J. Mater. Environ. Sci. 7 (8) (2016) 2689-2698
ISSN : 2028-2508
CODEN: JMESC

Detection of argan oil adulteration with olive oil using fluorescence spectroscopy and chemometrics tools

S. Addou, F. Fethi*, M. Chikri, A. Rrhioua
Laboratoire de Physique, de la Matière et de Rayonnements (LPMR).
Département de Physique, Université Mohamed Premier, Faculté des Sciences. Oujda, Maroc

Received 27 Nov 2015, Revised 17 Mar 2016, Accepted 23 Mar 2016
**Corresponding author. E-mail: fethi.fouad@yahoo.fr (F. Fethi); phone: +212676882599 /+212662442776*

Abstract
The Laser Induced Fluorescence technique for determining the adulteration of argan oil by olive oil was investigated. The data were analyzed without any preprocessing using multivariate calibration methods of Principal Component Analysis and Partial Least Square regression. Fluorescence spectra of different samples were measured using a laser beam of 532 nm wavelength. Spectral data are analyzed to build a best calibration model for predicting the percentage of olive oil added to pure argan oil. A model with a regression coefficient of r^2=0.992, a Standard Error of Prediction SEP=1.311 and bias=0.31 for adulteration has shown better prediction. A sensitive wavelength corresponding to the adulteration of argan oil by olive oil was proposed on the basis of loading by Principal Component Analysis. The results show that Laser Induced Fluorescence is a better, non-destructive and rapid optical method for analyzing argan oil adulterated by olive oil. The sensitivity detection of adulteration is possible starting from 0.43% olive oil mixed with 99.57% argan oil (w/w). The results of multivariate analysis reveal that argan oil contains enough information to detect doped argan.

Keywords: Authentication, Argan oil, Olive oil, Fluorescence spectroscopy, Chemometrics

1. Introduction
Argan oil is a plant oil produced from the kernels contained in the fruit of the argan tree (Argania spinosa) [1]. It is rich in vitamin E and essential fatty acids; also used for nutritive and for it's cosmetic properties. The tree is endemic to southwestern Morocco, mainly covers the provinces of Essaouira, Agadir, Chtouka Ait Baha, Tiznit and Taroudant [2]. It is protected by UNESCO and is extremely well adapted to drought and other harsh environmental typical conditions of southwestern Morocco. Its production is mainly run by berber women; it is an oleic linoleic oil [3] which has an original unsaponifiable composition [4], probably beneficial as a health product [5-6]. The brown-colored mash emits a pure characteristic odor and taste of hazelnut which makes it's a popular gourmet oil. At 20 °C (68 °F) it has a relative density ranging from 0.906 to 0.919 [7]. The price of argan oil for example, is high compared to other vegetable oils (soybean, sunflower, corn) and olive oil which makes it prone to temptation to falsify [8]. Depending on the extraction method, argan oil may be more resistant to oxidation than olive oil [9]. The use of vegetal oils is increasingly popular around the world which implies a wide production. However a large amount of oils is exposed to adulteration (doping). Adulteration is to substitute a product that has undergone the addition or deletion of a compound, by the original one. In order to control the authenticity of argan oils several studies have been carried out. They have as aim the detection of oil adulteration by different techniques such as Infrared spectroscopy Fourier Transform (FTIR) [10-13] and Raman spectroscopy [9], Near Infrared Spectroscopy (NIRS) [14-17], High-Performance Liquid Chromatography (HPLC) [18, 19], Inductively Coupled Plasma (ICP) optical emission spectrometry [20] and dielectric spectroscopy [21].

Because of its simplicity and rapidity, Fluorescence spectroscopy is an analytical tool increasingly used in various fields of food analysis [22-26]. It allows the detection of adulteration, for example, the addition of sunflower oil in virgin olive oil [25] and the addition of butter in original cocoa butter [27-29], through multivariate analysis methods.
In this task, the adulteration of argan oil with olive oil is studied by (LIF) combined with chemometric tools: the Principal Component Analysis (PCA) and Partial Least Square (PLS). This method is fast, efficient and robust for non-destructive analysis. The originality of this work is to detect the signal of chlorophyll fluorescence which is present in olive oil, and completely absent in argan oil. The objective of this work was to study the authenticity of pure argan oil using Laser Induced Fluorescence (LIF) and to demonstrate application of this technique, as a rapid analysis method to determine the quality of oils.

