

## Control of quality and silo storage of sunflower seeds using near infrared technology

By I. González-Martín<sup>a</sup>, V. Villaescusa-García<sup>a, b</sup>, F. López-González<sup>b</sup>, C. Oiz-Jiménez<sup>b</sup>,  
I.A. Lobos-Ortega<sup>a</sup>, B. Gordillo<sup>c</sup>, and J.M. Hernández-Hierro<sup>c, \*</sup>

<sup>a</sup> Departamento de Química Analítica, Nutrición y Bromatología, Facultad de Ciencias Químicas.  
Universidad de Salamanca. Spain

<sup>b</sup> Compañía Española Comercializadora de Oleaginosas, S.A. Nava del Rey-Valladolid, Spain

<sup>c</sup> Lab. Color y Calidad de Alimentos, Universidad de Sevilla, Área de Nutrición y Bromatología,  
Facultad de Farmacia, 41012 Sevilla, Spain

\* Corresponding author: jmhhierro@us.es

### RESUMEN

#### Control de calidad y de almacenamiento de semillas de girasol mediante tecnología de infrarrojo cercano

En este trabajo se evalúa la espectroscopía de infrarrojo cercano para su uso en el control de calidad y almacenamiento de semillas de girasol. Los resultados indican que el método analítico empleado puede utilizarse como método de determinación rápida de humedad, grasa y contenidos altos/bajos de ácido oleico. Los rangos de aplicación son comparables con los valores que se han determinado mediante métodos clásicos de análisis, encontrándose entre 4.6-21.4% la humedad, 38.4-49.6% la grasa y 60.0-93.1% de ácido oleico del total de los ácidos grasos. Además se ha utilizado un análisis discriminante lineal por pasos determinando las longitudes de onda más adecuadas para la clasificación de semillas de girasol en los grupos alto/bajo oleico. El modelo generado permitió la clasificación de semillas de girasol en los grupos alto y bajo oleico con unos porcentajes de muestras correctamente clasificadas de un 90.5% en validación interna y de un 89.4% en validación cruzada.

**PALABRAS CLAVE:** Ácido oleico – Control de calidad – Espectroscopía de infrarrojo cercano – Semillas de girasol.

### SUMMARY

#### Control of quality and silo storage of sunflower seeds using near infrared technology

This work assesses the application of near infrared spectroscopy technology for the quality control of sunflower seeds direct from farmers and from a storage silo. The results show that the analytical method employing near infrared spectroscopy can be used as a rapid and non-destructive tool for the determination of moisture, fat and high/low oleic acid contents in samples of sunflower seeds. The ranges obtained were comparable to those reported for classic chemical methods, and were between 4.6-21.4% for moisture; 38.4-49.6% for fat, and 60.0-93.1% for oleic acid expressed as percentage of total fatty acids. A stepwise discriminant analysis was performed to determine the most useful wavelengths for classifying sunflower seeds in terms of their (high/low) oleic acid composition. The discriminant model allows the classification of sunflower seeds with high or low oleic

acid contents, with a prediction rate of 90.5% for internal validation and of 89.4% for cross-validation.

**KEY-WORDS:** Near infrared spectroscopy – Oleic acid – Quality control – Sunflower seeds.

### 1. INTRODUCTION

Sunflower (*Helianthus annuus* L.) is one of the most widely cultivated crops in the world. Sunflower oil usually has a low proportion of saturated fatty acids (10%), high contents of unsaturated fatty acids (oleic acid, 15-30%; and linoleic acid, 55-75%) and also natural antioxidants such as  $\alpha$ -tocopherol. When the percentage of oleic acid is greater than 75% oleic acid expressed as percentage of total fatty acids, the term high-oleic acid sunflower oil is used. In this case, the linoleic acid contents are close to 9%.

