

# Fourier Transform Near-Infrared Spectroscopy to Predict the Gross Energy Content of Food Grade Legumes

Tamás Szigedi · Marietta Fodor · Dolores Pérez-Marin · Ana Garrido-Varo

Received: 18 April 2012 / Accepted: 1 October 2012 / Published online: 5 December 2012  
© Springer Science+Business Media New York 2012

**Abstract** The feasibility of Fourier transform near-infrared reflectance spectroscopy (FT-NIRS) for determining gross energy content in different food legumes has been investigated. Eighty food-grade legume samples were obtained from different retailers and local markets in Hungary and they included 42 common beans (*Phaseolus vulgaris* L.), 20 peas (*Pisum sativum* L.), 10 lentils (*Lens culinaris* L.), and 8 soya beans (*Glycine max* L.) both as full fat food and defatted. The samples were analyzed by an adiabatic bomb calorimeter and then scanned in a Bruker MPA FT-NIR Analyzer (800–2,500 nm). Two algorithms for spectral selection of calibration and validation samples, which represent variability encountered in the full population, were tested. Partial least squares regression were developed for the prediction of gross energy using four different spectral preprocessing methods (first and second derivative alone and combined with standard normal variation and multiplicative scatter correction). The results show that first derivative produced the most accurate results with very high coefficient of determinations in validation (<93 %) and with very low standard errors of validation (<0.025 kcal/g) as compared to the standard error of the reference method (0.204 kcal/g).

**Keywords** Food-grade pulses · FT-NIRS · Selection of spectral outliers

T. Szigedi · M. Fodor  
Department of Applied Chemistry,  
Corvinus University of Budapest,  
Villányi út 29-35,  
1118 Budapest, Hungary

T. Szigedi  
e-mail: [tamas.szigedi@uni-corvinus.hu](mailto:tamas.szigedi@uni-corvinus.hu)

D. Pérez-Marin · A. Garrido-Varo (✉)  
Non Destructive Spectral Sensors Unit, Faculty of Agriculture  
and Forestry Engineering, University of Cordoba,  
Campus of Rabanales,  
14071 Cordoba, Spain  
e-mail: [palgavaa@uco.es](mailto:palgavaa@uco.es)

## Introduction

The energy content of foods can be determined in different ways. One of the methods is the gross energy content determination using an adiabatic bomb calorimeter (Cohen and Schilken 1994; Santos 2010). It is well known that the human body cannot completely digest proteins and fibers of foods consumed. Therefore, the adjustments in gross energy values are required to reflect food energy content that can be utilized by humans (Kays and Barton 2002). In practical terms, whatever energy method is used, energy food labeling still depends on wet chemistry analysis. There is little prior research to guide key decision makers to provide food processing industries with rapid and affordable analytical methods to determine the energy of every package food put into circulation in the market.

Over the past 30 years, a plethora of papers showed how near-infrared spectroscopy (NIRS) applications in agriculture and food industry have outstripped all other techniques in terms of diversity, while maintaining an excellent reputation for accuracy and reproducibility in the agro-food sector from grains, forages, trees, shrubs, feeds, product of animal origin (milk, meat, eggs, protein, and fat animal byproducts), fruits and vegetables, baking products, snacks, beverages and liquors, coffee, chocolate, tea, tobacco, sugar cane, potatoes, sauces, medicinal and aromatic plants, to non-food such as soils, slurries, water, timbers, wood, paper, and a still incomplete and never ending list. From the extensive research done, it can be concluded that NIRS is ideal for quality and safety control in many industries and farms (Osborne 2006; Roberts et al. 2004).

NIRS is based on the absorption of electromagnetic radiation at wavelengths in the range 780–2,500 nm. NIR absorption bands are produced when NIR radiation at specific frequencies (wavelengths) resonates at the same frequency as a molecular bond in the test sample. This allows association of a specific wavelength with a specific chemical bond

vibration (C–H stretch, C–C and C–O–C stretch) generating a specific spectra that in turn is related to concentration of a specific feed component (Shenk and Westerhaus 1995). However, for most food samples, the spectrum is not only due to the chemical information but also to other physical properties such as the particle size of powders. This means that to extract relevant chemical information from the spectra NIRS is used as a secondary method requiring calibration against a reference method for the constituent of interest.

