Prospective of NIR Calibration transfer need for detecting peanut traces in powder foods: a transcontinental approach

Teresa R. Cuadrado Dominguez [1] , P. Barreiro Elorza*[2], L. Lleó García[2], Facundo Ruiz[3], Pablo Delgado Sánchez [1]

[1] Facultad de Agronomía y Veterinaria, UASLP, San Luis Potosí, México
[2] LPF_TAGRALIA Departamento de Ingeniería Agroforestal, UPM CEI Moncloa, Madrid, España
[3]Facultad de Ciencias, UASLP, San Luis Potosí, México
* Corresponding author. e-mail (pilar.barreiro@upm.es)

Abstract

The effective detection of the presence of trace amounts of allergens in processed foods is of crucial importance in the food industry. Traces of these allergens can cause adverse reactions, and may become very serious in certain sectors of the population. Previous work of our own, propose NIR models capable of detecting the presence of peanut and another dried fruits in powder matrices. Mentioned models include the band 1200 nm corresponding to the fats. The objective of this work is to show the need of developing transfer of calibration procedures between NIR spectroscopy equipment belonging to different laboratories in different communities of practices including countries (Spain-Mexico), for the detection and determination of allergens (peanut traces in food matrices). Two NIR equipment have been used: Hamamatsu 896 to 1686 nm (LPF-Tagralia UPM, Madrid, Spain), and an Ocean Optics model NIR512, in the range 900 to 1700 nm equipment (Faculty of Sciences UASLP; San Luis Potosí, México). Samples of different varieties of peanut and nuts as well as different food substances (milk powder, cocoa, seeds) have been considered for the determinations. A comparative analysis of the spectra (raw and normalized) has been faced on identical samples; then PCA models in both spectra sets have been computed independently. The set of spectra of the other instrument were interpolated to the corresponding wavelength ranges and then projected. Results show the possibilities of segregation of the spectra of nuts with respect to the other spectra. However, direct application of the model obtained from one instrument to the spectra obtained with another instrument, it is not at all optimal and requires prior correction procedures for increasing reproducibility.

Keywords: Calibration of transfer, NIR, espectroscopy, peanuts, allergens.

1. Introduction

Nuts can cause severe allergy reactions to sensitive population, with peanut being a main allergen for humans; it has been estimated that 1% of human population reacts to peanuts. However, peanut is often used in the food industry as an additive, increasing the risk for allergic patients, since small doses lead to severe clinical reactions. (Satyabrata Ghosh, 2015)

NIR spectroscopy has shown the ability to detect traces of peanuts (Puneet Mishra, 2015) in powder samples. NIR spectroscopy has multiples advantages: it is fast, sensitive, non-destructive and relatively inexpensive; it does not require prior sample preparation; and a massive amount of information related to the components present in a powder mixture can be obtained. (Satyabrata Ghosh, 2015). This technique allows the detection and quantification of fat and moisture in foods due to the presence of absorption bands of chemical bonds such asC-H, N-H, and O-H (Cozzolino et al., 2008). NIR spectroscopy is also able to predict the composition of feed, (Dolores C. Pérez-Marín, 2004) and seeds such as cereals and legumes (Gokhan Hacisalihoglu, 2015).). NIR has been opened new trends in applied research and innovation in the food industry, because it has an excellent ability to obtain accurate information (Piqueras, 2012).

However, when spectra are obtained from different instruments, even from the same manufacturer, signal shift are found (intensity and/or shape) arising from the configuration of the acquisition system (Robert N. Feudale, 2002). Signal drifts may also be due to changes on the instruments (light font, detector… etc.) over the time and due to external factors such as temperature and moisture. All these factors produce errors when the models are transferred directly between instruments (Greensill, 2001). On the other hand, in recent years new portable devices have been developed. They present multiple advantages such as small size, robustness, low cost ergonomic design, simple user interface and portability. These new devices allow new applications and online analysis of products (E. Zamora-Rojas, 2012) and thus, it is of utmost interest to transfer models from laptop to portable NIR equipment in order to make profit of previous research and experience. All the above mentioned facts show the importance of the development of calibration transfer models between different devices, which are essential for broad application of spectroscopy. (Olmo, 2002; Gokhan Hacisalihoglu, 2015).

