

Evaluation of Ginger adulteration with beans using Attenuated Total Reflectance Fourier Transform Infrared spectroscopy and multivariate analysis

W. Terouzi, A. Oussama

Abstract— Quantitative analysis of food adulterants is an important for health, wealth and economic issue that needs to be fast, simple and reliable. Attenuated Total Reflectance-Fourier Transform Mid Infrared Spectroscopy (ATR-FTMIR), combined with multivariate analysis, has been used to quantify the Beans content in a binary mixture with Ginger. Blends of Ginger with different percentages (0–30%) of Beans were measured using ATR-FTMIR spectroscopy. Spectral and reference data were firstly analyzed by principal component analysis (PCA) to check outliers samples. Partial least square regression (PLSR) was used to establish calibration model. Excellent correlation between ATR-FTIR analysis and studied blend samples was obtained $R^2 = 0.99$; with Root Mean Square Errors of Prediction < 1.102 , Limit of Detection 3.305%, and Relative Prediction Errors as low as 0.67. These results indicate that ATR-FTMIR spectroscopy combined with chemometrics (multivariate analysis) can be used for rapid prediction of Beans content in Ginger.

Index Terms—Adultration, Attenuated Total Reflectance-Fourier Transform Mid Infrared Spectroscopy, ginger, multivariate analysis.

I. INTRODUCTION

Spices play an important role as flavouring agents in the diet and are used throughout the world. Various phytochemicals present in spices have been recognized to have health promoting benefits and preventive role in chronic diseases [1], [2], [3].

In fact, food industry is facing challenges in preserving better quality of fruit and vegetable products after processing. Recently, many attentions have been drawn to ginger rhizome processing due to its numerous health promoting properties.

Ginger (*Zingiber officinale* Roscoe, Zingiberaceae) is one of the most commonly used spices around the world, originates in China and then spreads in India [4]. It is also an important medicine for treating cold, stomach upset, diarrhea, and nausea. Phytochemical studies show that ginger has antioxidant and anti-inflammatory activities, and some of them exhibit potential cancer preventive activity [5], [6], [7]. On the other side, food authenticity is a major issue worldwide. It has been the target of government authorities, and presents a huge importance for consumers, food

processors, and industries, in order to satisfy food quality and safety requirements [8], [9].

According to the literature, the most commonly used methods in the field of food fingerprinting are based on spectroscopic data, for example, generated by using nuclear magnetic resonance (NMR), near-infrared (NIR) or FTIR spectroscopy. These techniques offer the possibility to analyze relatively small amounts of sample or its extract in a non-destructive, easy, quick and direct (with or without minor sample preparation) way. Therefore, the application of these spectroscopic methods represents a suitable strategy for the characterization of complex biological systems such as foods, since they allow a simultaneous determination of a high number of compounds [10].

In recent years, thanks to chemometric tools, the increased specificity and sensitivity of the analytical instruments offered the feasibility of obtaining a wide range of information in one shot. This technological breakthrough became more and more attractive and thus a normal approach to studying foods, in terms of either quality or authenticity assessment [11].

Additionally, the authenticity and quality control of Ginger by MIR spectroscopy combined with chemometrics has not been reported so far, even though mid-infrared is a region used for quantitative and qualitative analysis of several products.

The current study presents an application of ATR-FTMIR spectroscopy coupled with chemometric methods for quantification analysis of the fraudulent addition of Beans in Ginger. This application was considered to develop improved and reliable regression model (PLSR) which could later be used as a quick and accurate analysis tool for quantifying the actual percentage of Beans in the binary mixture with Ginger.

II. MATERIALS AND METHODS

A. Samples preparation

In this study, to prepare the adulterated Ginger samples we used:

- One Kilogram of Pure Ginger was purchased in a local supermarket grinded with an electric grinder and preserved at 17°C until preparation of blends.
- Good quality crude beans was obtained from local market: ½ kilogram of Moroccan beans picked up in Beni-Mellal, was grinded with an electric grinder and preserved at 17°C until preparation of blends.

Samples were prepared by mixing Ginger powder (G) with Beans powder (B). Samples with a final mass of 10 g were prepared in different percentages in the 0–30 % weight ratio range of Beans. All the samples were stored in a dry and dark location at ambient temperature (25°C) until analysis.