NIRS has many advantages as compared to traditional wet chemistry methods (i.e., nondestructive, minimal or no sample preparation, instantaneous measurement, no use of reagents or chemical wastes, simultaneous analysis of several parameters, etc). However, it also has some limitations such as the moderately large upfront cost for the instrumentation needed, and the need to calibrate the instrument for each component of each type of sample to be analyzed. The calibration process can be resource-intensive. NIRS has been successfully applied to the determination of gross chemical composition of many small grains including pulses (Norris and Williams 2004). Furthermore, several studies have been published within the topic of the determination of energy content on cereal food products (Kays and Barton 2002) and homogenized meals (Kim et al. 2007). However, the ability of NIRS to predict gross energy has not been reported in food-grade legumes. The main goal of the present study is to demonstrate the feasibility of Fourier-transform NIRS and multivariate methods to predict gross energy content in a variety of food-grade legumes.

## Materials and Methods

### Samples

A total of 80 food-grade legume samples were used in the study. The samples were obtained from different retailers and local markets in Hungary and they included 42 common bean samples (*Phaseolus vulgaris* L.), 20 pea samples (*Pisum sativum* L.), 10 lentil samples (*Lens culinaris* L.), and 8 soya bean samples (*Glycine max* L.) both as full fat food and defatted.

The samples were dried with a Memmert drying oven (Mettler, Germany) at  $80 \pm 2$  °C during 16 h. Afterwards, the samples were ground in a laboratory mill, sieved through a 315 µm mesh size metal sieve and kept in airtight bags before analyses.

### Reference Analysis

Milled samples were analyzed in duplicate by an adiabatic bomb calorimeter (IKA Werke, Germany) using benzoic acid as the standard, according to the manufacturer's recommendations.

The standard error of the reference method (SEL) value was calculated according to the formula:

$$SEL = \sqrt{\sum_{i=1}^N \frac{(Y_{i1} - Y_{i2})^2}{2N}}$$

where  $N$  is the number of samples,  $Y_{i1}$  is the first laboratory data for sample  $i$ , and  $Y_{i2}$  is the second laboratory data for sample  $i$ .

### Spectroscopic Analysis

Milled samples were scanned with a Bruker MPA FT-NIR Analyzer (Bruker Optics, Ettlingen, Germany). The instrument scans the 800–2,500 nm wavelength range and it is equipped with a PbS detector. Petri dishes (internal diameter=93 mm, depth=14 mm) were used as sample holders to record the spectra collected in diffuse reflectance mode with the sphere microsample measurement channel with  $8 \text{ cm}^{-1}$  spectral resolution. Each spectrum was recorded as the average of 32 subsequent scans.

### Population Structuring and Detection of Spectral Outliers

Principal component analysis (PCA) was performed on the full set of samples ( $N=80$ ) in order to decompose and compress the data matrix. Before carrying out PCA, the spectra were pretreated by the standard normal variate (SNV) plus detrending algorithms (Barnes et al. 1989) in order to remove the multiplicative interferences of scatter, and finally, the first derivative was calculated (1,10,5,1); where the first digit is the order of the derivative, the second is the gap over which the derivative is calculated, the third is the number of data points in a running average or smoothing and the fourth is the second smoothing (ISI 2000). After PCA analysis, the center of the spectral population was determined in order to detect outlier samples. The Mahalanobis distance (GH) was calculated between each sample and the mean spectrum. Samples with a GH value higher than 3 were considered outliers (Shenk and Westerhaus 1996).

### Selection of Calibration and Validation Sets

After elimination of outlier spectra, calibration and validation samples were selected solely on the basis of spectral data. Two different calibration and validation sets (CAL1, VAL1 and CAL2, VAL2) were evaluated with the same number of initial samples. The two calibration and validation sets were selected using two different algorithms, differing in the procedure of sample selection based on spectral information. Therefore, they obviously arrive at different combination of samples. The main reason for doing that was to investigate optimal methods for sample selection, based on spectral information only and not in chemical information. We tested two strategies: a well-known first one that has been used by many researchers, based on the patented algorithm of Shenk

and Westerhaus (1991a, b) and the second one that was based on the possibilities supported by the software OPUS 6.5 (Bruker Optics, Ettlingen, Germany), an option definitely associated to the software of the Bruker instrument used in this work.

