Research Paper

Quantitative Detection of Cow Milk in Goat Milk by Chemometrics Analysis Based on Mid Infrared Spectroscopy

W.Terouzi[1], S. Omari[1], K. Boutoial[1], A.Oussama[1]

Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco

ABSTRACT: *The application of Attenuated Total Reflectance- Fourier Transform Mid Infrared Spectroscopy (ATR-FTMIR) associated with chemometric tools was evaluated as a rapid method for quantitative detection of milk adulteration. Blends of Goat milk with different percentages of Cow milk were measured using ATR-FTMIR spectroscopy. Spectral and reference data were firstly analyzed by principal component analysis (PCA). Partial least square regression (PLSR) was used to establish calibration model. Excellent correlation between ATR-FTIR analysis and studied milk blends was obtained $R^2 = 0.99$; with Root Mean Square Errors of Prediction < 1.41, Limit of Detection 4.202%, and Relative Prediction Errors as low as 0.0079. This result demonstrated the feasibility of ATR-FTMIR spectroscopy combined with chemometrics to quantify successfully binary mixtures of Goat milk in the 0–30 % weight ratio range of Cow milk with a reliable, rapid and inexpensive tool without the need for sample preparation.*
Keywords: Chemometrics, Cow milk, Goat milk, mid infrared spectroscopy, quantification.

I. INTRODUCTION

The authentication of agro-food products has become a crucial issue for regulatory agencies, consumers, food processors, and industries [1]. In fact, both intentional and accidental food adulteration may be considered a health risk in some cases. Several major food adulteration events were discovered. These events include the adulteration of protein in China (2007), where samples of wheat gluten were mixed with melamine to increase the protein content, the milk scandal in China (2009) , and in India (2012), where milk was adulterated with detergent, fat, and even urea [2], [3].

On the other side, Cows' milk proteins, even in low concentrations, are known to be among the most common food allergens, especially among children [4], [5]. Usually only the total elimination of cow milk from the diet prevents the allergy, but in the majority of cases the problem can be avoided by replacing it with the milk of some other species, as Soy milk and Goat milk. Therefore there is a growing demand for developing new, highly sensitive methods for species identification to prevent the risk of allergy.

According to literature, various analytical techniques have been tried and developed to ensure the quality of dairy products, and especially milk authenticity. As a tool for ensuring authenticity of milk, digital colour image analysis combined with chemometric methods has been successfully applied to detect adulterations in liquid milks [6] and discriminate adulterated milks from authenticated milks [7], [8]. Also, Near InfraRed Spectroscopy (NIRS), has been used in the authenticity of adulterated food [9], [10] and detection of the contents of adulterants in powdered or liquid milk [11], [12], [13]. Fourier Transform Mid Infrared (FTMIR) spectroscopy is a rapid biochemical fingerprinting technique [14]. It can be potentially applied to deliver results with the same accuracy and sensitivity as the reference methods in short time [15].

In this context, the aim of this current work was to explore the possibility of using ATR-FTMIR spectroscopy with chemometric tools for the quantitative detection of cow milk in Goat milk. In fact, 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 Cow milk in the adulterated milk samples.

*Corresponding Author: W.Terouzi[1]   1 | Page
Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology
Of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco*

II.   MATERIALS AND METHODS

2.1. Sampling

Samples of raw cow and goat milk were obtained from local farmers, beloging to Fkih ben Salah zone, in Beni Mellal-Khnifra region, Morocco. For the adulteration study, milk samples were prepared by mixing Goat milk (GM) with Cow milk (CM) in the range of 0–30%. The samples were analysed directly at ambient temperature.

There were 52 samples in total, among which 37 samples were randomly taken for establishing principal component analysis (PCA) and partial least square regression (PLSR) model. Other 15 samples were used for testing the reliability of the model.