2. Chemometrics

Chemometric is the use of mathematical or statistical methods for the treatment of chemical data with various objectives: The description of data under a synthetic shape, modeling, classification, experimental design and optimization. Exploratory methods correspond to the objective data description and include both elementary methods such as the calculation of the mean and standard deviation of data, and more sophisticated methods, such as (PCA) and (PLS) regression for example [30-31].

2.1. *Principal Component Analysis*

The Principal Component Analysis is a statistical projection method of multidimensional data, used to reduce the dimensionality of data. The purpose of the (PCA) is to condense the original data into new groups called new components, so that they no longer pose a correlation between them and are ordered in terms of percentage of variance provided by each component. Therefore, the first new component contains information about the maximum percentage of variance; the second contains information about the following percentage variance. The process is repeated until the obtention of the last new component. There is a limited number of principal components that can be derived from the data in the X-matrix. The largest number of components is either n (number of objects) or p (number of variables), depending on which is the smaller [31-34].

2.2. *Partial least square regression*

Partial Least Squares regression (PLS) is a statistical method that bears some relation to principal components regression; instead of finding hyperplanes of minimum variance between the response and independent variables, it finds a linear regression model by projecting the predicted variables and the observable variables to a new space. (PLS) is a multivariate analysis technique used in spectroscopy. It defines a linear relationship between two variables in a single phenomenon, it is applied when two continuous quantitative variables appear linked to each other by a linear relationship. A linear relationship between x and y (the variables) is of the form $y = \mathbf{a}x + \mathbf{b}$ where **a** and **b** are real numbers. The goal of regression is to explain the variations of values and one or more dependent variables (the "y") by the values and variations of one or more variables ("x"). This regression has a simple goal: to find the coefficients for each explanatory variable, which minimizes the difference for the dependent variable between the values estimated by the model and the observed values in practice for a given sample which is regressed. This is to minimize the sum of residuals (squared, simply to avoid the systematic compensation of positive and negative errors), or, in other words, to maximize the correlation coefficient. The (PLS) calibration model was developed by UNSCRAMBLER 7.6 software (CAMO, Oslo, Norway). The quality of the calibration model is usually estimated from the correlation (r^2), the Standard Error of Calibration (SEC) and the Standard Error of Prediction (SEP). Once the model has been established, it can be applied to unknown samples. The accuracy of the regression can be identified by samples that have not involved in the calibration by the square root of the quadratic error: calibration, validation and prediction (RMSEC) and (RMSEP). Models are acceptable, if the correlation coefficient (r^2) is near 1 and the (SEC), (RMSEC), (SEP) and (RMSEP) errors are close to the low values [31, 35-37].

The dominant Partial Least Square (PLS) application in spectroscopy is an indirect measurement and calibration. Hence, the aim is to replace costly reference measurements with predictions of fast and inexpensive spectroscopic measurements.

2.2.1. *Calculation of the calibration variance*

In the first step, we make a model based on X_{cal} and Y_{cal}. Then we feed the X_{cal} values back into the model to predict $\hat{Y}_{cal}$. Comparing the predicted and measured Y_{cal} values gives us an expression of the modelling error (i.e. $\hat{Y}_{cal}$ - Y_{cal}). This can be calculated for each sample. Summing the square differences and taking their mean value gives the calibration residual Y -variance:

$$\text{Residual variance}_{cal} = \frac{\sum(\hat{Y}_{cal} - Y_{cal})^2}{n} \qquad (1)$$

The square root of this equation (divided by the weights used for scaling at calibration) gives us RMSEC (Root Mean Square Error of Calibration). The modeling error expressed in original units is given by

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_{i,cal} - Y_{cal})^2}{n}} \qquad (2)$$

2.2.2. *Calculation of the validation variance*

When a model is established, both the calibration and validation variance are automatically calculated at each model step. The program put the test set (X_{val}) into the model and predicts $\hat{Y}_{val}$ for which X_{val} or Y_{val} of the set test are not involved in the calibration. Then we compare the predicted and measured Y_{val} values to get an expression of the prediction error:

$$\text{Prediction error} = \hat{Y}_{val} - Y_{val} \qquad (3)$$

By summing the squared differences and taking their mean value, we calculated the residual Y-variance such as:

$$\text{Residual variance}_{val} = \frac{\sum(\hat{Y}_{val} - Y_{val})^2}{n} \qquad (4)$$

The square root of this expression divided by the weights used for scaling in the calibration gives us RMSEP (Root Mean Square Error of Prediction).

