

Separation and characterization of the pomegranate plant cultivars by chemometric methods based on mid infrared spectroscopy

W. Terouzi¹, I. Taouaf², A. Boulli² & A. Oussama¹

¹ *Laboratory of Spectro- Chemometrics and environment, Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco.*

² *Team Environment and Agro Resources Valorization. Faculty of Science and Technology of Beni Mellal, University of Sultan Moulay slimane, 21000- Beni Mellal, Morocco.*

Abstract: *Ensure the quality and authenticity of the products has been a concern of several country. The goal of this work was to attempt characterization and classification of moroccan pomegranates according to their botanical origin based on the leaf analysis by Attenuated Total Reflectance-Fourier Transform Mid Infrared Spectroscopy (ATR-FTMIR) and find the markers of their authenticity. This study was focused on the pomegranate samples of five varieties and picked in same farm. The classification models were developed by principal component analysis (PCA) and support vector machines (SVMs). On the basis of a PCA, five distinct clusters were recognized. The SVMs procedure was then elaborated. The model resulted able to separate the five classes and classify new samples into the appropriate defined classes with a percentage prediction of 98.11%. This Result show the capability of MIR and the important role of chemometric tools in developing accurate models to identify rapidly pomegranates variety.*

Keywords: *Authenticity, chemometric tools, leaf, mid infrared spectroscopy, pomegranates.*

I. Introduction

Food authenticity is a problem with increasing importance for food processors, retailers, regulatory authorities and consumers [1]. In fact, the authenticity of food has become a worldwide problem, making it more necessary to detect the introduction of certain deceitfully labelled and low quality products, either for economical reasons or for public health matters [2], [3].

Food authenticity is a term which simply refers to whether the food purchased by the consumer matches its description. False description can occur in many forms, from the undeclared addition of water or other cheaper materials, or the wrong declaration of the amount of a particular ingredient in the product, to making false statements about the source of ingredients i.e., their geographic, plant, or animal origin [4]. A discussion on the concept of authenticity can be found in Ref. [5].

Generally, determination of authenticity is very important for the protection of wealth and health of consumers. Also, the authenticity of food and the methodology for testing are well-known topics in food science since many years [6].

In order to protect consumer interests and to combat the problems of food fraud and adulteration, Many of scientific expertise and technologies are constantly being developed to test the authenticity of foods [7]. Adulteration can occur for a variety or geographical origin of food, according to economic reasons, often linked to financial gain.

Additionally, in recent years, much research has been done on the different techniques to achieve a robust identification and characterization method, such as molecular markers [8], [9] and genetic profile [10], but these methods are not fast and cannot measure on site. In addition, they are also expensive and time-consuming.

On the other hand, The spectroscopic techniques to study the origin and differentiation of food products have been developed considerably in recent years. In particular, Fourier transform infrared spectroscopy (FTIR) has been successfully adopted to qualitative studies. For qualitative analysis, the mid-IR (MIR) spectroscopy has more applications, because the 'fingerprints' of functional groups can be displayed narrowly and intensely in the MIR region (4000–400 cm⁻¹). Fourier transform infrared (FTIR) spectroscopy with attenuated total reflectance (ATR) or transmission cell accessories has been used to characterize, identify or classify fats and oils [11], [12], [13], [14], [15].

FTIR data have been often treated with multivariate analytical techniques to develop methods of classification and characterization, through the building of relative models. This approach has demonstrated to be very useful in many applications, due to the ability in achieving the spectral resolution of the FTIR signals [16], [17], [18], [19], [20].

In this case, the goal of this study was to evaluate and evince the potential of ATR-FTIR combined with chemometrics to characterize and testing the authenticity of Moroccan pomegranate by using the ATR-FTIR data carried out from direct analysis of the pomegranate leaves. In fact, the application of FTIR spectroscopy which combined with chemometric methods is a relatively new approach in food characterization studies [21]. It is a non-destructive technique that offers a fast and effective method of food analysis [22]; simple and easy to apply in research, quality control laboratories and industrial settings.

To our knowledge, this is the first study on the characterization and classification of Moroccan pomegranate for five varieties by ATR-FTIR spectra of leaves.

II. Materials And Methods

2.1. sampling and procedures

The study area is confined to region of Beni mellal- Khnifra in central Morocco, expanded on a surface of nearly 33 208 Km². Investigation was focused on the pomegranate leaves samples of five varieties : Nagro Monstruoso (NM), Grenade Rouge (GR), Sefri (S), Wenderful (W) and Ounk Hmam (OH) ; picked up from trees belonging to the same farm in Beni Mellal area.

