VIS/NIR spectral signature for the identification of peanut contamination of powder foods

Satyabrata Ghosh[*1], Teresa R. Cuadrado Domínguez[1], Belén Diezma[*1], Lourdes Lleó[1], Pilar Barreiro[1*], Teresa García Lacarra[2], Jean-Michel Roger[3]

(1. Department of Ingeniería Rural, ETSI Agrónomos, Avda. Complutense s/n,Universidad Politécnica de Madrid (UPM), CEI Moncloa,28040 Madrid, Spain;

2. Nutrición, Bromatología and Tecnología de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, CEI Moncloa,28040 Madrid, Spain;

3.UMR ITAP, IRSTEA Montpellier, France)

Abstract: Visible-near infrared reflectance spectra are proposed for the characterization of IRMM 481 peanuts variety in comparison to powder food materials: wheat flour, milk and cocoa. Multidimensional analysis of reflectance spectra of powder samples shows a specific NIR band centred at 1200 nm that identifies peanut compared to the rest of food ingredients, regardless compaction level and temperature. Spectral range of 400-1000 nm is not robust for identification of blanched peanut. The visible range has shown to be reliable for the identification of pre-treatment and processing of unknown commercial peanut samples. A spectral index is proposed based on the combination of three wavelengths around 1200 nm that is 100% robust against pre-treatment (raw or blanched) and roasting (various temperatures and treatment duration).

Keywords: VIS/NIR spectroscopy; peanut; PCA

Citation: Ghosh, S., T. R. C. Domínguez, B. Diezma, L. Lleó, P. Barreiro, T. G. Lacarra, and J. M. Roger. 2015. VIS/NIR spectral signature for the identification of peanut contamination of powder foods. AgricEngInt: CIGR Journal, 17(2):310-329.

1 Introduction

Peanut (*Arachis hypogaea*) is a very popular food ingredient which is used in various commercial food materials such as biscuit, bread and confectionery product (Hird et al., 2003). In twenty century peanut has been considered as a most severe food allergen for the commercial food material (Hourihane et al., 1997). But the avoidance of peanut-containing foods can be difficult for peanut allergen sufferers and food producers, and thus reliable analytical methods for the detection of hidden allergens in foodstuff are required by the food industry and control agencies.

The allergenic proteins of peanut can be identified by a traditional protein detection method such as enzyme-linked immunosorbent assay (ELISA), which is based on the antigen-antibodies interaction (Platteau et al., 2011; Scaravelli et al., 2008). This method is very sensitive and the most commonly used by the industry and official food control agencies (Besler, 2001). However, sometimes peanut allergenic proteins are modified due to processing and may fail to detect the allergen protein (Immer, 2006; Taylor et al., 2009). Real Time Polymerase Chain Reaction (RT-PCR) is an alternative method in which the detection of allergens is made by means of DNA-based methods. The target molecules, DNA sequences, are amplified by the RT-PCR. For peanut allergen, several RT-PCR assays

Received date: 2015-01-21 **Accepted date:** 2015-03-05

***Corresponding author: Satyabrata Ghosh, Belén Diezma, Barreiro.** Department of Ingeniería Rural, ETSI Agrónomos, Avda. Complutense s/n,Universidad Politécnica de Madrid (UPM), CEI Moncloa,28040 Madrid, Spain. Tel: 34-91-3363260. Fax: 34-91-3365845. Email: satyabrataghosg23@gmail.com; pilar.barreiro@upm.es; belen.diezma@upm.es.

have been developed (Hird et al., 2003; Scaravelli et al., 2008).

López-Calleja et al. (2013), has recently reported the development of a real time PCR assay method for the detection of trace amounts of peanut in processed foods. This test was validated by means of peanut samples provided by the Institute for Reference Materials and Measurements (IRMM) of the European Commission, which consisted of peanut varieties samples from different geographical origins exposed to different treatments. The authors have also demonstrated through RT-PCR that 27 out of 133 commercial food products contained peanut traces while they did not declare their presences in the labelling. Still, the main drawback of RT-PCR is the requirement of skilled labour, while being time consuming and expensive as well.

The application of spectroscopy to evaluate product quality offers potential improvements in cost-efficiencies compared to other analytical procedures, especially where non-destructive techniques can be adapted to in-line sorting and processing (Phan-Thien et al., 2011). Techniques using near-infrared (NIR) spectroscopy are being applied in food processing and quality inspection (Shiroma and Rodriguez-Saona, 2009), producing several advantages over conventional physical and chemical analytical methods for food quality analysis: NIR is rapid, non destructive, and achieves large information about the components present in food products (Coates et al., 2008; Mauer et al., 2009; Rubio-Diaz et al., 2011; Lembe et al., 2013). Spectra measured in the NIR range contain absorbance bands that are mainly due to three chemical bonds: C–H, which is usually from fats and oil; O–H bond which are found in water; and N–H bonds, which are found in protein (Cozzolino et al., 2008). Shiroma et al. (2009) determined fat and moisture content of potato chips, achieving the differentiation of potato chips by source of frying oil. NIR is ideal for quantitatively determining oil, protein and moisture by deducing C–H, N–H and O–H bonds (Cozzolino et al., 2008). In addition, high scatter coefficients allow for excellent diffuse reflectance spectra of solids (Sundaram et al., 2010). NIR spectroscopy may be applied with minimal sample preparation and has been used to determine peanut fatty acid concentrations of individual peanut kernels (Tillman et al., 2005; Fox et al., 2006) and peanut oil (Panforda et al., 1990). NIR has also been used to predict the total oil and fatty acid concentrations of peanut pods (Sundaram et al., 2009a; Sundaram et al., 2009b). Therefore, this very common analytical method is now being used in a more commercial aspect. Sundaram et al. (2010) reported that NIR reflectance spectroscopy is used to quantify the total amount of oil and fatty acid concentration of Virginia and Valencia types of in-shell peanuts. In such work moisture content (MC) of intact kernels of grain and nuts could be determined by NIR reflectance spectrometry (Sundaram et al., 2012). Regarding powder, full spectra (VIS-NIR) have also been accomplished for rapid and non invasive quantification of two adulterants (flour and mung bean) in spirulina powder (a dietary supplement) with a limit of detection of 10% in mass when using non-spatially resolved spectroscopy. In such work three wavelength bands were identified as the most relevant: one in the visible and the other two in the near infrared range based on a PLS model (Wu et al., 2011).

Hyperspectral imaging system (HIS) technique is a kind of spectroscopic vision system which provides information about spatial distribution, shape, texture, and mixture homogeneity. Mixture homogeneity is essential with a view to obtaining products of a high quality and uniform content. All mixing processes should ideally provide a “perfect mixture”, where all components are uniformly distributed in the mass. In practice, this is usually very difficult, especially with powder mixtures, which can contain widely different components. The type of mixer used can also affect the uniformity of the product (Rosas and Blanco, 2012). The HIS imaging system has been used to monitor a powder flows leaving a dosing feeder, showing that changing the spatial resolution of the HIS enables to view the powder as either

homogeneous or heterogeneous (Scheiblhofer et al., 2012). HIS in the NIR region has already been used for process and quality monitoring in the pharmaceutical industry with special regard to assessing homogeneous distribution of dosage, which proves the concept to be ready for technology transfer towards the food industry (Gowen et al., 2008).

According to previous work of the authors (López-Calleja et al., 2012), it is possible to segregate peanut, milk, and flour in powder under limited condition using hyperspectral vision in the range from 400 nm to 1000 nm. However, it has not been proved the reliability of these results regardless the type and treatment of peanuts. Therefore, the specificity of such procedure remains unrevealed.

Hence, the goals of this work are: to identify the spectral range to segregate peanuts in powdered foods from other ingredients regardless treatment (no treatment, blanching and roasting) and to establish the spectral bands required for a multispectral system according to the sensitivity needed when using it as a complementary and screening technique for RT-PCR analytical tools.

2 Materials and methods:

2.1 References samples

The reference peanut samples were obtained from European Commission Joint Research Centre of IRMM (Brussels, Belgium) and are the same as those used by López-Calleja et al. (2013) for the validation of RT-PCR method (Table 1). The kit (IRMM-481) with six different vials contains non-salted peanut powder with a normal particle size from 500 μm to 1000 μm. Five of the vials were filled with approximately 2 g of each variety and treatment: vial IRMM-481a (RPA), variety Runners and origin Argentina, corresponded to blanched peanuts air- roasted at 140^0C for 20 min; vial IRMM-481b (RPB) variety Common Natal from South Africa refers to raw peanuts, air roasted at 160^0C for 13 min; vial IRMM-481c (RPC) variety Virginia and origin from USA, were blanched peanuts, oil roasted at 145^0C for 25 min; vial IRMM-481d (RPD) variety Virginia and origin from, China, also corresponded to blanched fruits, oil roasted at 140^0C for 9 min; vial IRMM-481e (RPE) variety Jumbo Runners and origin from, USA, were blanched peanuts without roasting. On the other hand vial IRMM-481f (RPF) was a mixture of all five peanut vials at the same ratio.