$$RMSEP = \sqrt{\frac{\sum(\hat{Y}_{val} - Y_{val})^2}{n}} \qquad (5)$$

The Standard Error of Performance (SEP) and Bias are two measurements closely connected to (RMSEP). Bias is the averaged difference between predicted and measured Y-values for all samples in the validation set.

$$Bias = \frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)}{n} \qquad (6)$$

Bias is also used to check if there is a systematic difference between the average values of the training set and the validation set. If there is no such difference, the Bias will be zero. The (SEP), on the other hand expresses the precision of the results, corrected for the Bias:

$$SEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i - Bias)^2}{n-1}} \qquad (7)$$

3. Materials and method

3.1. *Samples*

For this study, pure argan oil samples were prepared in the Agadir region (southwest of Morocco). Oil extraction was performed using the traditional method, from which we have prepared 56 samples adulterated by olive oil in different percentages, varying from pure argan oil (sample N°1) to (sample N°56), which is a mixture of 31.8% olive oil and 68.2% of argan oil (w/w). The samples were stored in a dark room at ambient temperature until the day of analysis. The calibration model was prepared with 37 samples in order to determin the level of olive oil in the range from 0% to 27% of olive oil in argan oil. The validation of the model was performed with the same calibration samples and the prediction was performed with others 19 samples in the range from 2% to 26%.

3.2. *Measurement method*

3.2.1. *Instrumentation*

Laser-induced fluorescence (LIF) spectra are obtained by irradiating the sample by a continuous laser beam (YAG) (Yttrium Aluminium Garnet) (532 nm) (figure 1), which produces an output power of 100 mW. The recipient holding the sample was placed at 12 cm from the output laser. A multimode fiber (SMA905) with a diameter core of 400 µm, a length of 2 m and a numerical aperture of 0.22 is used to guide the fluorescence light to the detector spectrometer. The optic fiber was positioned at right angle from the incidence light in order to collect only the induced fluorescence and to have a minimum noise. The signal was then recorded with an optical spectrum analyzer (AVS-USB2000, Netherland) for analyzing and displaying the fluorescence in 500-1000 nm range with the Avantes software. All measurements were made in the dark and at room temperature with time integrating 100 ms with an average spectrum equal to 1.

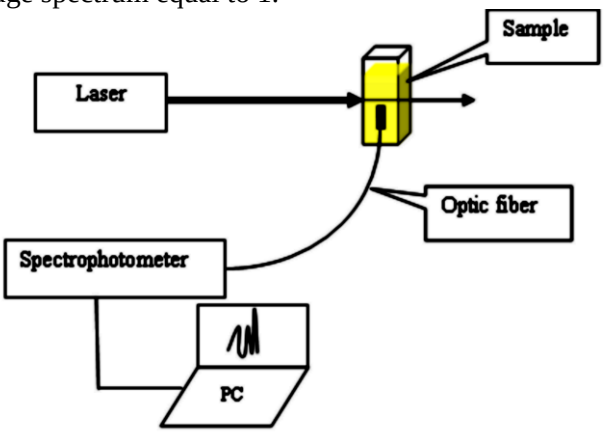

Figure 1: Schematic diagram of the induced fluorescence system.

4. Results and Discussion

Figure 2 shows the laser-induced fluorescence spectra of two samples in the spectral range 520-1100 nm. The (AO) spectrum is pure argan oil's fluorescence which exhibits a maximum reading in the yellow-orange region 570 to 630 nm, the (OO) spectrum corresponds to a pure olive oil sample with two bands as well as the red and far-red chlorophyll fluorescence near 675 nm and 730 nm considered a weak shoulder. We notice a net difference between these spectra situated around the wavelengths 675 nm [38, 39] and 720 nm which belongs to the chemical bond C=O in the olive oil [25]. The comparison between these spectra shows a high concentration of the chlorophyll in olive oil, while it is completely absent in argan oil. This large difference in bands spectral is very important in order to detect very low concentrations of olive oil added to argan oil.

The following spectra (figure 3) show different fluorescence of the argan and olive oils mixture. One can observe that the intensity of chlorophyll fluorescence in the spectral range 656-725 nm increases with increasing olive oil's concentration in the mixture. The spectrum with high peak at 675 nm corresponds to a sample mixture of 30% olive oil and 70% of argan oil (w/w), where the fluorescence signal of chlorophyll is around 230 arbitrary units (a. u). On the other side the fluorescence intensity in this wavelength of pure argan oil is too low.