A series of 150 pomegranate leaves were sampled during the harvested season 2015. In fact, we created two different collections of samples. The first includes 80 leaves were manually harvested from the farm belong to Beni-Mellal area. The second consists of 70 leaves were randomly collected from same farm after 20 days of first collection.

In total 150 samples chosen from the harvested leaves based on their size and the absence of surface defects. Similar size leaves were chosen to minimize the effect of size on spectral measurements.

A series of 53 samples (from second collection) was used as an external validation set (prediction set). This last series was used to establish the robustness of the SVMs model. whilst the remaining 97 samples were selected to build up the calibration models.

The leave samples were kept in cold storage (7°C) during the nights between the days of measurements. Spectroscopic measurements were taken from the laeves after they had been brought into equilibrium with the room temperature of 25°C.

2.2. Instrumentation and 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, 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 600cm⁻¹ and the data were handled with OPUS logiciel. Leaf of pomegranate samples were directly deposited between two well-polished KBr plates, without preparation on an Attenuated Total Reflectance cell provided with a diamond crystal. For ATR-FTIR measurements, it was necessary to keep a controlled pressure, to ensure good contact between the sample and the diamond surface. 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. Statistical treatment of data

2.3.1. Principal Component Analysis (PCA)

Principal component analysis (PCA) is a non-supervised statistical tool commonly used for classification of data [23]. It is a method of data reduction [24], constructs new uncorrelated variables called principal components (PCs) that are linear combination of the original measurement variables [25], [26], [27]. Then, PCs can define most of variation in data [28]. In classification case, PCs are oriented according to the largest variance, not according to the largest class separation ability [29]. So, PCA is a linear projection method, based on variance, reduces the dimensionality of the original data matrix retaining the maximum amount of variability [30], [31]. Results of PCA are visualised by scores and loading plots. The scores vectors describe the relationship between the samples and allow checking if they are similar or dissimilar, typical or outlier, while the loadings vectors describe the importance of each variable [32], [33]. PCA was used in this work to perform pomegranates characterization according to their botanical origin.

2.3.2. Support Vector Machines (SVMs)

Support Vector Machines (SVMs) is a supervised classification technique [34], [35]. SVMs are kernel based learning algorithms introduced by Vapnik [36], [37]. It uses linear or non-linear kernel-functions to project the data into a high-dimensional feature space. Correlation is then performed in this hyperspace based on the structural risk minimization principle; i.e., aiming to increase the generalization ability of a model [38], [39].

SVMs as multivariate analysis of classification show effective results in various applications [40], [41], [42], [43].

2.3.3. Data pre-processing procedures and Software

PCA and SVMs were carried out on the spectral data. ATR-FTIR spectra were standardized by using mean centre pre-treatment to remove or minimise any unwanted baseline contribution. To apply classification methods, a series of pre-treatment elaborations were tested on the data matrix. The Savitzky–Golay [44] and Norris gap [45] algorithms were tested for data derivatization. Standard normal variate (SNV) and multiple scatter correction (MSC) [46] were also tested.

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

III. Results And Discussion

3.1. Data acquisition

ATR-FTIR spectra of 150 leaf samples were recorded and divided in two sets: a calibration set of 97 samples and an external validation set of 53 samples. A mean spectrum was calculated for each class of calibration set. The resultant mean spectra of five classes are shown in Fig. 1.

Fig.1 shows the mean FTIR spectra of the studied pomegranate leaves. The differences among them were clearly small and occurred only in limited regions of the spectra. The obtained spectra are dominated by significant bands of water are clearly visible in the leaf spectra at 3400 and 1640 cm^{-1} . The two bands at 2920 cm^{-1} and 2848 cm^{-1} are characteristic of lipids composition. The typical infrared pattern of sugar is observed in the region 1200 - 900 cm^{-1} , while the range 2400 - 2300 cm^{-1} is due to CO_2 [46], [47], [48], [49].

The use of single peaks or narrow wavelength ranges to obtain information useful to distinguish the pomegranates according to their botanical origins seemed very hard. These data were so conveniently handled by multivariate statistical techniques. With the aim to obtain more information from the ATR-FTIR spectral data, the spectra were firstly subjected to mathematical elaboration. In particular, derivative transformations were applied [50]. The best improvement in data variance was reached when the derivative function through the Gap algorithm was used. The derivative parameters were optimised and so fixed: 2nd order, gap size 17 with centered data. After this pre-treatment, the data were conveniently handled by appropriate multivariate statistical tools.