The aim of this study is to demonstrate the need of developing calibration transfer models between different NIR spectroscopy equipment for the detection and identification of traces of food allergens (traces of nuts and peanuts with respect to various types of foods). In the present study these devices belong to two laboratories, from Spain and Mexico. First, we try to show the similarities and differences between groups of spectra acquired with both equipment and the

models derived from them. Second, we show the result of the direct application of the models obtained from the spectra of one device to the spectra of the other device.

2. Materials and Methods

The materials considered for spectral measurement from Spain were five varieties of peanuts of different origins and subjected to different treatments; provided by the Institute for Reference Materials EC (Table 1). The remainder food samples were milk powder, cocoa, wheat flour and peanuts from different Spanish trademarks. The samples from Mexico were bought in various commercial establishments; they were milk powder, cocoa, wheat flour, different types of flours, three types of pecans, two types of almonds, sesame seeds, two types of peanuts, pinion, corn flour, pistachio, hazelnut Macadamia Coquito Brazil, and cocoa.

Table 1: * The vial IRMU- 481 (RPF) mass mixture of all five peanut vials of the same ratio.

Variety IRMM-484	Origen aria	Treatment	Particle size (μm)	Weight (g)
Runner (RPA)	Argentina	Blanched, hot air, 140ºC / 20 min	500 – 1000	2
Common Natal (RPB)	South Africa	Raw, hot air, 160ºC /13 min	500 – 1000	2
Virginia (RPC)	USA	Blanched, hot oil 145ºC /25 min	500 – 1000	2
Virginia (RPD)	China	Blanched, hot oil 140ºC / 9 min	500 – 1000	2
Jumbo Runners (RPE)	USA	Blanched	500 – 1000	2
Mezcla (RPF)		*	500 – 1000	5

Specifications of reference samples of peanut provided by the Institute for Reference Materials and Measurements (IRMM-481). These samples constitute a kit (IRMM – 481) with six different bottles containing unsalted peanut powder with a particle size between 500 to 1000 .nm.

Non reference samples (cocoa, flour, milk) were submitted to a screening process with a set of sieves of 212/160/100 microns in order to obtain homogeneous granulometry. For this study samples retained in sieves of 100 and 160 micron were selected, so that the size of both samples corresponds to 125 – 100 microns and to 212 - 160 microns. According to the AOAC 965.22 standard, 98% of flour and semolina corresponded to 100-micron sieve. Furthermore, commercial peanuts analyzed in Spain were crushed right through a mechanical mill, and two particle sizes were considered: one above 200 nm and the other below 1000 nm. Nuts used in the experiment of Mexico were crushed with a mortar and then sieved manually with a sieve of 850 microns, the organic toasted sesame seed was not screened, it was measured directly.

The protocol for sample preparation for the measurements in Spain consisted firstly in filling the sample holder (white plastic cylindrical containers of 30 mm of diameter and 10 mm deep) with 1 g of foodstuff and then applying pressure on the entire surface of $1{,}4 \cdot 10^4$ N/m^2 to achieve a smooth surface. For all materials they were acquired NIR spectra (896 – 1686 nm, n = 1110 of which milk spectra were n =210, wheat flour n = 20, cocoa n = , n = 298 commercial peanuts, n=183 for reference peanuts samples from Spain; for the samples from Mexico 900-1700 nm the total acquired spectra were n=55, of which n=2 cocoa, n=13 nuts, n=32 flour, n=4 milk, n=3 semolina, n= seeds).

The protocol for sample preparation from Mexico consisted firstly in filling the sample holder (black steel containers with a hole, rectangular cylindrical shape of 5 mm diameters by 3 mm deep on a rectangular base 30 mm x 10 mm). Afterwards, a uniform pressure was applied by means of a spatula to achieve a smooth surface.

The two spectrophotometers considered for the experiments were Hamamatsu PMAII, C8147 – 4, (896 - 1686 mm, (LPF – Tagralia UPM, Madrid) and Ocean Optics NIR5P4 model (Faculty of Sciences at the UASLP, Mexico).The characteristics of the equipment are listed in table 2.