The oleic acid content of sunflower seeds can be increased significantly by genetic improvement, affording plants a higher proportion of the enzyme oleil-CoA-desaturase (Cantamutto and Poverene, 2007). These seeds are appreciated by the food industry oil-producing companies because their low level of unsaturated fatty acids elicits less auto-oxidation than oil from animal fat or palm oil (Frankel, 1998) and hence greater stability than the oil from traditional hybrid seeds. For these reasons, high-oleic acid cultivars have been recently developed whose oil has greater oxidative stability and better dietary properties than the oil from standard genotypes (Flagela *et al.*, 2002). The main uses of these oils are in cooking and industrially as a biofuel (biodiesel), which represents a viable alternative to gasoline and diesel fuel (Mosser, 2008; Pereyra-Irujo *et al.*, 2009).

Traditionally, the basic parameters that determine the quality of sunflower seeds (% fat, % moisture and % of oleic acid) are based on wet analysis and chromatographic techniques (Hajimahmoodi *et al.*, 2005), which are not only slow and expensive methods, but also involve the use of large amounts of reagents, with the consequent generation of important quantities of residues.

NIRS (near infrared spectroscopy) technology has been proven to be an efficient tool for the rapid determination of fatty acids in cereals and oleaginous crops such as soy (Roberts *et al.*, 2006) maize (Weinstock *et al.*, 2006) and olives (Maggio *et al.*, 2009) and for the determination of designation of origin (López-Feria *et al.*, 2008). It also allows for determining the oxidative index of edible oils (Le Dréau *et al.*, 2009).

Regarding sunflower seeds, NIRS technology has just been satisfactorily applied to assess the content of fat, moisture and crude protein (Fassio and Cozzolino, 2004), the total content of fatty acids (Moschner and Biskupek-Korell, 2006); the content of oleic acid (Cantarelli *et al.*, 2009), and to the discrimination between sunflower and wheat stubble and weed species (Jurado-Expósito *et al.*, 2003).

The drying process of sunflower seeds and the way in which they are stored (Christensen, 1982; Kent, 1975; Pomeranz, 1978; Salunkhe *et al.*, 1991) strongly influences their quality. In technological terms, it is the procedure that requires the most attention since it prevents the decomposition and deterioration of the fat material contained in the seeds and the oil extracted from them.

In light of the above, the aim of the present work was to study the use of NIRS technology for a rapid control of the quality of sunflower seeds supplied by farmers and for the control of such seeds in a silo in order to determine the physicochemical characteristics (moisture, fat, oleic acid content) of the raw materials and to discriminate between high- and low-oleic acid sunflower seeds.

## 2. MATERIALS AND METHODS

### 2.1. Samples

The seeds used in the experiment, harvested in September, October and November, were collected from Cecosa S.A (Compañía Española Comercializadora de Oleaginosas, S.A), devoted to the industrial drying of oleaginous seeds and maize, and to storing such products coming directly from the field. Samples from the field (103) correspond to seeds taken at random from different plots and farmers. The samples from silo (160) were conserved for 6 months prior to sale. A total of 263 samples was analyzed.

### 2.2. Reference methods

Moisture was measured by gravimetry (AOAC, 1990), following the UNE-EN-ISO-665-2000 norm. Fat was measured using the EN ISO 10565:1988 method, by means of pulsed nuclear magnetic resonance (NMR), using an Oxford 4000 NMR device from Oxford Analytical Instruments Ltd. (Oxon, England). Chromatographic analyses were carried out following the ISO 5590 method for the determination of fatty acid methyl esters (ISO, 1978). This was performed by Enoquisa,

an external laboratory. Samples were allocated into high/low oleic acid groups according to their percentage of oleic acid expressed as percentage of total fatty acids. Samples with percentages greater than 75% oleic acid were allocated into the high-oleic group (Codex alimentarius, 1999).