W. Terouzi, Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco, (212) 0610437333.

A. Oussama, Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco, (212) 0661573763.

The final data base consists of 91 samples, containing spectroscopic and compositional information of the analyzed mixtures. Among which 66 samples (calibration set) were randomly selected for establishing principal component analysis (PCA) and partial least square regression (PLSR) models. Other 25 samples were used to test the applicability of the regression model (prediction set).

B. ATR-FTIR analysis

ATR-FTIR spectra were obtained using a Vector 22 Bruker FTIR Spectrophotometer equipped with an attenuated total reflectance accessory (ATR single reflexion, Diamond, incident angle 45°, Pike Miracle, Pike Technologies, Madison, USA) with DTGS detector, Global (MIR) Source and KBr Germanium separator, with a resolution of 4 cm⁻¹ at 80 scans. Spectra were scanned in the absorbance mode from 4000 to 600cm⁻¹ and the data were handled with OPUS logiciel. About 1g of each binary blend powder samples of Ginger and Beans were directly placed, without preparation on an Attenuated Total Reflectance cell provided with a diamond crystal. Analyses were carried out at room temperature (25°C). The background was collected before every sample was measured. Between spectra, the ATR plate was cleaned in situ by scrubbing with ethanol solution, enabling to dry the ATR.

C. Data pre-processing procedures

In this study, a series of pre-processing elaborations were tested on the spectral data prior to the multivariate calibration. In fact, several pre-processing methods were applied before calibration development in order to find regression model with as high a predictive power as possible. The Savitzky-Golay [12] and Norris gap [13] algorithms were tested for data derivatisation. Standard normal variate (SNV) and multiple scatter correction (MSC) [14] were also tested. For data pre-treatment giving best result is the derivative function. In all PCA and PLSR models, second derivative through the Gap algorithm has been applied as preprocessing technique with centered data, in order to correct the spectrum by separating overlapping peaks and to enhance spectral differences.

D. Chemometric methods

• Principal Component Analysis (PCA)

Principal component analysis (PCA) is an unsupervised technique commonly used for quantification, characterization and classification of data. It is based on variance, transforms the original measurement variables into new uncorrelated variables called principal components [15], [16]. It maps samples through scores and variables by the loadings in a new space defined by the principal components. The PCs are a simple linear combination of original variables. The scores vectors describe the relationship between the samples and allow checking if they are similar or dissimilar, typical or outlier. It provides a reduction in data set dimensionality and allows linear combinations of the original independent variables that are used to explain the maximum of data set variance [17].

• Partial least squares regression (PLSR)

Partial least squares regression (PLSR) [18] is popular and the most commonly used multivariate calibration chemometrics methods. It is able to resolve overlapping spectral responses [19]. It assumes a linear relationship

between the measured sample parameters (for example, concentration or content) and the experimentally measured spectra.

PLSR attempts to maximize the covariance between X and y data blocks as it searches for the factor subspace most congruent to both data blocks. A new matrix of weights (reflecting the covariance structure between the X and y) is calculated and provided rich factor interpretation information [20].

In this study, the collected ATR-FTMIR spectra will be used as the X matrix, and the Beans compositions of the different samples will be used as the Y vector.

• Software

The pre-treatment procedures and all chemometric models were performed by using the Unscrambler X software version 10.2 from Computer Aided Modelling (CAMO, Trondheim, Norway).

III. RESULTS AND DISCUSSION

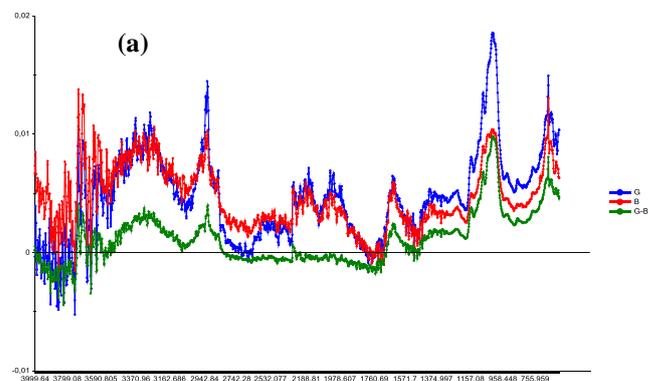
A. Data acquisition

In the first step, ATR-Fourier transform mid infrared (ATR-FTMIR) spectra of pure Ginger (G) and Beans (B) were obtained. One spectrum is the average of 80 scans of the same sample on ATR-FTMIR. The average spectra of all considered samples are presented in Fig.1. (a)

In the second step, ATR-FTMIR spectra of 91 samples of the adulterated Ginger were recorded in triplicate and a mean spectrum was calculated for studied samples. The resultant mean spectrum of binary mixtures (G-B) is shown in Fig.1. (a).