2.2. Spectroscopic measurements

ATR-FTMIR spectra were obtained using a PerkinElmer spectrum, Version 10.5.1 equipped with an attenuated total reflectance accessory with DTGS detector, Globar (MIR) Source and KBr Germanium separator, with a resolution of 4 cm^{-1} at 60 scans. Spectra were scanned in the absorbance mode from 4000 to 450 cm^{-1} and the data were handled with PerkinElmer logiciel. The adulterated milk samples 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.

2.3. Data pre-processing procedures and software

In this study, a series of pre-processing elaborations were tested on the spectral data prior to the multivariate calibration. The Savitzky–Golay [16] and Norris gap [17] algorithms were tested for data derivatisation. Standard normal variate (SNV) and multiple scatter correction (MSC) [18] were also tested.

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).

2.4. Chemometrics analysis

2.4.1. Principal Component Analysis (PCA)

PCA, as a linear projection method is based on variance, transforms the original measurement variables into new uncorrelated variables called principal components [19], [20]. 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 [21].]. 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.

2.4.2. Partial least squares regression (PLSR)

Partial least squares regression (PLSR) [22] is popular and the most commonly used multivariate calibration chemometric methods. It is able to resolve overlapping spectral responses [23]. 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 [24].

In this work, the collected ATR-FTMIR spectra will be used as the X matrix, and the Cow milk compositions of the different adulterated milk samples will be used as the Y vector.

III.   RESULTS AND DISCUSSION

3.1. Spectral diversity

In the first step, ATR-Fourier transform mid infrared (ATR-FTMIR) spectra of pure Goat milk (GM) and Cow milk (Adult. CM) were obtained. One spectrum is the average of 60 scans of the same sample of milk on FT-MIR. The average spectra of all considered samples are presented in **Fig.1. (a)**

In the second step, ATR- FTMIR spectra of 52 samples of the adulterated milks were recorded in triplicate and a mean spectrum was calculated for studied samples. The resultant mean spectrum of binary mixtures (GM-CM) is shown in **Fig.1. (a)**
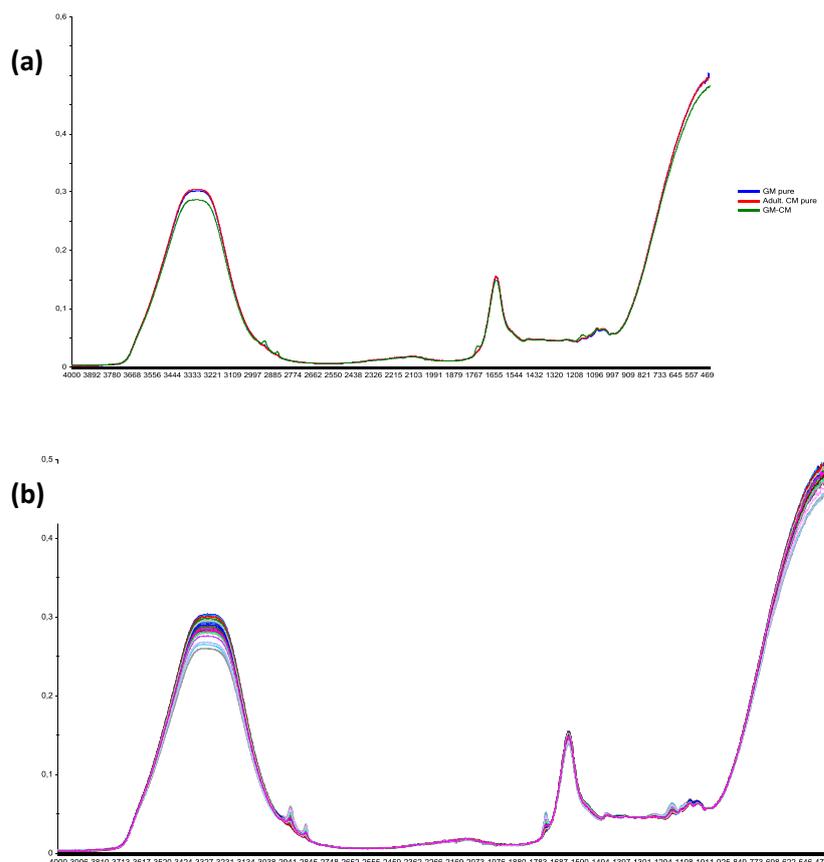
Fig.1. ATR-FTMIR spectra of: (a) Goat Milk (GM), Cow milk (Adult. CM) and binary mixtures GM-CM; (b) the binary mixture (Goat milk – Cow milk) samples of calibration set in the 0–30 % weight ratio range, at MIR region of 4000–450 cm^{-1}