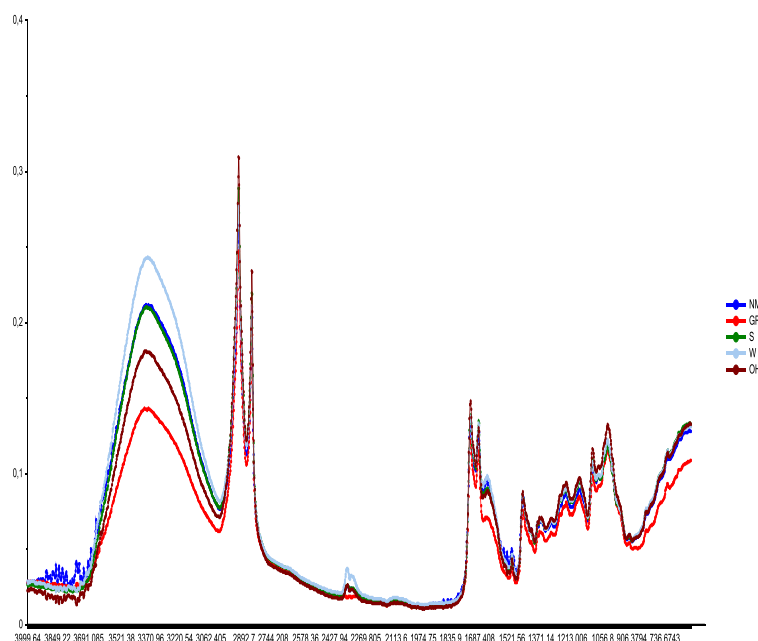


Fig. 1. The mean spectra calculated for each class: Nagro Monstruoso (NM), Grenade Rouge (GR), Sefri (S), Wenderful (W) and Ounk Hmam (OH)

3.2. PCA modeling

The PCA model was built with the NIPALS algorithm, by considering all frequency intervals 4000–600 cm^{-1} , as X variables, and the model was validated by full Cross validation. With full Cross validation, the

same samples are used both for model estimation and testing: leaves out only one sample at a time [51]. Results of PCA are visualised by scores and loading plots.

According to the score plot (**Fig.2**) the separation among the five classes is clearly shown when the plot is in 3D. The explained variance (%) obtained from the full cross-validation of the PCA model and three PCs were selected for a complete description of the variance in the spectral data set: 93% and are able to cluster the samples in the five classes.

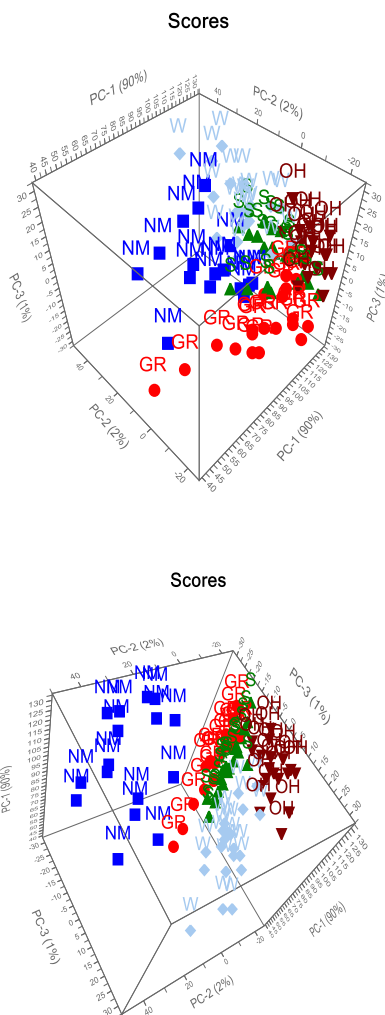


Fig. 2. 3D- Score plot by PCA analysis on the calibration set : Nagro Monstruoso (NM), Grenade Rouge (GR), Sefri (S), Wenderful (W) and Ounk Hmam (OH)

In **Fig.3**, it is 1-D correlation loading plot. The 1-D correlation loading plots are especially useful when interpreting important wavelengths in the analysis of spectroscopic. In the 1-D correlation loadings plot there are red dashed lines for the bounds of the 50 and 100% explained variance for the given latent variable. Values that lie within the upper and lower bounds of the plot are modelled by that latent variable. Those that lie between the two inner bounds are not.

ATR-FTIR spectra of leaf samples were examined by loadings plot (**Fig.3**) of PCA, to determine the bandwidths, which can be considered as the finger-prints of discrimination between studied samples : Typical fingerprint regions of carbohydrates ($1200\text{--}900\text{ cm}^{-1}$), proteins ($1680\text{--}1540\text{ cm}^{-1}$), lipids ($3000\text{--}1700\text{ cm}^{-1}$) and water (3400 cm^{-1} and 1641 cm^{-1}).