Spectra of Mexico were interpolated considering the wavelength range of Madrid and then PCA was computed Madrid samples with following projection of Mexico spectra. Alternatively, Spectra of Madrid were interpolated considering the wavelength range of Mexico and then projected into a PCA Mexico model. This procedure was applied to raw and pre-processed spectra by SNV. Summarizing four groups of models were computed with and without SNV, for Madrid and for Mexico spectra.

Table 2. Characteristics of the equipment.

	Equipment from Spain	**Equipment from MÉXICO**
Equipment NIR	Hamamatsu photonic multi-channel analyzer (Japón)	Ocean Optics
Model	PMA – 11C8147-34	NIR512
Sensor/Detector	Indium arsenide gal of (InGaAs)	Hamamatsu G9204 512 (InGaAs)
Wavelength	244 Wavelength 896-1486 nm	512 Wavelength 852.71-1741.68 nm
Measuring system	Interactance	
Lamp	Monolight Optical Spectrum Analyzer MIO-6262-100 Tungsten Halogen model L5E9611 12/100w	Tungsten Halogen LS1 de Ocean Optics
Light Transmission	Fiber optic model MIO-6134-SS/N6794	Bifurcate fiber Optic and random model Lab-Grade Reflection Probes.
White	Neighborhood sulfate Tablet	Win- 3 type aluminium oxid

3. Results and Discussion

Figure 1 left, shows the raw spectra from Spain (red, n=1110, 896 - 1700 nm, intensity level up to 7 x 104). There appears the absorption band of lipids around 1200 nm in the spectra of nuts, in both groups of spectra,. This fact shows the potential of both instruments for the detection of allergens. However, the shape of the spectra differs significantly; Madrid spectra generally show a downward trend with the wavelength and those of Mexico are shown with higher intensity level and horizontal trend. It is also very neat the multiplicative effect in both groups of spectra.

Figure 1 right, shows the two groups of spectra of Spain (red) and Mexico (cyan) pre-processed by Standard Normal Variate procedure (SNV). The multiplicative effect was drastically reduced thanks to this procedure. Furthermore, two absorption bands are observed on the two sets of spectra; one around 1200 nm (lipids), (Ghosh, S. et al., 2015) and the other around 1440 nm (water and carbohydrates) (Alomar, D., 1998).

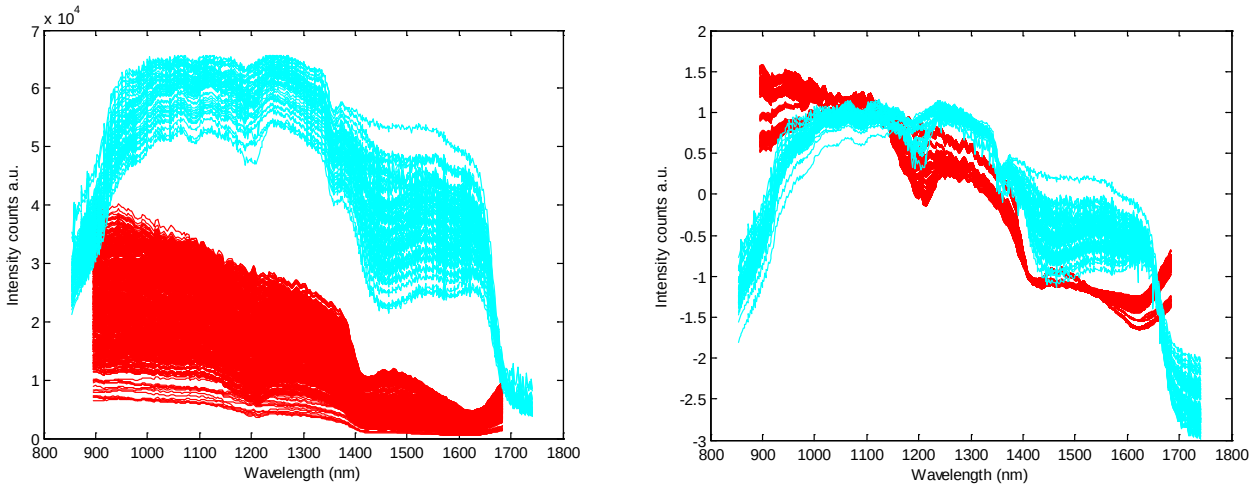

Figure 1. Left: raw spectra of food materials from Spain experiment (red, n=110) and cyan spectra (n=55) experiment from Mexico. Right the same spectra that have been subjected to pre-processed SNV.