Table 1 Five of the vials were filled with approximately 2 g of each variety and treatment

Vail No.	Variety of IRMM-484	Variety symbol	Variety name	Origin aria	Correspond to peanuts	Type of roasted	Rate of Air roasted
1	A	RPA	Runners	Argentina	blanched	Air	140 ^{0}C / 20 min
2	B	RPB	Common Natal	South Africa	raw	Air	160 ^{0}C / 13 min
3	C	RPC	Virginia	USA	blanched	Oil	145 ^{0}C / 25 min
4	D	RPD	Virginia	China	blanched	Oil	140 ^{0}C / 9 min
5	E	RPE	Jumbo Runners	USA	blanched	without	-
6	F	RPF*					

* The vial IRMM-481f (RPF) was a mixture of all five peanut vials at the same ratio.

2.2 Commercial samples

Commercial samples of peanut (MP, MP1 and MP2), skimmed milk powder (MM), wheat flour (MF), and cocoa powder (MCC) were obtained from local market (Madrid, Spain). The manufactures of wheat flour, milk powder and cocoa were Nomen, Tarragona (Spain), Central Lechera Asturiana, Asturias (Spain) and Valor Reposter á, Alicante (Spain) respectively (Table 2). In-shell peanut was widely commercially available in the market Madrid, Spain and manufactured by Itac China.

Food ingredients (MF, MM and MCC) were subjected to screening process to characterize particle size provided 965.22 more than 98% MF must pass through the sieve of 212 μm.

Table 2 Specification and characterization of commercial samples (MM, MF, MCC and MP)

Product	Brand	Nutritional value / 100 g	Grit, μm	Temperature
Milk powder (MM)	Central Lechera Asturiana Asturias (Spain)	Energy value: 2050 kJ Protein: 25 g Carbohydrates: 39 g Fat: 26 g Calcium: 1200 mg Sodium: 0.5 g	212 > MM > 160 and 125 > MM > 100	8 ℃ and 25 ℃
Wheat flour (MF)	Nomem Tarragona (Spain)	Energy value: 1426 kJ Protein: 9.5 g Carbohydrates: 72 g Fat: 1.1 g	212 > MF> 160 and 125 > MF > 100	5 ℃ and 19 ℃
Cocoa (MCC)	Valor Alicante (Spain)	Energy value: 1303 kJ Protein: 25.5 g Fat: 16 g Sugars: 0.7 g Sodium: 0.0128 g	212 > MCC > 160 and 125 > MCC > 100	8 ℃ and 20 ℃
Peanuts (MP)	Itac (China)		MP < 1000 and 2000 > MP	5 ℃ and 19 ℃

that according to the standard for wheat flour AOAC

For the screening of MF, MCC and MM, a sifter was used: ARESA ORTO with 10 vibration rates, all employed sequentially upstream. This device is provided with a battery of sieves calibrated at: 160 /125 /100 /80 /63 /50 /40 μm. For this study we selected the particle retained in the first and third sieves, respectively, so that the size of both samples corresponds to 212-160 μm and 125-100 μm. Commercial peanuts were crushed by mechanical grinder and two particle size considered: above 2000 μm and below 1000 μm. All samples were analysed at two ranges of temperature 5 ℃-10 ℃ and 19 ℃-25 ℃.

2.3 Samples preparation

Peanut samples were kept in an air tight container. The mass of each sample was 1 g, placed inside a round plastic container for the spectroscopic measurement. Material was pressed with a Chatillon (DISMAE, Model-DPP) to achieve 1.41 kg/cm^2 (or 98 N with a 30 mm flat plate) (Figure 1).

2.4 Spectroscopic instruments and measurements

Extended visible (VIS) and near infrared (NIR) spectral measurements were performed using a Hamamatsu photonic multi-channel spectrometer (Japan): C7473 and PMA-1 respectively. The optical system consisted of a bifurcated optical fiber, (Monolight Optical Spectrum Analyser, United Kingdom) that leads the incident light of a 100 W Tungsten lamp to the sample and reflected to the detector.

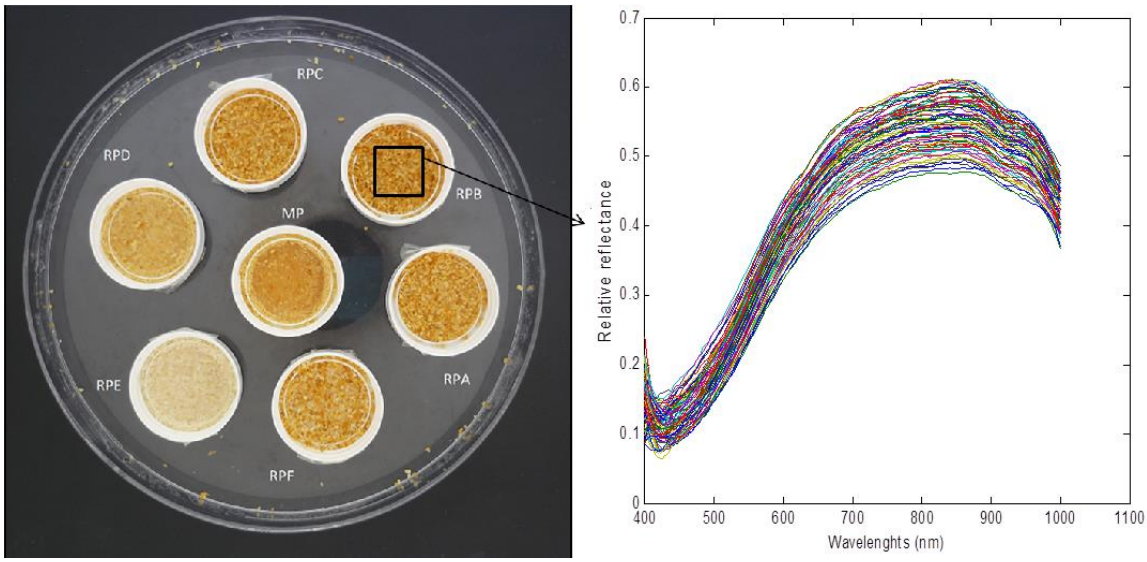


Figure 1 Reference peanuts samples (RPA-RPF) and commercial peanut (MP)

The detector of VIS and NIR equipment had wavelength ranges between 196-958 nm and 896 - 1686 nm respectively. The light source was focused on the sample in order to interact with it, and then the reflectance spectra was collected and recorded. Each measurement averaged nine spectra and were repeated three times for varying integration times: visible spectra (VISS) from 20 ms to 40 ms and NIR from 50 ms to 80 ms, and for the latter with two temperatures ranges as well: 5℃-10℃, and 19℃-25℃. For VISS, relative reflectance spectra have been considered for further analysis; for that white reference (barium sulphate plate) and dark current spectra were taken before acquiring measurements of the samples, and then the relative reflectance was computed subtracting the dark current to each raw spectrum and dividing this result by the white reference minus the dark current spectrum. For NIR, the raw spectra were considered, that is, the intensity level at each wavelength without considering the white reference.

2.5 Hyper-Spectral Measurements

A pushbroom hyperspectral camera (Hyperspec VNIR C-Series G4-131, USA) has been used with a wavelength range between 400-1000 nm. It is equipped with a progressive line-by-line scan spectrograph with an interchangeable slit of 25 μm. Hyperspectral imaging system (HIS) is surrounded by a rectangular tent made from black wood to prevent other lightning interference, and it is composed of the following components: an illumination unit which consists of a single halogen lamp adjusted at an angle of approximately 45° to illuminate the camera's field of view (FOV); a sample conveying translation stage driven by a stepping motor with movement synchronized with the image acquisition by mean of the PC supported HyperspecTM software.

Relative reflectance spectra were computed for each pixel. The selected spectral resolution was 3.2 dpi (189 wavelengths). The setup of the camera allowed adjusting the size of the pixel at 69.7 μm ×69.7 μm. Hyperspectral images were used in this study in order to achieve a high number of spectra containing spatial variability; i.e. the images were considered such as a source of spectra, and the spatial information was not analysed in the present research. Therefore, manually selected region of interests (ROI) from images were set for the spectral analysis of commercial peanut along with peanut references. Similarly, ROI were manually selected on previous hyperspectral images (López-Calleja et al., 2012) of MF, MM and MCC (particle size in MF and MM between 125 μm and 100 μm, higher than 160 μm in MCC) and MP that were included for spectral comparison and projection onto multidimensional models (Figure 2).

