 Micklander, E., Brimer, L., Engelsen, S. B., 2002. Noninvasive assay for cyanogenic
387 constituents in plants by raman spectroscopy: content and distribution of amygdalin
388 in bitter almond (*Prunus amygdalus*). Appl. Spectrosc., 56, 1139–1146. DOI:
389 10.1366/000370202760295368.
390 Rodriguez-Campos, J., Escalona-Buendía, H. B., Orozco-Avila, I., Lugo-Cervantes, E.,
391 Jaramillo-Flores, M. E., 2011. Dynamics of volatile and non-volatile compounds in
392 cocoa (*Theobroma cacao L.*) during fermentation and drying processes using
393 principal components analysis. Food Res. Int., 44, 250–258. DOI:
394 10.1016/j.foodres.2010.10.028.
395 Rodriguez-Saona, L. E., Fry, F. S., McLaughlin, A., Calvey, E. M., 2001. Rapid analysis
396 of sugars in fruit juices by FT-NIR spectroscopy. Carbohydr. Res., 336, 63–74. DOI:
397 10.1016/S0008-6215(01)00244-0.
398 Rohman, A., Erwanto, Y., Man, Y. B. C., 2011. Analysis of pork adulteration in beef
399 meatball using Fourier transform infrared (FTIR) spectroscopy. Meat Sci., 88, 91–
400 95. DOI: 10.1016/j.meatsci.2010.12.007.
401 Sádecká, J., Jakubíková, M., Májek, P., Kleinová, A., 2016. Classification of plum spirit
402 drinks by synchronous fluorescence spectroscopy. Food Chem., 196, 783–790. DOI:
403 10.1016/j.foodchem.2015.10.001.
404 Sánchez-Pérez, R., Jørgensen, K., Olsen, C. E., Dicenta, F., & Møller, B. L., 2008.
405 Bitterness in Almonds. Plant Physiol., 146, 1040–1052. DOI:
406 10.1104/pp.107.112979.
407 Saraiva, C., Vasconcelos, H., de Almeida, J. M., 2017. A chemometrics approach applied
408 to Fourier transform infrared spectroscopy (FTIR) for monitoring the spoilage of

409 fresh salmon (*Salmo salar*) stored under modified atmospheres. Int. J. Food
410 Microbiol., 241, 331–339. DOI: 10.1016/j.ijfoodmicro.2016.10.038.
411 Savitzky, A., Golay, M. J. E., 1964. Smoothing and differentiation of data by simplified
412 squares procedures. Anal. Chem., 36, 1627–1639. DOI: 10.1021/ac60214a047.
413 Soares, S. F. C., Gomes, A. A., Galvão Filho, A. R., Araújo, M. C. U., Galvão, R. K. H.,
414 2013. The successive projections algorithm. Trends Anal. Chem., 42, 84–98. DOI:
415 10.1016/j.trac.2012.09.006.
416 Vahur, S., Knuutinen, U., Leito, I., 2009. ATR-FT-IR spectroscopy in the region of 500–
417 230 cm^{-1} for identification of inorganic red pigments. Spectroc. Acta Pt. A- Molec.
418 Biomolec. Spectr., 73, 764–771. DOI: 10.1016/j.saa.2009.12.056.
419 Valdés, A., Beltrán, A., Garrigós, M. C., 2013. Characterization and classification of
420 almond cultivars by using spectroscopic and thermal techniques. J. Food Sci., 78,
421 138-144. DOI: 10.1111/1750-3841.12031.
422 Velasco, J. F., Aznar, J.A. Boletín Económico de ICE, 2016, 3079, 77–87.
423 Vlachos, N., Skopelitis, Y., Psaroudaki, M., Konstantinidou, V., Chatzilazarou, A.,
424 Tegou, E., 2006. Applications of Fourier transform-infrared spectroscopy to edible
425 oils. Anal. Chim. Acta, 573, 459–465. DOI: 10.1016/j.aca.2006.05.034.
426 Wu, B., Abbott, T., Fishman, D., McMurray, W., Mor, G., Stone, K., Ward, D., Williams,
427 K., Zhao, H., 2003. Comparison of statistical methods for classification of ovarian
428 cancer using mass spectrometry data. Bioinformatics, 19, 1636–1643. DOI:
429 10.1093/bioinformatics/btg210.


Sweet
Bitter
Absorbance
0.35
0.34
0.33
0.32
0.31
0.30
0.29
0.28
4000
3600
3200
2800
2400
2000
1600
1200
800
Wavenumber (cm-1)


(a)
Sweet almonds
Bitter almonds
PC2 (17 %)
PC1 (69 %)
(b)
X-loading
2930
2860
1740
2350
1218
PC1
PC2
Wavenumber (cm-1)


Amygdalin content, Predicted
Amygdalin content, Measured
0
1
2
3
4
5
6
Sweet
Bitter


(a)
Discriminant function 2
-30
-25
-20
-15
-10
-5
0
0
-5
-10
-15
-20
-25
-30
-35
Discriminant function 1
(b)
35
30
25
20
15
10
5
0
Discriminant function 2
0
5
10
15
20
25
30
35
Discriminant function 1
(c)
2
1.5
1
0.5
0
-0.5
-1
Y-Predicted
0
10
20
30
40
50
60
Samples
Sweet
Bitter
Sweet misclassified
Bitter misclassified