PC2 and PC3 computed on the Spanish set with and without SNV (Figure 2, first row) shows a relevant contribution of the loadings in the area of lipids from 1207 to 1210 nm as stated by Ghosh et al. in 2015. Both PC2 (with and without SNV) are very similar in shape and they show also a relevant absorption band at around 1400 nm.

Figure 2, second row, shows the loadings of PC2 and PC3 in the spectra from Mexico, left and right with SNV and without SNV respectively. PC3 captures 0.3% of the total variance and shows a noisy pattern probably due to the fact that the number of spectra for the PCA computation is low (n=55). Despite the fact, PC3 presents relevant areas in their chemical meaning around 1200 nm, 1440 nm, and 1650 nm that may correspond with lipids, water, carbohydrates and lipids respectively (Daniel Alomar, 1998). In addition, the shape of both PC3 (Mexico) are similar

to both PC2 (w and wo SNV) from Spain. This is in accordance to the coefficient of correlation between the Mexico and Madrid loadings (table 3) showing the highest values between Madrid-PC2 and Mejxico-PC3 regardless signal pre-processing (Raw and SNV). It is worth noting that the correlation coefficient between Madrid-PC2 and Mexico-PC3 is always lower for SNV (0.63) compared to raw spectra (0.80). The fact that SNV performs a non-linear transformation could be in the base of such differences.

