

Data fusion strategy in quantitative analysis of spectroscopy relevant to olive oil adulteration

Yang Li, Yanmei Xiong*, Shungeng Min*

College of Science, China Agricultural University, Beijing, 100193, PR China

ARTICLE INFO

Keywords:

NIR
MIR
Data fusion
SPA
Olive oil
Adulteration

ABSTRACT

Olive oil adulteration with various less expensive edible oils represents a great danger for consumers. Spectrometry has been used to detect olive oil adulteration with other oil, but we need more robust and accurate model. Therefore, this work investigated the combination of infrared (NIR) and mid infrared (MIR) spectroscopy for the quantification of rapeseed oil in olive oil blends. Furthermore, a partial least squares (PLS) model was established to predict the concentration of the adulterant. Models constructed using baseline correction by combination of standard normal variate (SNV), SG smoothing and vector normalization pretreatments, respectively. Three data fusion strategies (low, mid and high-level) have been applied to take advantage of the synergistic effect of the information obtained from NIR and MIR. We chose algorithm (SPA) to extract spectral features for mid-level data fusion. Binary linear regression used in high-level data fusion. We selected the best pretreatment for final evaluation according to the evaluation parameters (R^2 of calibration and validation, RMSECV and RMSEP). NIR, MIR and data fusion models were evaluated by comparing the R^2 of validation and RMSEP (root mean square error of prediction). The RMSEP of low-level (3.44), high-level (2.86) data fusion were better than NIR(7.09), MIR(4.04), mid-level(6.09) and the validation coefficient of determination R^2 of low-level data fusion (0.975) and high-level data fusion (0.988) are better than the NIR (0.896) and MIR (0.966). Results showed that: (1) NIR and MIR are fast and non-destructive testing tools to detect the extra-virgin olive oil adulteration with rapeseed oil. (2) Low-level data fusion can effectively improve model prediction accuracy. (3) SPA reduced the number of variables, but it did not improved the results. (4) High-level data fusion strategy can be used as a reliable tool for quantitative analysis.

1. Introduction

Virgin olive oil (VOO) is a valuable and highly appreciated vegetable oil [1]. It is extracted from fresh and healthy olive fruits. Extra virgin olive oil is popular because it contains aldehydes, alcohols, esters, hydrocarbons, ketones furans and other volatile substances [2]. However, oils of higher quality can be intentionally blended with cheaper vegetable oils or olive oil grades and The Food Fraud Database shows that oils such as corn, sunflower, sesame, and nuts are often used as adulterants, for this reason, the authenticity of olive oil has currently become an important issue for producers, consumers, and quality control evaluators. Various chemical and biochemical techniques have been developed for determining the authenticity of olive oil. A number of technologies, including vibrational spectroscopic techniques [3–5], nuclear magnetic resonance spectroscopy [5], headspace-mass spectrometry [6], rapid synchronous fluorescence method [3], and high-performance liquid chromatography have been used to detect the

adulteration of olive oil [7]. Vibrational spectroscopy, such as NIR, MIR and Raman, is an alternative to quantify the liquid systems due to its simplicity, rapidness and affordability [8,9]. There are many studies on olive oil adulteration by infrared spectroscopy combined with chemometrics methods. Thiago O. Mendes et al. (2015) used NIR, MIR, and Raman spectroscopy associated with PLS for the quantitative analysis of olive oil blends with soybean oils. Li Wang et al. [10] used attenuated total reflectance MIR and fiber optic diffuse reflectance NIR combined with SIMCA and PLS for the qualitative and quantitative analysis of adulterated oil. Jović et al. [11] developed dwPLS for quantitative prediction and variable-selection on spectroscopic data, MIR-ATR combined with dwPLS could be implemented to quantitative determination of edible-oil adulteration. NIR and MIR spectra offer significant ‘fingerprint’ information about individual components in complex samples. The peaks in the NIR region ($10,000\text{ cm}^{-1}$ to 4000 cm^{-1}) are broad and weak, and are combinations and overtones of the sample functional groups. The MIR spectra region (4000 cm^{-1} to 400 cm^{-1}),

* Corresponding authors.

E-mail addresses: xiongyim@cau.edu.cn (Y. Xiong), minsg@cau.edu.cn (S. Min).