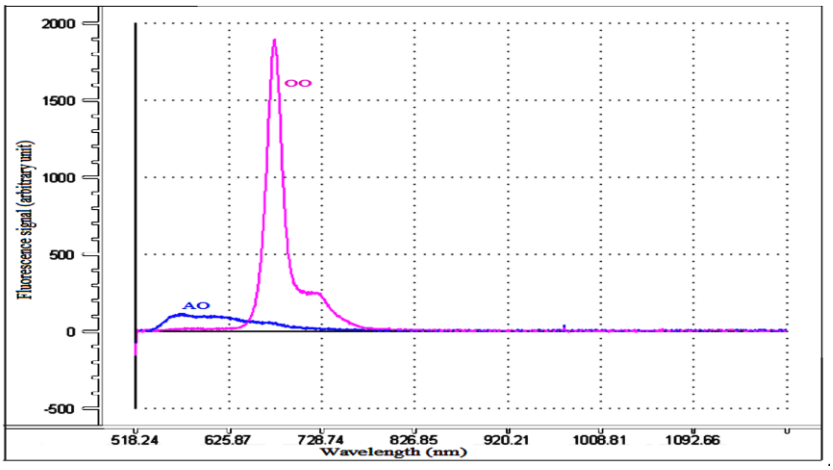

Figure 2: Laser induced fluorescence spectra of pure argan oil (AO) and pure olive oil (OO)

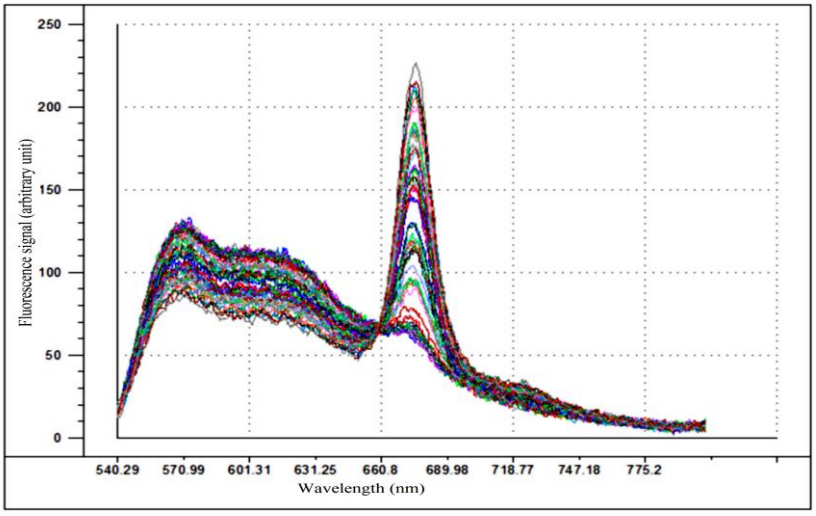

Figure 3: Green light-radiation induced fluorescence emission spectrum of argan oil mixed with olive oil in different percentages (λ_{ex}=532 nm). The fluorescence bands are in the green-orange (λ=570 nm), red (λ=675 nm and far-red (λ=740 nm) regions.

4.1. *Analysis method*

In the first step, the set of fluorescence spectra are analyzed by principal components analysis (PCA) without any pretreatment, since the spectra have very low noise and no offset between them. This analysis can significantly reduce the dimensions of data spectra by using the principal components (PCs), the induced fluorescence spectra can then be represented by several main PCs.

4.1.1. *Score plots and interpretation*

The score plot "map of samples" is one of the most powerful tools that "Principal Component" based methods can offer us. A score plot is simply score vectors plotted against each other figures 4 (a) and (b). They show the bidimensional and tridimensional score respectively. Plotting the score vectors corresponds to plotting the samples in PC-space. The PC-space may be multidimensional; therefore it can't always to be fully visualized in one plot. The most commonly used plot in multivariate data analysis is the score vector for PC1 versus the score for PC2. These are the two directions along which the data swarm exhibits the largest and the second largest "spread". In figure 4(a), scores for PC1 are along the "x-axis" and the scores for PC2 are along the "y-axis". The first one PC1 contains 97% of information of samples repartition in space; however the second vector does not

contain any information. The rest of the information is divided into other vectors such as: PC3, PC4 ... but the most important information is carried on the first vectors. Note that in the PC1 axis, the samples are classified in ascending order from left to right. This preferred direction is the direction of the increasing percentage of added olive oil in argan oil. The maximum variance of addition olive oil in argan oil is carried by the PC1 vector and all information 97% is distributed along PC1 direction.