PCA **WITHOUT** SNV PCA **WITH** SNV

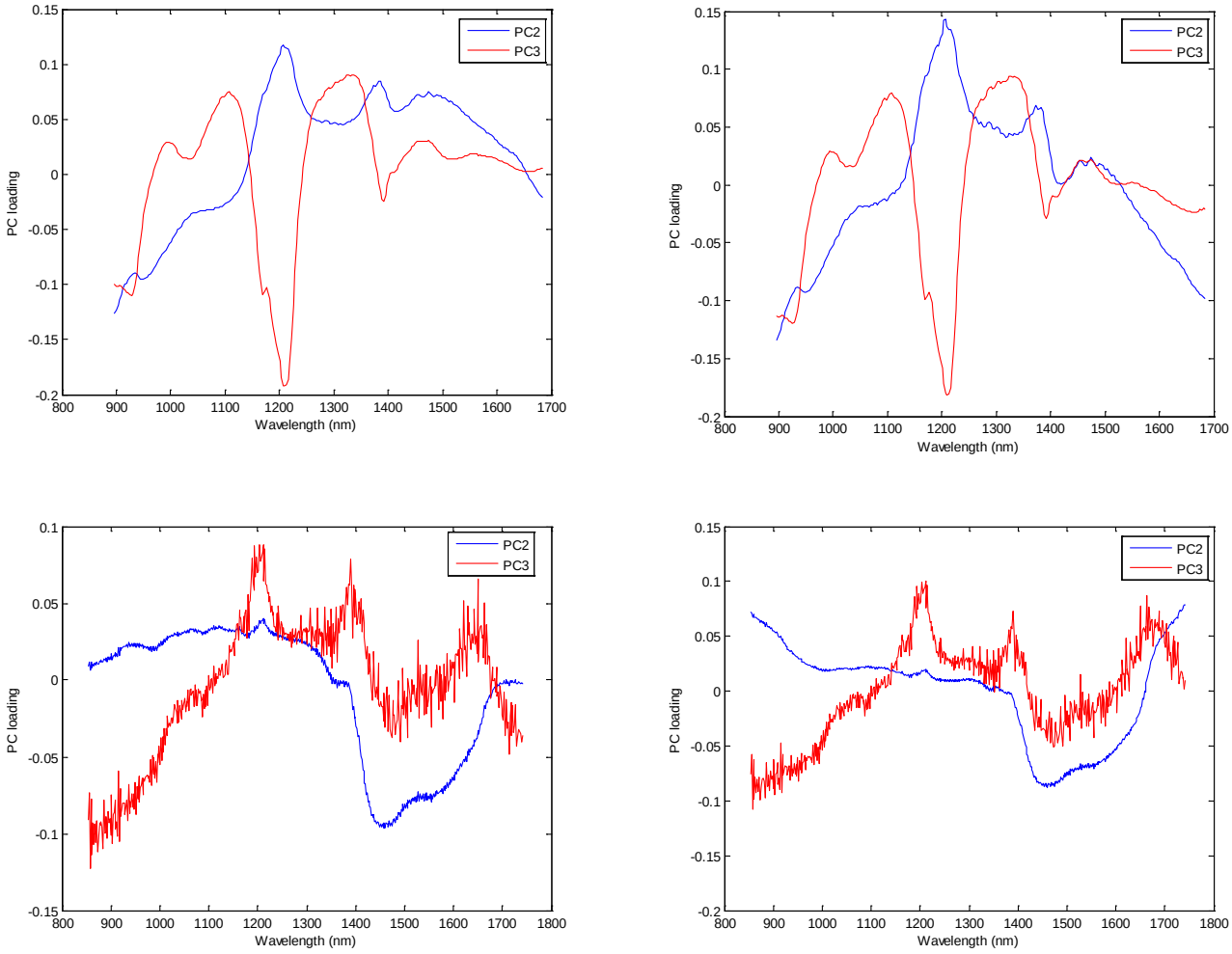

Figure 2. First row: Spectral loadings for PC2 blue and PC3 red, from PCA computed on Spain data set; without SNV left and with SNV right. **Second row** shows the loadings for PC2 in blue and PC3 red of Mexico, without SNV left and SNV right.

Table 3. Correlation coefficient between the PC's computed on the spectra of Madrid and Mexico.

			Mexico			
			PCA on raw spectra		PCA on SNV spectra	
			PC2	PC3	PC2	PC3
Madrid	PCA on RAW spectra	PC2	-0,37	**0,80**	-0,55	**0,65**
		PC3	-0,17	-0,02	-0,23	-0,13
	PCA on SNV spectra	PC2	0,18	**0,74**	-0,07	**0,63**
		PC3	-0,04	0,03	-0,13	-0,08

Figure 3 shows the distribution of scores on PC2 vs PC3 plane corresponding to PCA computed on raw spectra of Madrid (n=1110, first row), and on raw spectra of Madrid (n=55, second row). PC1 scores were not plotted in any case since PC1s did not allow distinguishing the different food groups.
Model computed on Mexico spectra n=55 (first row) shows segregation by countries and a not perfect segregation by food substances; the different groups of scores showed certain overlapping (i.e. scores of nuts and milk). Clearly PCA plane of Madrid (second row) allows segregating peanuts (red markers) from the other ingredients, milk, flour and cocoa (other markers). All scores of Madrid are distributed radially (second row, left); where external positions correspond to higher integration time. Scores for different substances that correspond to high integration time were better separated from each other and therefore high integration time showed greater power of segregation. The group of scores of Mexico appears far from Madrid group. Considering a straight line (in black, close to diagonal direction of the PC2 vs PC3 plane) nut scores of Madrid and Mexico are plotted over the line and appear separated from the rest of food ingredients (bellow the line). This fact shows that this model is able to separate nuts from the rest; however this segregation needs to be improved. The direct application of the model obtained from raw spectra of Madrid and/ or of Mexico, on spectra obtained with the other instrument, is not optimal and requires, prior application, a procedure of correcting them.