<https://doi.org/10.1016/j.vibspec.2018.12.009>

Received 30 September 2018; Received in revised form 7 November 2018; Accepted 20 December 2018

Available online 02 January 2019

0924-2031/ © 2018 Published by Elsevier B.V.

mainly provides information from frequencies of fundamental molecular vibrations and generally exhibits sharp absorption bands and distinct spectral features. NIR and MIR may provide different chemical information, and data fusion full use of those information and combine the advantages of various kinds of instrument technology [12,13]. In order to fully use the advantages of both near-infrared (NIR) and mid-infrared (MIR) spectra, different data fusion methods were adopted in this study. We evaluated the low-, mid-, and high-level data fusion strategies by comparing the determination coefficient (R^2) of validation and the RMSEP.

Data fusion systems are now widely used in various areas such as sensor networks [14], robotics [15], image processing [16], intelligent system design [17] and food analysis [18], etc. The data fusion strategy is increasingly used in the fields of food quality assessment, such as food adulteration [19,20]. It can be used to combine the outputs and complementary information provided by different instruments to generate better inferences than that from a single technique. Basically, it can be done at three levels: low-level, mid-level, and high-level [21]. As a case study, A. Dankowska et al. [22] applied fluorescence, UV–vis spectroscopies as well as the low- and mid-level data fusion of both spectroscopies for the quantification of concentrations of roasted *Coffea arabica* and *Coffea canephora* var. *robusta* in coffee blends. Mid-level data fusion of UV–vis and fluorescence gave quantitative information, and low-level data fusion model have good qualitative effect. [Our study found that low-level data fusion has better quantitative effect than mid-level data fusion, the results were influenced by the feature selection method. Sun et al. [13] fused the data obtained from two spectroscopic techniques (MIR and NIR spectroscopy) at low level and mid-level to distinguish official rhubarbs and unofficial rhubarbs. Márquez et al. [19] used two data fusion strategies (high and mid-levels) combined with a multivariate classification approach to identify the adulteration of hazelnut paste with almond. All the data fusion strategies effectively improved the performance compared with the use of individual techniques. In the quantitative research on food issues, few researchers used high-level data fusion strategies to improve the accuracy of quantitative models.

The aim of our study was, on the one hand, to investigate the spectroscopic techniques combined with chemometrics methods of quantitative analysis of the olive oil adulteration, and on the other hand, to evaluate the three data fusion methods. We implemented high-level data fusion, which communication bandwidth savings and improved decision accuracy [23]. To date, few researchers have used high-level data fusion to improve the accuracy of quantitative research on food issues.

2. Material and methods

2.1. Sample preparation

Two different brands of rapeseed oils and two different brands of extra-virgin olive oils were bought from local supermarkets. The authenticity of the samples were approved and authorized by the local governmental food regulatory agencies. The final rapeseed oil was a mixture of two equal volumes of two different brands of rapeseed oils. The final extra-virgin olive oil was a mixture of two equal volumes of two different brands of Olive oil. The amount of rapeseed oil as the adulterate in adulteration samples ranged from 1 to 80% (w/w). (Table 1). One sample was configured for each proportion.

2.2. Instrumentation

2.2.1. NIR

Near infrared spectra were acquired using a NIR Spectrometer (PerkinElmer, Spectrum One NTS, Perkinelmer business management, Shanghai, China), over 4000–10,000 cm^{-1} , the resolution was 16 cm^{-1} and every spectrum was scanned 32 times. Each sample was recorded

Table 1

Percentage composition(% w/w) of rapeseed oil.

Sample number	mass(rapeseed oil)	mass(Olive oil)	ratio
1	0.15	14.85	1%
2	0.45	14.55	3%
3	0.75	14.25	5%
4	1.2	13.8	8%
5	1.5	13.5	10%
6	1.95	13.05	13%
7	2.25	12.75	15%
8	2.7	12.3	18%
9	3	12	20%
10	3.45	11.55	23%
11	3.75	11.25	25%
12	4.05	10.95	27%
13	4.5	10.5	30%
14	4.8	10.2	32%
15	5.25	9.75	35%
16	5.55	9.45	37%
17	6	9	40%
18	6.3	8.7	42%
19	6.75	8.25	45%
20	7.05	7.95	47%
21	7.5	7.5	50%
22	7.95	7.05	53%
23	8.25	6.75	55%
24	8.55	6.45	57%
25	8.85	6.15	59%
26	9.15	5.85	61%
27	9.45	5.55	63%
28	9.75	5.25	65%
29	10.05	4.95	67%
30	10.35	4.65	69%
31	10.65	4.35	71%
32	10.95	4.05	73%
33	11.25	3.75	75%
34	11.55	3.45	77%
35	11.85	3.15	79%
36	12	3	80%

Table 2

Characteristics of calibration and validation sets for rapeseed oil content in adulteration oil.

