Microchemical Journal 133 (2017) 669–675

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Transfer of multivariate classification models applied to digital images and fluorescence spectroscopy data

Karla Danielle Tavares Melo Milanez [a], Thiago César Araújo Nóbrega [a], Danielle Silva Nascimento [b], Matías Insausti [b], Márcio José Coelho Pontes [a,*]

[a] Departamento de Química, Universidade Federal da Paraíba, João Pessoa, PB, Brazil
[b] Laboratorio FIA, INQUISUR-CONICET, Departamento de Química, Universidad Nacional del Sur, Av. Alem 1253, B8000CPB Bahía Blanca, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 15 December 2016
Received in revised form 23 February 2017
Accepted 4 March 2017
Available online 7 March 2017

Keywords:
Extra virgin olive oil
Adulteration
Fluorescence spectroscopy
Digital images
Classification transfer

ABSTRACT

This work evaluates the use of transfer of classification models for identifying adulteration of extra virgin olive oil (EVOO) samples involving, separately, two analytical techniques: fluorescence spectroscopy and digital imaging. The chemometric procedures, including development of classification models and application of classification transfer methods, were performed individually for each analytical technique. Methods of direct standardization (DS) and piecewise direct standardization (PDS) were applied to transfer samples sets in order to estimate an adjustment function and apply it to a samples set measured by the secondary instrument. For purposes of comparison, classification models were built based on linear discriminant analysis (LDA) with previous selection of variables by the successive projections algorithm (SPA), and partial least squares discriminant analysis (PLS-DA). The performance of the classification models was evaluated according to the number of errors and correct classification rate (CCR) for the prediction set measured by the secondary instrument. Before standardization, SPA-LDA and PLS-DA models achieved the same CCR using two analytical techniques: 54% for fluorescence emission spectra and 47% for histograms of digital images. After the standardization, a substantial increase of the CCR was observed. For the SPA-LDA models, a CCR of 88% was obtained for the fluorescence emission spectra and 82% for the histograms of the digital images. The PLS-DA classification models reached 85% and 76% of CCR for the fluorescence and imaging data, respectively, after standardization. These results demonstrate the efficiency of standardization procedures applied to multivariate classification models developed from fluorescence spectroscopy and digital images.

1. Introduction

Multivariate classification models have been widely used in qualitative analysis involving conformity assessment and identification of adulteration in a variety of matrices [1–6]. Usually, the information obtained is used to discriminate samples into categories, according to similarity standards [7].

A number of supervised classification methods are reported in literature, including: soft independent modeling of class analogy (SIMCA) [8], k-nearest neighbors (k-NN) [9], partial least squares discriminant analysis (PLS-DA) [10] and linear discriminant analysis (LDA) [11]. Among these, LDA and PLS-DA have been constantly employed in classification problems [12–15]. Complications emerge when the classification models built from an instrument or laboratory condition cannot correctly classify samples analyzed under new conditions or by a different instrument [16]. The differences between the spectral responses may be related to changes in the chemical and/or physical composition of the samples or changes in the instrumental response, normally caused by maintenance irregularities, repairs, changes in the environment of the instrument or even natural wear [17].

A solution to this problem is a full model recalibration from a data set acquired under the new analytical conditions. However, this requires time, money and rigorous control of the experimental conditions [18]. Alternatively, the classification transfer could be undertaken using standardization methods of spectral responses.

Standardization methods include direct standardization (DS) and piecewise direct standardization (PDS) [16]. In both cases, the response from a secondary instrument is adapted to correspond to the response from a primary instrument [19]. For this purpose, spectra of a representative samples set, known as *transfer samples*, are recorded in the two instruments and used in the standardization procedures [20].

The mathematical manipulations involved in transfer procedure can be carried out along the entire spectral range (in direct standardization)

* Corresponding author at: Universidade Federal da Paraíba, Departamento de Química – Laboratório de Automação e Instrumentação em Química Analítica/Quimiometria (LAQA), CEP 58051-970 João Pessoa, PB, Brazil.
E-mail address: marciocoelho@quimica.ufpb.br (M.J.C. Pontes).

http://dx.doi.org/10.1016/j.microc.2017.03.004
0026-265X/

or in windows of variables within the spectrum (in piecewise direct standardization) [20]. Considering the matrices $\mathbf{X}^{P}$ ($N_{transf} \times q$) and $\mathbf{X}^{S}$ ($N_{transf} \times p$) containing the spectra of the N_{transf} transfer samples recorded in the primary and secondary instruments, respectively, standardization tries to relate them linearly as shown below:

$$\mathbf{X}^{P} = \mathbf{X}^{S} \cdot \mathbf{F} \quad (1)$$

where **F** is a transformation matrix of dimension ($p \times q$). In the DS method, **F** is obtained as follows:

$$\mathbf{F} = \left(\mathbf{X}^{S}\right)^{-1} \cdot \mathbf{X}^{P} \quad (2)$$

where ($\mathbf{X}^{S}$) is the generalized inverse of $\mathbf{X}^{S}$. Once the number of spectral variables is greater than the number of transfer samples and $\mathbf{X}^{S}$ is a rectangular matrix, **F** is normally determined using principal component regression (PCR) or partial least square (PLS). In these methods, the $\mathbf{X}^{P}$ and $\mathbf{X}^{S}$ spectral matrices are projected onto the column space of $\mathbf{X}^{P}$, calculated from singular value decomposition. The scores obtained in these projections are used in the standardization and calibration procedures [21].