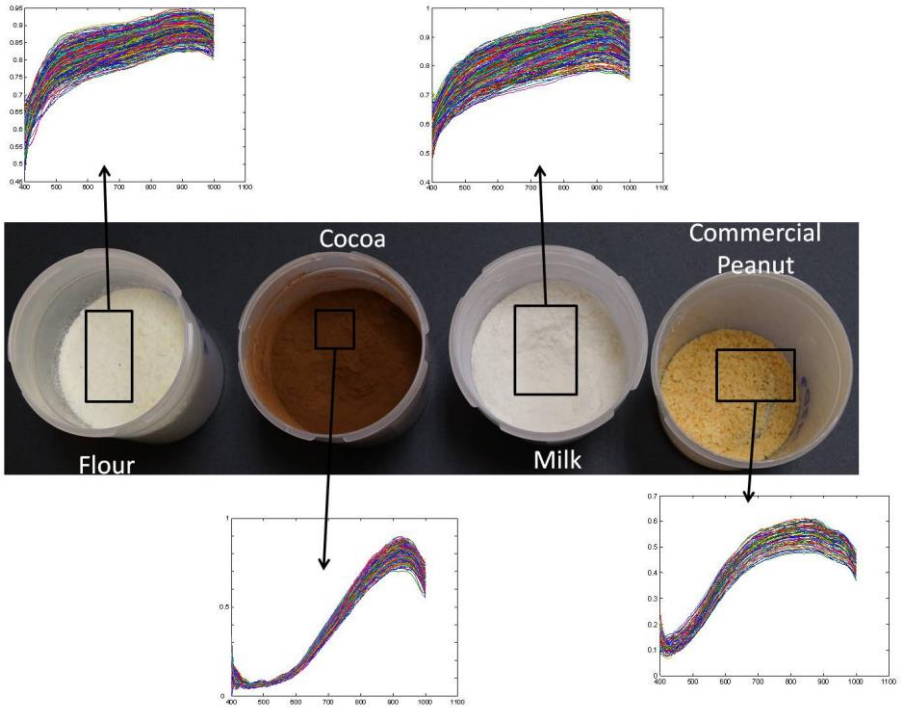


Figure 2 Particle size of commercial samples with hyper-spectral images. Particle size: 125-100 μm and >160 μm (MM, MF and MCC); <1000 μm and >2000 μm (MP)

2.6 Spectral data analysis

VIS and NIR spectra were analysed by using multivariate data analysis software (Matlab R2011 with statistical toolboxes Natick, Massachusetts, U.S.A). Three sets of data were considered in this study as calibration:

- HIS (400 -1000nm) with n = 8788 corresponding to RPA-F (n = 6670) and MP (n = 2118),
- VISS (400 - 1000 nm) with n = 117 corresponding to RPA-F (n = 108) and MP (n = 9)
- NIR (896 – 1600 nm) with n = 1110, corresponding to RPA-F (n = 158), MP (n = 323), MM (n = 210), MF (n = 215), and MCC (n = 204).

These three sets (Table 3) were used independently to perform three principal component analysis (PCA) in order to define the spectral response of the food ingredients.

Beside, a fourth spectral HIS data set from a previous research (López-Calleja et al., 2012) composed by MF, MM, MCC and MP were projected onto multidimensional models computed with the HIS calibration set.

Table 3 Calibration set of NIR, HIS and VISS

	Number of reference peanut samples						Number of commercial samples				
	RPA	RPB	RPC	RPD	RPE	RPF	MP	MM	MCC	MF	TS*
NIR	26	21	24	33	31	23	323	210	204	215	1110
HIS	758	860	685	1548	1308	1511	2118	366	-	316	8788
VSS	18	18	18	18	18	18	9	-	-	-	117

*Total number of spectra.

PCA on HIS and VISS were conducted on centred data set to assess the feasibility of segregating peanut reference samples based on peanut processing: blanching and roasting, while PCA on NIR aims at defining specific wavelength ranges for the identification of peanuts regardless treatment. Several spectral indexes were proposed to segregate the peanuts from the other powder foods. These indexes were defined based on the spectral patterns and on the most relevant wavelengths selected from the loadings of PCA. Several Analysis of Variance (ANOVA) were computed in order to compare the performance of each proposed index and the scores from PCA.

Additionally, for each spectrum of the VISS and HIS data it was computed the sum of the relative reflectance at each wavelength such as a global measurement of the intensity level of the spectrum, it was called the spectral sum (SS). Similarly, for the NIR data the spectral sum was computed for each spectrum. The normalization of each VISS, HIS and NIR spectrum was carried out in order to avoid global scattering, dividing the intensity level of each wavelength of the spectrum by SS. Then PCA was computed on these sets of spectra.

3 Results and discussion

3.1 Extented visible spectra

The analysis of extended visible spectra obtained by the VISS and HIS are presented in this section. At a first step some considerations are given with regard to relative reflectance spectra, followed by the results of principal component analysis.

Figures 3 shows that all the average reflectance spectra of the VISS and HIS are very similar. In both cases, the average reflectance spectrum from peanut RPE has higher reflectance in the visible range and it is well separated from the rest of the reference peanuts (RPA, RPB, RPC, RPD and RPF). The peanut RPE corresponds to blanched peanuts (more white) without roasting, while the rest of peanuts are all roasted either from blanched or raw peanut.

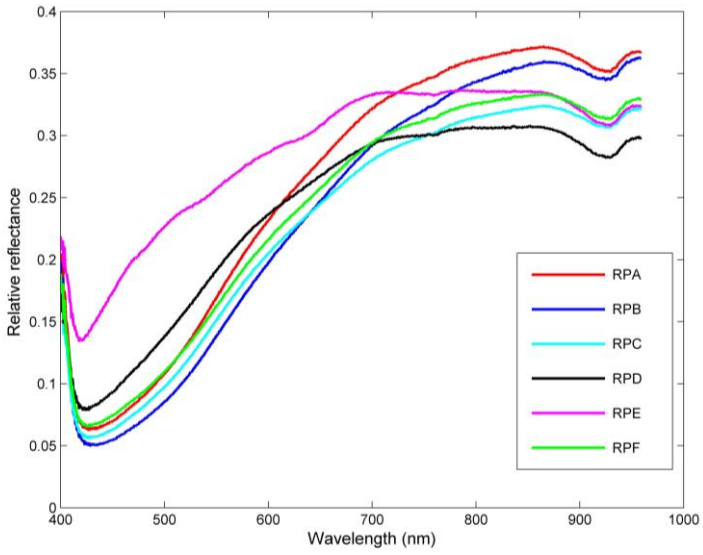


Figure 3a Average relative reflectance spectra from VISS

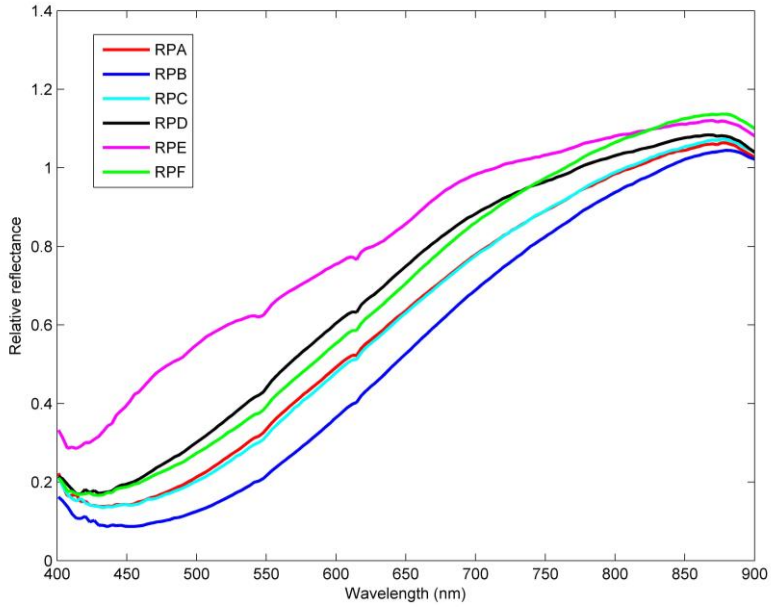


Figure 3b HIS; labels RPA to RPF correspond to IRMM 481 kit for peanuts.

