International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-5, Issue-3, July 2016

Quantification of cow milk in soy milk using Attenuated Total Reflectance Fourier Transform Mid Infrared spectroscopy (ATR–FTMIR) with chemometrics

S. Omari, W. Terouzi, K. Boutoial, A. Oussama

***Abstract*— Soy Milk (SM) is an important and healthy substitute for people who are allergic to cow milk protein and lactose. The aim of this work was to propose a new method for the quantitative analysis of Cow milk "as adulterant" (Adult.) in a binary mixture with Soy milk by applying Attenuated Total Reflectance-Fourier Transform Mid InfraRed Spectroscopy (ATR-FTMIR) associated with chemometric methods. Blends of Soy milk with different percentages of Cow milk were measured using ATR-FTIR 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 R2 = 0.99; with Root Mean Square Errors of Prediction < 2.31, Limit of Detection 6.923%. This result demonstrated the feasibility of ATR-FTIR spectroscopy combined with chemometrics to quantify successfully binary mixtures of Soy milk in the 0–40 % weight ratio range of Cow milk with a reliable, rapid and inexpensive tool without the need for sample preparation.**

***Index Terms*— Chemometric methods, Mid Infrared Spectroscopy, Quantification, Soy milk.**

I. INTRODUCTION

Milk and dairy product consumption is recommended by most nutritional experts because of their beneficial effects or calcium uptake and bone mineralization and as a source of valuable protein [1], [2].

Soymilk (SM) is often used as an alternate of dairy milk due to quite similar protein as of cow milk, except sulphur containing amino acids, in which SM is deficient [3].

In fact, soy milk consumption has been increasing in Morocco who imports this vegetable product for diet and persons are allergic to cow milk protein and lactose. However, in spite of its nutritional merits, it has not gained much popularity due to its flavor and its higher prices compared to cow milk.

Additionally, the authenticity of raw materials and food products presents a huge importance for regulatory agencies, consumers, food processors, and industries, in order to satisfy food quality and safety requirements [4], [5]. In this case, it is extremely important to develop an effective, convenient and quick method to detect and authenticate of milk products.

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 chemometrics 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 objective of this current study was to explore the possibility of using ATR-FTMIR spectroscopy with chemometric tools for the detection and quantitative prediction of cow's milk in Soy milk.

II. MATERIALS AND METHODS

A. Samples preparation

Soy milk and Cow milk were purchased in a local supermarket. For the adulteration study, milk samples were prepared by mixing Soy milk (SM) with Cow milk in the range of 0–40%. The samples were analysed directly at ambient temperature.

There were 45 samples in total, among which 30 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.

B. ATR-FTIR analysis

ATR-FTIR spectra were obtained using a PerkinElmer spectrum, Version 10.5.1 equipped with an attenuated total

S. OMARI, Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco, (212) 0668341303.

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.

K. BOUTOIAL, Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco, (212) 0624864101.

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.

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.

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 [16] and Norris gap [17] algorithms were tested for data derivatisation. Standard normal variate (SNV) and multiple scatter correction (MSC) [18] 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 segment 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 tools

• *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 [19], [20]. 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 [21].

• *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 study, the collected MIR spectra will be used as the X matrix, and the Cow milk compositions of the different milk samples will be used as the Y vector.

• *Software*

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

III. RESULTS AND DISCUSSION

A. ATR-FTMIR spectral analysis

In the first step, ATR-Fourier transform mid infrared (ATR-FTMIR) spectra of pure Soy milk (SM) and Cow milk (Adult.) 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**.

In the second step, ATR- FTMIR spectra of 45 samples of the adulterated milk were recorded in triplicate and a mean spectrum was calculated for studied samples. The resultant mean spectrum of binary mixtures (SM-Adult.) is shown in **Fig.1**.

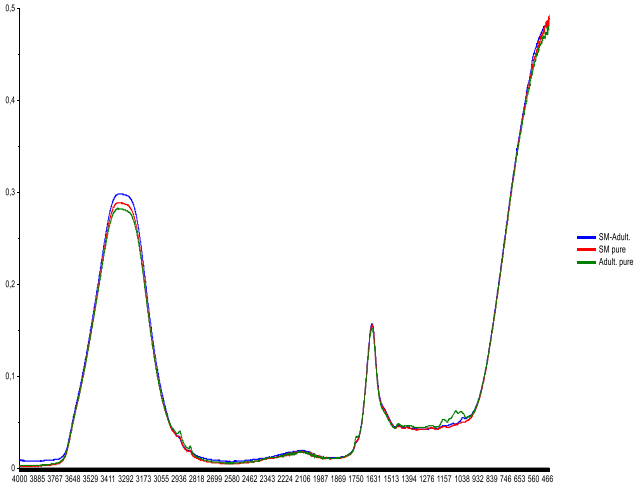

Fig.1. ATR-FTMIR spectra of Soy Milk (SM), Cow milk (Adult.) and binary mixtures SM-Adult. at MIR region of 4000–450 cm^{-1}

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 [25].

In fact, 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 theATR-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. Best results were obtained by fixing the following parameters: 2nd order, gap size 17 and segment size 15, with centered data.

B. Statistical 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 [26]. **Fig.2** shows the score plot obtained by PCA model in calibration set of adulterated milks.