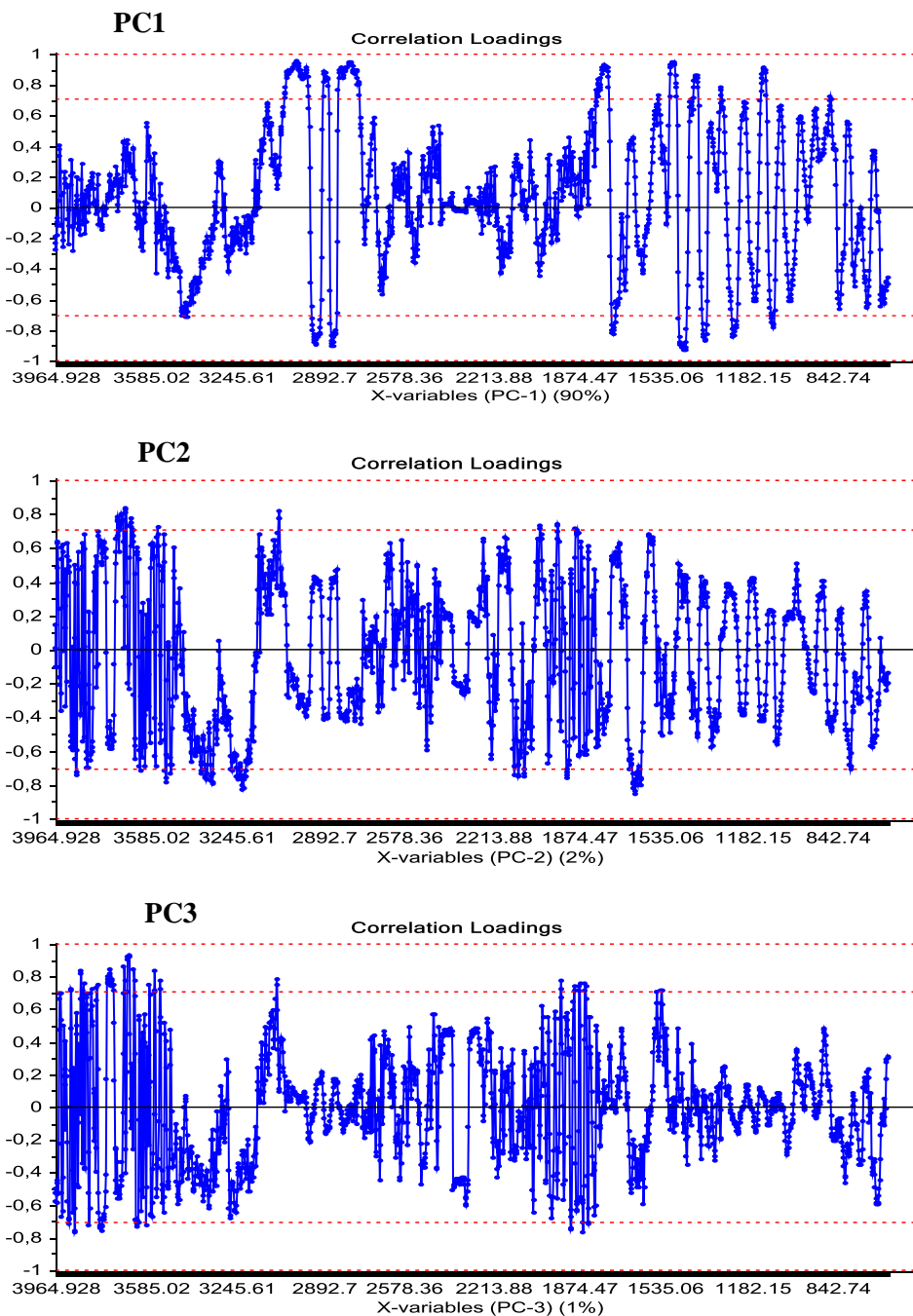


Fig. 3. 1-D correlation loadings of the first three principal components (PC1, PC2 and PC3)

From the loadings plot corresponding to the first three components, it is possible to determine peaks responsible of this distinction. The results from **Fig.3.** indicated that the spectral regions ($2960.197\text{-}2761.564\text{ cm}^{-1}$) and ($1756.833\text{-}1091.512\text{ cm}^{-1}$) associated with lipids and sugar composition were the best ones for leaf samples discrimination by PCA.