Another important observation from the average raw spectra is that peanut RPD remains between peanut RPE and peanut RPA-RPF, which points to peanut RPD as an intermediate status, confirmed by it was blanched and roasted during only 9 min compared to the rest (mostly between 20 min and 25 min). It has been reported that roasting treatment has an effect on peanut properties like moisture and hexanal compound which are responsible for colour and flavour (Macdeniel, 2011). This type of physical or chemical properties has been changed in our peanuts also due to the different blanching and roasting treatment.

PCA performed on HIS data of reference and commercial peanut samples showed that PC1 represents 99.31% variance of the relative reflectance spectra. The determination coefficient (r^2) between PC1 scores and spectral sum (SS) is 96.3% which indicates that almost 96% of total spectral variance is due to the global intensity of the relative reflectance spectra.

In spite of the plane PC2 / PC3 retains only 0.64% of total variance, PC2 and PC3 scores are particularly explicative for the quality or treatment of the product while PC1 is related to the signal intensity of the relative reflectance spectra.

As it expected, the normalization procedure corrected the scattering effect and consequently PC1 and PC2 of the normalized spectra were directly related to treatments of peanuts (PC1 and PC2 retained 98.83% of the total variance).

In the same way PCA was performed on VISS data of reference and commercial peanut samples; PC1 represented 94.2% of the variance of the relative reflectance spectra and the determination coefficient (r^2) between PC1 and SS is 99.54%.

Figure 4 shows the loadings of PC2 and PC3 obtained from the PCA performed on HIS and VISS non normalized spectra of reference and commercial peanut samples. Vertical lines indicate the most relevant wavelengths corresponding to highest loads values. A very high correspondence is found between the principal component generated from VISS and HIS as expected, since the spectral range is similar in both cases. The most relevant wavelengths (those with maximum and minimum loading values) are highlighted by vertical lines:

469 nm, 550 nm, and 650 nm, all of them related to colour of the sample.

case raw material, roasting and blanching treatments. Thus, peanut RPB (IRMM-481B) is the only raw material

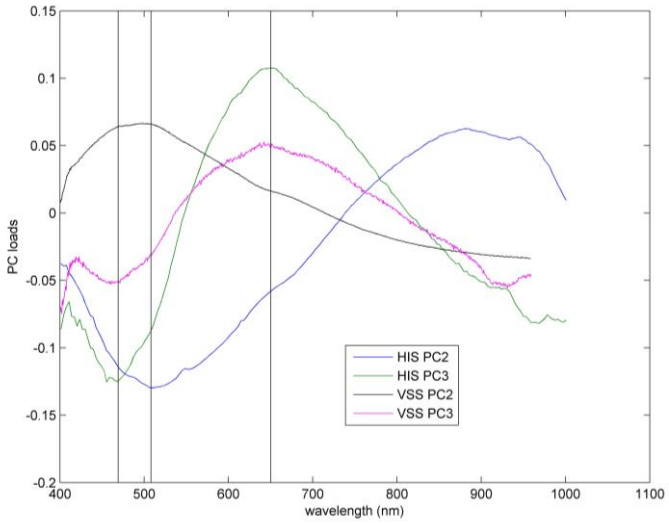


Figure 4 PC2 and PC3 loadings according to HIS and VISS; PCAs computed on peanuts.

Figure5 shows PCA based on hyperspectral data (400 -1000 nm) for segregation of reference peanuts and commercial peanut based on processing treatments (blanching and roasting). Labels RPA to RPF correspond to IRMM 481 kit for reference peanuts and MP to commercial. PC2 and PC3 scores of non-normalized (Figure 5a) and PC1 and PC2 scores of normalized HIS (Figure 5b) showed similar pattern and allow segregating between RPA-F and along with MP, which suggests the convenience of the performed normalization. Both planes of scores are situated in an orthogonal pattern which refers to the existence of unrelated factors, in this and is clearly segregated from the rest, as it happens with peanut RPE (IRMM-481 E) which is blanched without roasting. The commercial peanut spectra labelled as MP were projected onto the planes PC2/PC3 (non-normalized) and PC1/PC2 (normalized) generated with the HIS spectra of reference peanut samples, and it can be observed that it overlays on IRMM-481 RPD which origin is from China, aspect that is also confirmed from the product information in the commercial sample. A major conclusion from this graph is that there are significant differences in the visible spectra among peanuts due to blanching and roasting treatment which makes it difficult to develop a universal segregation

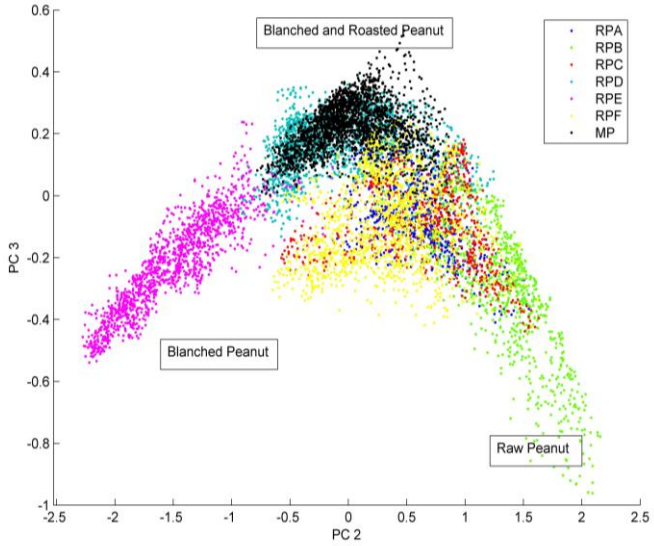


Figure 5a HIS score plots of PC2 vs. PC3 of non-normalized spectra

procedure based on the visible region of the spectrum.

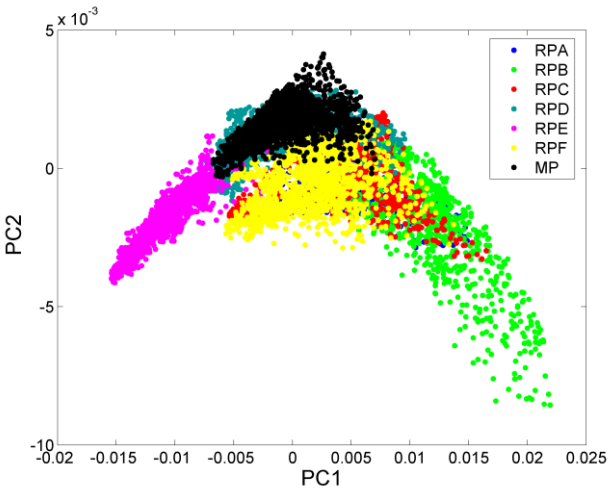


Figure 5b HIS score plots of PC1 vs. PC2 normalized spectra

Figure 6 shows the projection of spectra data (MF, MM and MP) from a previous research work of the authors López-Calleja et al. (2012) into the PC plane generated with non normalized spectra of reference and commercial peanut samples. PC2 and PC3 scores do not allow differentiating among MM (cyan colour n = 366), and MF (black colour n = 316), being also mixed with some reference peanuts (blue points, n = 8788 calibration data set) mainly blanched samples (IRMM-481, RPE). The MP from previous research (red colour n = 2118) overlay on the mixed peanut region as expected. The wide variability in previous experiment could be related to the use of totally un-pressed and disperse powder particle. A major feature extracted from Figure 6 is that PC2 and PC3 scores from VISS do not provide enough information to segregate MF and MM from all types of peanut samples (RPA-F/MP) and thus other spectral ranges are investigated, in this case NIR 896 – 1686 nm. Analogously, it was performed the projection of the corresponding normalized spectra of the same food samples onto the plane PC1 vs. PC2, and similar distribution and results were obtained (data not shown).

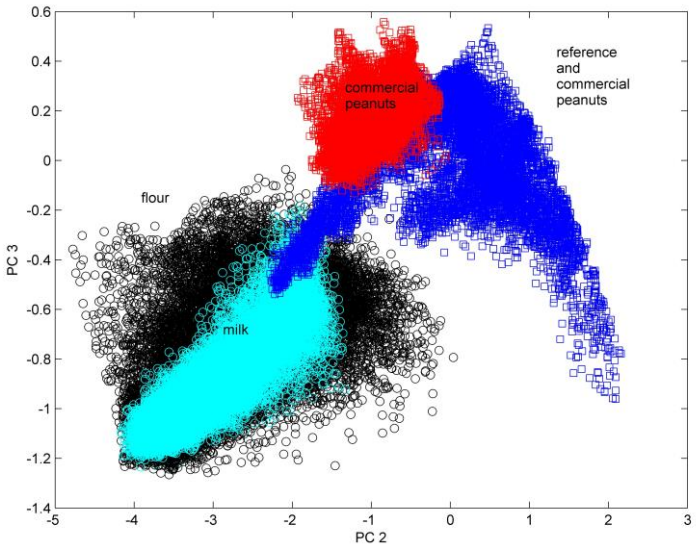


Figure 6 PC2 vs PC3 score plot of the non-normalized spectra of HIS validation dataset

3.2 NIR spectra

As for extended visible spectra, this section is divided into considerations about raw spectra and principal component analysis performed on both reference and commercial samples. Some considerations are also given regarding to the definition of a spectral index.