ATR-FTIR spectra of 91 samples of the studied binary mixtures were recorded and divided in two sets: a calibration set of 61 samples and an external validation set of 25 samples. One spectrum is the average of 80 scans of the same sample of blend. The average spectra of all considered samples in calibration set are presented in Fig.1. (b).

MIR spectroscopy is a fingerprint technique, allows differentiating between authentic milks and those adulterated with others by observing the spectra changes due to the adulteration. According to Fig.1, the MIR spectra obtained of the studied samples (pure or adulterated) to be similar. The detection of adulteration is more difficult, especially when the adulterant has similar chemical composition to that of the original one. In this case, multivariate analysis appeared to be ideal to provide an effective solution, as they allow extracting of unspecific analytical information from the full-spectra or large regions of them.



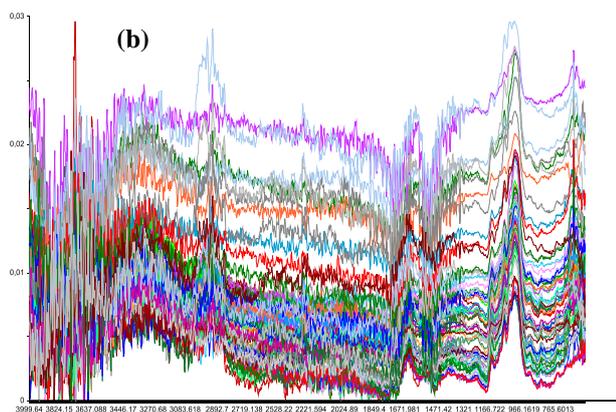


Fig.1. ATR-FTMIR spectra of: (a) pure Ginger (G), pure Beans (B) and binary mixtures G-B; (b) the binary mixture (G – B) samples of calibration set in the 0–30 % weight ratio range, at MIR region of 4000–600 cm^{-1}

With the aim to obtain more information from the ATR-FTMIR spectral data, the spectra were firstly subjected to mathematical elaboration. The best improvement in data variance was reached when the derivative function through the Gap algorithm was used. Different mathematical parameters in the derivative procedure were tested and results were optimized when the following parameters were selected: 2nd order, gap size 17; with centered data.

B. Multivariate analysis

• PCA modeling

Principal component analysis was carried out to detect the presence of any spectral outliers in the spectral data, prior to develop a prediction model using PLS regression. Many studies indicate that PCA is a useful tool for the identification of spectral outliers in the absorbance spectra of the samples and can be employed to increase the quality of the prediction-model [21]. Fig.2 shows the score plot obtained by PCA model in calibration set of adulterated samples.

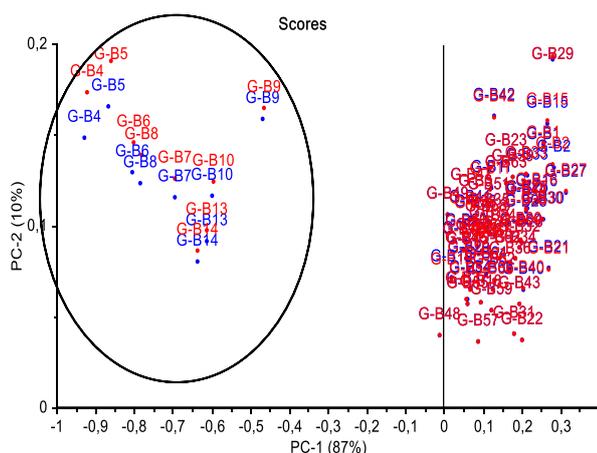


Fig.2. PC1 / PC2 Score plot by PCA analysis on the calibration set of binary mixtures (Ginger-Beans) samples. According to Fig.2 of PCA score plot, the data set contained nine spectral « outlier »: G-B4; G-B5; G-B6; G-B7; G-B8; G-B9; G-B10; G-B13 and G-B14. However, at first, the prediction model (PLSR) was building with all samples including this sample to insure his nature (outlier or extreme sample).

• PLSR modeling

In general, the modeling consists of two steps: (1) calibration, where data characteristics (Calibration and internal validation samples) are investigated to find a model for their behavior; and (2) External validation, where data that did not participate in the calibration step (external validation samples) are used to evaluate the model adequacy and capability.