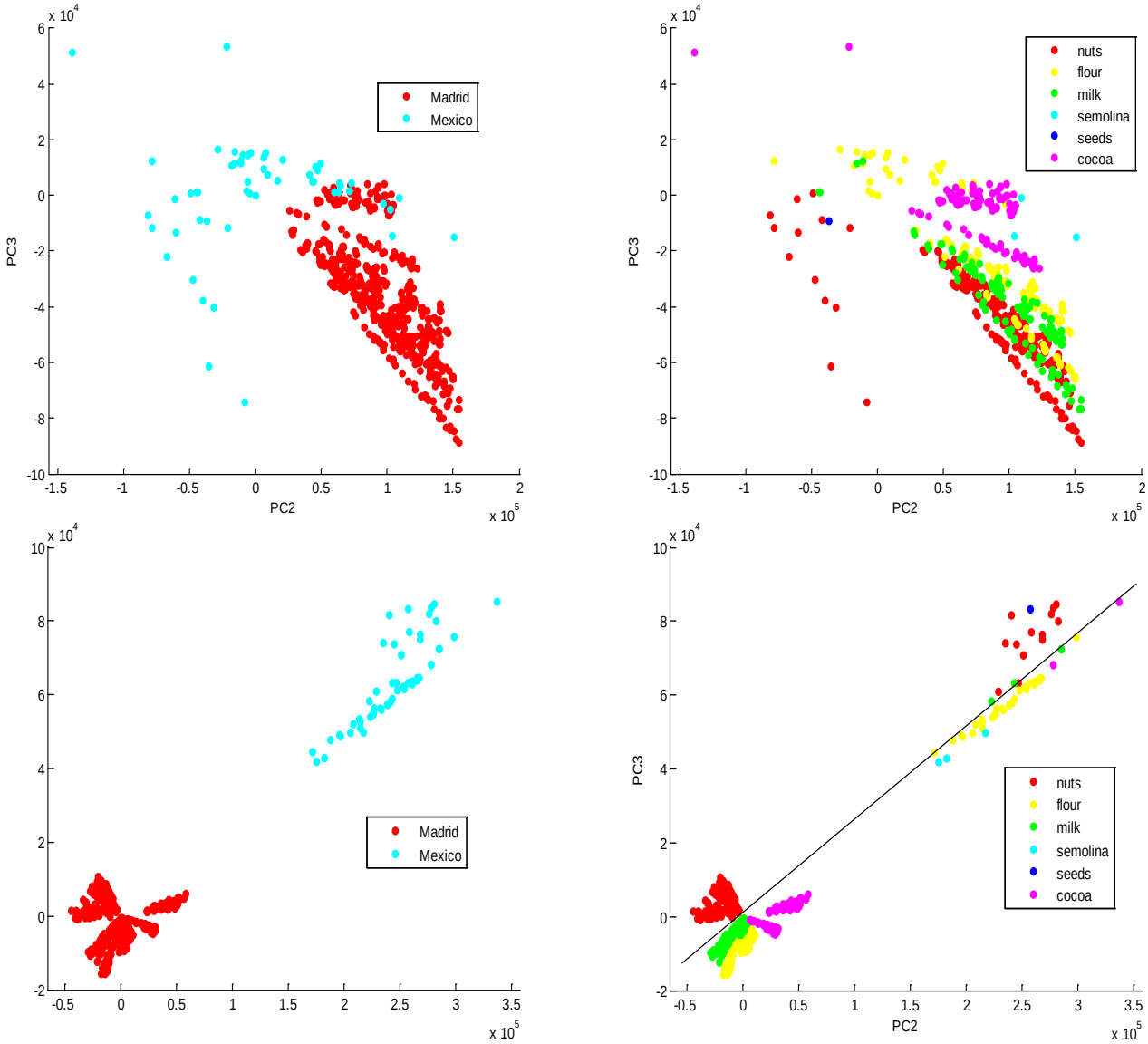

Figure 3. First row, distribution of scores PC2 *vs*. PC3 of PCA computed on the set of spectra of Mexico, for each type of sample (nuts, flour, milk, semolina, seeds and cocoa). On the left the distribution of scores for the samples of each country. On the right for each type of sample food. Second row: distribution of scores PC2 vs. PC3 of PCA computed on the set of spectra of Madrid for the two experimental data sets together (Madrid and Mexico). On the left the distribution of scores for the samples of each country. On

the right for each type of sample food.

Figure 4 show similar plots of scores than those of Figure 3, but considering spectra with SNV. First row show the scores for the PC2 vs PC3 of Madrid (left, by countries and right by food substances). Again both country scores are separated. This plot shows no perfect segregation of foods, with some overlapping of nuts and milk scores. Second row in Figure 4 shows the results of the projection onto Mexico model (PC2 vs PC3), again separation was observed by the type of material on the direction perpendicular to diagonal of the plane. Comparing these plots with those of Figure 3, the initial structure of the data (related to the integration time) is lost as due to the correction of SNV on the spectra.