The PDS method considers that spectral correlations may be limited to small regions. Thus, a moving window is used to relate each wavelength j of $\mathbf{X}^{P}$ to wavelengths contained in a small region around j of the $\mathbf{X}^{S}$ spectrum [16]. In PDS method, **F** is estimated by calculating regression vectors $\boldsymbol{f}$ for each window using PCR and PLS [21]. The regression vectors $\boldsymbol{f}$ are arranged in a diagonal matrix **F** according to:

$$\mathbf{F} = \mathrm{diag}\left(\boldsymbol{f}_1^{T},\ \boldsymbol{f}_2^{T}, \ldots, \boldsymbol{f}_j^{T}, \ldots,\ \boldsymbol{f}_k^{T}\right) \quad (3)$$

where $\boldsymbol{k}$ is the number of wavelengths. The use of the moving window may cause edge effects, which can be corrected truncating the ends of the spectrum or estimating them by extrapolation [21].

With the transformation matrix **F**, the spectrum of a new sample measured on the secondary instrument can be adjusted to $\mathbf{X}^{a}$, so as to resemble the spectra obtained by the primary instrument as expressed in Eq. (4). The correction of the spectrum obtained in the secondary instrument allows it to be properly classified in the model developed for the primary instrument.

$$\mathbf{X}^{a} = \mathbf{X}^{S} \cdot \mathbf{F} \quad (4)$$

Few studies involving standardization methods have been directed towards classification problems [22–25]. Our work evaluates the use of transfer of multivariate classification models for identifying adulteration of extra virgin olive oils. For this, were employed two data sets, one obtained by fluorescence spectroscopy and the other by digital imaging. SPA-LDA and PLS-DA classification models were built and their results were compared before and after standardization procedures in terms of the number of misclassification and correct classification rate (CCR) for the prediction set as measured by the secondary instrument. Direct standardization (DS) and piecewise direct standardization (PDS) were performed using different numbers of transfer samples and window sizes.

2. Materials and methods

2.1. Samples

Two classes of extra virgin olive oil (EVOO) samples were used in this work: unadulterated and adulterated. The unadulterated samples were acquired in local commerce with different batches from the same manufacturer. The manufacturer was chosen based on an investigation performed by a Brazilian Association of Consumer Protection [26], which evaluates the quality of commercially available products (more details regarding this investigation can be found in http://www.proteste.org.br/azeite).

In order to evaluate the authenticity of the unadulterated EVOO samples used in this study, the determination of specific extinction (K) by absorption in the ultraviolet region (AOCS Official Method Ch 5-91) was performed using cyclohexane of spectrophotometric grade as solvent, as described in [27].

This method uses the ultraviolet region to obtain information on the quality, state of conservation and changes by processing. The conjugated diene and triene systems, whose presence in EVOO can be attributed to oxidation and/or refining, are evaluated at the wavelengths of 232 and 270 nm [27]. The values of the specific extinction (K_{232} and K_{270}) and specific extinction variation (ΔK) obtained at these wavelengths are compared with the previously established limits and can be used to differentiate the EVOO from the other oil categories [28].

The results demonstrated that all unadulterated EVOO samples presented values of specific extinction at the wavelengths of 232 and 270 nm (K_{232} and K_{270}) and specific extinction variation (ΔK) equal or lower than the limits established by International Olive Council [28], reinforcing the results obtained by [26].

All samples were stored in amber glass bottles, protected from light and kept at a temperature of approximately 23 ± 2 °C until time of analysis. No sample pretreatment was performed.

2.2. Fluorescence spectroscopy

The first data set consisted of 88 extra virgin olive oil (EVOO) samples (48 unadulterated samples and 40 adulterated samples). The adulterations were prepared by addition of soybean oil at different levels (1.0%, 5.0%, 10.0%, 15.0%, 20.0%, 25.0% and 30.0% mm^{-1}).

The fluorescence spectra were recorded under the same experimental conditions using a Jasco FP-6500 Spectrofluorometer, which was considered as the primary instrument, and an Aminco Bowman Serie 2 Spectrofluorometer, used as the secondary instrument. In both instruments, the slit width was 3 nm for excitation and 5 nm for emission. The acquisition interval and integration time was maintained at 1 nm and 0.5 s, respectively. A PMT lamp 400 W and a quartz cell 10 × 10 × 45 mm were used for right-angle geometry.

The fluorescence emission spectra were collected between 350 and 600 nm. The excitation wavelength of 330 nm was selected for development of the multivariate classification models. The excitation wavelength range from 280 to 480 nm was evaluated in 10 nm steps.

2.3. Digital images

A total of 54 EVOO samples were measured by digital images (unadulterated samples: 21 and adulterated samples: 33). In this case, the adulterations were prepared by addition of soybean oil in the proportion of 0.5%, 1.0%, 2.0%, 3.0%, 4.0%, 5.0%, 6.0%, 7.0%, 8.0%, 9.0% and 10.0% (mm^{-1}).

The images were recorded using a webcam with HD (1280 × 720) resolution (primary instrument) and a smartphone with resolution of 1536 × 2560 (secondary instrument). The system employed in the digital images acquisition and the procedures for obtaining the histograms are described in [13]. The smartphone captured images at a distance of 20 cm from the quartz flow cell.

