

Assessment of non-destructive spectroscopy and chemometrics tools for the development of green analytical methods to determine the shelf-life of olive oils

Aimen El Orche^{1,*}, Mohamed Mbarki¹, Amine Cheikh³, and Mustapha Bouatia².

¹Laboratory of Chemical Processes and Applied Materials University of Sultan MoulaySlimane, Beni-Mellal, Morocco.

²Laboratory of Analytical Chemistry, Faculty of Medicine and Pharmacy, Mohammed V University, Rabat, Morocco

³Faculty of Medicine, Abulcasis University, Rabat, Morocco.

Abstract. The development of sustainable and environmentally friendly analytical methods for agri-food products and the modification of reference methods is an essential issue to be treated in green analytical chemistry. The potential application of non-destructive spectroscopic techniques with chemometrics tools to achieve these principles are examined in this work. In this study a new sustainable analytical approach based on the use of fluorescence spectroscopy and multivariate analysis methods of Machine-Learning (Support Vector Machine regression) and chemometrics (Partial Least Square regression) have been developed to control the quality of virgin olive oils in Morocco according to their shelf life. The spectral data of 45 samples were first analyzed by principal component analysis method (PCA), the PCA method shows an important classification of the three groups of olive oil according to their shelf life. The use of the regression methods SVM and PLS shows a high ability to predict the quality of olive oils, this ability is shown by the high value of R-square and the low value of root mean square error of calibration and cross-validation (RMSEC, RMSECV), the validation of these models by cross-validation shows the potential of this sustainable analytical approach in the determination of the quality of virgin olive oils.

1 Introduction

Olive oil is an essential vegetable oil in the Mediterranean countries; nowadays, this nutrient is attracting the interest of many consumers around the world, thanks to its significant nutritional and organoleptic properties and its contribution to the protection of human well-being [1]. These qualities are particularly related to its composition rich in fatty acids, in particular oleic and linoleic acid [2], and its high content of minor compounds with bioactive characteristics, mainly phenolic compounds, tocopherol and chlorophyll. These compounds are also considerably reduced during the storage of olive oil, although new products are created due to the oxidation process. In numerous markets, the storage of olive oil may range from 6 to 24 months, resulting in a deterioration of the quality of olive oil. [1].

Determining the shelf life of olive oil constitutes an important challenge to protect consumers against fraud because many sellers expose in the market oil stored for a long time as much as freshly produced oil. Currently, different methods are used to assess the oxidative deterioration of olive oil. Routine methods include the peroxide value (PV), which determines the quantity of primary oxidation products, and the acid value, which determines the acidity of olive oil and the extinction coefficients at 232 and 270 nm, which measure the formation of primary and secondary oxidation products [2]. Other analytical method based on the use of high performance liquid chromatography HPLC and gas chromatography GC which consists in determining the molecular composition of oils [3,4].

Nevertheless, these techniques are complex to implement and interpret, costly and time-consuming, and use

1 Corresponding author: aimen.elorche@gmail.com

expensive reagents of both organic and inorganic nature, since these reagents have an impact on the environment. For this reason it is necessary to develop innovative and sustainable analytical tools capable of carrying out rapid and reliable quality controls of virgin olive oil in order to determine the shelf life of olive oil and protect consumers against fraud.

The combination of chemometric and spectroscopic techniques to develop methods according to the concept of green chemistry has become an important challenge over the last decades. A wide variety of chemometric methods have been used so far in the field of analytical chemistry and studying food products [5-7], These techniques are characterized by the rapidity and do not involve the use of reagents which leads to the destruction of the ecosystem.

The objective of this report is to show the importance of fluorescence spectroscopy coupled with recognition algorithms in the development of rapid and sustainable analytical methods for the control of virgin olive oil according to its shelf life.

2 Materials and methods

2.1. Sampling

A total of 45 samples of freshly produced virgin olive oil and two groups of expired olive oil stored for a period of time fluctuate between 12 and 24 months. These samples were generally stored in a dark place with a temperature interval of $10^{\circ}\text{C} \pm 1$.