	Samples	Rapeseed oil content(%)	Mean	stdev ^a
Calibration set	24	1%-80%	46.54	24.77
Validation set	12	3%-73%	34.42	22.95

^a Standard deviation.

three times and the corresponding average spectra were calculated and used for all data treatment. Between samples, the quartz cuvette was flushed once with acetone then purged by air to dryness. The experimental temperature during the whole experiment was maintained at room temperature. The background spectra were collected on an empty quartz cuvette. The background spectrum was collected before every sample was measured. Software used for spectral collected is QUANTC Software.

2.2.2. MIR

MIR spectra were acquired using a PerkinElmer, Spectrum 100 MIR spectrometer (Perkinelmer business management, Shanghai, China) equipped with a DTG detector and flat zinc selenide single reflectance crystal using blunt thin glass rod as described in the reference. Each spectrum was recorded in the region of 4000–650 cm^{-1} using an average of 64 scans at a resolution of 8 cm^{-1} . The experimental temperature during the whole experiment was maintained at room temperature. Each sample was recorded three times and the corresponding average spectra were calculated and used for data treatment. Between samples, the crystal was cleaned with dichloromethane, ethanol, distilled water and then purged with air to dryness. The background was

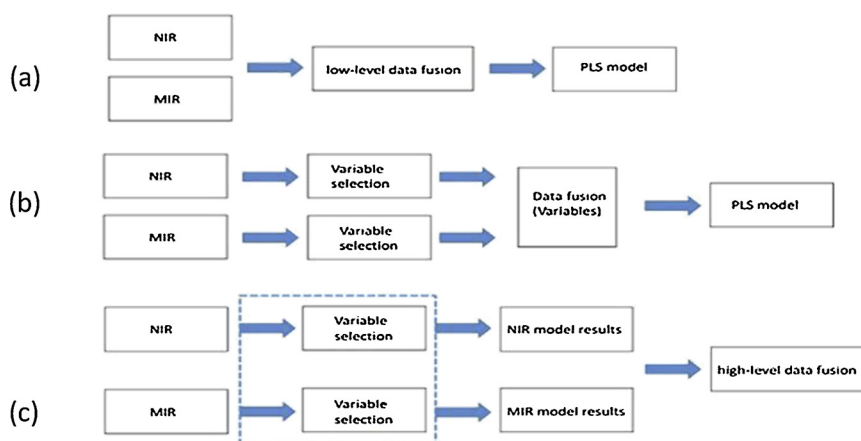


Fig. 1. Scheme of the data fusion process: (a)low-level data fusion; (b)mid-level data fusion; (c)high-level data fusion.

Table 3

Predicted values for validation data set by PLS regression (All values represent a content of addition of rapeseed oil in adulteration oil).

Validation set	NIR	MIR	Low-level	Mid-level(SPA)	High-level
3.0	16.2	0.4	7.4	−4.8	10.9
8.0	16.4	3.9	5.1	−4.6	14.5
15.0	11.5	7.5	10.0	10.0	21.3
20.0	3.2	23.0	16.7	25.6	26.7
25.0	23.4	24.7	26.6	20.0	29.2
27.0	25.0	20.3	20.8	17.7	33.5
32.0	34.4	25.3	31.4	29.9	37.2
37.0	34.2	33.8	32.8	33.6	42.8
45.0	50.2	44.5	46.2	47.6	48.6
61.0	60.3	62.0	63.4	55.1	60.2
67.0	65.2	66.1	63.6	66.8	65.7
73.0	69.5	75.2	72.3	72.6	69.4

collected before every sample was measured. Software used for spectral collection is PerkinElmer Spectrum 10.