ATR-FTIR spectra of 52 samples of the studied binary mixtures were recorded and divided in two sets: a calibration set of 37 samples and an external validation set of 15 samples. One spectrum is the average of 60 scans of the same sample of milk blend. The average spectra of all considered samples in calibration set are presented in Fig.1. (b).

In Fig.1, the obtained spectra are dominated by the significant bands of water are clearly visible in the studied milks spectra at 3400 cm^{-1}. The band of aromatic ring stretch of lignin should appear at 1604 cm^{-1}. However, this region was obscured by the strong water deformation band centered at 1638 cm^{-1}. The typical infrared pattern of sugar is observed in the region 1200 - 900 cm^{-1}. The two small bands at 2927 cm^{-1} and 2856 cm^{-1} are characteristic of fatty acids.

In fact, the main signals in the mid-IR region are in 1800–1500 cm^{-1}. There is a band at about 1680 cm^{-1} which is associated with the C=O stretching of proteins. On the other hand, C=O stretching band of amide I and N–H bending of amide II are both located in this spectral region [23].

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 milks (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, chemometric methods 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.

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 segment 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.

3.2. Statistical analysis
3.2.1.   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 [25]. **Fig.2** shows the score plot obtained by PCA model in calibration set of adulterated milks.
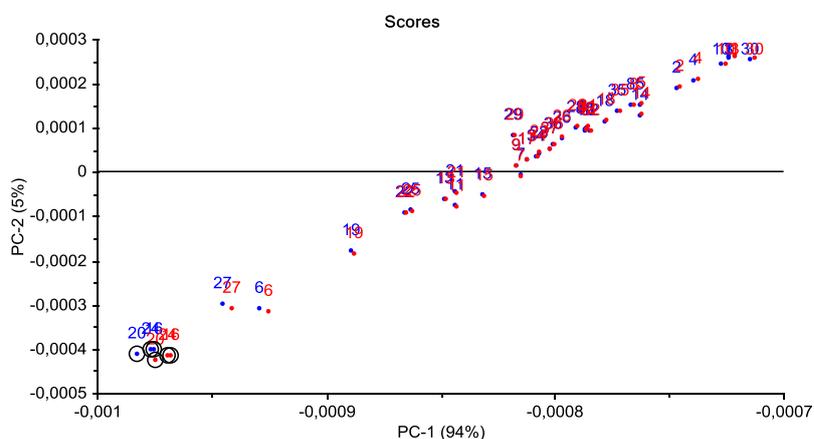
According to **Fig.2** of PCA score plot, the data set contained in total three spectral « outliers »: 16, 20 and 24.

To verify the true nature (outliers or extremes) of the three samples, we appealed to the Hotelling T² statistics plot (**Fig.3**).