The first group was selected using the CENTER algorithm included in the WinISI II (Infrasoft International, Port Matilda, PA, USA). The algorithm performs a PCA, thus reducing the original spectral information ( $\log(1/R)$  values) to a small number of linearly independent variables to facilitate the calculation of spectral distances. These new variables were used to calculate the center of the spectral population and the distance (expressed as the Mahalanobis “GH” distance) of each sample in the calibration set from the center (Shenk and Westerhaus 1991a; b). After ordering the sample set by spectral distances (from the smallest to the greatest distance to the center), a total of 52 samples were selected.

The second group of samples used for calibration was selected with the option “automatic selection of subset” available in OPUS. Sixty-six percent of the samples were selected in this case (OPUS 2007). This selection method selects the samples based on the scores for the two first PCs. The closer the samples are in the scores plot, the more similar they are with respect to PC1 and PC2. Conversely, samples far away from each other are different from each other. The “automatic selection subset” procedure selects the samples by removing redundant samples while preserving the variability existing in the original population.

#### Calibration Development

Multivariate analysis was used for quantitative analysis. Partial least squares (PLS) algorithm provided by the OPUS 6.5 and Bruker Quant 2 software packs was used in the present study. One of the main advantages of PLS is that the resulting spectral vectors are directly related to the constituents of interest. Moreover PLS also offers reliable results when analyzing systems with widely varied constituent concentrations. In the PLS regression, the spectral and concentration data are first encoded in matrix form, then reduced to few factors and finally, the optimum number of factors is determined by lowest root mean square error of cross-validation (RMSECV) values. Four different spectral preprocessing methods were used to develop calibration models: first and second derivatives, first derivative combined with SNV, and first derivative combined with MSC. The first derivative calculates the first derivative of the spectrum. This method emphasizes steep edges of a peak. It is used to emphasize pronounced, but small features over a broad background. Spectral noise is also enhanced. Second derivative is similar to the first derivative, but with a more drastic result.

SNV normalizes a spectrum by first calculating the average intensity value and subsequent subtraction of this value from the spectrum, then the sum of the squared intensities is calculated and the spectrum is divided by the square root of

this sum. This method is used to account for, e.g., different sample thickness values.

MSC performs a linear transformation of each spectrum to best fit the mean spectrum of the whole set. This method is often used for spectra measured in diffuse reflection (OPUS 2007). The performance of the final PLS model was evaluated in terms of RMSECV for cross-validation and the coefficient of determination ( $R^2$ ). For RMSECV, a leave-one-out cross-validation is performed: the spectrum of one sample of the calibration set is deleted from this set and a PLS model is built with the remaining spectra of the calibration set. The left-out sample is predicted with this model and the procedure is repeated with leaving out each of the samples of the calibration set. The number of PLS vectors used is defined in the OPUS software by the size of the “RANK”. The optimum PLS factors are arranged in the correct sequence, according to their relevance to predict the component values. The first factor explains the most drastic changes of the spectrum.

The residual (Res) is the difference between the true and the fitted value. The sum of squared errors (SSE) is of the quadratic summation of the Res values.

$$\text{SSE} = \sum (\text{Res}_i)^2$$

The determination coefficient,  $R^2$  gives the percentage of variance present in the true component values, which is reproduced in the regression.  $R^2$  approaches 100 % as the fitted concentration values approach the true values.

$$R^2 = \left( 1 - \frac{\text{SSE}}{\sum (y_i - y_m)^2} \right) \times 100,$$

where  $y_i$  is the reference value and  $y_m$  is the mean of the reference results for all samples.

In case of cross-validation, the RMSECV is calculated using the following formula:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (\bar{y}_i - y_i)^2}{n}}$$

where  $n$  is the number of samples in the training set,  $y_i$  is the reference measurement result for the sample  $i$  and  $\bar{y}_i$  is the estimated result for sample  $i$  when the model is constructed with the sample  $i$  removed. The number of PLS factors included in the model is chosen according to the lowest RMSECV. This procedure is repeated for each of the pre-processed spectra.