2.2.3. Software

Unscramble (version 9.7 CAMO Software AS, Camo Analytics, Norwegian, Kongeriket Norge) was used for NIR and MIR data pretreatment. Matlab (version R2014a, The MathWorks, USA) was used for data processing.

2.2.4. Chemometrics methods

Quantification of rapeseed oil adulteration levels was calculated by PLS regression analysis in Matlab. For chemometric evaluation, the samples of all the adulteration oils were divided into a calibration and a validation set. To enable comparison of the spectral techniques and

data fusion strategy the same sets of samples were used for both spectral techniques. The division into sets was done in order to obtain similar mean values and standard deviations so that both sets spanned the full range of rapeseed oil contents [24] (Table 2).

Several spectral pretreatments including baseline correction, baseline correction + Standard normal variate (SNV), baseline correction + Savitzky-Golay(SG) smoothing(9) and baseline correction + Vector normalization were investigated for the optimization of the calibration model [25]. Savitzky-Golay (SG) smoothing is able to attenuate noise signals and preserve the shape, height, width, and valuable features of the chemical analyte signals [26,27]. SNV has commonly been used to remove the multiplicative interferences of scattering and particle size from IR spectra [27]. Normalization handles the data mapping to the range of 0–1; the step makes the modeling faster and more convenient [28].

The optimum number of PLS factors was determined by five-fold cross-validation employing cancellation one standard at a time by plotting the number of factors against the root mean square error of cross validation (RMSECV) and determining the minimum. The relative performance of the established model was accessed by the required number of factors, the root mean square error of calibration (RMSEC) and the RMSECV, and its predictive ability was evaluated from the root mean square error of prediction (RMSEP) [29].

The scheme of data fusion in this experiment is shown in Fig. 1. In low-level fusion, raw data from all sources are straight forwardly concatenated and then processed as a particular fingerprint of the samples [21]. NIR and MIR data were simply concatenated and the same pre-processing previously mentioned was used for each data set.

In mid-level fusion, also called “feature-level fusion”, informative features of the raw data from each instrument are separately extracted

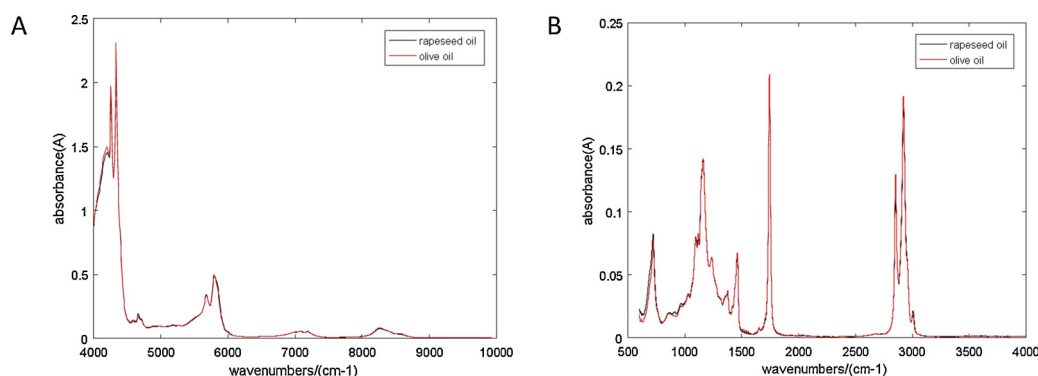


Fig. 2. Spectra of olive oil (red line) and rapeseed oil (black line). A NIR spectra with absorbance value at 10,000–4000 cm^{-1} , B MIR spectra with transmittance value at 4000–650 cm^{-1} .

Table 4
Main components of olive oil and Rapeseed oil(w/w).

	Olive oil	Rapeseed oil
Oleic acid (%)	75.0	40.6
Linoleic acid (%)	9.0	22.8
Flax acid (%)	1.0	3.2
Stearic acid, Palm acid and others (%)	15	5.7

by several kinds of variable selection or feature-extraction algorithm, then aligned and concatenated into a single matrix that is used for multivariate analysis [30]. We used SPA to extract feature from the full NIR and MIR spectra with the same pretreatment, and the mid-level PLS model was established by fusing the spectral points selected by SPA [31]. The number of SPA selected points from NIR and MIR were both 19, the distribution of characteristic variables selected by SPA is shown in Fig. 4.