(a)

(b)
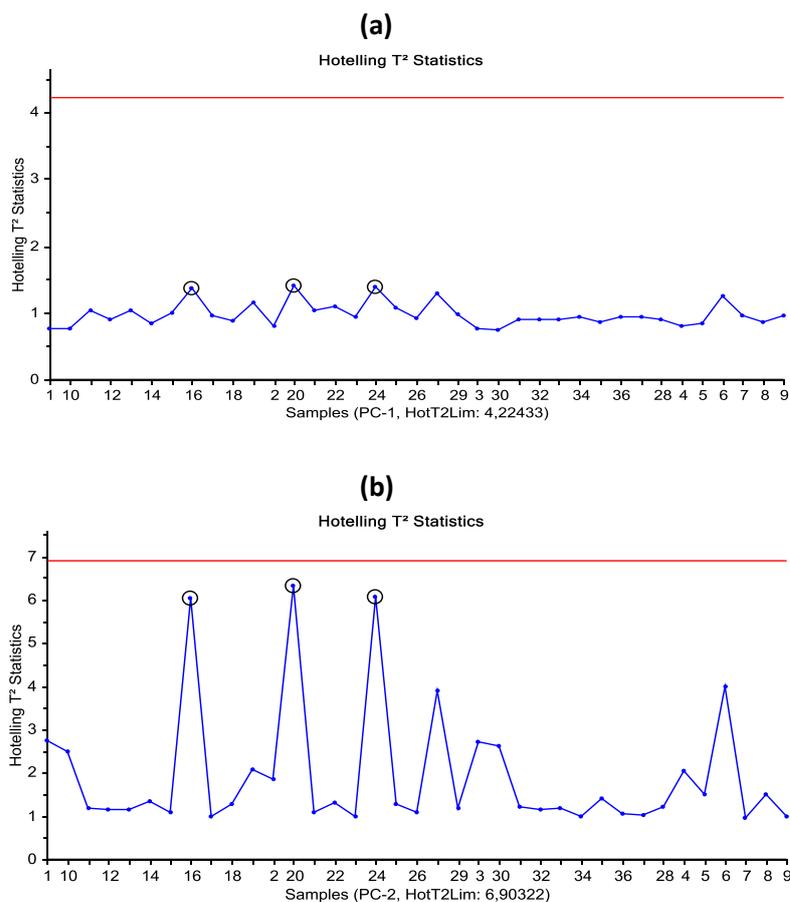
Fig.3. Hotelling T² statistics plot of calibration set in: (a) PC1; (b) PC2

---

*Corresponding Author: W.Terouzi[1]

Fig.3 is the Hotelling T² statistics plot. It is an alternative to plotting sample leverages. The plot displays the Hotelling T² statistic for each sample as a line plot. The associated critical limit (with a default p-value of 5%) is displayed as a red line.

The Hotelling T² statistic has a linear relationship to the leverage for a given sample. Its critical limit is based on an F-test. Use it to identify outliers or detect situations where a process is operating outside normal conditions.

According to **Fig.3** there's no « outliers » in the data set. Then, the prediction model (PLSR) was building with all samples of calibration set.

3.2.2. PLSR modeling

The quantification of Cow milk as adulterant in Goat milk was carried out using PLS algorithm. The PLSR model is built by considering the all spectra range 4000–450 cm^{-1} with X as variable and the Y variables is associated to different percentages of the adulterant. The data set contained 37 adulterated milk samples.

The PLSR model was evaluated using coefficient of determination (R^2) in calibration, root-mean-square error of calibration (RMSEC) and 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) [26], [27].

The resulting regression model seems to be able to predict the percentage of Cow milk, as adulerant in the adulterated milk samples (**Fig.4**).

The PLSR model is validated by full cross validation. This validation was also called "leave-one-out method". In this technique, one of the calibration samples is removed. The value for the removed sample was then predicted, and the prediction residual computed. The process was repeated with another sample of the data set, and so on, until every sample had been left out once [28].

The obtained statistical parameters RMSEC, RMSECV and R^2 are summarized in **Fig.4**. The coefficient of determination (R^2) of 0.99, RMSEC of 0.9432 and RMSECV of 1.1259, could be considered satisfactory.