Figure 7 shows average NIR raw spectra for each food ingredient. The average spectrum of MM appears clearly differentiated from the rest ingredients spectra in the range comprised between 1150 nm and 1700 nm. The raw spectra of all types of peanuts (RPA-RPF) show a clear valley around 1200 nm, which is related to one of the absorption peaks of lipids (Tsai et al., 2001), while the average spectra of MM, MF and MCC do not present such absorption band. The MM presented a high reflectance value at 1200 nm; this observation is congruent with the fact that is skimmed milk.

3.2.1 Spectral index based on NIR spectra

$$SI = R_{1141}+R_{1250}- 2R_{1207} \quad (1)$$

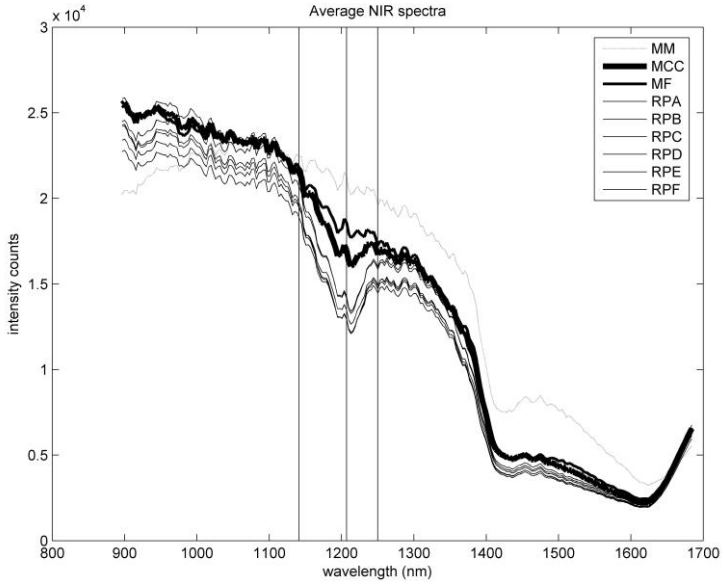


Figure 7 Raw average NIR spectra for different food ingredient

Considering the NIR spectral patterns observed and previously comented, a spectral index (SI) based on several wavelengths around 1200 nm was proposed in order to segregate between a) MF, MM, MCC and b) MP, RPA-F. Equation (1) is a linear combination of 1141 nm, 1207 nm and 1250 nm, which is an approximation to the depth of the absorption peak at 1207 nm (Equation 1).

Figure 8 shows the values of the SI with regard to the spectral sum (SS). SI allows segregating between MM, MF, MCC and peanuts (RPA-F or MP), but it shows to be largely affected by SS, which refers to the total global intensity of the spectra. Therefore, it is also interesting to compute a normalized spectral index (NSI) dividing by SS to correct the global scattering effect.

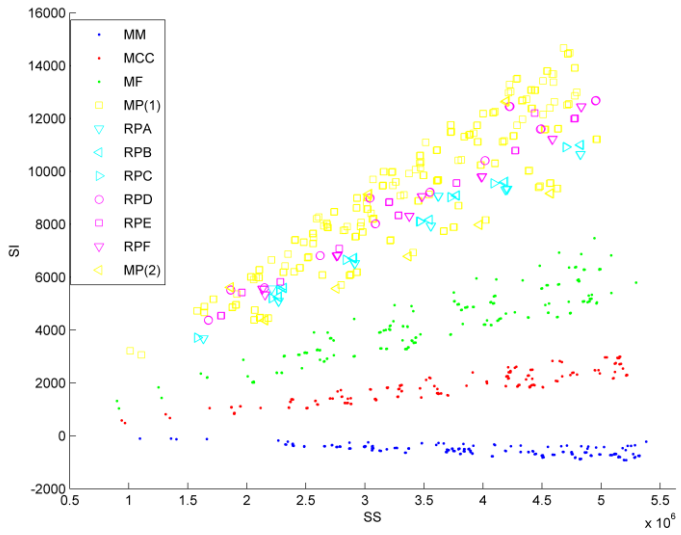


Figure 8 Values of SI (*Y*-axis) vs. SS (*X*-axis)

Considering the multiple comparisons of means for SI and NSI taking into account the ingredients and the integration times, it was observed that, generally both indexes distinguish between MM, MF, MCC and RPA-F or MP. In the case of SI, the differences are more accentuated for high integration time (80 ms), while for 50 ms the values of MF and some of the peanuts present overlap (data not shown). In NSI, also some differences between types of peanuts could be found but in general, they are not significant. Regarding NSI, the effect of the integration time has been removed; the same value of NSI was obtained for each ingredient for all the integration times. Within the peanuts (that appear very separated of the rest of the ingredients) two groups can be distinguished RPA, RPB and RPC, with low NSI values, and RPD and RPE with high NSI values, while RPF shows an intermediate position, which is expected since RPF is a mixture of all the previous samples; so far, we do not find the features that share RPA, RPB and RPC compared to RPD and RPE (considering treatments and / or origin). Similar finding will be further discussed with the spectral indexes generated from PCA in Figure 13.

3.2.2 Principal components

PCA were also been performed on non-normalized and normalized NIR spectra. As in previous cases for non-normalized spectra, PC1 is mainly related to the intensity level of the raw spectra (SS) while PC2 and PC3 scores provide the features for segregating among food materials. PC1 represent 99.37% of total spectral variance. The determination coefficient (r^2) is 98.9%, which means that SS explains almost 99.4% of total spectral variance (Figure not show). PC2 and PC3 represented 0.58% and 0.03% of the spectral variance respectively.

When looking at the NIR spectral loadings for PC2 and PC3 (Figure 9) a very large contribution of 1207-1210 nm is found. Wavelength 1145 nm and 1259 nm provide intersection points between the loading curves of PC2 and PC3 which are very closed to zero loading value. This fact points to the possibility of using such spectral wavelengths, for base line correction and therefore it is decided to compute another spectral indexes (SI2 and NSI2) based on wavelengths 1145 nm, 1207 nm, 1210 nm and 1259 nm (Equation 2). The wavelengths selected based on PC loadings are highly congruent with those addressed by an expert eye on the raw NIR spectra (mentioned above).

$$SI = R_{1145}+R_{1259}-R_{1207}-R_{1210} \quad (2)$$

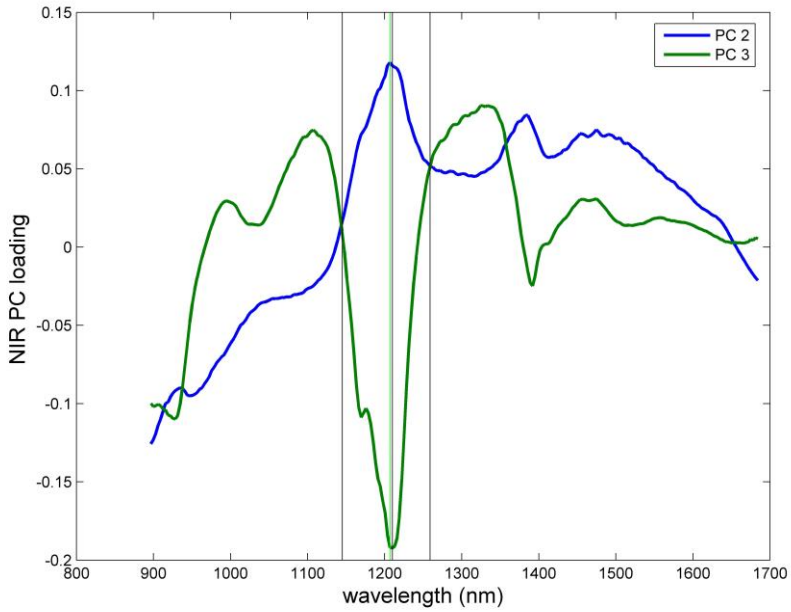


Figure 9 PC2 and PC3 loadings for NIR spectra.