In high-level fusion, the fusion occurs at the decision level. Each source datum is calculated by a separate multivariate analysis model, and the results of each individual model are integrated to obtain an “ensemble decision” [25]. We used binary linear regression to solve the advanced fusion problem [32]. The binary linear regression model was as follows, the least square method is used for parameter estimation:

$$y_{i, \text{actual value of calibration}} = b_0 + b_1 x_{\text{NIR, predicted of calibration}} + b_2 x_{\text{MIR, predicted of calibration}}$$

$$y_{i, \text{high-level data fusion predicted value of calibration}} = b_0 + b_1 x_{\text{NIR, predicted value of calibration}} + b_2 x_{\text{MIR, predicted of value of calibration}}$$

$$y_{i, \text{high-level data fusion predicted value of validation}} = b_0 + b_1 x_{\text{NIR, predicted of validation}} + b_2 x_{\text{MIR, predicted of validation}}$$

$y_{i, \text{actual}}$ is the true value of calibration set;

$y_{i, \text{high-level data fusion predicted value of validation}}$ is the high-level data fusion prediction value;

$x_{\text{NIR, predicted of calibration}}$ is the prediction value of NIR calibration set from 5 fold cross validation;

$x_{\text{MIR, predicted of calibration}}$ is the prediction value of MIR calibration set from 5 fold cross validation;

$x_{\text{NIR, predicted of validation}}$ is the prediction value of NIR validation set;

$x_{\text{MIR, predicted of validation}}$ is the prediction value of MIR validation set;

b_1 is the weight of NIR;

b_2 is the weight of MIR;

b_0 is the intercept of binary linear regression;

The SPA feature extraction did not improve the model results, therefore, we used actual content and predicted content of calibration sets from the full spectrum modeling results (PLS model pretreatment is baseline correction + SNV) to achieve binary linear regression. The predicted results of low-content samples were not good, therefore the data with poor prediction results were removed. In this experiment, we selected data with the concentration more than 20% for high-level data fusion, we obtained the results of the binary linear regression model : b_0 , b_1 , b_2 , then used them to calculated the fusion predicted based on the value of validation set from NIR and MIR. The high-level RMSEP and the high-level R^2 of validation was calculated. The caculation method of the high-level RMSEC and the high-level R^2 of calibration are the same as above. The PLS regression model results of NIR, MIR and three data fusion strategies are shown in Fig. 3 and Table 3 and Table 5.

3. Results and discussion

3.1. Spectra features of NIR

The typical NIR spectra of pure rapeseed oil and olive oil are shown in Fig. 2A. Absorption maxima are clearly observed at 8261, 7073, 5793, 5677, 4335, 4258, 4196 cm^{-1} . Bands around 8261 cm^{-1} arise from the second overtones of CH stretching vibrations, while those at 5793, 5677 cm^{-1} are attributable to the first overtones of CH stretching