The analysis of the samples was performed using fluorescence spectroscopy, the excitation wavelength was fixed at 400 nm and the observation interval between 415 nm and 785 nm.

2.2. Multivariate data analysis

Firstly the spectral data are analyzed by principal component analysis (PCA) in order to reduce the dimensionality and to exploit the dataset [8].

To develop training models able to predict the quality of olive oil the partial least square PLS method and support vector machine SVM have been applied to the spectral data set.

Partial least squares regression, or PLS regression, is a fast, efficient and optimal method for a well-controlled covariance minimization criterion. Its use is recommended when a large number of explanatory variables are used, or when there are strong collinearities between variables [9].

The SVM method is one of the supervised learning algorithms. This method solves the problem of pattern recognition. It aims to find an optimal separator that optimizes the margin between two classes of data, using a limited set of training sequences [10].

3 Results and discussion

3.1. Principal component analysis

A preliminary examination of the spectra was carried out by the PCA applied to the fluorescence spectra. Figure 1 shows a score plot of the first two principal components. From the score plot PC1-PC2 it can be seen that the first component accounts for all of the total variability in the data set (96% of the total variance) and the second principal component accounts for 2%. These two components show a classification of olive oils according to their freshness.

The results obtained by the exploratory analysis method PCA show that the spectral fluorescence data provide information on the molecular properties of the different olive oils, allowing their classification according to their shelf life.

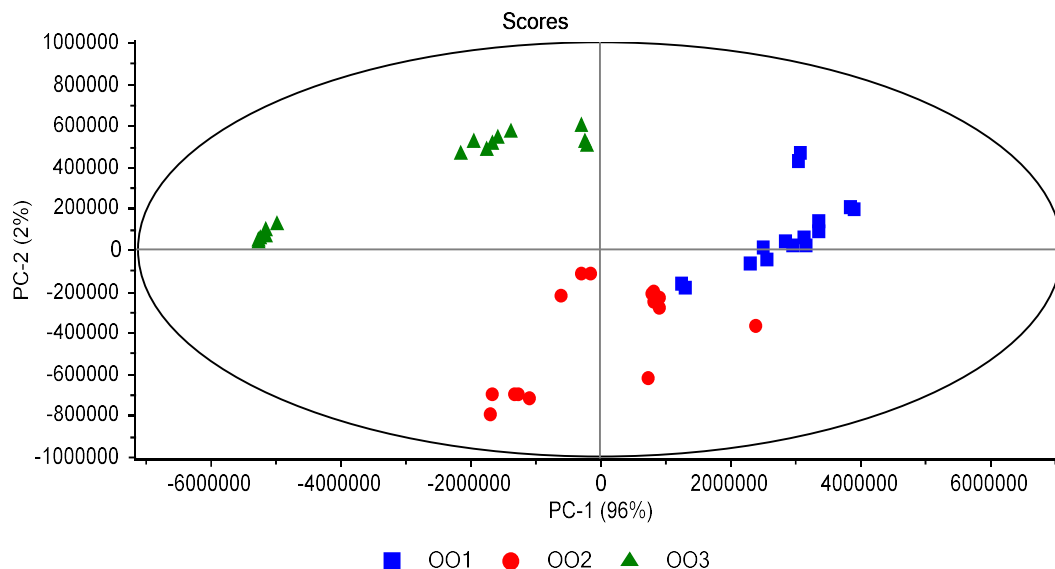


Fig.1: Score plot of the first two PCs obtained by PCA applied to the fluorescence data of olive oils(OO1: olive oil fresh, OO2: olive oil stored 1 year, OO3: olive oil stored 2 years).

For the development of sustainable analytical methods capable of determining the shelf life of olive oils during the storage process, it is necessary to choose statistical methods of multivariate analysis which are part of the regression type learning methods, these methods will be applied directly on the spectral data obtained by fluorescence spectroscopy.

3.2. Partial Least Square regression

In order to develop models capable of predicting the quality of olive oils (shelf-life), the PLS regression has been applied to the spectral data.