Figure 10a, corresponding to PC2 and PC3 scores, clearly segregates RPA-F and MP (unfilled markers) from the rest of the food ingredients MM, MF and MCC (bold), with little distance among peanut types (RPA-F to MP). This fact is very relevant showing that NIR is far less sensitive to differences among peanuts in comparison to visible spectra, while enlarging distances with other food ingredients. Also in Figure 10a shows all food stuff scores are allocated in radios of the circle, where length of the radios increase correspond to higher integration time. Scores for the different substances corresponding to high integration time are more separated one another, and therefore exhibiting higher segregation power. This fact shows that PC2 and PC3 are still affected by the global intensity of the spectra, justifying the normalization of spectra before the computation of PCA. Arrows in Figure 10a indicate the projection of the spectral indexes: SI, SI2, NSI, NSI2 onto the PC2-PC3 plane, and show to be clearly aligned with peanuts, as opposite to the rest of ingredients.

As it was expected, the normalization procedure corrected the scattering effect and consequently PC1 and PC2 of the normalized spectra were able to segregate peanuts from the other foods. The corresponding scores (Figure 10b) showed some similarities with PC2 and PC3 scores of non normalized spectra. PC1 (of normalized spectra) discriminated between MM and the rest of foods, while PC2 segregated MCC and MF from the other ingredients. It could be observed some effects due to integration times: scores of MM, MCC and MF are distributed along the diagonal of PC1-PC2 plane.

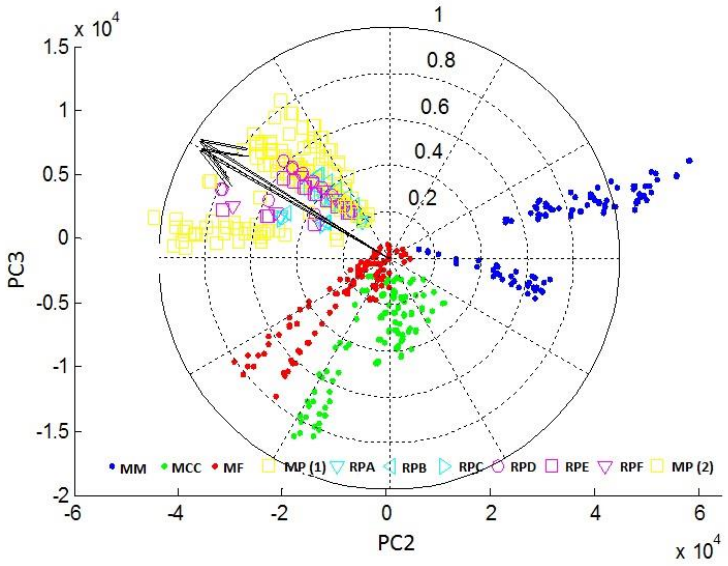


Figure 10a Representation of PC2 (*X*-axis) vs. PC3 (*Y*-axis) scores of the food ingredients spectra

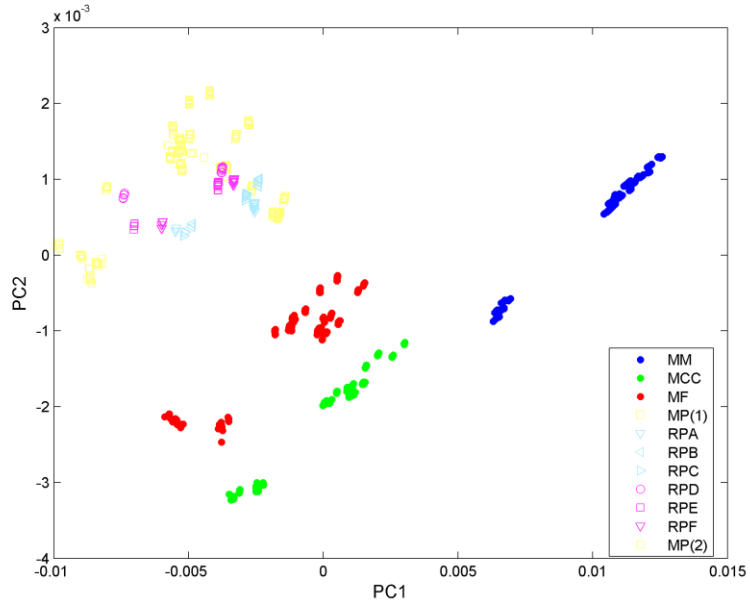
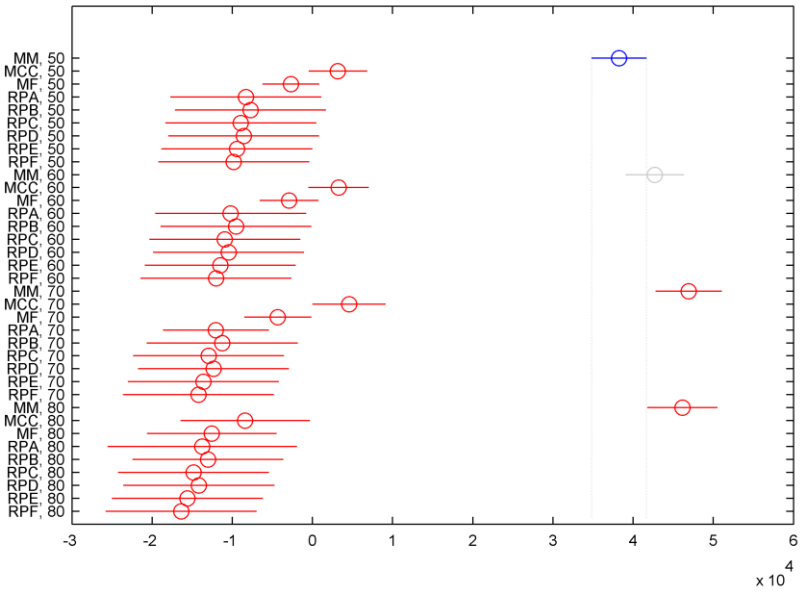


Figure 10b Representation of PC1 (*X*-axis) vs. PC2 (*Y*-axis) scores of the food ingredients spectra

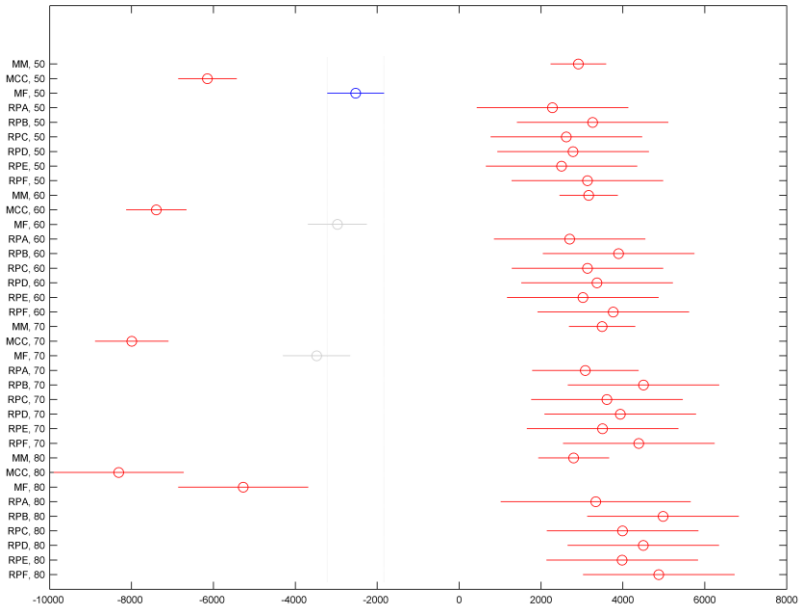
Results of the multiple comparisons of means applied to the scores of PC2 and PC3 (corresponding to non-normalized spectra) are included in Figure 11. PC2 scores are able of segregating MM from the rest of the food ingredients; while PC3 scores (Figure 11b) distinguish three groups: 1) MCC, 2) MF and 3) RPA-F and MM. Scores values of both PC are affected by the integration time, as explained before. None of both PC's alone is able to segregate peanuts from the rest foodstuffs. Figure 12 shows the results of multiple comparisons applied to scores of PC1 and PC2 of the normalized spectra. Regarding segregation of foods it could be observed similar behaviours between PC1 of normalized spectra and PC2 of non-normalized spectra, and PC2 (normalized spectra) and PC3 (non-normalized spectra).

Figure 13 includes the results of multiple comparisons for SI2 and NSI2, showing a similar behaviour than SI

and NSI (discussed in previous paragraph). It can be found that the segregation performance of SI2 is affected by the integration time (Figure 13a), while such effect disappears in NSI2 (Figure 13b).