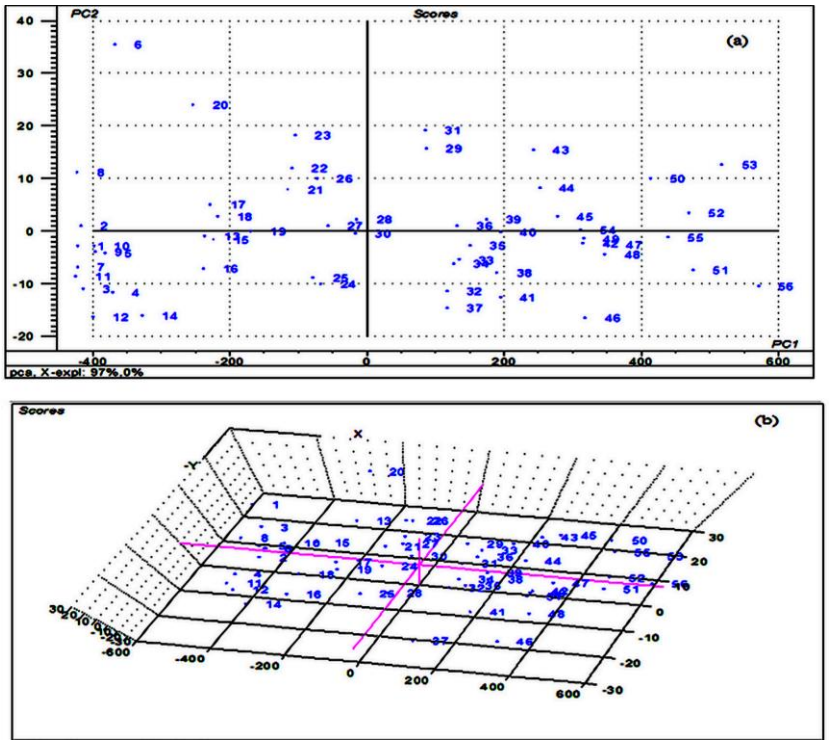

Figure 4: (a) Bi-dimensional score plot of PC1 and PC2 vectors. (b) Tri-dimensional score plot of PC1, PC2 and PC3 vectors of 56 samples of argan oil adulterated by olive oil

4.1.2. *Loading plot*

Figure 5(a) shows the loading plot for PC1 vector, it is referred to as loading spectra. Large loading imply wavelengths in which there is significant absorption related to the constituent of interest. This is a source of information to help us understanding the chemistry of the samples. It looks like the original spectra in wavelength range 656-748 nm. The one dimensional loading plot is very good for the assignment of spectral bands. As can be seen from figure 5(a), such a loading plot shows great similarity to a spectrum of figure 2. This loading plot is obtained from a data set consisting of induced fluorescence spectra of a system with 56 samples. The loading in this figure belongs to vector PC1 that can be related to the presence of chlorophyll in the sample. For comparison between figure 3 (original spectrum) and loading plot figure 5(a), we can see that the spectral features in the lower wavelength region can be recognized in the loading plot. The loading plot shows large loadings (97%), i.e. important variables in the lower wavelength region. Since the chlorophyll molecule emits in this region, we conclude that PC1 provides a model of chlorophyll presence which is present in the olive oil only. This result confirms that the wavelengths have a major influence to distribute samples in the vector space of PC1. We can discern that the most striking change is around the wavelength 675 nm. It represents the amount of olive oil or chlorophyll in the sample. On the other side, figure 5(b) shows that PC2 and PC3 loading and certainly other PCs oscillate around zero, which confirms that PC1 contains all information about the samples dispersion. We can conclude that PC1 describes the most important structural variation in the data set but not for other components of PCs.