The RGB model (R - red, G - green and B - blue) for pixel color was used in this study. Each color component of the model is composed of 256 color levels, varying from 0 to 255 for each channel. The data matrix is formed by samples located in rows, while the columns are constituted of variables corresponding to the color levels obtained for each color component. All color components were employed to develop the SPA-LDA and PLS-DA classification models.

Table 1
Number of training and prediction samples in each class for the two data sets. *N* indicates the number of samples in each class.

Class	*N*	Fluorescence data set		*N*	Digital images data set	
		Subsets			Subsets	
		Training	Prediction		Training	Prediction
Unadulterated	48	34	14	21	14	7
Adulterated	40	28	12	33	23	10
Total	88	62	26	54	37	17

2.4. Chemometric procedures and software

The chemometric procedures described below were performed individually for each analytical technique. For both data sets, the samples analyzed in the primary instrument were divided into training (70%) and prediction (30%) sets using the classic Kennard-Stone (KS) algorithm [29]. Table 1 presents the number of training and prediction samples in each class for both data sets.

The training samples were used in the modeling procedure (cross-validation), including SPA variable selection for LDA and determination of factors (latent variables) in PLS-DA, whereas the prediction set (measured on both instruments) were used in the final evaluation of the classification models. The LDA models were developed based on the variables selected by the successive projections algorithm (SPA) adapted for internal validation [30]. For the PLS-DA model, a threshold value of 0.5 was adopted. When a value above 0.5 is predicted, the sample is considered as belonging to the class under study, while a value below 0.5 indicates that the sample does not belong.

In order to select subsets of transfer samples from the training set measured on both instruments, the KS algorithm was also used. Different numbers of transfer samples (4, 6, 8 and 10) were investigated.

Two strategies of standardization were employed: direct standardization (DS) and piecewise direct standardization (PDS) of the prediction set measured on the secondary instruments. The PDS algorithm was run with different window sizes (3, 5, 7, 9 and 11). DS and PDS algorithms were performed using PLS Toolbox (version 3.5) for Matlab.

The results obtained by the SPA-LDA and PLS-DA models developed from the primary instrument were compared in terms of the number of errors and correct classification rate (CCR) for the prediction set measured on secondary instruments before and after standardization. In addition, the numbers of transfer samples employed in each model were also evaluated.

In the discussions of the results, the abbreviation prediction $^{P}_{P}$ will be used to represent the prediction set measured on the primary instrument, while prediction $^{S}_{P}$ to prediction set measured on the secondary instrument data.

The analytical information contained in the histograms of the digital images was obtained using a graphic interface coded in Matlab® 7.1. KS and SPA-LDA algorithms were coded in Matlab (Mathworks, USA). PLS-DA was carried out using Unscrambler X.1 (CAMO S/A).

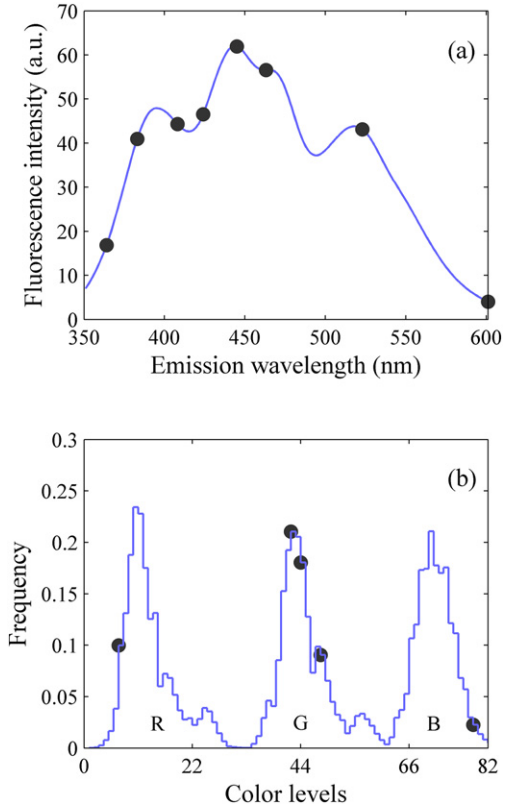

Fig. 2. (a) Average spectrum with the wavelengths selected by the SPA for fluorescence data. (b) Average histograms with the color levels selected by the SPA for digital images data.

Table 2
Classification results obtained by SPA-LDA and PLS-DA models in the prediction set. The numbers of variables employed in the SPA-LDA model and latent variables used in the PLS-DA model are indicated in parenthesis. *N* indicates the number of prediction samples employed in this study.

Parameters	*N*	Classification results for *fluorescence data*	
		SPA-LDA (8)	PLS-DA (4)
Prediction $^{P}_{P}$ errors	26	3	3
Prediction $^{S}_{P}$ errors (raw data)	26	12	12
Prediction $^{S}_{P}$ errors (preprocessed data)	26	12	12

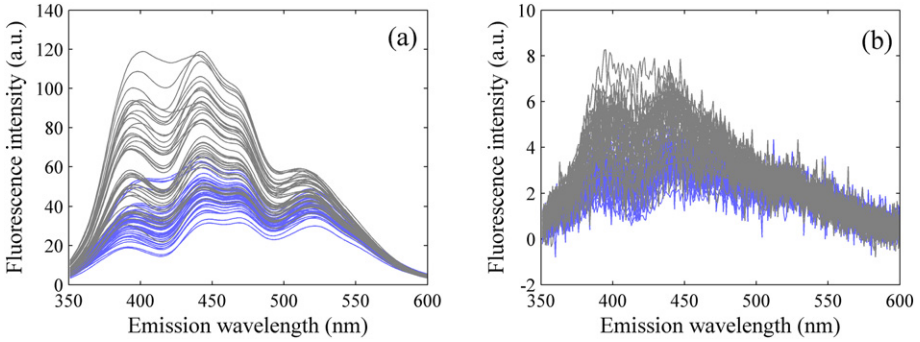

Fig. 1. Fluorescence spectra of the EVOO samples acquired in the (a) primary and (b) secondary instruments. Unadulterated (—) and adulterated (—) samples.