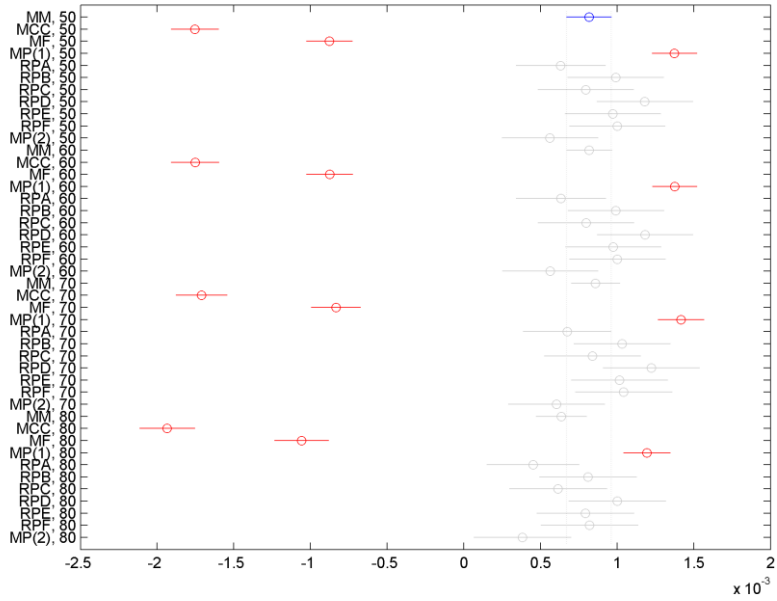


(a)

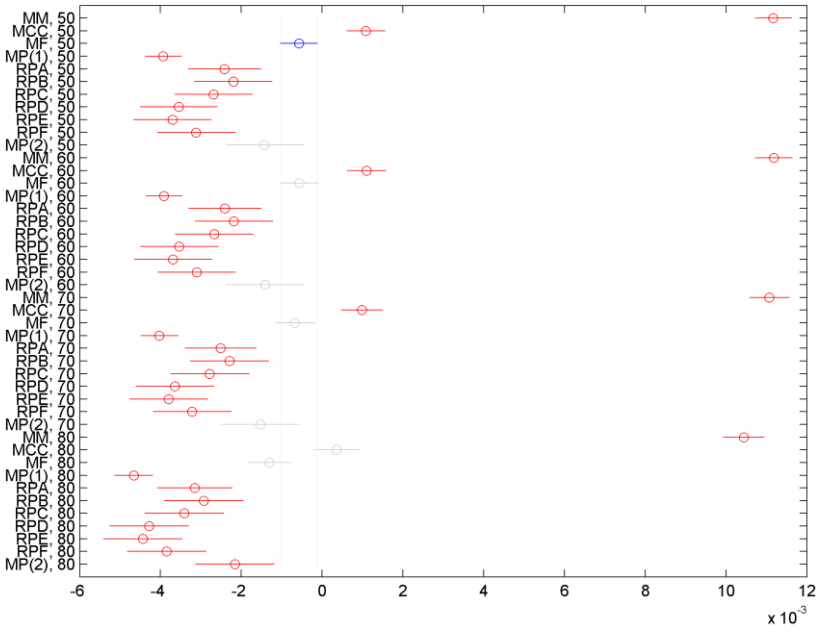


(b)

Figure 11 Multiple comparisons of scores of PC2 (a) and PC3 (b) from NIR PCA categorized by food ingredient (MM, MF, MCC and IRMM 481 RPA–RPF) and integration time (50 ms, 60 ms, 70 ms and 80 ms). The points represent the mean value and the horizontal lines the range considering the standard error of the mean.

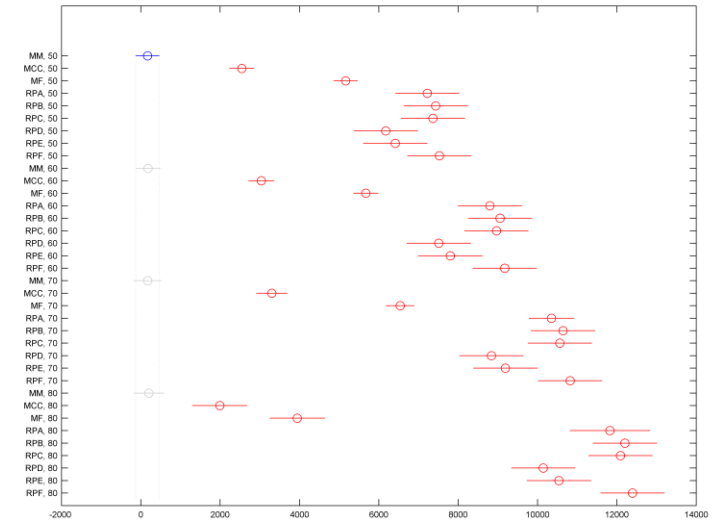


(a)

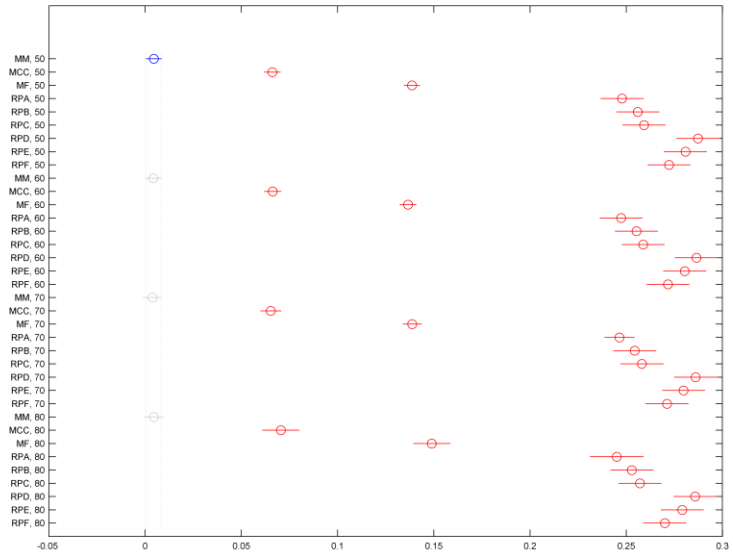


(b)

Figure 12 Multiple comparisons of scores of PC2 (a) and PC3 (b) from normalized NIR PCA categorized by food ingredient (MM, MF, MCC, IRMM481 RPA–RPF peanuts) and integration time (50 ms, 60 ms, 70 ms and 80 ms). The points represent the mean value and the horizontal lines the range considering the standard error of the mean.



(a)



(b)

Figure 13 Multiple comparisons of SI2 (a) and NSI2 (b) categorized by food ingredient (MM, MF, MCC, IRMM481 RPA–RPF peanuts) and integration time (50 ms, 60 ms, 70 ms and 80 ms). The points represent the mean value and the horizontal lines the range considering the standard error of the mean.

3.2.3 Discrimination performance of indexes

Analysis of variance have been performed on SI, NSI, SI2 and NSI2 regarding the ingredient type with only two groups (peanuts and the rest of the ingredients) and computed for two extreme integration times (50 ms and 80 ms; Table 4). SI and SI2, non normalized indexes, show more discrimination ability for high integration time (F= 860 and 969 respectively) than for low (F= 223 and 207). SI2 is more sensitive to integration time than SI. However, for normalized indexes (NSI and NSI2) the F values are similar for both integration times (Table 4). For high integration times the separation between the two groups of foodstuffs is higher with the non normalized indexes compared to the normalized one,

which could be due to a decrease in the signal / noise ratio for the normalized indexes. However, the normalized indexes would allow removing the uncontrolled variability of light that can appear along experimental works.

Table 4 Results of analysis of variance performed on all indexes considering two groups: peanuts and the rest of the ingredients (*n*=176 samples in total)

Index	Integration time, ms	*F*	Sum Sq. Factor	d.f.	Mean Sq.	Sum Sq. Error	d.f.	Mean Sq.
SI	80	860	1.7e+9	1	1.7e+9	1.2e+8	62	2.0e+6
SI	50	223	8.5e+8	1	8.5e+8	6.7e+8	175	3.8e+6
SI2	80	969	1.8e+9	1	1.8e+9	1.2e+8	62	1.9e+6
SI2	50	207	8.3e+8	1	8.3e+8	7.0e+8	175	4.0e+6
NSI	80	535	1.007	1	1.007	0.117	62	0.002
NSI	50	626	1.707	1	1.707	0.002	175	0.002
NSI2	80	558	1.073	1	1.073	0.119	62	0.002
NSI2	50	648	1.818	1	1.818	0.491	175	0.002

4 Conclusions:

VIS-NIR spectra were studied for the characterization of a wide variety of reference peanuts (Kit IRMM 481a) in comparison to powder food materials: MF, MM and MCC, in order to define a specific spectral index robust against pre-treatment (raw or blanched) and roasting (various temperatures and treatment duration).