### 2.3. NIR spectra

Prior to the collection of the spectra, the seeds were ground in order to homogenize the seed coat and internal fruit. The ground samples were placed in a vial and the spectra were recorded in the 1445-2348 nm range using a 19 filter instrument. The nineteen available wavelengths were 1445, 1680, 1722, 1734, 1759, 1778, 1818, 1940, 1982, 2100, 2139, 2180, 2190, 2208, 2230, 2270, 2310, 2336 and 2348 nm. Each spectrum was the mean of 30 scans obtained using an InfrAlyzer 2000 spectrometer from Bran+Luebbe GmbH (Norderstedt, Germany) in reflectance mode. The device was controlled via the Sesame v.3.10 program (Norderstedt, Germany). Figure 1 shows the spectra corresponding to four seed samples.

### 2.4. Statistical analysis

NIR spectra typically contain unselective, extensively overlapping bands that require the use of multivariate chemometric analytical tools (Massart *et al.*, 1998). To perform the qualitative analyses, the LDA method was employed using the SPSS 13.0 for Windows software package (SPSS, Inc., Chicago, IL). A stepwise discriminant analysis was performed to determine the wavelengths most useful for classifying the samples in terms of high or low oleic acid composition, using Wilk's lambda as a statistical selection criterion for the variables.

To build the quantitative calibration models of the different parameters the samples were divided into two sets; one for calibration and the other one for validation. The samples belonging to each set were selected from a Principal Component Analysis

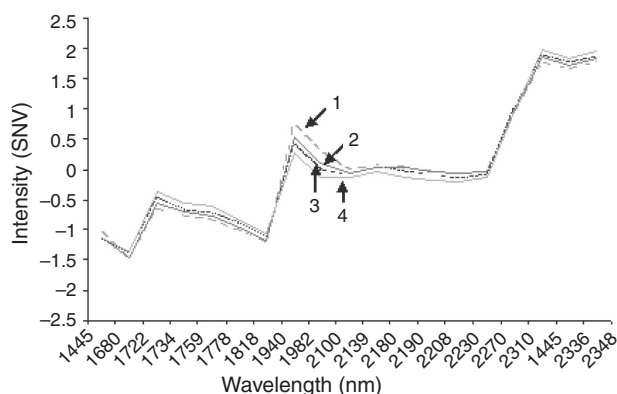


Figure 1

NIR spectra corresponding to sunflower seed samples: (1) Fat, 38.69%, Moisture, 11.75%, High oleic; (2) Fat, 41.33%, Moisture, 11.25%, Low oleic; (3) Fat 42.16%, Moisture, 9.33%, Low oleic; (4) Fat, 48.89%, Moisture, 7.71%, High oleic.

(PCA) of the NIR spectra. In the quantitative analysis, the multiple linear regression (MLR) calibration model was used, which was compiled using the Combination Search Algorithm, where the best wavelengths are determined by trying all possible combinations for the number of search wavelengths entered. For each combination, the Multiple Correlation Coefficient,  $R$ , was calculated by comparing the reference property values with the values calculated from the regression equation:

$$y = F_0 + F_1x_1(\lambda_1) + F_2x_2(\lambda_2) + F_3x_3(\lambda_3) + \dots + F_kx_k(\lambda_k)$$

where  $y$  is the property value;  $F_0$  is the interception of the regression line;  $F_n$  are the regression coefficients, and  $x$  is the measured or transformed value at a specific wavelength. The best wavelength combination to model the calibration set is thus the one where  $R$  is the highest. To construct the models, the spectra were used untreated, using the aforementioned nineteen wavelengths in the 1445-2348 nm spectral range.

The relative standard error (RSE, %) was the parameter used to assess the quality of calibration (RSEC) and prediction (RSEP). To evaluate the quality of the model, the values predicted by NIR were plotted against the reference values, and the confidence intervals of the slope and the ordinate at the origin were calculated for a significance level of 95%. The multivariate calibration models, the spectral pre-treatments and the principal components analysis used in sample selection were run with Sesame v.3.10 (Norderstedt, Germany) and Unscrambler v.9.1 from CAMO software (Trondheim, Norway).