The quantification of Beans (Adult.%) in adulterated Ginger samples was carried out using PLS algorithm. The PLSR model is built by considering the spectra range 4000–2400 cm^{-1} and 2300–600 cm^{-1} with X as variable and the Y variables is associated to the different percentages of the Beans. The range 2400–2300 cm^{-1} was deleted prior to calculations because of its low signal-to-noise ratio and the presence of fluctuations independent from sample composition [22].

The PLSR model was evaluated using coefficient of determination (R^2) in calibration, root-mean-square error of calibration (RMSEC) and cross validation (RMSECV).

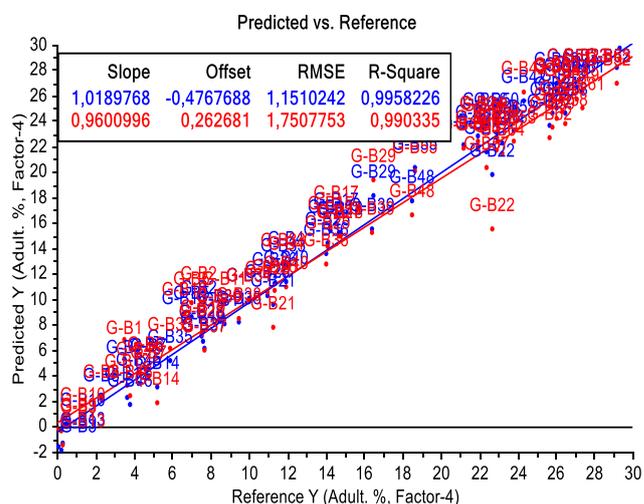


Fig.3. Measured vs. Predicted values for Beans in the studied binary mixtures obtained from the final PLSR model developed from the ATR-FTMIR spectra

Fig.3 shows the PLSR model which correlates the « actual » and « predicted » values of Beans percentages obtained from ATR-FTMIR spectra. The term « actual » refers to the known percentage of Beans. The « predicted » refers to a value calculated by the PLSR model using spectral data. The difference between the actual and the predicted percentage is relatively small with coefficient of determination (R^2) values 0.9958 with calibration set and 0.9903 with internal validation. The low value RMSEC (< 1.16) indicates the good performance of PLSR model [23].

Additionally, validity of the model was checked by running several diagnostics including R^2 , root mean standard error of calibration (RMSEC) and root mean standard error of cross validation (RMSECV). Root mean square error of cross-validation (RMSECV), recovery percentage and coefficient of determination (R^2) were used as parameters to determine appropriate number of latent variables (LV) [24], [25].

The determination of latent variables number was based on the statistical parameters that they offer the highest values of

R^2 and the lowest values of error, either in calibration or in prediction models. The statistical parameters RMSEC, RMSECV and R^2 are summarized in Fig.3. The coefficient of determination (R^2) of 0.99, RMSEC lower than 1.16 and RMSECV lower than 1.751, could be considered satisfactory. The four latent variables (factors) were sufficient for describing PLSR model, with explained variances above 99% (Fig.4).

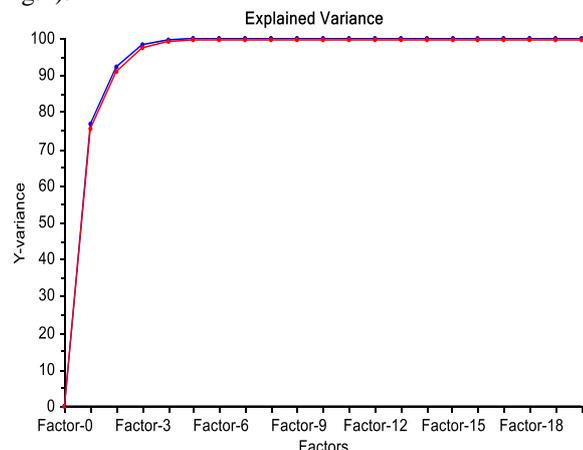


Fig.4. Plot of explained variance of factors describing PLSR model.

•Prediction of Beans content in the new binary blend samples (External validation)

In order to verify the applicability, performance and how reliable this model in estimating the percentage of Beans in binary mixtures with Ginger, the external validation process was carried out.

PLSR model is used to predict percentage of Beans in new blend samples. The new samples were prepared within the range considered by the original database (0-30%). These samples have the same matrix effects as samples of calibration set. In this step, the model was subdued to validation procedure by quantifying the new objects.