The application of PLS regression shows a very high capacity, this capacity was represented by the value of R-square that reach 99% and the low value of RMSEC that represent 0.08 using the first four latent variables.

To assess the predictive capacity of the model, the cross-validation method was applied, the results found by the cross-validation are considered performant, this performance was demonstrated by the value of R-square and low value of RMSECV which represents successively 91% and 0.23 as shown in the figure 2.

The application of the PLS method on the spectral data of different olive oil categories shows important and reliable results (figure 2), this reliability is explained by the high value of R-square coefficients which reaches 99% and a low error of 0.08 (RMSEC), this high value shows that there is a strong dependency between the reference value and the value predicted by the model. The predictive evaluation of the model developed by the PLS algorithm using the cross-validation shows that the developed model is able to predict the shelf life of the stored olive oil with high reliability explained by the R-square value and the error value which represent respectively 91% and 0.23%.

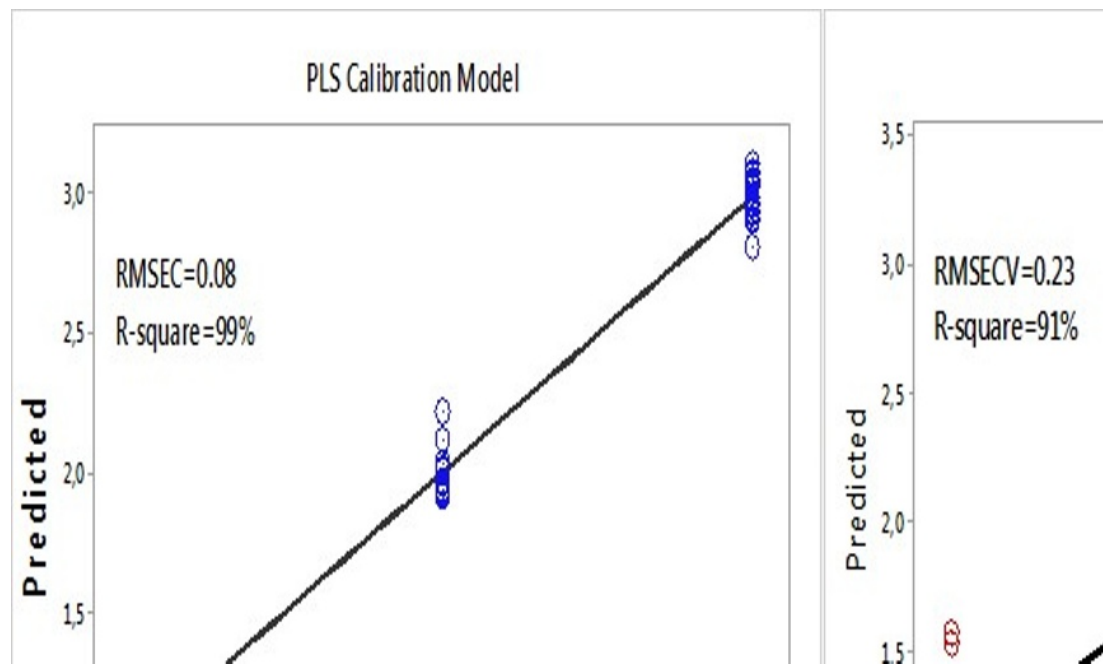


Fig.2: PLS regression results of calibration and cross-validation using 4 latent variables (LV).

3.3. Support Vector Machine regression

The analysis of the spectral results by the SVM regression also shows a very high capacity represented by the high value of R-square 99.99% and the low value of RMSEC 0.0005 as shown in the figure 3.

on the other hand, the application of the SVM regression method on the spectral data of the three categories of olive oil also shows an important capacity of the constructed model, this capacity is represented by the strong linearity between the reference value and the value predicted by the model expressed by the R-square value and the RMSEC error which represent respectively 99% and 0.0005 as shown in the figure.

the application of cross-validation for the evaluation of the constructed SVM model shows the high capability of this model in the determination of the shelf life of olive oils, this capability is shown by the value of R-square and the error which represent 97% and 0.19 respectively.