### 3. RESULTS AND DISCUSSION

#### 3.1. Qualitative analysis

Discrimination of the high- and low- content oleic sunflower samples was performed with LDA. A stepwise discriminant analyses was performed

to determine the wavelengths most useful in classifying the samples in terms of high- or low-oleic composition, using Wilk's lambda as a statistical selection criterion for the variables. In this case, five wavelengths (1680, 1722, 1734, 1940 and 2139 nm) were required for the discriminant analysis.

Only one model was developed using all the 263 samples (direct from the farmers and from the silo). The results of the qualitative analysis are shown in Table 1. In the analysis, leave-one-out cross-validation was performed on all the samples. In cross-validation, each case was classified by the discriminant function. A percentage of 90.5% of the cases was correctly classified and 89.4% of cases were correctly classified in leave-one-out cross-validation. The results obtained with both types of samples (high/low oleic) were acceptable.

#### 3.2. Quantitative analysis

To build the calibration models for the different parameters (fat, moisture, oleic acid) the samples were divided into two sets; one for calibration and the other one for validation. The calibration samples should have the greatest spectral variability possible and in turn should cover the complete range of values of the parameter to be measured. These samples were chosen from a scatterplot (PC1 vs. PC2 scores) from a PCA of the spectra using the aforementioned nineteen wavelengths in the 1445-2348 nm spectral range and applying the SNV (Standard Normal Variate) as the only spectral treatment. The variance explained by the first two factors was greater than 99%. Figure 2 shows the score plot for in silo samples as an example.

The models for the determination of the percentage of moisture and fat in the samples provided by the farmers –i.e. samples obtained directly from the field (in September, October and November)– and the models for the determination of the percentage of moisture, fat and oleic acid in samples stored in the silo were optimized. The models were developed individually with samples

Table 1  
Qualitative analysis: discrimination of low/high oleic acid levels

		Classification Results			
		Predicted Group Membership			Total
			Low oleic	High oleic	
Original	Count	Low oleic	109	11	120
		High oleic	14	129	143
	%	Low oleic	90.8	9.2	100
		High oleic	9.8	90.2	100
Leave-one-out cross-validation	Count	Low oleic	107	13	120
		High oleic	15	128	143
	%	Low oleic	89.2	10.8	100
		High oleic	10.5	89.5	100

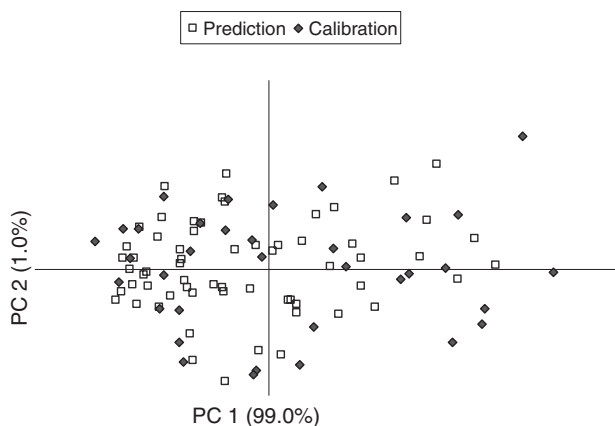


Figure 2  
Selection of samples (used in calibration and prediction) by PCA. In silo samples.

from the farmers and from the silo because the moisture range was very different and this could have strongly affected the NIR analyses.

Untreated spectra were used to build all the calibration models, employing the use of the aforementioned nineteen wavelengths in the 1445-2348 nm spectral range. The most relevant parameters for both the calibration and prediction curves are shown in Table 2. The maximum relative standard errors of prediction (RSEP, %) in the samples from the farmers and from the silo were 5.8% and 2.3% for moisture; 1.3% and 1.6% for fat respectively, and 1.6% for oleic acid in the silo samples.

Figures 3 and 4 show the parameters of the calibration and prediction curves for moisture and fat in the samples from the farmers, and moisture, fat and oleic acid for the silo samples, comparing the reference data with those predicted by the NIR models. The obtained results presented an adequate capacity for determining the aforementioned parameters.