The performance of the PLSR models on the independent validation set (External validation) is assessed using R^2 , RMSEP and the residual prediction deviation (RPD). Here, the criteria of classifying RPD values [26] is adopted as follows: an RPD value below 1.5 indicates that the calibration is not usable; an RPD value between 1.5 and 2.0 indicates the possibility of differentiating between high and low values; an RPD value between 2.0 and 2.5 makes possible approximate quantitative predictions. For RPD value between 2.5 and 3.0 and beyond 3.0, the prediction is classified as good and excellent, respectively. Generally, a good model should have high values of R^2 and RPD, and low values of RMSEC, RMSECV and RMSEP.

The PLSR model was applied to a group of external samples (25 samples), the results are shown in Fig.5.

Fig.5 shows the PLSR model reconstructed by external validation samples, following the same previous pre-treatments. This PLSR model correlates the « actual » and « predicted » values of Beans percentages obtained from ATR-FTMIR spectra. The difference between the actual and the predicted percentage is relatively small.

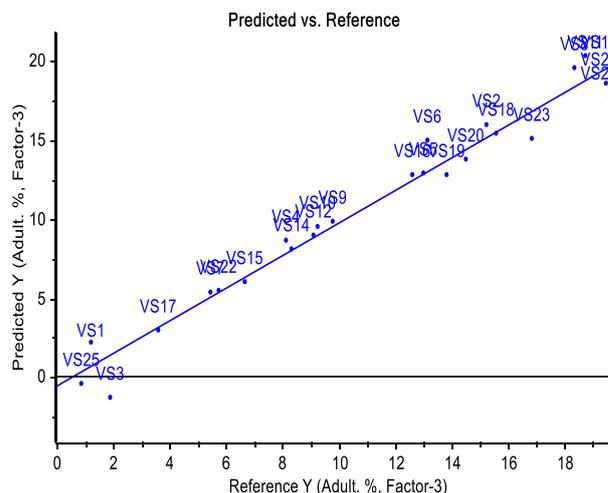


Fig.5. Measured vs. Predicted values for Beans in binary mixtures Ginger-Beans of external validation set.

Figures of merit of the calibration graphs are summarized in Table1. As can be seen, PLSR model offered good values for the different multivariate parameters.

Table1. Statistical parameters carried out by external validation on PLSR

LVs	R_p^2	RMSE P	Bias
4	0.992	1.1019	-0.2091
SEP	REP %	RPD	LD%
1.104	0.6656	5.5057	3.3057
2			

IV. CONCLUSION

In order to ensure the quality and authenticity of the Ginger, the productive sector and the regulatory agencies require a rapid, robust and accurate analytical method. Multivariate methods based on mid infrared spectroscopy have been proposed as an alternative for quality control analysis of Ginger.

According to the statistical results, it has been proved that the proposed method allow the correct quantification of Beans in the studied binary blends with Ginger. The PLSR model obtained from ATR-FTMIR spectra gave correlation coefficients of 0.99 and root mean square errors of prediction (RMSEP) value of 1.1019.

In general, the developed method presented better results that prove their performance and robustness for routine analysis in quality control monitoring by food industry and regulatory agencies.

REFERENCES

- [1] Codex Alimentarius Commission. Code of Hygienic Practices for Spices and Dried Aromatic Plants. 2014. CAC/RCP 42-1995.
- [2] Kaefer CM, Milner JA. The role of herbs and spices in cancer prevention. J Nutr Biochem. 19, 2008, 347-361.
- [3] Ferrucci LM, Daniel CR, Kapur K, et al. Measurement of spices and seasonings in india: opportunities for cancer epidemiology and prevention. Asian Pacific J Cancer Prev. 11, 2010,1621-1629.
- [4] Ali B, Blunden G, Tanira M, et al. Some phytochemical, pharmacological and toxicological properties of ginger (Zingiber