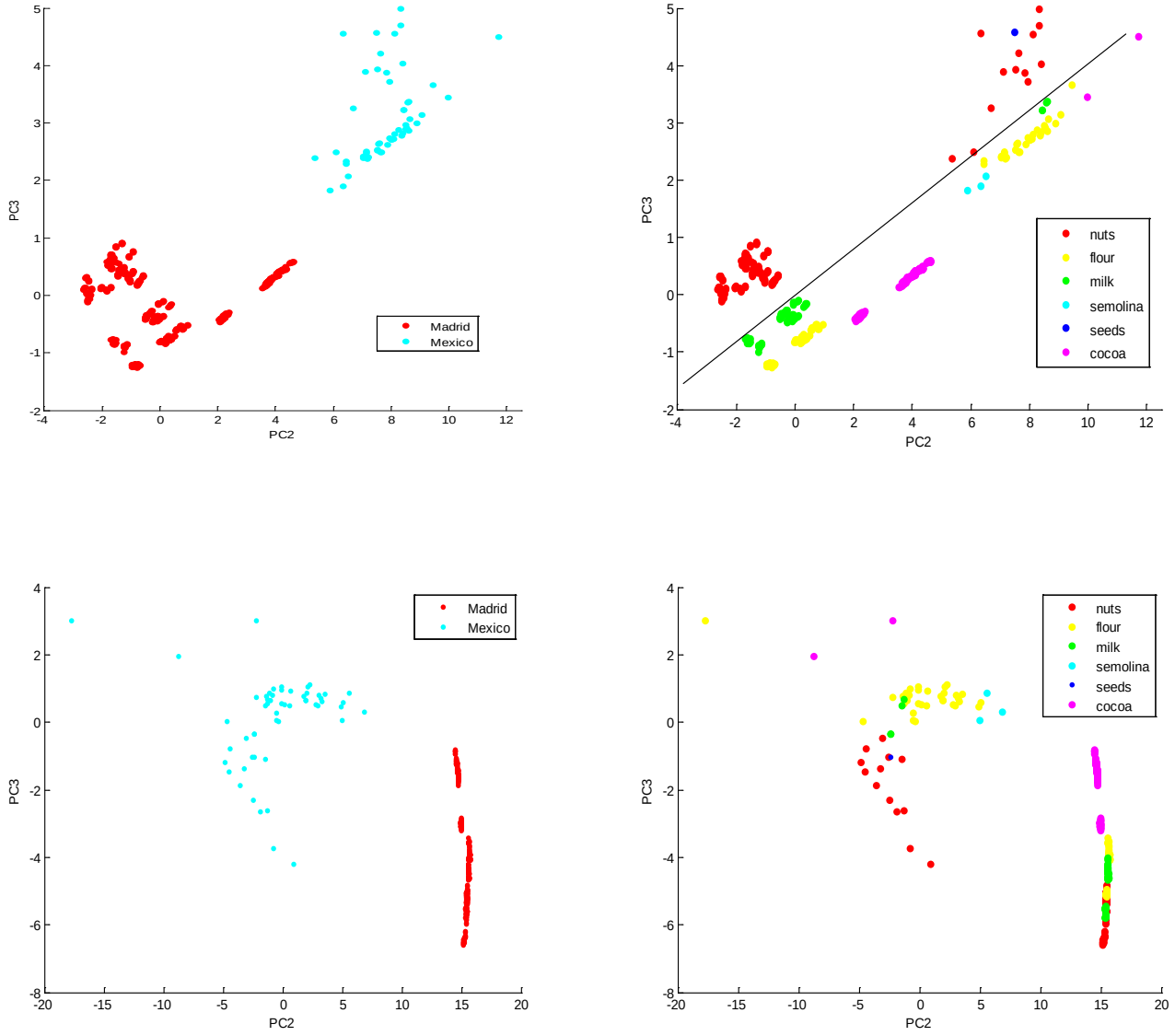

Figure 4. First row, distribution of scores PC2 *vs*. PC3 of PCA computed on the set of SNV spectra of Mexico, for each type of sample (nuts, flour, milk, semolina, seeds and cocoa). On the left the distribution of scores for the samples of each country. On the right for each type of sample food. Second row: distribution of scores PC2 vs. PC3 of PCA computed on the set of SNV spectra of Madrid for the two experimental data sets together (Madrid and Mexico). Spectra of Mexico were interpolated and projected on to the plane. On the left the distribution of scores for the samples of each country. On the right for each type of sample food.