3.3. Classification model : SVMs modeling

The SVMs model was built by considering, as X variables, all spectra ; and the classification model was validated by Cross validation with segment = 97. We constructed SVMs model with a radial SVM-kernel classification, different kernels have been tested on these data, and the results showed that the best choice is the linear kernel, to determine the hyperplane that give best separates the classes. The optimal parameter for “gamma” is selected as the value that give the maximum correct classification rate, “gamma”= 0.0005668934. Consequently, 94 of calibration samples are retained as support vectors, where 15 samples of Nagro Monstruoso

(NM) variety, 19 of Grenade Rouge (GR) samples, 22 of Sefri (S) samples, 19 of Wenderful (W) samples and 19 samples of Ounk Hmam (OH) variety.

The obtained SVMs model was able to separate and discriminate all samples of the calibration set with a correct classification of 100%. This result can be seen in the score plot of **Fig. 4**.

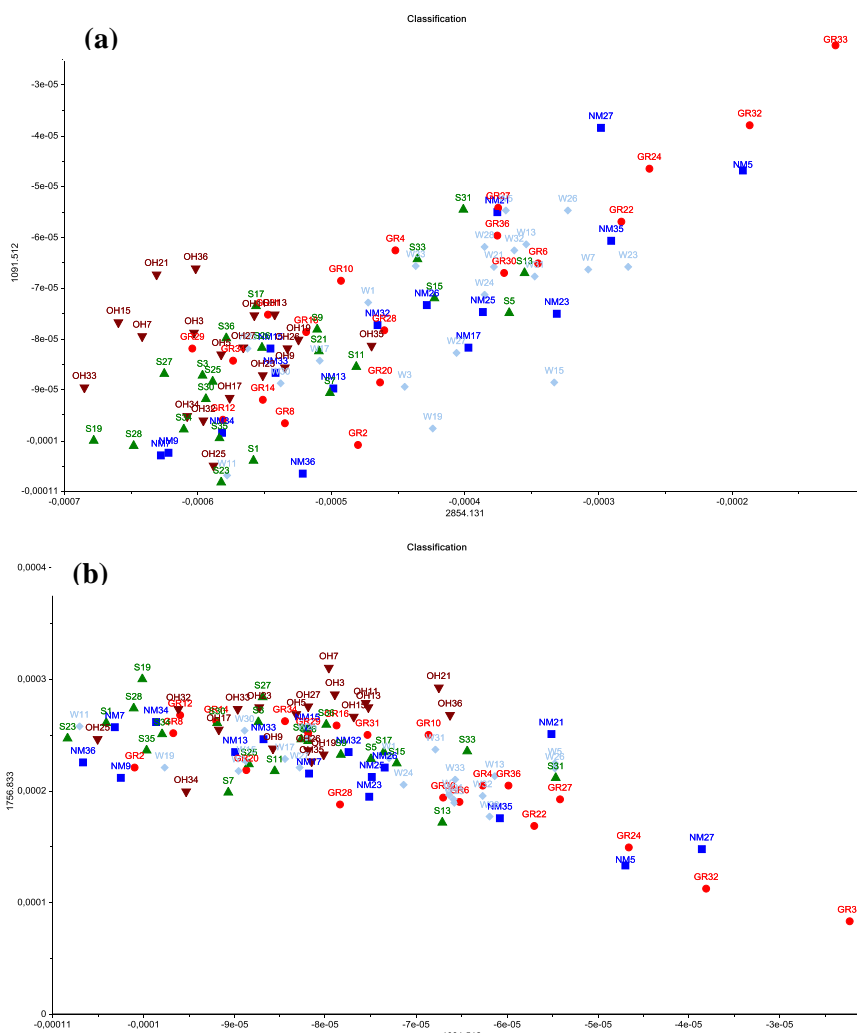


Fig.4. 2D score plot of classification results by SVMs on calibration : (a) with 2 wavelenths 2854.131 / 1091.512 cm-1; (b) with 2 wavelenths 1091.512 / 1756.833 cm-1

This result is confirmed by confusion matrix (**Table.1**). The confusion matrix is a matrix used for visualization for classification results from supervised method, SVMs. It carries information about the predicted and actual classifications of samples, with each row showing the instances in a predicted class, and each column representing the instances in an actual class. **Table.1** shows that all samples are well classified.