Visible range allows classifying reference peanut samples and shows orthogonal influences of pre-treatment: roasting and blanching.

The projection of the spectra of powder food materials such as MF and MM with a granulometry from 100 μm to 160 μm, allows confirming that blanched peanuts cannot be distinguished from other food ingredients in the visible range, and thus other spectral ranges (NIR) were inspected. A specific band for peanut identification with regard to MF, MM and MCC powder has been found centred at 1200 nm that corresponds to a band of lipids absorption. Therefore, spectral indexes based on the combination of three wavelengths around 1200 (1141 nm, 1200 nm and 1250 nm) are proposed and compared. Once the indexes are proposed a much cheaper system, multispectral, could be employed in order to compute the false colour images of indexes and to attain a screening system that would operate in conjunction with a RT-PCR procedure. In order to quantitatively assess the nature of powder mixtures at a ppm level, based on powder size and ingredient nature, a proper combination of spatial resolution (70 μm) and field of view size (above 70000 μm^2 to inspect above 1 M particles), together with spectral range (only NIR seems to be enough sensitive and specific) has to be validated, making profit of chemometric and image texture analysis tools.

Acknowledgements

Research has been financially supported by IDEAS project of Erasmus Mundus and PICATA of the Moncloa Campus of International Excellence (UCM-UPM, 2012).

References

Besler, M. 2001. Determination of allergens in food. *Trends in Analytical Chemistry*, 20: 662-672.

Coates, D. B., and R. M. Dixon. 2008. Development of near infrared analysis of faeces to estimate non-grass proportions in diets selected by cattle grazing tropical pastures. *Journal of Near Infrared Spectroscopy*, 16: 471–480.

Cozzolino, D. M., J. Kwiatkowski, R. G. Dambergs, W. U. Cynkar, L. J. Janik, G. Skouroumounis, and G. Gishen. 2008. Analysis of elements in wine using near infrared

spectroscopy and partial least squares regression. *Talanta*, 74: 711–716.

Fox, G., and A. Cruickshank. 2005. Near infrared reflectance as a rapid and inexpensive surrogate measure for fatty acid composition and oil content of peanuts (Arachis hypogaea L.). *Journal of Near Infrared Spectroscopy*, 13: 287–291.

Gowen, A. A, C. P. O'Donnell, P. J. Cullen, and S. E. Bell. 2008. Recent applications of Chemical Imaging to Pharmaceutical process monitoring and quality control. *European Journal of Pharmaceutical and biopharmaceutics: official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik*, 69: 10-22.

Hird, H., J. Lloyd, R. Goodier, J. Brown, and P. Reece. 2003. Detection of peanut using real-time polymerase chain reaction. *European Food Research and Technology*, 217: 265-268.

Hourihane, J. O. B., S. A. Kilburn, J. A. Nordlee, S. L. Hefle, and S. L. Taylor. 1997. An evaluation of the sensitivity of subjects with peanut allergy to very low doses of peanut protein: a randomized, double-blind, placebo-controlled food challenge study. *Journal of Allergy and Clinical Immunology*, 100: 596-600.

Immer, U. 2006. *Detecting allergens in food.* Cambridge: Wood head Publishing Limited pp. 330-347.

Lembe, S. M., L. O. Umezuruike, A. T. Leon, L. Sandra, J. R. C. Paul, H. N. H él`ene, H. Arthur, S. Wouter, and M. N. Bart. 2013. Evaluation of Fourier transform-NIR spectroscopy for integrated external and internal quality assessment of Valencia oranges. *Journal of Food Composition and Analysis*, http://dx.doi.org/10.1016/j.jfca.2013.05.007.

López Calleja, L., L. Lunadei, S. de la Cruz, P. Barreiro, and T. Garc á. 2012. Hyperspectral images and polymerase chain reaction (PCR) for the detection of allergen (peanuts traces) in powder foods. *EFFoST Annual Meeting. Montpellier*.

López-Calleja, I. M., S. de la Cruz, N. Pegels, I. Gonz ález, T. Garc á, and R. Mart ín. 2013. Development of a real time PCR assay for detection of allergenic trace amounts of peanut (Arachis hypogaea) in processed foods. *Food Control*, 30: 480-490.

Mauer, L. J., A. A. Chernyshova, A. Hiatt, A. Deering, and R. Davis. 2009. Melamine detection in infant formula powder using near- and mid-infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, 57: 3974–3980.

Mcdaniel, K., M. Jean-Christophe, S. Timothy, and D. Jack. 2011. Effect of Different Time/Temperature Roast Combinations on Peanut Flavor-Descriptive Sensory. Electronic Nose and Electronic Eye Characterization. *Institute of Food Technology* (conference)

Panforda, J. A., and J. M. De Manb. 1990. Determination of oil concentration of seeds by NIR: influence of fatty acid composition on wavelength selection. *Journal of the American Oil Chemists Society*, 67: 473–482.

Phan-Thien, K.Y., G. Golic, G. C. Wright, and N. A. Lee. 2011. Feasibility of estimating peanut essential minerals by near infrared reflectance spectroscopy. *Sensory & Instrumentation for Food Quality*, 5: 43–49.

Platteau, C., M. De Loose, B. De Meulenaer, and I. Taverniers. 2011. Detection of allergenic ingredients using real-time PCR: a case study on hazelnut (Corylus avellana) and soy (Glicine max). *Journal of Agricultural and Food Chemistry*, 59: 10803-10814.

Rosas, J. G., and M. Blanco. 2012. A criterion fpr assessing homogenicity distribution in hyper spectral images. Part1: Homogenity index bases and blending process. *Journal of Pharmaceutical and Biomedical Analysis*, 70: 680- 690.

Rubio-Diaz, D. E., D. M. Francis, and L. E. Rodriguez-Saona. 2011. External calibration models for the measurement of tomato carotenoids by infrared spectroscopy. *Journal of Food Composition and Analysis*, 24: 121–126.

Scaravelli, E., M. Brohee, and R. Marchelli, and A. Van Hengel. 2008. Development of three real-time PCR assays to detect peanut allergen residue in processed food products. *European Food Research and Technology*, 227: 857-869.

Scheiblhofer, O., D. M. Koller, P. Kerschhaggl, and J. G. Khinast. 2012. Continuous powder flow monitoring via near infrared hyperspectral Imaging, Instrumentation and Measurement Technology Conference (I2MTC). *IEEE International*: 13-16 May 2012, Conference Publications pp : 748 – 753

Shiroma, C., and L. Rodriguez-Saona. 2009. Application of NIR and MIR spectroscopy in quality control of potato chips. *Journal of Food Composition and Analysis*, 22: 596–605.

Sundaram, J., C. V. Kandala, K. N. Govindarajan, and J. Subbiah. 2012. Sensing of moisture content of in-shell peanuts by nir reflectance spectroscopy. *Journal of Sensor Technology*, 2: 1-7.

Sundaram, J., C. V. Kandala, R. A. Holser, C. L. Butts, and W. R. Windham. 2010. Determination of in-shell peanut oil and fatty acid composition using near-infrared reflectance spectroscopy. *Journal American Oil Chemists Society*, 87: 1103–1114.

Sundaram, J., C. V. Kandala, R. A. Holser, W. R. Windham, S. E. Kays, C. L. Butts, and M. C. Lamb. 2009a. Non destructive estimation of oil and moisture content using NIR spectroscopy in Valencia and Virginia peanuts. In: AOCS Annual Meeting, Orlando, FL, USA, 3–6.

Sundaram, J., C. V. Kandala, R. A. Holser, W. R. Windham, and C. L. Butts. 2009b. Estimating oil and fatty acids contents of in-shell peanuts using NIR spectroscopy. In: ASABE Georgia and Florida Section, Daytona Beach, FL, USA, 11–13.

Taylor, S. L., J. A. Nordlee, L. M. Niemann, and D. M. Lambrecht. 2009. Allergen immune assays considerations for use of naturally incurred standards. *Analytical and Bioanalytical Chemistry*, 395: 83-92.

Tillman, B. L., D. W. Gorbet, and G. Person. 2006. Predicting oleic and linoleic acid concentration of single peanut seeds using near infrared reflectance spectroscopy. *Crop Science*, 46: 2121–2126.

Tsai, C. L., J. C. Chen, and W. J. Wang. 2001. Near-infrared absorption property of biological soft tissue constituents. *Journal of Medical and Biological Engineering*, 21: 7-14.

Wu, D., P. Nie, J. Cuello, Y. He, Z. Wang, and H. Wue. 2011. Application of visible and near infrared spectroscopy for rapid and non-invasive quantification of common adulterants in Spirulina powder. *Journal of Food Engineering*, 102: 278–286.