- officinales Rosc): a review of recent research. *Food Chem Toxicol.* 46, 2008,409-420.
- [5] Shukla, Y., & Singh, M. Cancer preventive properties of ginger: A brief review. *Food and Chemical Toxicology*, 45(5), 2007, 683–690.
- [6] Stoilova, I., Krastanov, A., Stoyanova, A., Denev, P., & Gargova, S. Antioxidant activity of a ginger extract (*Zingiber officinale*). *Food Chemistry*, 102(3), 2007, 764–770
- [7] Thomson, M., Al-Qattan, K. K., Al-Sawan, S. M., Alnaqeeb, M. A., Khan, I., & Ali, M. The use of ginger (*Zingiber officinale* Rosc.) as a potential antiinflammatory and antithrombotic agent. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 67(6), 2002, 475–478.
- [8] Singhal, R.S., Kulkarni, P.R., Rege, D.V. *Handbook of Indices of Food Quality and Authenticity*, Woodhead Publishing Ltd., Cambridge, UK, 1997.
- [9] Lees, M. (Ed.), *Food Authenticity and Traceability*, Woodhead Publishing Ltd., Cambridge, UK, 2003.
- [10] Esslinger, S., Riedl, J., Fahl-Hassek, C. Potential and limitations of nontargeted fingerprinting for authentication of food in official control. *Food Research International* 60, 2014, 189–204.
- [11] Consonni, R. and Cagliani, L. R. Nuclear Magnetic Resonance and Chemometrics to Assess Geographical Origin and Quality of Traditional Food Products. *Advances in Food and Nutrition Research*, Volume 59, ISSN 1043-4526, DOI: 10.1016/S1043-4526(10)59004-1. 2010 Elsevier Inc.
- [12] Savitzky, A. & Golay, M.J.E. Smoothing and differentiation of data by simplified least-squares procedures. *Analytical Chemistry*, 36, 1964, 1627–1639.
- [13] Norris, K.H. & Williams, P.C. Optimization of mathematical treatments of raw near infrared signal in the measurement of protein in hard Red Spring wheat, I: influence of particle size. *Cereal Chemistry*, 62, 1984, 158–165.
- [14] Iñón, F. A., Garrigues, J. M., Garrigues, S., Molina, A., and de la Guardia, M. Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance-Fourier transform infrared spectroscopy. *Analytica Chimica Acta*, 489, 2003, 59–75.
- [15] Vandeginste, B.G.M., Massart, D.L., Buydens, L.M.C., Jong, S.D.E., Lewi, P.J., Smeyers Verbeke, J. *Handbook of Chemometrics and Qualimetrics: Part B*, Elsevier, Amsterdam, The Netherlands, 1998, p. 88.
- [16] Wold, S., Esbensen, K. & Geladi, P. Principalcomponent analysis. *Chemom. Intell. Lab. Syst.*, 2, 1987, 37-52.
- [17] Brown, S.D., Tauler, R., Walczak, B. *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, vol. 2. Elsevier, Amsterdam, 2009.
- [18] Martens, H. & Naes, T. *Multivariate calibration*. Chichester: Wiley. 1989.
- [19] Khanmohammadi, M., Bagheri Garmarudi, A., Ghasemi, K., Garrigues, S., De la Guardia, M. Artificial neural network for quantitative determination of total protein in yogurt by infrared spectrometry. *Microchemical Journal* 91, 2009, 47–52.
- [20] Geladi, P. & Kowalski, B. Partial least-squares regression: a tutorial. *Anal Chim Acta*, 185, 1986, 1–17.
- [21] Pirie, A.; Singh, B. & Islam, K. Ultra-violet, visible, near-infrared, and mid-infrared diffuse reflectance spectroscopic techniques to predict several soil properties. *Aust J Soil Res*, 43, 2005, 713-721.
- [22] Maggio, R.M., Kaufman, T.S., Del Carlo, M. et al. Monitoring of fatty acid composition in virgin olive oil by Fourier transformed infrared spectroscopy coupled with partial least squares. *Food Chemistry*, 114, 2009, 1549–1554.
- [23] Rohman, A., & Che Man, Y.B. The use of Fourier transform mid infrared (FT-MIR) spectroscopy for detection and quantification of adulteration in virgin coconut oil. *Food Chemistry* 129, 2011, 583–588
- [24] Durante, C., Cocchi, M., Grandi, M., Marchetti, A., & Bro, R. Application of N-PLS to gas chromatographic and sensory data of traditional balsamic vinegars of Modena. *Chemometrics and Intelligent Laboratory Systems*, 83(1), 2006, 54–65.
- [25] Rezaei, B., Khayamian, T., & Mokhtari, A. Simultaneous determination of codeine and noscapine by flow-injection chemiluminescence method using N-PLS regression. *Journal of pharmaceutical and biomedical analysis*, 49(2), 2009, 234–239.
- [26] Mouazen, A.M., Baerdemaeker, J.D. and Ramon, H. Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 14 (3), 2006, 189-99.