Table1. Confusion matrix of calibration set, carried out by SVMs

Confusion matrix	NM	GR	S	W	OH
Predicted	1	2	3	4	5
NM	16	0	0	0	0
GR	0	20	0	0	0
S	0	0	22	0	0
W	0	0	0	20	0
OH	0	0	0	0	19

3.4. Classification of new samples by SVMs model

Classification model of SVMs is trained with calibration samples from five classes and is used to classify future samples (prediction set) into one of these classes. The SVMs model was applied to a group of

unknown samples from different pomegranate leaves of the five varieties (53 external pomegranate leaf samples), the results are listed in **Table 2**.

Table2. Classification of pomegranate leaf samples of the prediction set by using SVMs model

Samples	Predicted	Reference class
NM1	NM	NM
NM2	NM	NM
NM3	NM	NM
NM4	S	NM
NM5	NM	NM
NM6	NM	NM
NM7	NM	NM
NM8	NM	NM
GR1	GR	GR
GR2	GR	GR
GR3	GR	GR
GR4	GR	GR
GR5	GR	GR
GR6	GR	GR
GR7	GR	GR
GR8	GR	GR
GR9	GR	GR
GR10	GR	GR
GR11	GR	GR
GR12	GR	GR
S1	S	S
S2	S	S
S3	S	S
S4	S	S
S5	S	S
S6	S	S
S7	S	S
S8	S	S
S9	S	S
S10	S	S
S11	S	S
S12	S	S
W1	W	W
W2	W	W
W3	W	W
W4	W	W
W5	W	W
W6	W	W
W7	W	W
W8	W	W
W9	W	W
OH1	OH	OH
OH2	OH	OH
OH3	OH	OH
OH4	OH	OH
OH5	OH	OH
OH6	OH	OH
OH7	OH	OH
OH8	OH	OH
OH9	OH	OH
OH10	OH	OH
OH11	OH	OH
OH12	OH	OH

Table.2 shows the classification results with the comparison between the predicted results of each class and the theoretical reference classes. The rate of correct classification was 98.11% within the test set. In particular 52 samples were safely assigned in the five classes, while one samples (NM4) resulted classified in another class.

IV. Conclusions

In this work, we arrived to separate pomegranate leaf samples of five varieties by differences in their ATR-FTIR spectra. MIR spectroscopy combined with chemometric tools PCA and SVMs are successfully

applied to the characterization and classification of pomegranates according to their botanical origin (percentage of correct classification is 98.11%).

These results indicate that ATR-FTIR with chemometrics can be a useful tool for rapid control of pomegranate authenticity. Its application was rapid and simple because no chemical treatment of samples was required. Also, it can be used in agricultural-food industry for the reliable, cheap and fast quality control of raw material.