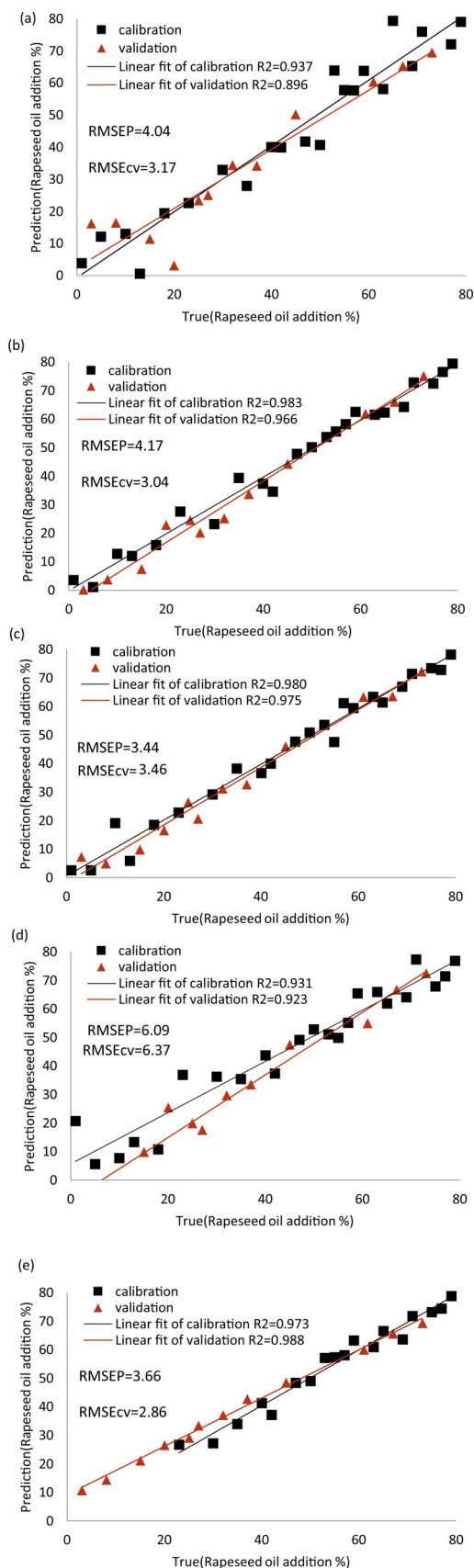


Fig. 3. PLS regression of predicted vs actual rapeseed oil content (a) NIR (Baseline + SNV pretreatment), (b) MIR (Baseline + Normalization pretreatment), (c) low-level data fusion, [NIR (Baseline + SG Smoothing) + MIR (Baseline + SG Smoothing)], (d) mid-level data fusion [NIR SPA (Baseline + SNV) + MIR SPA (Baseline + SNV)], (e) high-level data fusion model predictive results [NIR (Baseline + SNV) + MIR (Baseline + SNV)].

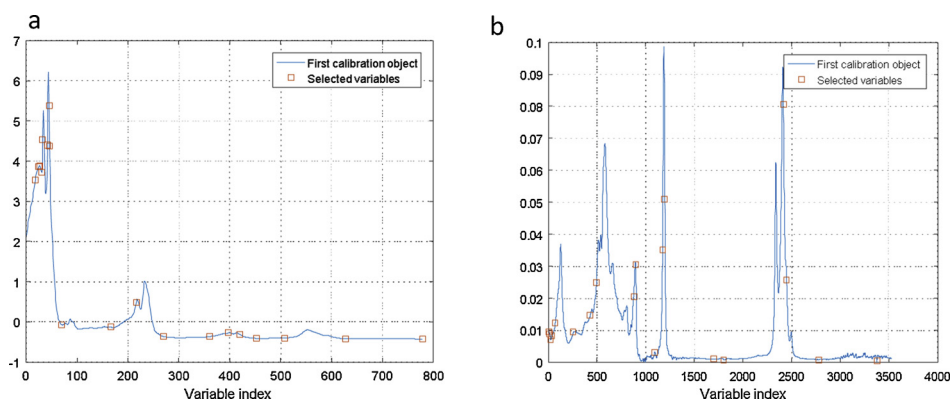


Fig. 4. SPA extracted characteristics of variable a NIR(Baseline + SNV), b MIR(Baseline + SNV).

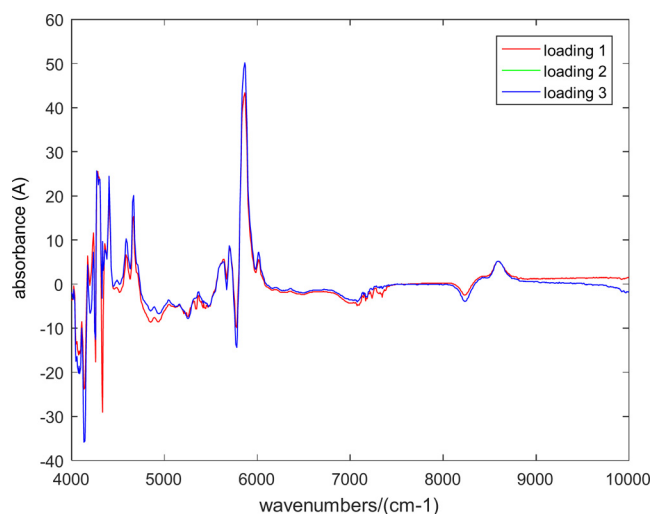


Fig. 5. PLS loading 1, 2 and 3 for NIR (Baseline + SNV pretreatment).