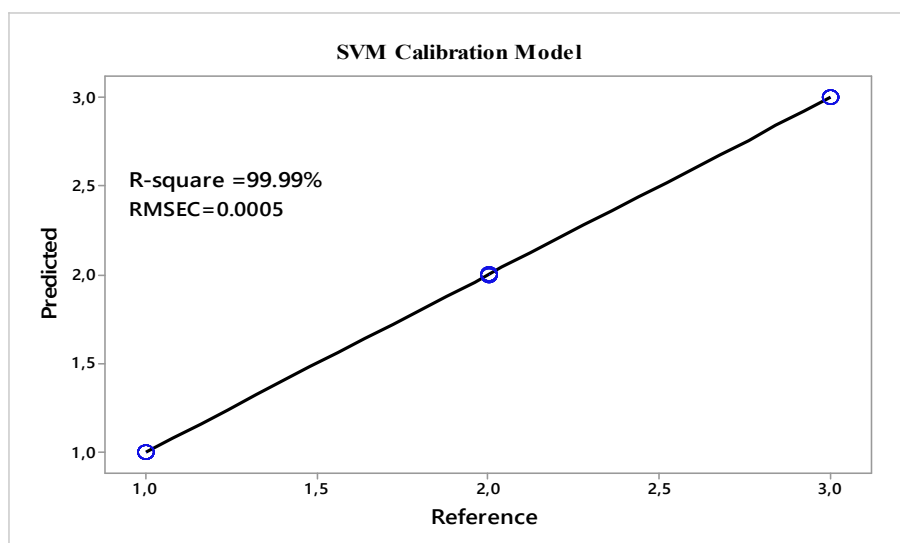


Fig. 3. SVM regression calibration model using linear kernel type algorithm.

According to the results found, the two multivariate statistical methods PLS and SVM give efficient results, these results show that the application of the SVM regression has a higher R-square and lower RMSE compared to the results found by the PLS regression.

4 Conclusion

The present work shows the capability of fluorescence spectroscopy combined with PLS and SVM multivariate regression analysis in the determination of the quality of olive oils according to their storage life.

The advantage of this technique is that it doesn't use environmentally harmful reagents and it is characterized by the rapidity of its use.

Last but not least the development of new environment-friendly analytical techniques and the modification of pre-existing methods are key points in advancing the concept of green analytical chemistry.

Reference

1. K. Bouarroudj, A. Tamendjari, R. Larbat, *Ind Crops Prod*, **8** (2016).
2. E.G. Shendi, D.S. Ozay, M.T. Ozkaya, N.F. Ustunel, *OCL - Oilseeds and fats, Crops and Lipids*, **6** (2020).
3. M. Farrés-Cebrián, R. Seró, J. Saurina, O. Núñez, *J Separations*, **13** (2016).
4. S. Vichi, A.I. Castellote, L. Pizzale, L.S. Conte, S. Buxaderas, E. López-Tamames, *J Chromatogr A*, **15**(2003).
5. N. Sinelli, M.S. Cosio, C. Gigliotti, E. Casiraghi, *Anal Chim Acta*, **7**(2007).
6. A. Gredilla, S. Fdez-Ortiz de Vallejuelo, N. A. Elejoste, de Diego, J.M. Madariaga, *TrAC Trends Anal Chem*, **39** (2016)
7. A. El Orche, M. Bouatia, S. Yanisse, H. Labjar, M. Mouhsin, M. Bouha, M. Oubenali, M. Mbarki, *J Spectrosc*, **9** (2020)
8. A. El Orche, M. Bouatia, M. Mbarki, *JAMC*, **9** (2020)
9. W. Svante, S. Michael, E. Lennart, *J Chemometrics*, **22** (2001).
10. A. El Orche, M. Bouatia, H. Labjar, M. Maaouni, M. Mbarki, *IEEE 6th International Conference on Optimization and Applications, Beni Mellal, Morocco*, **4** (2020).