4. Conclusions

This papers provides a first step approach in the reuse and transfer of a PCA model from a Hamamatsu NIR equipment (244 wavelengths, 800-1600nm), towards an Ocean Optics NIR device (512, same wavelength range) for the detection of peanut traces in powder foods. This case is selected as an example in a general purpose strategy. The

identification of calibration transfer needs and the definition of adequate evaluation and processing algorithms for the NIR spectra are of utmost importance to make profit of spectra databases taken with a varied type of instruments. Instrument reproducibility within the same brand, as well as comparison among different commercial devices becomes crucial in this context.

In this paper after the obvious requirement of wavelength interpolation, an analysis is performed to assess whether the inner correlation structure in the data remains regardless the type of instrument. It has been outlined that non-linear spectra preprocessing such as SNV alters the spectra of both equipment in a slightly different way. Therefore it has been decided to procced with calibration transfer based on raw rather than on pre-processed spectra.

However large the effect of the instrument, peanuts can still be segregated from other powder food material like milk flour or cocoa based on a common model and on an identical feature which strongly relates in both instruments with 1200nm band related to the presence of lipids, which sound rather reasonable and has been validated several times by our group and some others.

Stepwise direct standardization will be applied and in the near future based on this set of equipment, also in the aim of maintaining the largest spectra resolution which in this case correspond to the Ocean Optic instrument (1.4nm) compared to 3nm in the Hamamatsu.

Acknowledgements

Acknowledgements, Funded by Comunidad de Madrid, S2013/ABI-2747 (TAVS-CM) and Fondos Estructurales de la Unión Europea

References

A, G., Barreiro, P., Soto, D., León, L., & N., H.-S. (s.f.). Desarrollo de un espectrofotometro de campo para su aplicación en un programa de mejora genética.

Dolores C. Pérez-Marín, A. G.-V.-G.-C. (2004). Near-infrared reflectance spectroscopy (NIRS) for the mandatory labelling of compound feedingstuffs: chemical composition and open-declaration. Animal Feed Science and Technology ELSERVIER, 333-349.

E. Zamora-Rojas, D. P.-M.-S.-G.-V. (2012). Handheld NIRS analysis for routine meat quality control: Database transfer from. Chemometrics and Intelligent Laboratory Systems ELSERVIER, 30-35.

Gokhan Hacisalihoglu, J. L. (2015). Enhanced Single Seed Trait Predictions in Soybean (Glycine max) and Robust Calibration Model Transfer with Near-Infrared Reflectance Spectroscopy. Agricultural and food chemistry, 1079-1086.

Greensill, C. W. (2001). Calibration transfer between PDA-based NIR spectrometers in the NIR assessment of melon soluble solids content. Appl. Spectrosc., 55, 647–653.

Nicolaî, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys W, Theron, K., y otros. (2007). Nondestuctivemeasument of fruit and vegetable quality by means of NIR SPETROSCOPY. Postharvest and Biol. Technol , 46:99-118.

Olmo, J. G. (2002). Clasificación y autentificación de canales de Cerdo Ibérico mediante Espectroscopía en el Infrarrojo Cercano (NIRS). Cordoba: Servicio de Publicaciones de la Universidad de Córdoba.

Piqueras, S. B. (2012). Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution. chemometrics and intelligent laboratory systems, 169-182.

Robert N. Feudale, N. A. (2002). Transfer of multivariate calibration models: a review. Chemometrics and Intelligent Laboratory Systems ELSERVIER, 181-192.

Satyabrata Ghosh, T. R.-M. (2015). VIS/NIR spectral signature for the identification of peanut contamination of powder foods. AgricEngInt, 310-329.