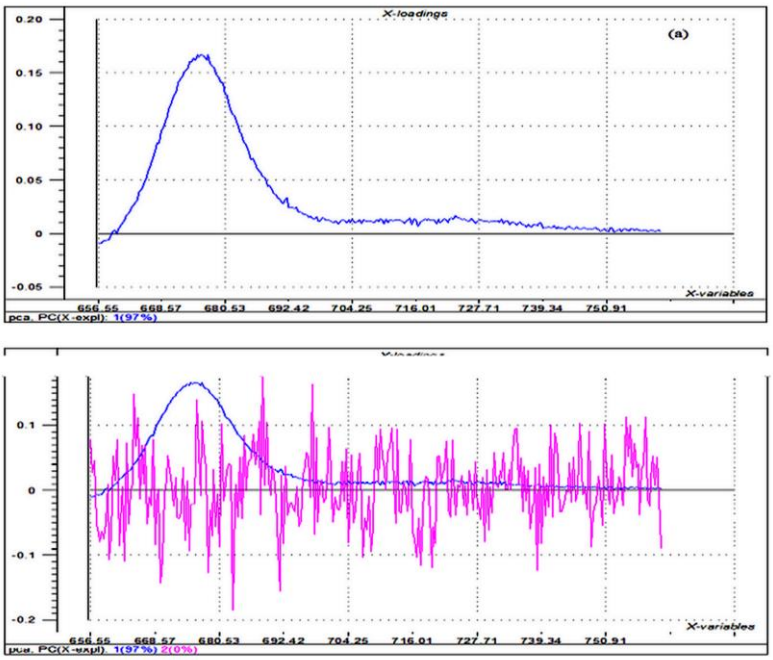

Figure 5: (a) One-dimensional loading plot for PC1, (b) Loading plot for PC1 and PC2 from PCA of data set of fluorescence

4.1.3. *Calibration model*

The results from the discriminate analysis were further confirmed and assessed using PLS algorithm. The validity of the resulting and final calibration model built on the calibration set of 37 samples. RMSEC is the corresponding measure for the fit model, calculated from the calibration samples. In the competing we would like RMSEC to be as small as possible. In our case REMSEC = 0.396 is lower (figure 6).

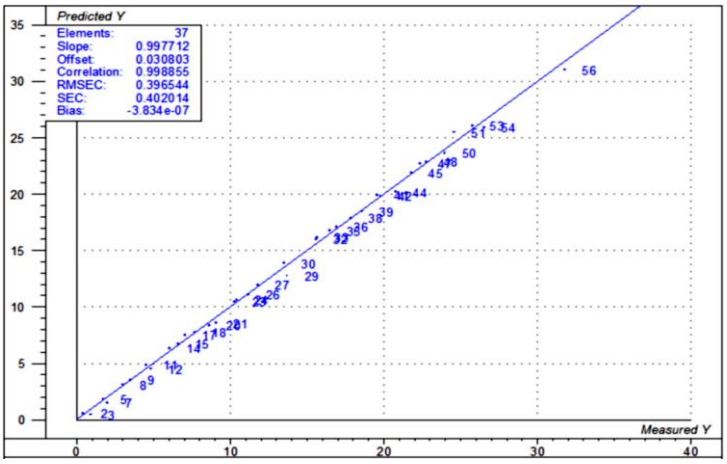

Figure 6: Calibration curve obtained with 37 samples

In this work, we used Partial Least Squares (PLS) regression to predict adulteration of 19 samples of argan oil with olive oil. The best (PLS) calibration model for determining the amount of olive oil in argan oil has been developed and optimized in the spectral range 656-725 nm. The selection of this band is explained by the presence of chlorophyll pigments located in this spectral region. Good correlation r^2 = 0.99 and good linearity were obtained from the values calibration by Laser Induced Fluorescence (figure 6).

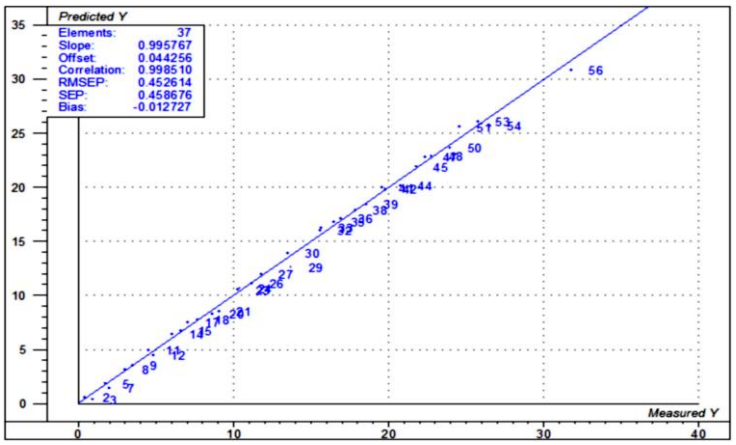

Figure 7: Validation curve obtained with 37 samples

The validation test (figure 7) shows a better correlation coefficient and a slope right around a ≈ 1 of a linear equation in the form y = ax. The RMSEP and SEP errors are small: 0.452 and 0.458 respectively and are the same order of magnitude (table 1).