References

- [1]. Downey, G. Review: Authentication of food and food ingredients by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, Volume 4 Issue 1, 1996, 47–61
- [2]. Mackie, I. M. Authenticity of fish. In P. R. Ashurt & M. J. Dennis (Eds.), *Food authentication*. London: Blackie Academic and Professional. 1996, 140–170
- [3]. Simone, S. Souza, Adriano G. Cruz, Eduardo H.M. Walter, Jose A.F. Faria, Renata M.S. Celeghini, Márcia M.C. Ferreira, Daniel Granato, Anderson de S. Sant'Ana. Monitoring the authenticity of Brazilian UHT milk: A chemometric approach. *Food Chemistry*, 124, 2011, 692–695
- [4]. Kvasnička, F. Capillary electrophoresis in food authenticity. *Journal of Separation Science*, Volume 28, Issue 9-10, 2005, 813–825
- [5]. Schellenberg, A. Authenticity of food. *Journal Deutsche Lebensmittel-Rundschau*, Vol. 106, No. 3, 2010, 152-155
- [6]. Jürg Lüthy. Detection strategies for food authenticity and genetically modified foods. *Food Control*, 10, 1999, 359-361
- [7]. Grundy, H.H., Kelly, S.D., Charlton, A.J., Donarski, J.A., Hird, S.J., Hird H.J. & Collins, M.J. Food Authenticity and Food Fraud Research: Achievements and Emerging Issues. *Journal of the Association of Public Analysts (Online)*, 2012, 40, 65-68
- [8]. Cooke R.J. In : Wrigley C.W ed. Identification of food-grain varieties. Introduction : the reasons for variety identification. American Association of Cereal Chemists, St Paul. 1995a, 1-17.
- [9]. Cooke R.J. Varietal identification of crop plants. *New Diagnostics in Crop Sciences*. In : Skerritt J.H, Appels R (eds). CAB International, Wallingford. 1995b, 33-63.
- [10]. Ancillo G, Gadea J, Forment J, Guerri J, Navarro L. Class prediction of closely related plant varieties using gene expression profiling. *Journal of Experimental Botany*, 58, 2007, 1927-1933.
- [11]. Lai, Y. W., Kemsley, E. K., & Wilson, R. H. Potential of Fourier transform infrared spectroscopy for the authentication of vegetable oils. *Journal of Agricultural and Food Chemistry*, 42, 1994, 1154–1159.
- [12]. Dahlberg, D. B., Lee, S. M., Wenger, S. J., & Vargo, J. A. Classification of vegetable oils by FT-IR. *Applied Spectroscopy*, 51, 1997, 1118–1124.
- [13]. Dupuy, N., Duponchel, L., Huvenne, J. P., Sombret, B., & Legrand, P. Classification of edible fats and oils by principal component analysis of Fourier transform infrared spectra. *Food Chemistry*, 57, 1996, 245–251.
- [14]. Ozen, B. F., Weiss, I., & Mauer, L. J. Dietary supplement oil classification and detection of adulteration using Fourier transform infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, 57, 2003, 5871–5876
- [15]. Safar, M., Bertrand, D., Robert, P., Devaux, M. F., & Genot, C. Characterization of edible oils, butters and margarines by Fourier transform infrared spectrometry with attenuated total reflectance. *Journal of the American Oil Chemists_ Society*, 71, 1994, 371–377.
- [16]. Zagonel, G.F., Peralta-Zamora, P., Ramos, L.P. Multivariate Monitoring of Soybean Oil Ethanolysis by FTIR. *Talanta*, 63, 2004, 1021-1025
- [17]. De Luca, M., Terouzi, W., Ioele, G., Kzaiber, F., Oussama, A., Oliverio, F., Tauler, R., Ragno, G. Derivative FTIR spectroscopy for cluster analysis and classification of morocco olive oils. *Food Chem.* 2011, doi: 10.1016/j.foodchem.2010.07.010.
- [18]. De Luca, M., Terouzi, W., Kzaiber, F., Ioele, G., Oussama, A., Ragno, G. Classification of moroccan olive cultivars by linear discriminant analysis applied to ATR–FTIR spectra of endocarps. *International Journal of Food Science and Technology*, 47, 2012, 1286–1292.
- [19]. Pinheiro, P. B. M. and da Silva, J. Chemometric classification of olives from three Portuguese cultivars of *Olea europaea*. L. *Anal. Chim. Acta.*, 544, 2005, 229–235.
- [20]. Terouzi, W., De Luca, M., Bolli, A., Oussama, A., Patumi, M., Ioele, G., Ragno, G. A discriminant method for classification of Moroccan olive varieties by using direct FT-IR analysis of the mesocarp section, *Vib. Spectrosc.*, 2011, doi:10.1016/j.vibspec.2011.01.004.
- [21]. Yang, H., Irudayaraj, J., Paradkar, M.M. Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy. *Food Chemistry*, 93, 2005, 25–32.
- [22]. Karoui, R., Dufour, E., Pillonel, L., Picque, D., Cattenoz, T., & Bosset, J. O. Fluorescence and infrared spectroscopies: a tool for the determination of the geographic origin of Emmental cheeses manufactured during summer. *Lait*, 84, 2004, 359–374.
- [23]. Scott, I.M., Lin, W., Liakata, M., Wood, J.E., Vermeer, C.P., Allaway, D., Ward, J.L., Draper, J., Beale, M.H., Corol, D.I., Baker, J.M., King, R.D. Merits of random forests emerge in evaluation of chemometric classifiers by external validation. *Analytica Chimica Acta* 801, 2013, 22– 33.
- [24]. Campitelli, P., Ceppi, S. Chemical, physical and biological compost and vermicompost characterization: A chemometric study. *Chemometrics and Intelligent Laboratory Systems* 90, 2008, 64–71.
- [25]. Zhao, H., Guo, B., Wei, Y., Zhang, B. Effects of grown origin, genotype, harvest year, and their interactions of wheat kernels on near infrared spectral fingerprints for geographical traceability. *Food Chemistry* 152, 2014, 316–322.
- [26]. Wold, S., Esbensen, K. & Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* 2, 1987, 37-52.
- [27]. 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.
- [28]. Rajalahti, T. & Kvalheim, O.M. Multivariate data analysis in pharmaceuticals: a tutorial review. *Int. J. Pharm.* 417, 2011, 280–290.
- [29]. Arjmand, M., Kompany-Zareh, M., Vasighi, M., Parvizzadeh, N., Zamani, Z., Nazgooei, F. Nuclear magnetic resonance-based screening of thalassemia and quantification of some hematological parameters using chemometric methods. *Talanta* 81, 2010, 229–1236.
- [30]. Ivosev, G., Burton, L. & Bonner, R. Dimensionality reduction and visualization in principal component analysis. *Anal. Chem.* 80,13, 2008, 4933-44.
- [31]. Garrido-Delgado, R., Arce, L., Guamán, A.V., Pardo, A., Marco, S., Valcárcel, M. Direct coupling of a gas–liquid separator to an ion mobility spectrometer for the classification of different white wines using chemometrics tools. *Talanta* 84, 2011, 471–479.