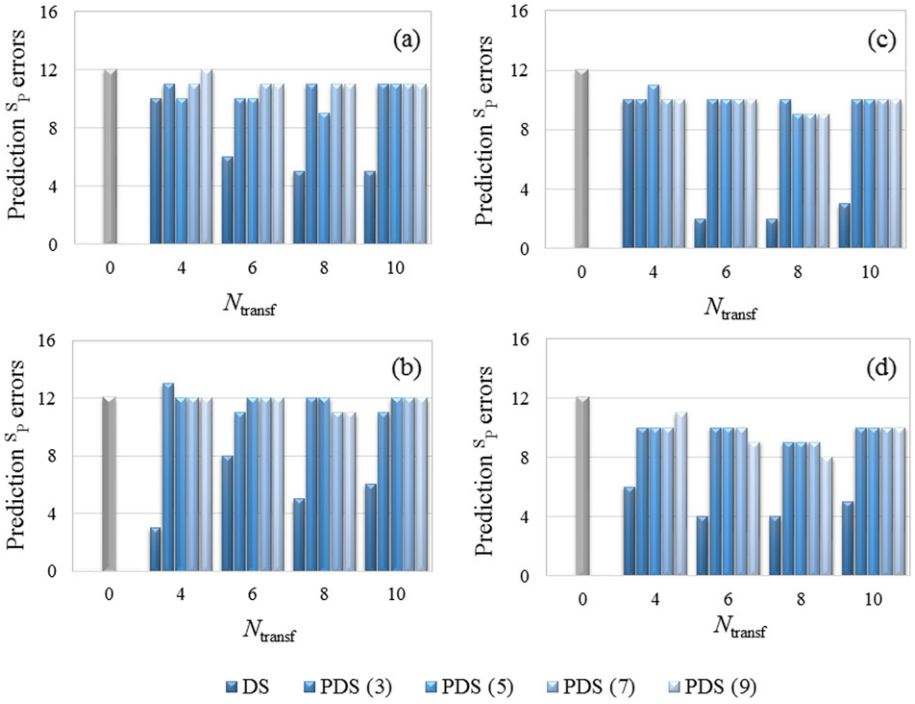

Fig. 3. Results obtained by SPA-LDA (a and b) and PLS-DA (c and d) models applied to the raw (a and c) and preprocessed (b and d) spectra of the prediction $^{S}_{P}$ samples before ($N_{transf} = 0$) and after performing the DS and PDS, respectively. The window sizes employed in the PDS standardization method is indicated in parenthesis.

3. Results and discussion

3.1. Fluorescence spectroscopy

Fig. 1 presents the fluorescence spectra of EVOO samples acquired in each instrument. Comparing the two spectra (Fig. 1a and b), we can observe a substantial difference in signal quality, which for the secondary instrument (Fig. 1b) is noisy and has low signal intensity. In the spectral range investigated, the fluorescence of EVOOs has been attributed mainly to the oxidation products of fatty acids (two smooth peaks at 445 and 475 nm) and vitamin E (525 nm) [31].