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i - \text{BIAS})^2}{n - 1}}$$

where,  $n$  is the number of samples in the test set,  $y_i$  the reference measurement result for test set sample  $i$  and  $\hat{y}_i$  is

the estimated result of the model for test sample  $i$ . SEP is corrected by BIAS.

For the test set, the root mean square error of prediction (RMSEP) is calculated as follows

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

where,  $n$  is the number of samples in the test set,  $y_i$  the reference measurement result for test set sample  $i$ , and  $\hat{y}_i$  is the estimated result of the model for test sample  $i$  (OPUS 2007).

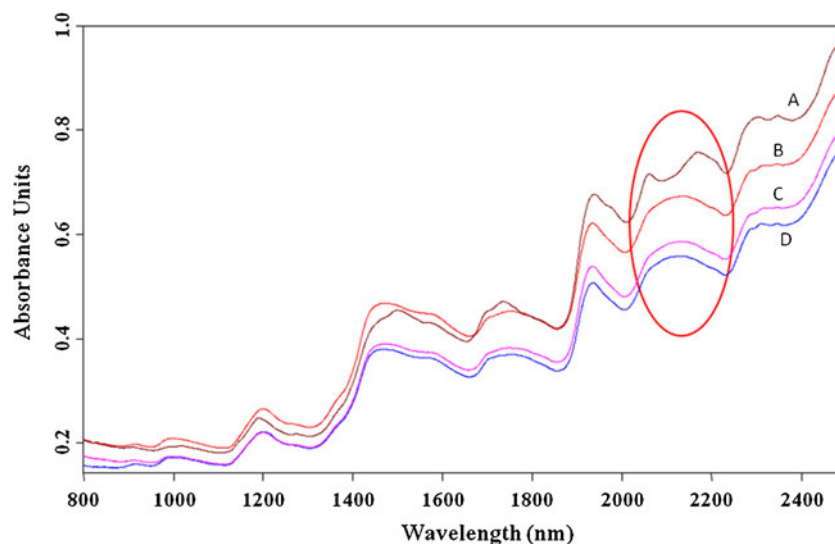
In addition to the statistics implemented in OPUS 6.5, other statistics used were the RPD (SD/RMSECV, RMSEP); RER=range/ RMSECV or RMSEP and CV=(RMSECV or RMSEP /Mean)×100 (Williams 2001).

## Results and Discussion

### Spectral Features and Population Structure

Figure 1 shows the spectra of (a) soybean, (b) lentil, (c) common bean, and (d) pea samples. The spectral pattern of the mean of the four legumes is similar with a clear exception of the shape of the curve for the mean soybean spectrum (C) in the region compressed between 2,000 and 2,200 nm. The peaks at 2,054 and 2,172 and 2,094 have been associated with a strong absorption of protein and starch, respectively (Williams 2001). The differences observed in the soybean spectrum in relation to the others legumes spectra may be attributed to the changes in physicochemical composition (decrease in the content of free amino acids, Maillard reaction between protein and starch, etc.) caused by the technological processing of the raw soy beans into the so-called new generation soy foods (Boge et al. 2009).