Three VLs are necessary to have a good PLSR performance. **Table1** lists the explained variances from the developed model.
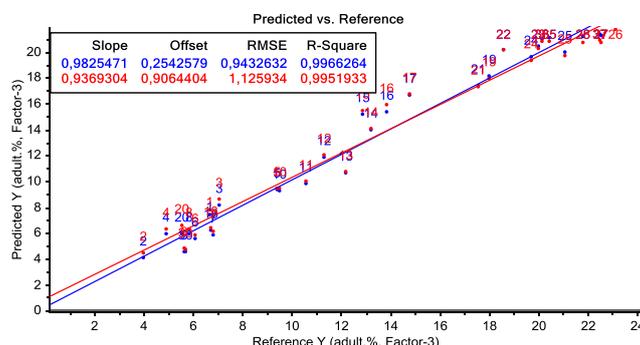
Fig.4. The relationship between actual and estimated percentages of adulterant in Goat milk, obtained from the final PLSR model developed from the ATR-FTMIR spectra

Table1. Explained variances (%) of LVs used in the PLSR model.

	Explained	Factor 1	Factor 2	Factor 3
PLSR	Calibration	81. 92378	98. 97132	99. 66264
	Validation	80. 88881	98. 45364	99.51933

3.2.2.1 Determination of Cow milk content in the new adulterated milk samples (External validation)

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

PLSR model is used to predict percentage of Cow milk 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 [29] 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 were applied to a group of external samples, the results are listed in **Table 2**. The deviations of prediction of Cow milk content in the blend samples by ATR-FTMIR spectroscopy were between 0.9266 and 1.4627, which were very satisfied (**Table2**).

Table2. Prediction results of Cow milk content (%) in the adulterated milk samples by ATR-FTMIR spectroscopy coupled with PLSR.

Samples of Validation	Predicted	Deviation	Reference
Val1	9.2970	1.0312	9.5000
Val2	11.8297	1.1011	11.3000
Val3	13.9995	1.1041	13.2100
Val4	16.6424	1.1798	14.8000
Val5	18.1076	1.2363	18.0200
Val6	4.1032	0.9386	2.1900
Val7	20.4000	0.9910	20.0100
Val8	19.9953	0.9266	21.0800
Val9	21.2434	1.4627	24.8500
Val10	24.9897	1.4627	22.0100
Val11	8.1631	1.0462	9.0000
Val12	4.5410	1.0212	5.4200
Val13	23.9638	0.9666	24.3000
Val14	20.9541	1.0570	21.4100
Val15	5.8881	0.9415	4.2000

Figures of merit of the calibration graphs are summarized in **Table3**. As can be seen, PLSR model offered good values for the different multivariate parameters.

Table3. Statistical parameters carried out by external validation on PLSR

External validation	LVs	Rp^2	RMSEP	Bias	SEP	REP %	RPD	LD%
	2	0.991	1.40078	0.03069	1.5906	0.0078	4.4624	4.202

IV. CONCLUSION

Quantitative analysis of food adulterants is an important for health, wealth and economic issue that needs to be fast, simple and reliable. In this work, we arrived to develop a new method based on ATR-FTMIR analysis associated with PLSR technique as a rapid, inexpensive and non destructive adulteration measuring tool, useful to determine the percentage of Cow milk in the binary mixture with Goat milk.

The PLSR model obtained from transformed infrared spectra gave correlation coefficients of 0.99 and root mean square errors of prediction (RMSEP) value of 1.40078. This result demonstrated that proposed method guarantee good prediction of the percentage of Cow milk, as adulterant in Goat milk without sample preparation. Then, it can be used in dairy industry for the reliable, cheap and fast quality control of raw material, ensuring a rapid authentication of final products to be commercialized.