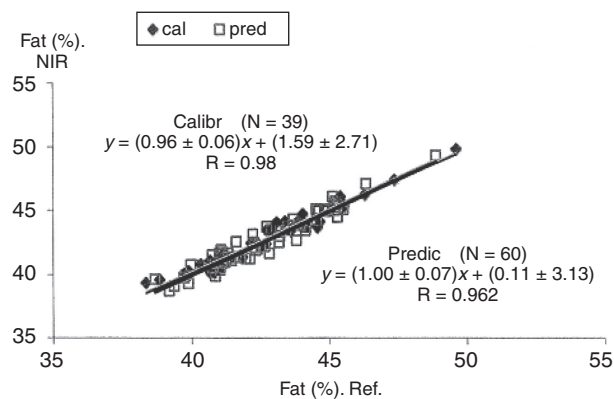
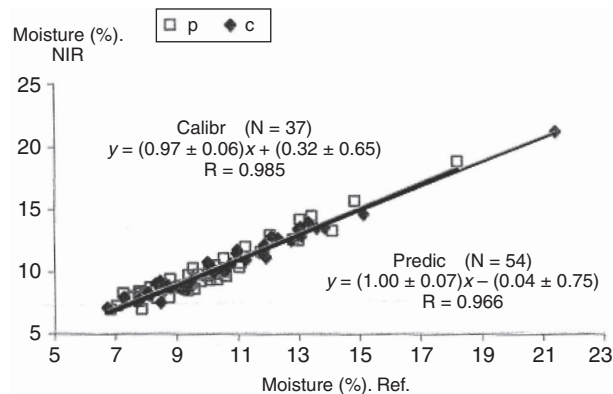


Figure 3  
Parameters of the calibration and prediction curves for the percentage of moisture and fat. Farmers' samples.

### 3.3. Predictive capacity of the models developed

The calibration models obtained were applied to unknown samples for routine analysis with a view to determining their predictive capacity: a Student's t test of residuals (level of significance 95%) of all the samples revealed that the residuals obtained with the five MLR models were not significantly

Table 2  
Statistical parameters for regression line of predicted NIR values vs. reference values for each parameter

			Samples	Range (%)	RSE (%)
Farmers	Moisture	Calibration	37	6.7-21.4	4.2
		Prediction	54	6.8-18.2	5.8
	Fat	Calibration	39	38.4-49.6	1.1
		Prediction	60	38.7-48.9	1.3
Company Silo	Moisture	Calibration	31	4.6-6.8	1.1
		Prediction	52	4.7-6.8	2.3
	Fat	Calibration	30	38.7-48.8	1.0
		Prediction	34	39.4-47.9	1.6
	Oleic acid <sup>a</sup>	Calibration	17	60.0-93.1 <sup>a</sup>	0.6
		Prediction	32	61.5-89.5 <sup>a</sup>	1.6

<sup>a</sup> Oleic acid expressed as percentage of total fatty acids.