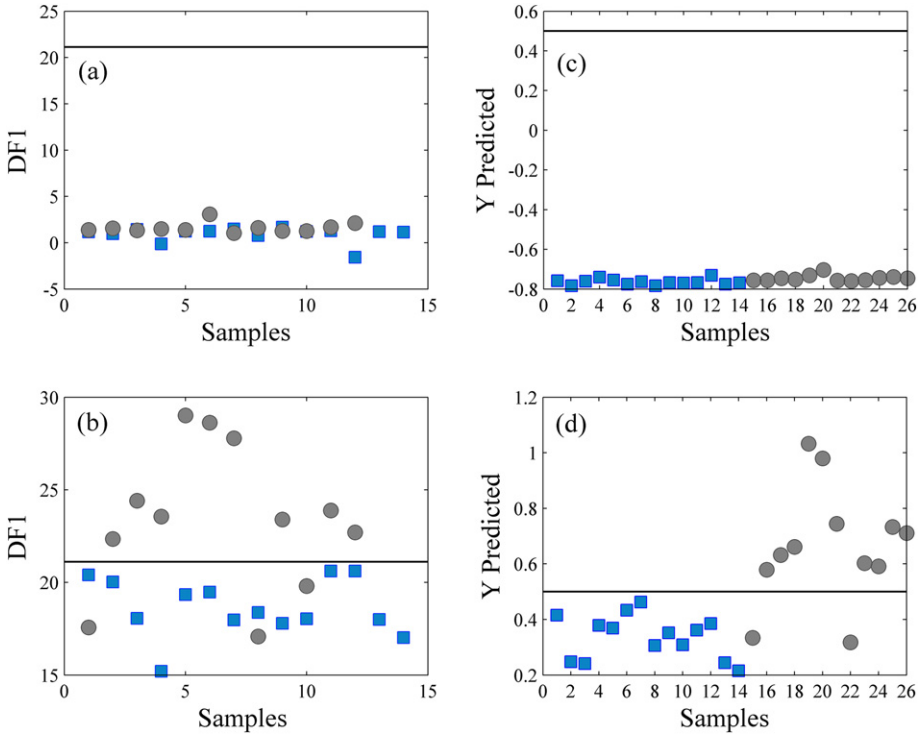

Fig. 4. Score plots obtained by SPA–LDA model for the prediction $^{S}_{P}$ samples before (a) and after (b) DS of the preprocessed data using $N_{transf} = 4$. Values estimated by the PLS–DA model for the prediction $^{S}_{P}$ samples before (c) and after (d) DS of the raw data using $N_{transf} = 6$. Unadulterated (■) and adulterated (●) samples measured by fluorescence spectroscopy. The horizontal lines are the decision boundaries established by SPA-LDA and PLS-DA models.

Fig. 2a shows the wavelengths selected by SPA algorithm applied to the data set of the primary instrument. These wavelengths were used to build models based on linear discriminant analysis (LDA). As can be observed in Fig. 2a, wavelengths were selected along the entire spectral range used.

Table 2 shows the number of errors obtained by SPA-LDA and PLS-DA models applied to the prediction set measured on both instruments. For the prediction set measured by a secondary instrument, prediction $^{S}_{P}$, the results were evaluated for the raw and preprocessed data. The pre-processing performed was the first-derivative using smoothing by a Savitzky–Golay filter [32] with a second-order polynomial and a 25-point window.

As can be seen in Table 2, when the primary models are applied to the secondary instrument data, there is an increase of the number of prediction $^{S}_{P}$ errors both in the raw and preprocessed data, indicating the need for a classification transfer procedure.

Fig. 3 presents the results, in terms of the number of errors, obtained by the SPA-LDA (Fig. 3a and b) and PLS-DA (Fig. 3c and d) models applied to the raw (Fig. 3a and c) and preprocessed (Fig. 3b and d) spectra of the prediction $^{S}_{P}$ samples before and after performing the standardization procedures. The bars indicate the number of prediction $^{S}_{P}$ errors. The first bar, in $N_{transf} = 0$, correspond to classification results obtained before standardization. The remaining bars correspond to results obtained by DS and PDS, respectively. In this case study, the PDS algorithm was run with window sizes 3, 5, 7 and 9.

As can be seen in Fig. 3, for both models, different results were obtained when different N_{transf} and window sizes were employed. For SPA-LDA model, the best performance was obtained after the DS of the preprocessed data (Fig. 3b) using $N_{transf} = 4$, with only 3 samples incorrectly classified. The PLS-DA model achieved the best classification result (2 misclassifications) for the raw data (Fig. 3c) also after the DS, with $N_{transf} = 6$ and $N_{transf} = 8$.

Fig. 4 compares the score plots obtained by the SPA–LDA (Fig. 4a and b) and estimated values by PLS–DA models (Fig. 4c and d) applied to the prediction set measured by the secondary instrument before and after standardization. Fig. 4b and d correspond to DS of the preprocessed (for SPA-LDA) and raw (for PLS-DA) data using $N_{transf} = 4$ and $N_{transf} = 6$, respectively. As can be seen, the classification models performed better when standardized spectra were employed.

Table 3
Classification results obtained by SPA-LDA and PLS-DA models in the prediction set. The numbers of variables employed in the SPA-LDA model and latent variables used in the PLS-DA model are indicated in parenthesis. *N* indicates the number of prediction samples employed in this study.

Parameters	*N*	Classification results for *digital images data*	
		SPA-LDA (5)	PLS-DA (6)
Prediction $^{P}_{P}$ errors	17	2	2
Prediction $^{S}_{P}$ errors	17	9	9

3.2. Digital images

Fig. 5 shows the RGB histograms obtained for the 54 EVOO samples measured on both devices. For each color component, 256 color levels can be obtained varying from 0 to 255. However, many of these presented a response equal to zero (not participating in the image color composition). These color levels were removed from the data set so that only the resulting histograms are shown (Fig. 5a and b) and used in the chemometric procedure. In Fig. 5, the total number of color levels after removal is presented on the x-axis, while their corresponding values on the 0–255 scale are listed for the RGB components.

As can be seen in Fig. 5, different numbers of color levels were obtained in each instrument. In addition, it is possible to observe that the difference between the two samples classes is more evident in the primary instrument, especially for R and G components.

Fig. 2b shows the color levels selected by SPA algorithm applied to the data set of the primary instrument. As can be seen in Fig. 2b, most of the selected color levels are in the G component, where the difference between the two classes of samples is more evident.

Table 3 presents the detailed classification results obtained by SPA-LDA and PLS-DA models applied to the prediction set measured on both instruments. Comparing the prediction $^{P}_{P}$ errors with prediction $^{S}_{P}$ errors, an increase of the value when the models are applied to the secondary instrument data can be observed.

Fig. 6 presents the results in terms of the number of errors obtained by the SPA-LDA and PLS-DA models applied to prediction $^{S}_{P}$ samples before ($N_{transf} = 0$) and after performing the standardization procedures.