**Fig. 1** Typical average FT-NIR spectra of the four legumes: A soybean, B lentil, C common bean, D pea

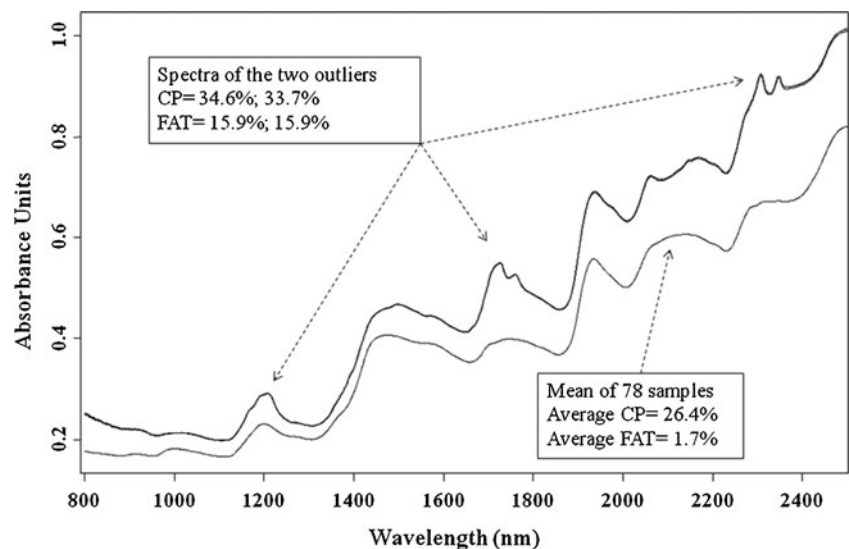


Detection of spectral outliers is critical in any calibration development process. Therefore, the full set of spectral data was examined using PCA to find patterns in the group of spectra that contribute the most to the variation among the spectra. During the PCA analysis two samples were detected as spectral outliers. A detailed analysis of the spectra of these two samples as compared to the mean of the other 78 spectra (Fig. 2 clearly) shows that the outliers have well defined peaks in the regions around 1,700 and 2,300 nm. In a review paper, Garrido-Varo et al. confirmed that the peaks at 1,734, 1,765, 2,310, and 2,345 nm are due to absorption of fatty acids (Garrido-Varo et al. 2004). Available data of the fat content of the full set used in this paper, determined by wet chemistry (data not shown) confirmed that the outlier samples contained much more fat than rest of legume samples (Fig. 2).

These results confirm the suitability of NIRS technology to develop calibration equations to predict the chemical composition of food-grade legumes. Also, the results indicate that the spectra alone contain a considerable amount of information that can help the soya food processing industry to control the production process by online monitoring in specific wavelengths regions if well managed with an appropriate multivariate data treatment.

At the time of the calibration development for gross energy prediction, a decision should be made on whether to keep the two outliers samples in the calibration file or to remove them. It may be noted that despite that NIRS literature declares samples with high Mahalanobis' distance values as outliers, we should consider them extremely important in the long run to build a robust NIRS calibration. However, as these samples are unique ones, it was decided to remove them from the calibration set and to wait until more samples of these characteristics can be added to the spectral library file.

**Fig. 2** Spectra of the two outlier samples and average spectrum of the full set



### Determination of Gross Energy by an Adiabatic Bomb Calorimeter

The next step in the process of calibration development is to analyze samples by a precise reference method, and afterwards, to select appropriate calibration and validation sets. Table 1 shows the mean, range, and standard deviation values of the gross energy determined by the reference method for the full set and for the two calibrations and validation sets obtained using different strategies.

The range of gross energy measured by the adiabatic bomb calorimetric method was 4.149–5.511 kcal/g with a standard deviation of the method of 0.204 kcal/g and with the SEL value of 0.015 kcal/g. It is recommended (Mark and Campbell 2008) that the ratio of the calibration range and the standard deviation of the reference method should be at least 10, with the larger value the better. In our study, this ratio reaches a value as high as 20.1. This preliminary information confirms that the highly accurate determination of the gross energy values determined by an adiabatic bomb calorimeter together with a proper methodology for selection of the calibration and validation sets could ensure highly accurate equations.

### Selection of Calibration and Validation Sets

One of the key issues in NIRS calibration development is the sample selection for calibration, validation, and further

enlargement. The OPUS software provides a routine protocol that can automatically select a subset of samples based on the scores for the two first PCs. However, the CENTER algorithm implemented in WinISI uses PCA followed by the calculation of the Mahalanobis distance. When both algorithms were applied to the full population of  $N=78$  samples, they produced very similar results considering the statistics (mean, range, and SD) of gross energy (GE) values (Table 1). Furthermore, the calibration and validation sets are almost identical to the full set showing that the two methods based on spectral selections are useful to obtain subsets of samples representative of the variation encountered in the full population.

### Calibration Development

The development of calibrations for the measurement of gross energy was done by the OPUS 6.5 software applying PLS regression. Every calibration sets has a total of 52 samples. For calibration the 800–2,500 nm wavelength range was used. However, during the calibration process using the OPUS software, the regions 1,110.3–1,337.1 and 1,638.8.2–2,175.5 nm regions were selected as the most important for GE.