REFERENCES

[1] Cotton, P. A., Subar, A. F., Friday, J. E., Cook, A., Dietary sources of nutrients among US adults, 1994 to 1996. J. Am. Diet. Assoc., 104, 2004, 921–930.
[2] Lakshmi, V. Food adulteration. International Journal of Science Inventions Today, 1(2), 2012, 101-113.
[3] Editorial. Melamine and food safety in China. The Lancet, 373, 2009, 353-353.
[4] Sampson, H.A. Food allergy. J. Allergy Clin. Immunol. 111, 2003, 540–547.
[5] Wal, J.M., 2004. Bovine milk allergenicity. Ann. Allergy Asthma Immunol. 93, 2004, 2-11.
[6] Santos, P. M. D., Wentzell, P. D., & Pereia-Filho, E. R. Scanner digital images combined with color parameters: A case study to detect adulterations in liquid Cow milk. Food Analytical Methods, 5, 2012, 89–95.
[7] Santos, P. M., Pereira-Filho, E. R., & Rodrigues-Saona, L. E. Rapid detection and quantification of milk adulteration using infrared microspectroscopy and chemometrics analysis. Food Chemistry, 138, 2013, 19–24.
[8] Santos, P. M., Pereira-Filho, E. R., & Rodrigues-Saona, L. E. Application of handheld and portable infrared spectrometers in bovine milk analysis. Journal of Agricultural and Food Chemistry, 61, 2013, 1205–1211.
[9] Downey, G., McIntyre, P., & Davies, A. N. Detecting and quantifying sunflower oil adulteration in extra virgin olive oils from the eastern mediterranean by visible and near-infrared spectroscopy. Journal of Agricultural and Food Chemistry, 50, 2002, 5520–5525.
[10] Downey, G., Fouratier, V., & Kelly, J. D. Detection of honey adulteration by addition of fructose and glucose using near infrared transflectance spectroscopy. Journal of Near Infrared Spectroscopy, 11, 2003, 447–456.
[11] Borin, A., Ferrao, M., Mello, C., Maretto, D., & Poppi, R. Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk. Analytica Chimica Acta, 579, 2006, 25–32.

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

[13]   Zhang, L-G., Zhang, X., Ni, L-J.,Xue, Z-B., Gu, X.,Huang, S-X. Rapid identification of adulterated cow milk by non-linear pattern recognition methods based on near infrared spectroscopy. Food Chemistry, 145, 2014, 342–348

[14]   Nicolaou, N., Xu, Y., Goodacre, R. Fourier transform infrared spectroscopy and multivariate analysis for the detection and quantification of different milk species. J. Dairy Sci., 93, 2010, 5651-5660.

[15]   Nicolaou, N., and Goodacre, R. Rapid and quantitative detection of the microbial spoilage in milk using Fourier transform infrared spectroscopy and chemometrics. Analyst, 133, 2008, 1424-1431.

[16]   Savitzky, A. & Golay, M.J.E. Smoothing and differentiation of data by simplified least-squares procedures. Analytical Chemistry, 36, 1964, 1627–1639.

[17]   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.

[18]   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.

[19]   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.

[20]   Wold, S., Esbensen, K. & Geladi, P. Principalcomponent analysis. Chemom. Intell. Lab. Syst., 2, 1987, 37-52.

[21]   Brown, S.D., Tauler, R., Walczak, B. Comprehensive Chemometrics: Chemical and Biochemical Data Analysis, vol. 2. Elsevier, Amsterdam, 2009.

[22]   Martens, H. & Naes, T. Multivariate calibration. Chichester: Wiley. 1989.

[23]   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.

[24]   Geladi, P. & Kowalski, B. (1986) Partial least-squares regression: a tutorial. Anal Chim Acta, 185, 1986, 1–17.

[25]   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.

[26]   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.

[27]   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.

[28]   Gurdeniz, G., Tokatli, F., & Ozen, B. Differentiation of mixtures of monovarietal olive oils by mid-infrared spectroscopy and chemometrics. European Journal of Lipid Science and Technology, 109, 2007, 1194−1202.

[29]   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.