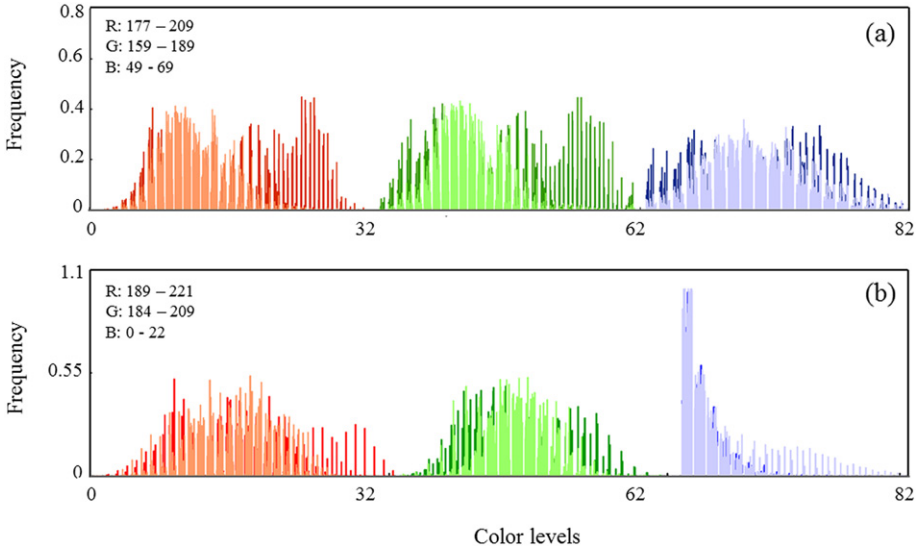

Fig. 5. RGB histograms of the EVOO samples acquired in the (a) primary and (b) secondary instruments. The unadulterated samples are represented by darker shades of each color component, while the adulterated are represented by lighter shades. The values listed for the RGB components correspond to the color levels present in the histograms.

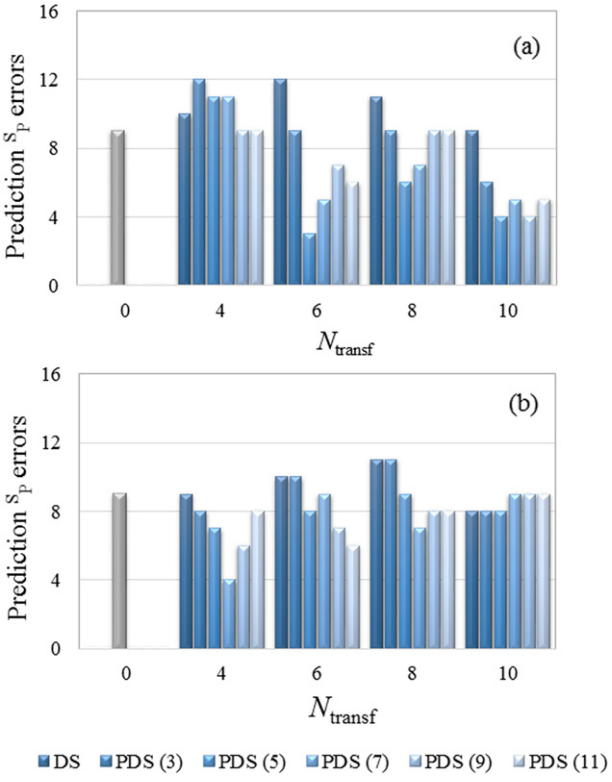

Fig. 6. Results in terms of prediction $^{S}_{P}$ errors obtained by the SPA–LDA (a) and PLS–DA (b) models before ($N_{transf} = 0$) and after performing the DS and PDS, respectively. The window sizes employed in the PDS standardization method is indicated in parenthesis.

In this case study, the PDS method was evaluate for window sizes 3, 5, 7, 9 and 11. Different from the previous data set, the best classification result involving digital images was obtained after piecewise direct standardization (PDS), with $N_{transf} = 6$ and a 5-point window for the SPA-LDA model. The least number of errors obtained after the PDS for the digital images can be related to the fact that this data set is formed by discrete variables. Given this, the standardization of the windows of the variables proved to be more efficient.

Fig. 7 shows the score plots of the discriminant function (DF1) obtained by SPA-LDA (Fig. 7a and b) and estimated values by PLS-DA (Fig. 7c and d) models applied to the prediction $^{S}_{P}$ set before and after standardization. As can be seen, the classification models performed better when standardized spectra were employed.

Table 4 summarizes the classification results obtained by models applied to the prediction $^{S}_{P}$ samples before and after standardization procedure. When the transformed spectra and histograms are employed, a substantial increase in the correct classification rate (CCR) can be observed for both models.

4. Conclusion

In this paper, we evaluated the use of transfer multivariate classification models to identify adulteration of extra virgin olive oils by fluorescence spectroscopy and digital imaging. The final results of the SPA-LDA and PLS-DA classification models were compared before and after standardization procedures in terms of the correct classification rate (CCR) for the prediction set measured by the secondary instrument. Direct standardization (DS) and piecewise direct standardization (PDS) were performed using different number of transfer samples and window sizes.

A CCR for the SPA-LDA classification models, before standardization, achieved 54% for fluorescence emission spectra and 47% for histograms of digital images. After the standardization, a substantial increase of the CCR was observed, obtaining 88% for the fluorescence emission spectra and 82% for the histograms of the digital images.