In the application of the PLS algorithm, it is generally known that the spectral preprocessing methods and the number of PLS factors are critical parameters. The optimum number of factors is determined by the lowest RMSECV and

**Table 1** Gross energy values (kcal/g) statistics for the full set ( $N=78$ ) and for the three calibration and validation sets studied

	Full set $N=78$	Cal 1	Val 1	Cal 2	Val 2
Mean (kcal/g)	4.249	4.241	4.266	4.240	4.267
Range (kcal/g)	4.149–4.511	4.149–4.510	4.174–4.511	4.149–4.511	4.164–4.510
SD	0.097	0.091	0.107	0.092	0.105

**Table 2** Cross-validation results for the two Cal sets

		Calibration set Cal 1.				Calibration set Cal 2.			
		First derivative	Second derivative	First derivative +MSC	First derivative +SNV	First derivative	Second derivative	First derivative +MSC	First derivative +SNV
Calibration	$R^2$	98.32	95.67	97.63	98.16	96.70	98.26	95.82	96.02
	RMSEC	0.0124	0.0195	0.0147	0.0128	0.0156	0.0123	0.0201	0.0184
	RPD	7.7	4.8	6.5	7.4	5.5	7.6	4.9	5.0
Cross-validation	$R^2_{CV}$	95.44	91.01	94.69	95.68	94.24	93.68	94.03	92.48
	RMSECV	0.0187	0.0263	0.0201	0.0182	0.0192	0.0217	0.0226	0.0235
	RPD	4.7	3.3	4.3	4.8	4.2	4.0	4.1	3.7
	Number of PLS factors	7	5	7	6	5	6	5	6

highest value for  $R^2_{CV}$ . Table 2 shows the cross-validation results obtained with the two datasets evaluated.

The results of cross-validation for the CAL1 set show that the first derivative alone or combined with MSC or SNV can produce better results than those obtained with the second derivative. The calibrations explain between 91.01 and 95.68 % of the variation encountered in GE and they produced RMSECV values very low as compared to the SEL. However, when using CAL2, the second derivative preprocessing technique gave the highest  $R^2$  (0.98) and lowest RMSEC (0.0123); but these results were not confirmed in the cross-validation ( $R^2_{CV}=93.68$  % and  $RMSECV=0.0217$ ), where the best model was obtained by using the first derivative. The comparison of the RMSECV values to the SEL value for the reference method (adiabatic bomb calorimeter) for the best two models obtained with CAL1 and CAL 2 demonstrates that the NIRS prediction errors for gross energy were of 1.2 (CAL1)- and 1.28 (CAL2)-fold higher than the accuracy of the reference method (0.015), which is considered as a rather high accuracy for analytical purposes (Williams 2001).

Kays and Barton (2002) using MSC plus second derivative for the prediction of gross energy reported a standard error of cross-validation of 0.053 kcal/for a range of values of 4.05–5.49 kcal/g. To compare the results obtained in cereal foods by Kays et al., we calculated the RPD (5.02), RER (19.61), and CV (0.43 %) values for gross energy for the best calibration model (CAL1) reported in Table 2. Kays et al. reported values of 7.35 for RPD, 25.71 for RER, and 1.08 % for CV. The highest RPD and RER values found by Kays et al. are clear consequences of the highest SD and range values of the calibration set used by those authors. However, the lowest CV obtained in the actual study for food legumes could be explained by the very low value of the SEC. Williams (2001) reported that CV for NIRS chemical predicted values between 0.6 and 1.0 should be considered as exceptional and those between 1.1 and 2.0 % as excellent. That is, the low CV obtained in our work confirm that NIR spectroscopy can show an extremely high

accuracy for the prediction of gross energy in food-grade legumes.

### Validation

Once the equations were obtained, the predictive capabilities of calibrations were also evaluated using the 26 samples comprising the two validation sets. The statistics for the prediction using first derivatives are shown in Table 3.

The GE laboratory and NIRS-predicted values for the validation sets were highly correlated ( $r^2=96.6$  and 93.72 for VAL1 and VAL2, respectively). These values suggest accurate predictability. For all the pretreatments evaluated, the equations developed with CAL1 show the lowest SEP and the highest  $R^2$  values compared to those obtained with VAL2.

The model developed with VAL1 and using a first derivative plus SNV, had the best overall performance statistics. The SEP value (0.0253) was quite comparable to the benchmark standard error of the reference laboratory test (0.015) resulting in a  $RPD_{RMSEP}$  value of 4.2 and an exceptional CV of 0.58 %.