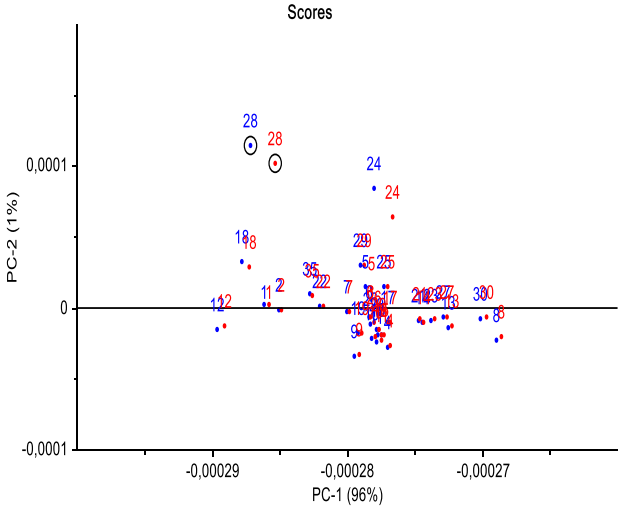

Fig.2. PC1 / PC2 Score plot by PCA analysis on the calibration set of binary mixtures (Soy milk- Cow milk) samples

According to **Fig.2** of PCA score plot, the data set contained one spectral « outlier » (28). 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*

The quantification of Cow milk as adulterant in Soy 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 30 milk samples including the spectral of number 28, the « outlier » sample identified by PCA (**Fig.2**) because it is considered extreme by PLS.

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) [27], [28].

The resulting regression model seems to be able to predict the percentage of Cow milk, as adulerant in the milk samples (**Fig.3**).

The PLSR model is validated by full cross validation. The obtained statistical parameters RMSEC, RMSECV and R^2 are summarized in **Fig.3**. The coefficient of determination (R^2) of 0.99, RMSEC of 1.32 and RMSECV of 2.267, could be considered satisfactory.

Four VLs are necessary to have a good PLSR performance. **Table1** lists the explained variances from the developed model.

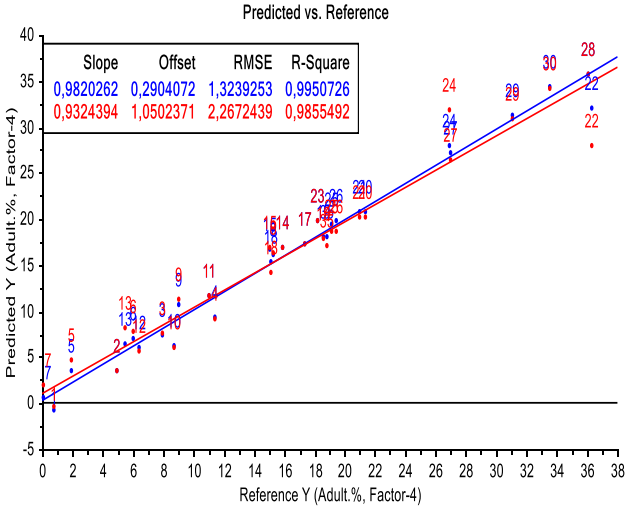

Fig.3.The relationship between actual and estimated percentages of adulterant in Soy milk, obtained from the final PLSR model developed from the ATR-FTMIR spectra

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

Explained	LV 1	LV2	LV3	LV4
Calibration	72.17697	93. 20123	98.11885	99. 50726
Validation	70.11115	90. 75676	96. 56287	98. 55492

•*Prediction of Cow milk percentage 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 Soy 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 (0-40%). 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 PLSR model was applied to a group of external samples (15 samples), the results are shown in **Fig.4**.

Fig.4 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 Cow milk percentages obtained from ATR-FTMIR spectra. The difference between the actual and the predicted percentage is relatively small.

Figures of merit of the calibration graphs are summarized in **Table2**. As can be seen, PLSR model offered good values for the different multivariate parameters.

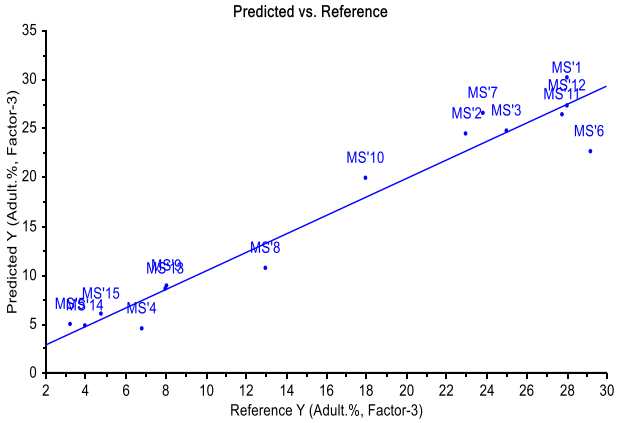

Fig.4. Measured vs. Predicted values for Cow milk in binary mixtures Soy milk-Cow milk of external validation set.

Table2. Statistical parameters carried out by external validation on PLSR

LVs	Rp^2	RMSEP	Bias	SEP	LD%
3	0.9858	2.3078	-0.00238	2.3888	6. 923

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 study, 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 Soy 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 2.3078. This result demonstrated that proposed method guarantee good prediction of the percentage of Cow milk, as adulterant in Soy 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] Melnik, B. C., Milk–the promoter of chronic Western diseases. Med. Hypotheses 2009, 72, 631–639.

[3] Chaiwanon, P., Puwastein, P., Nitithamyong, A., & Sirichakwal, P. P. Calcium fortification in Soybean Milk and In Vitro Bioavailability. Journal of Food Composition and Analysis, 13, 2000, 319-327.

[4] Singhal, R.S., Kulkarni, P.R., Rege, D.V. Handbook of Indices of Food Quality and Authenticity, Woodhead Publishing Ltd., Cambridge, UK, 1997.

[5] Lees, M. (Ed.), Food Authenticity and Traceability, Woodhead Publishing Ltd., Cambridge, UK, 2003.

[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] 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.

[26] 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.

[27] 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.

[28] 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.