The PLS-DA classification models found the same CCR as obtained by the SPA-LDA models before standardization using the two analytical

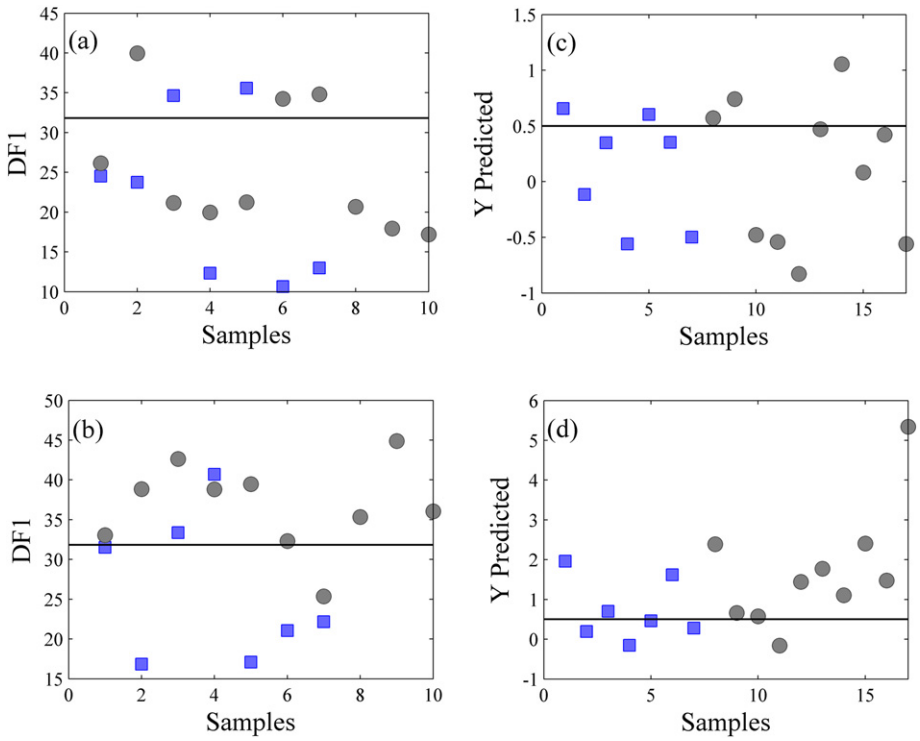

Fig. 7. Score plots obtained by SPA–LDA model for the prediction $^{S}_{P}$ samples before (a) and after (b) PDS with 5-point window using $N_{transf} = 6$. Values estimated by the PLS–DA model for the prediction $^{S}_{P}$ samples before (c) and after (d) PDS with 7-point window using $N_{transf} = 4$. Unadulterated (■) and adulterated (●) samples measured by digital images. The horizontal lines are the decision boundaries established by SPA-LDA and PLS-DA models.

Table 4
Correct classification rate obtained by SPA–LDA and PLS–DA models applied to prediction $^{S}_{P}$ set measured by the secondary instruments before and after standardization.

Models	Correct classification rate			
	Fluorescence data set		Digital images data set	
	Before standardization	After standardization	Before standardization	After standardization
SPA-LDA	54%	88%	47%	82%
PLS-DA	54%	92%	47%	76%

techniques. After standardization, the values were exceeded, reaching 85% and 76% CCR for the fluorescence and image data, respectively.

These results demonstrate the efficiency of standardization procedures applied to multivariate classification models developed from fluorescence spectroscopy and digital images.

Acknowledgements

The authors acknowledge the support of CNPq (303649/2015-1 and 162930/2013-5), UNS – Argentina (Universidad Nacional del Sur, PGI-UNS 24Q054) and CONICET – Argentina (Consejo Nacional de Investigaciones Científicas y Técnicas, PIP code 11220120100625).