- [32]. Ferreira, M. C. M., Morgano, M. A., Queiroz, S. C. N., & Mantovani, D. M. B. Relationships of the mineral and fatty acid contents in processed Turkey meat products. *Meat Science*, 69(3), 2000, 259–265.
- [33]. Granato, D., Castro, I.A., Katayama, F.U. Assessing the association between phenolic compounds and the antioxidant activity of Brazilian red wines using chemometrics. *LWT-Food Science and Technology*, 2010, in press, doi:10.1016/j.lwt.2010.05.031.
- [34]. Luna, A.S., da Silva, A.P., Ferré, J., Boqué, R. Classification of edible oils and modeling of their physico-chemical properties by chemometric methods using mid-IR spectroscopy. *Spectrochimica Acta Part A* 100, 2013, 109–114.
- [35]. Gavicho Uarrota, V., Moresco, R., Coelho, B., da Costa Nune, E., Martins Peruch, L.A., de Oliveira Neubert, E., Rocha, M., Maraschin, M. Metabolomics combined with chemometric tools (PCA, HCA, PLS-DA and SVM) for screening cassava (*Manihot esculenta* Crantz) roots during postharvest physiological deterioration. *Food Chemistry* 161, 2014, 67–78.
- [36]. Vapnik, V. *Chervonenkis, Theory of Pattern Recogn*, Nauka, Moscow, 1974.
- [37]. Vapnik, V. *The Nature of Statistical Learning Theory*, Springer Verlag, 1995.
- [38]. Vapnik, V. *The Nature of Statistical Learning Theory*. Second edition. New York: Springer, 1999.
- [39]. Drucker, H., Burges, C.J., Kaufman, L., Smola, A. and Vapnik, V. Support Vector Regression Machines. *Adv Neural Inf Process Syst*, 9, 1997, 155-161.
- [40]. Terouzi, W., Platikanov, S., Capdevila, A.J. & Oussama, A. Classification of olives from Moroccan regions by using direct FT-IR analysis: Application of support vector machines (SVM). *International Journal of Innovation and Applied Studies*, ISSN 2028-9324 Vol. 3 No. 2, 2013, 493-503
- [41]. Belousov AI, Verzakov SA, von Frese J. Support Vector Machines: A Versatile and Powerful Approach to Data Analysis, Poster at the Gordon Conf. On Statistics and Chemical Engineering, Williamstown, MA, 2001. [Online] Available: <http://www.Amstat.org/sections/spes/GRC2001.htm> (30 July 2004)
- [42]. Belousov AI, Verzakov SA, von Frese J. Applicational aspects of support vector machines. *J. Chemometrics*, 16, 2002, 482–489.
- [43]. Ding, C.Q. and Dubchak, I. Multi-class protein fold recognition using support vector machines and neural network. *Bioinformatics*, vol 17, n° 4, 2001, 349-358.
- [44]. Savitzky, A. & Golay, M.J.E. Smoothing and differentiation of data by simplified least-squares procedures. *Analytical Chemistry*, 36, 1964, 1627–1639.
- [45]. 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.
- [46]. 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.
- [47]. Hafidi, M., Amir, S., Revel, J. C. Structural characterization of olive mill waste-water after aerobic digestion using elemental analysis, FTIR and ¹³CNMR. *Process Biochem.* 40, 2005, 2615–2622.
- [48]. Droussi, Z., D’Orazio, V., Provenzano, M. R., Hafidi, M., Ouattmane, A. Study of the biodegradation and transformation of olive-mill residues during composting using FTIR spectroscopy and differential scanning calorimetry. *J. Hazard Mater.* 164, 2009, 1281–1285.
- [49]. Scibisz, I., Reich, M., Bureau, S., Gouble, B., Causse, M., Bertrand, D., Renard, C. Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. *Food Chemistry*, 125, 2011, 1390–1397.
- [50]. Maggio, R. M., Kaufman, T. S., Del Carlo, M., Cerretani, L., Bendini, A., Cichelli, A., 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.
- [51]. Esbensen KH. *Multivariate data analysis-in practice. An introduction to multivariate data analysis and experimental desing*. Oslo: CAMO, 2000