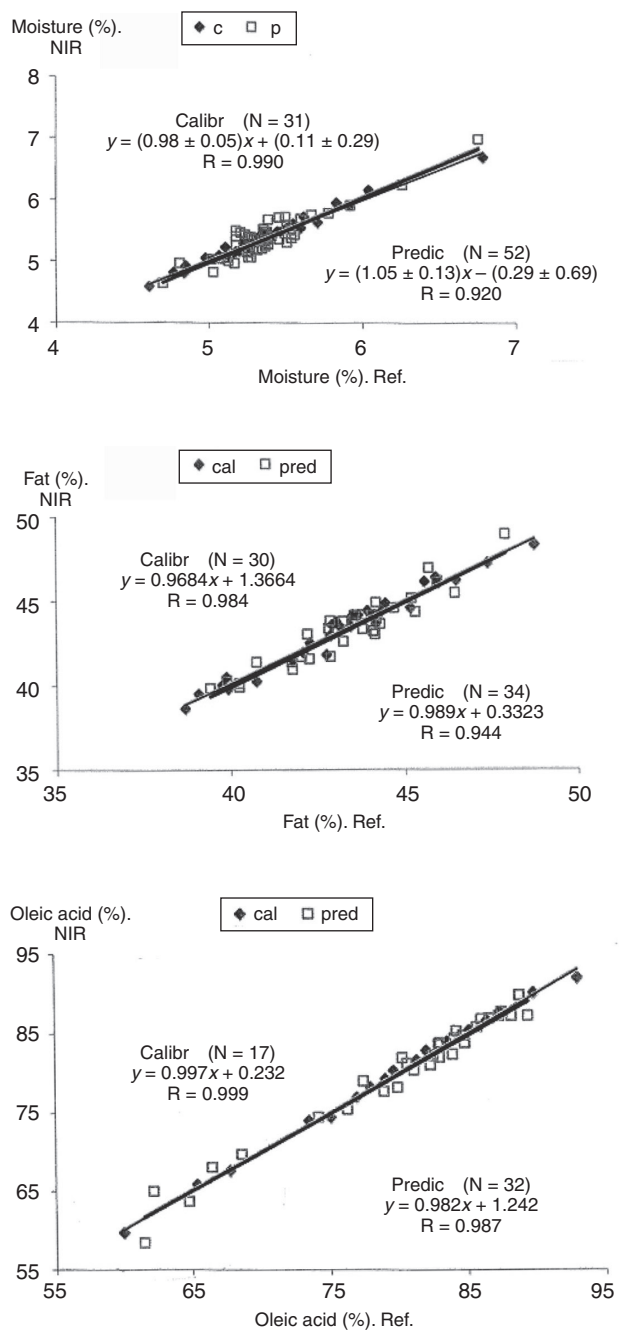


Figure 4  
 Parameters of the calibration and prediction curves for the percentage of moisture fat and oleic acid (expressed as percentage of total fatty acids). In silo samples.

different from 0 (Table 3). This confirms that they are adequate for establishing the values of the parameters used in the assessment of the quality of sunflower seeds.

#### 4. CONCLUSIONS

The results of the present work show that the proposed analytical method, employing the near infrared spectroscopy technique, can be used for the determination of moisture, fat, and high/low oleic acid levels in samples of sunflower seeds in a range between 4.6 and 21.4% for moisture; 38.4 and 49.6% for fat, and 60.0-93.1% for oleic acid, with results comparable to those obtained with classic chemical methods. Moreover, the discrimination model allows for the classification of sunflower seeds with high or low oleic acid levels, with a prediction rate of 90.5% in internal validation and 89.4% in leave-one-out cross-validation. The method developed allows for a control of the raw materials from farmers and monitoring of their storage in silos.

#### ACKNOWLEDGEMENTS

The authors are grateful to Cecosa SA for their collaboration in providing the samples and reference data, and the research project JACN of Castilla y León. J.M. Hernández-Hierro thanks to the Spanish MICINN for the Juan de la Cierva contract (JCI-2011-09201).

#### REFERENCES

AOAC (1990). Official Methods of Analysis, 15th ed. Association Official Analytical Chemists, Arlington, VA.  
 Cantamutto M, Poverene, M. 2007. Genetically modified sunflower release: Opportunities and risks. *Field Crop Res.* **101**, 133-144.  
 Cantarelli MA, Funes IG, Marchevsky EJ, Camiña JM. 2009. Determination of oleic acid in sunflower seeds by infrared spectroscopy and multivariate calibration method. *Talanta.* **80**, 489-492.  
 Christensen CM. 1982. Storage of Cereal Grains and their Products. AACC, St. Paul, Minnesota, USA.