References

[1] M.R. Almeida, C.H.V. Fidelis, L.E.S. Barata, R.J. Poppi, Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation, Talanta 117 (2013) 305–311.
[2] O.M.O. Kruse, J.M. Prats-Montalbán, U.G. Indahl, K. Kvaal, A. Ferrer, C.M. Futsaether, Pixel classification methods for identifying and quantifying leaf surface injury from digital images, Comput. Electron. Agric. 108 (2014) 155–165.
[3] V.A.G. Silva, M. Talhavini, I.C.F. Peixoto, J.J. Zaca, A.O. Maldaner, J.W.B. Braga, Non-destructive identification of different types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis, Microchem. J. 116 (2014) 235–243.
[4] W.T.S. Vilar, R.M. Aranha, E.P. Medeiros, M.J.C. Pontes, Classification of individual castor seeds using digital imaging and multivariate analysis, J. Braz. Chem. Soc. 26 (2015) 102–109.
[5] M.I.S. Gonçalves, W.T.S. Vilar, E.P. Medeiros, M.J.C. Pontes, A novel strategy for the classification of naturally colored cotton fibers based on digital imaging and pattern recognition techniques, Anal. Methods 7 (2015) 5869–5875.
[6] A.L.B. Brito, D.A. Araújo, M.J.C. Pontes, L.F.B.L. Pontes, Near infrared reflectance spectrometry classification of lettuce using linear discriminant analysis, Anal. Methods 7 (2015) 1890–1895.
[7] A.G. González, Use and misuse of supervised pattern recognition methods for interpreting compositional data, J. Chromatogr. A 1158 (2007) 215–225.
[8] S. Wold, Pattern recognition by means of disjoint principal component models, Pattern Recogn. 8 (1976) 127–139.
[9] Y. Gao, B. Zheng, G. Chen, W. Lee, K. Lee, Q. Li, IEEE Trans. Knowl. Data Eng. (2009) 1314–1327.
[10] R. Brereton, Chemometrics for Pattern Recognition, John Wiley & Sons, Chichester, 2007 (ISBN: 978-0-470-98725-4).
[11] R.A. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugenics 7 (1936) 179–188.
[12] U.T.C.P. Souto, M.F. Barbosa, H.V. Dantas, A.S. Pontes, W.S. Lyra, P.H.G.D. Diniz, M.C.U. Araújo, E.C. Silva, Identification of adulteration in ground roasted coffees using UV-Vis spectroscopy and SPA-LDA, LWT Food Sci. Technol. 63 (2015) 1037–1041.
[13] K.D.T.M. Milanez, M.J.C. Pontes, Classification of extra virgin olive oil and verification of adulteration using digital images and discriminant analysis, Anal. Methods 7 (2015) 8839–8846.
[14] M.R. Almeida, D.N. Correa, W.F.C. Rocha, F.J.O. Scafi, R.J. Poppi, Discrimination between authentic and counterfeit banknotes using Raman spectroscopy and PLS-DA with uncertainty estimation, Microchem. J. 109 (2013) 170–177.
[15] J.F. Povey, C.J. O'Malley, T. Root, E.B. Martin, G.A. Montague, M. Feary, C. Trim, D.A. Lang, R. Alldread, A.J. Racher, C.M. Smales, Rapid high-throughput characterisation, classification and selection of recombinant mammalian cell line phenotypes using intact cell MALDI-ToF mass spectrometry fingerprinting and PLS-DA modelling, J. Biotechnol. 184 (2014) 84–93.
[16] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferré, Transfer of multivariate calibration models: a review, Chemom. Intell. Lab. Syst. 64 (2002) 181–192.
[17] E. Bouveresse, D.L. Massart, Standardisation of near-infrared spectrometric instruments: a review, Vib. Spectrosc. 11 (1996) 3–15.
[18] Y. Hu, S. Peng, Y. Bi, L. Tang, Calibration transfer based on maximum margin criterion for qualitative analysis using Fourier transform infrared spectroscopy, Analyst 137 (2012) 5913–5918.
[19] F.A. Honorato, R.K.H. Galvão, M.F. Pimentel, B. de Barros Neto, M.C.U. Araújo, F.R. Carvalho, Robust modeling for multivariate calibration transfer by the successive projections algorithm, Chemom. Intell. Lab. Syst. 76 (2005) 65–72.
[20] R.K.H. Galvão, S.F.C. Soares, M.N. Martins, M.F. Pimentel, M.C.U. Araújo, Calibration transfer employing univariate correction and robust regression, Anal. Chim. Acta 864 (2015) 1–8.
[21] F.A. Honorato, B. de Barros Neto, M.N. Martins, R.K.H. Galvão, M.F. Pimentel, Transferência de calibração em métodos multivariados, Quim. Nova 30 (2007) 1301–1312.
[22] A.J. Myles, T.A. Zimmerman, S.D. Brown, Transfer of multivariate classification models between laboratory and process near-infrared spectrometers for the discrimination of green Arabica and Robusta coffee beans, Appl. Spectrosc. 60 (2006) 1198–1203.
[23] C.V. Di Anibal, I. Ruisánchez, M. Fernández, R. Forteza, V. Cerdà, M.P. Callao, Standardization of UV–visible data in a food adulteration classification problem, Food Chem. 134 (2012) 2326–2331.
[24] N.C. Silva, M.F. Pimentel, R.S. Honorato, M. Talhavini, A.O. Maldaner, F.A. Honorato, Classification of Brazilian and foreign gasolines adulterated with alcohol using infrared spectroscopy, Forensic Sci. Int. 253 (2015) 33–42.
[25] K.D.T.M. Milanez, A.C. Silva, J.E.M. Paz, E.P. Medeiros, M.J.C. Pontes, Standardization of NIR data to identify adulteration in ethanol fuel, Microchem. J. 124 (2016) 121–126.
[26] PROTESTE - Associação Brasileira de Defesa do Consumidor, http://www.proteste.org.br/azeite 2016 (accessed 09/16).
[27] O. Zenebon, N.S. Pascuet, P. Tiglea, Métodos físico-químicos para análise de alimentos, fourth ed. Instituto Adolfo Lutz, São Paulo, 2008.
[28] International Olive Oil Council, Trade standard applying to olive oils and olive-pomace oils, COI/T.15/NC N°3/Rev. 8, http://www.internationaloliveoil.org/estaticos/view/222-standards 2015 (accessed 03/15).
[29] R.W. Kennard, L.A. Stone, Computer aided design of experiments, Technometrics 11 (1969) 137–148.
[30] S.F.C. Soares, R.K.H. Galvão, M.J.C. Pontes, M.C.U. Araújo, A new validation criterion for guiding the selection of variables by the successive projections algorithm in classification problems, J. Braz. Chem. Soc. 25 (2014) 176–181.
[31] E. Guzman, V. Baeten, J.A.F. Pierna, J.A. Garcia-Mesa, Evaluation of the overall quality of olive oil using fluorescence spectroscopy, Food Chem. 173 (2015) 927–934.
[32] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639.