**Table 3** Validation results for the two Val sets

	Validation set Val 1. First derivative	Validation set Val 2. First derivative
$r^2$	96.6	93.72
SEP	0.0253	0.0264
RMSEP	0.0248	0.0259
RPD	4.2	4.0
RPDc	4.9	2.8
RER	16.69	13.72
CV	0.58	0.61

## Conclusions

It can be concluded that the FT-NIRS supported with multivariate analyses has high potential to estimate the gross energy of food-grade legumes in a nondestructive way and with high degree of accuracy. On the basis of the positive results obtained on this feasibility study, further research using a more comprehensive calibration with larger set of samples is justified and can be required to be suitable for industrial or commercial applications. The algorithm based on PCA and Mahalanobis distance seems to be a better procedure than the one based only on the two first PCA for selecting representative samples for enlarging the calibration set.

**Acknowledgments** The first author gratefully acknowledges the support of the TÁMOP-4.2.1/B-09/1/KMR-2010-0005 grant to spend a stay at the University of Cordoba where this research was undertaken. TSZ is also grateful to Mr. Antonio López (UCO), for technical assistance with the gross energy analyses conducted with the adiabatic bomb calorimeter.

## References

- Barnes RJ, Dhanoa MS, Lister SJ (1989) Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl Spectrosc* 43:772–777
- Boge EL, Boylston TD, Wilson LA (2009) Effect of cultivar and roasting method on composition of roasted soybeans. *J Sci Food Agric* 89:821–826
- Cohen BL, Schilken CA (1994) Calorie content of foods—a laboratory experiment introducing measuring by calorimeter. *J Chem Educ* 71:342–345
- Garrido-Varo A, Garcia Olmo J, Perez-Marin D (2004) Application in the analysis of fat and oils. In: Roberts C, Workman J, Reeves JB (eds). *NIR spectroscopy in agriculture*
- ISI (2000) The complete software solution using a single screen for routine analysis, robust calibrations and networking. Infrasoftware International Sylver Spring MD, USA
- Kays SE, Barton FE (2002) Rapid prediction of gross energy, and utilizable energy in cereal food products using near-infrared reflectance spectroscopy. *J Agric Food Chem* 50:1284–1289
- Kim Y, Singh M, Kays SE (2007) Near-infrared spectroscopic analysis of macronutrients and energy in homogenized meals. *Food Chem* 105:1248–1255
- Mark H, Campbell B (2008) An introduction to near infrared spectroscopy and associated chemometrics. *The Near Infrared Research Corporation*
- Norris K, Williams PC (eds) (2004) *Near-infrared technology in the agricultural and food industries*. AACC, Inc, St. Paul
- OPUS (2007) User manual. Ettlingen, Germany
- Osborne BG (ed) (2006) *Near-infrared spectroscopy in food analysis*. doi:10.1002/9780470027318.a1018
- Roberts CA, Reeves JB, Workman J (eds) (2004) *Near-Infrared Spectroscopy in Agriculture (Agronomy)*. Soil Science Society of America, Inc. Publishers, Madison WI
- Santos RC (2010) Food energy values. Example of an experimental determination, using combustion calorimetry. O valor energético dos alimentos. Exemplo de uma determinação experimental. *Usando Calorimetria De Combustão* 33:220–224
- Shenk JS, Westerhaus MO (1991a) Population definition, sample selection, and calibration procedures for near-infrared reflectance spectroscopy. *Crop Sci* 31:469–474
- Shenk JS, Westerhaus MO (1991b) Population structuring of near infrared spectra and modified partial least squares regression. *Crop Sci* 31:1548–1555
- Shenk JS, Westerhaus MO (1995) Analysis of agriculture and food products by near infrared reflectance spectroscopy. *Monograph Infra Int*. Port Matilda, PA
- Shenk JS, Westerhaus MO (1996) Calibration the ISI way. In: Davies AMC, Williams PC (eds) *Near infrared spectroscopy. The future waves NIR publications*, Chichester, pp 198–202
- Williams PC (2001) Implementation of near-infrared technology. In: Williams PC, Norris K (eds) *Near-infrared technology in the agricultural and food industries*. AACC, Inc, St. Paul, pp 145–169