Table 3  
 Students' test of residual ( $\alpha = 0.05$ ) of the unknown samples

		N	average	SD	t exp	t crit	Confidence level
Farmers	% Fat	60	-0.07	0.57	1.01	2.00	0.15
	% Moisture	54	-0.04	0.59	0.54	2.00	0.16
Company Silos	% Fat	34	-0.07	0.71	0.82	2.00	0.16
	% Oleic acid	32	-0.20	1.30	1.10	2.04	0.50
	% Moisture	52	-0.01	0.14	0.61	2.00	0.40

- Codex alimentarius. 1999. *Norma del CODEX para aceites vegetales especificados* (Codex STAN 210-1999). Revisión 2011.
- Fassio A, Cozzolino D. 2004. Non-destructive prediction of chemical composition in sunflower seeds by near infrared spectroscopy. *Ind. Crop Prod.* **20**, 321-329.
- Flagela Z, Rotunno T, Tarantino E, Di Caterina R, De Caro A. 2002. *Eur. J. Agr.* **17**, 221-230.
- Franke EN. 1998. In Lipid oxidation, The Oily Press Ltd. Scotland Dundee, U.K.
- Hajimahmoodi MY, Vander Heyden Y, Sadeghi NB, Jannat B, Oveisi MR, Shahbazian S. 2005. Gas-chromatographic fatty-acid fingerprints and partial least squares modeling as a basis for the simultaneous determination of edible oil mixtures. *Talanta* **66**, 1108-1116.
- ISO. International Organization for Standardization (1978). International Standard ISO 5509.
- Jurado-Expósito M, López-Granados F, Atenciano S, García-Torres L, González-Andújar JL. (2003). Discrimination of weed seedlings, wheat (*Triticum aestivum*) stubble and sunflower (*Helianthus annuus*) by near-infrared reflectance spectroscopy (NIRS). *Crop Prot.* **22**, 1177-1180.
- Kent NL. 1975. *Technology of Cereals* Pergamon Press UK, USA.
- Le Dréau Y, Dupuy N, Artaud J, Ollivier D, Kister J. 2009. Infrared study of aging of edible oils by oxidative spectroscopic index and MCR-ALS chemometric method. *Talanta*, **77**, 1748-1756.
- López-Feria S, Cárdenas S, García-Mesa JA, Valcárcel M. 2008. Classification of extra virgin olive oils according to the protected designation of origin, olive variety and geographical origin. *Talanta*, **75**, 937-943.
- Maggio RR, Kaufman TS, Carlo MD, Cerretani L, Bendini A, Cichelli A, Compagnone D. 2009. Monitoring of fatty acid composition in virgin olive oil by Fourier transformed infrared spectroscopy coupled with partial least squares. *Food Chem.* **114**, 1549-1554.
- Massart DL, Vandeginste BGM, Buydens LMC, De Jong SPJ, Lewi PJ, Smeyers-Verbeke S. 1998. *Handbook of Chemometrics and Qualimetrics: Part B*, Elsevier, The Netherlands.
- Moschne CR, Biskupek-Korell B. 2006. Estimating the content of free fatty acids in high-oleic sunflower seeds by near-infrared spectroscopy. *Eur. J. Lipid Sci. Tech.* **108**, 606-613.
- Mosser B. 2008. *Energy Fuel* **22**, 4301.
- Pereyra-Irujo GA, Izquierdo NG, Covi M, Nolasco SM, Quiroz F, Luis AN, Aguirrezábal LAN. 2009. Variability in sunflower oil quality for biodiesel production: A simulation study. *Biomass Bioenerg.* **3**, 459-468.
- Pomeranz Y. 1978. *Advances in Cereal Science and Technology*, vol II. AACC, St. Paul, Minn. USA.
- Roberts CA, Ren C, Beuselinck PR, Benedict HR, Bilyeu K. 2006. Fatty acid profiling of soybean cotyledons by near-infrared spectroscopy. *Appl. Spectrosc.* **60**, 1328-1333.
- Salunkhe DK, Chavan JK, Adsule RN, Kadam SS. 1991. *World Oilseeds: Chemistry, Technology and utilization*. Publish Van Nostrand Reinhold New York, USA.
- Weinstock BA, Janni J, Hagen L, Wrigth S. 2006. Prediction of oil and oleic acid concentrations in individual corn (*Zea mays* L.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. *Appl. Spectrosc.* **60**, 9-16.

Recibido: 26/9/12  
Aceptado: 8/10/12