Table 1: Result of calibration and validation LIF

		Calibration			Validation		
Adulteration	n	RMSEC	SEC	r^2	RMSEP	SEP	r^2
Olive oil (0-31.8 %)*	37	0.302	0.308	0.99	0.351	0.357	0.99

n: number of samples
*The percentages of olive oil (OO) in argan oil (AO) vary between 0 and 31.8%.
RMSEC: Root Mean Square Error of Calibration
SEC: Standard Error of Calibration
$r^{2:}$ Correlation coefficient
RMSEP: Root Mean Square Error of Prediction
SEP: Standard Error of Prediction

Figure 8 shows the prediction of the olive oil rate in argan oil. We find that the model is ideal since the correlation $r^2 = 0.992$ and RMSEP and SEP errors are almost equal 1.314 and 1.311 respectively.

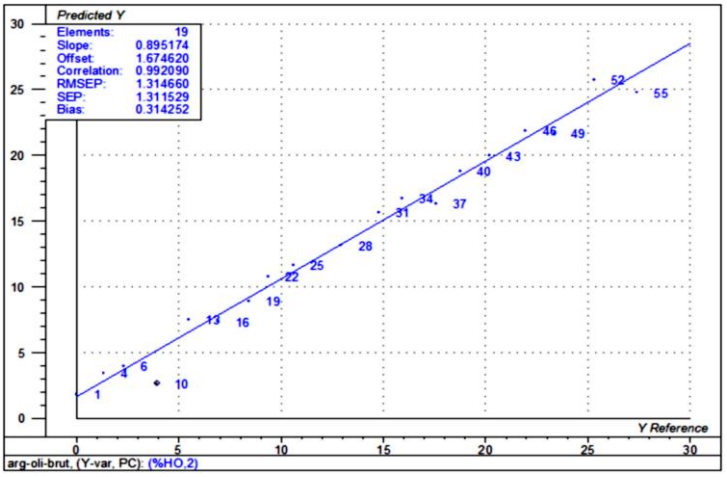

Figure 8: Prediction curve obtained with 19 samples

Conclusion

Laser induced fluorescence spectroscopy used in this original work has been very successful at evaluating the adulteration of argan oil by olive oil. This method is combined with Chemometrics tools by analyzing the data set without any preprocessing. In the first step, the principal component analysis (PCA) is applied to show the existence of spectral differences and to discriminate spectral data in relation with the adulteration of argan oil by olive oil. In the second step, a partial least squares (PLS) model has been established to predict the percentage of olive oil as adulterant in argan oil in the calibration range between 0% and 31.8% (w/w) with good prediction performances. The statistical results obtained are: r^2=0.992, REMSEP= 1.314 and SEP= 1.311. This study provided valuable results that could be applied to consumer protection by fraud, because the demand for high quality and safety in food production obviously calls for high standards for quality and process control. The laser induced fluorescence spectroscopy is highly desirable for analysis of food components because it often requires minimal or no sample preparation, efficient, non-destructive, provides rapid and on-line analysis, and has the potential to run multiple tests on a single sample. It seems to be very successful at evaluating the food quality.

References

1. Charrouf Z., Guillaume D., Ethnoeconomical, ethnomedical and phytochemical study of spinosa (L.) Skeels, *J. Ethnopharmacol.* 67 (1999) 7–14.
2. Farines M., Soulier J., Charrouf M., Soulier R., Étude de l'huile des graines d'Argania spinosa (L.) Sapotaceae. I) la fraction glycéridique. *Rev. Fr. Corps Gras [ZDB]*. 31 (1984) 283-6.
3. Charrouf Z., Valorisation des produits de l'arganier pour une gestion durable des zones arides du sud-ouest marocain, Collin, G. et Garneau. *F.-X.* (dir.) (1999) 195-209.
4. Farines M., Soulier J., *Rev. Fr. Corps Gras [ZDB]*. 31 (1984) 443-8.
5. Bennani H., Fiet J., Adlounia A., Impact de l'huile d'argan sur le cancer de la prostate : étude de l'effet antiprolifératif des polyphénols, *Revue Francophone des Laboratoires* 2009 Supplément au N°416, 23-26.
6. El Monfaloutia H., Guillaume D., Denhez C., Charrouf Z., *J. Pharm. Pharmacol.* 62 (2010) 1669-1675.
7. Charrouf Z., Guillaume D., *Eur. J. Lipid Sci. Technol.* 110(7) (2008) 632.
8. Oussama A., Elabadi F., Devos O., *Spectrosc Lett*. (2012) 11-15.
9. Yang H., Irudayaraj J., *J. Am. Oil Chem. Soc*. 78(9) (2001) 889-895.
10. Vardin H., Tay A., Ozen B., Mauer L., *Food Chem*. 108(2) (2008) 742–748.
11. Setiowaty G., Che Man Y B., *J. Am. Oil Chem. Soc. (JAOCS)*. 79(11) (2002) 1081-1084.
12. Azizian H., Karmer John K C., *Lipids.* 40(8) (2005) 855-867.
13. Isail A. A., Van de Voort F R., Emo G., Sedman J., *J. Am. Oil Chem. Soc.* 70(40) (1993) 335-341.
14. Özdemir D., Öztürk B L., *J. Food Drug Anal*. 15(1) (2007) 40-47.
15. Kasemsumran S., Kang N., Christy A., Ozaki Y., *Spectrosc Lett. 38(6) (2005) 839-851.*
16. Wesley I.J., Barnes R.J, McGill A.E.J, *JAOCS*. 72(3) (1995) 289-292.
17. Bellon V., Vigneau J. L., Leclercq M., *Appl. Spectrosc*. 47(7) (1993) 1079-1083.
18. Priego Capote F., Ruiz Jiménez J., Luque de Castro M. D., *Anal Bioanal Chem*. 388 (2007) 1859-1865.
19. Salghi R., Armbruster W., Schwack W., *Food Chem.* (2014) 387–392.
20. Gonzálvez A., Armenta S., de la Guardia M., *Food Chem.* 121 (2010), 878–886.
21. Hu Lizhi. Toyoda K., Ihara I., *J. Food Eng*. 96 (2010) 167–171.
22. Yongjiang Dong, Xuan liu, Chao Feng, *J. Food Eng.* 137 (2014), 95-100.
23. Baunsgaard A., Munck L., Norgaard L., *Appl. Spectrosc*. 54(3) (2000) 438-444.
24. Ntakatsane M P., Liu X. M., Zhou P., *J. Dairy Sci.* 96(4) (2013) 2130–2136.
25. Poulli K. I., Mousdis G. A., Georgiou C. A., *Anal Bioanal. Chem*. 386 (2006) 1571-1575.
26. Ammari F., Cordella Christophe B. Y., Boughanmi N., *Chemometr. Intell. Lab.Syst.* 113 (2012) 32–42.

27. Fethi F., Etude de qualité de l'eau, du chocolat et des huiles par les techniques spectroscopiques suivantes : LIBS, SPIR et FIL, thèse d'état, Faculté des Science Oujda, 2002.
28. Royston Goodacre, Elke Anklam, *JAOCS*, 78(10) (2001) 993-1000.
29. Lipp M., Anklam E., *Food Chemistry*, 62(1) (1998) 99-108.
30. Massart D. L., Vandeginste L. M., Buydens M. C., Chemometrics: A textbook, *Elsevier Science Publisher*. 23 (1988), 395–397.
31. Esbensen K.H., Multivariate Data Analysis – In Practice 5th Edition, An Introduction to Multivariate Data Analysis and Experimental Design, ISBN 82-993330-3-2, 1994 – 2002 CAMO Process AS All rights reserved. 5th edition. Re-print December 2004, Dominique Guyot, Frank Westad, Lars P. Houmoller.
32. Lebart L., Morineau A., Piro M., L'Analyse en Composantes Principales. In: statistiques exploratoires multidimensionnels, *Dunod, Bordas*. Paris, (1997) 32.
33. Wold S., Esbensen K., Geladi P., Chemometr. Intell. Lab. Sys., 2 (1987) 37-52.
34. Bro R., Workman J. J., Mobley P. R., Kowalski B. R., (1997). "Overview of chemometrics applied to spectroscopy: 1985–95, Part 3—Multiway analysis", *Applied Spectroscopy Reviews*, 32 (3) 237–261.
35. Wold S., Sjostrom M., Eriksson L., *Chemometr. Intell. Lab. Sys.*, 58 (2001) 109–130.
36. Tenenhans M., la regression pls, theorie et pratique, édition technip, paris, ISBN 2-7 108-0735-1, 1998.
37. Henseler J., *Comput Stat*. 25 (2010) 107–120.
38. Guimet F., Ferré J., Boqué, R., *Analytica Chimica Acta*, 544 (2005) 143-152.
39. Ourrach I., Rada M., Pérez-Camino M.C., Benaissa, M. Guinda Á., *Grasas Aceites*. 63(4) (2012) 355-364.