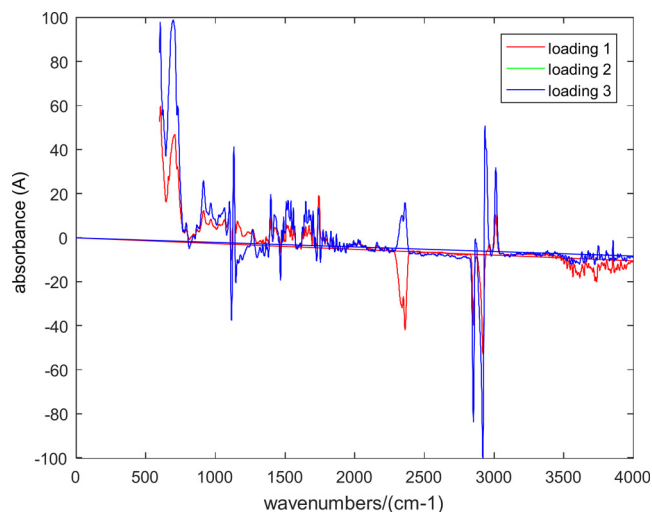


Fig. 6. PLS loading 1, 2 and 3 for MIR (Baseline + Normalization pretreatment).

vibrations of $-\text{CH}_3$, $-\text{CH}_2$ and $-\text{HC}=\text{CH}-$. Absorption at 4335 cm^{-1} arises from combination bands of CH stretching vibrations of $-\text{CH}_3$ and $-\text{CH}_2$. The rapeseed oil was different from olive oil at 5793 and 5677 cm^{-1} (Fig. 2A), this is caused by the different amount of unsaturated fatty acids in olive and rapeseed oil, as shown in Table 4.

3.2. Spectra features of MIR

The typical MIR spectra of pure rapeseed oil and olive oil (Fig. 2B) appear similar to the naked eye, due to their similar chemical composition. The assignments of prominent peaks include: C–H stretching mode in the wavenumber region of $2800\text{--}3100\text{ cm}^{-1}$, C=O stretching in the region of $1700\text{--}1800\text{ cm}^{-1}$, and C–O–C stretching and C–H bending in the region of $900\text{--}1400\text{ cm}^{-1}$.

3.3. NIR, MIR and three data fusion strategies

Quantification of the rapeseed oil contents in the adulterated oil samples was performed using PLS algorithm. The spectral regions used for discriminate analysis were also used for PLS models. Prediction models of the percentage of rapeseed oil added to olive oil from NIR spectra (Fig. 3a), MIR spectra (Fig. 3b), low-level data fusion (Fig. 3c), mid-level data fusion (Fig. 3d) and high-level data fusion (Fig. 3e) were obtained. The calibration models were developed from a set of 24 samples, or spectra, with 1 to 80% of addition rapeseed oil added to olive oil, and the other 12 samples were used to perform the validation of the model.

The results of the true values of addition of rapeseed oil to olive oil, the values provided by the respective calibration curves are shown in Table 3. The determination coefficients R^2 obtained for the NIR, MIR and three data fusion strategies coupled to PLS multivariate models can be used as alternative methods because they distinguish a wide range of frauds and mixtures based on the prediction of the percentage of added rapeseed oil in olive oil (See Figs. 5 and 6).

The R^2 coefficients were achieved from Fig. 3. The same set of samples for calibration and validation was used in the data fusion strategies. All sets of samples showed determination coefficients R^2 greater than 0.85 for both the calibration and validation data sets. However, the use of R^2 to evaluate if the fit model is satisfactory may not be the most appropriate. There are other parameters that could be investigated. The coefficients root mean squared error of prediction (RMSEP), and root mean square error computed from the selected round cross validation (RMSEcv) can give information about the global error of the calibration and validation models.

Table 3 shows the data used to obtain the relationship between true values and predicted values by the PLS model for the validation set. The first column exhibits the values of rapeseed oil addition in olive oil used for the validation data, set while the remaining columns indicate the predicted values from the PLS models for the NIR, MIR, low-level data fusion, mid-level data fusion and high-level data fusion. Table 4 also reveals that the prediction model for mid-level data fusion is most distant from expected values when compared with the results obtained by low-level data fusion and high-level data fusion. Table 4 also reveals that the prediction model for mid-level data fusion is most distant from expected values when compared with the results obtained by others.

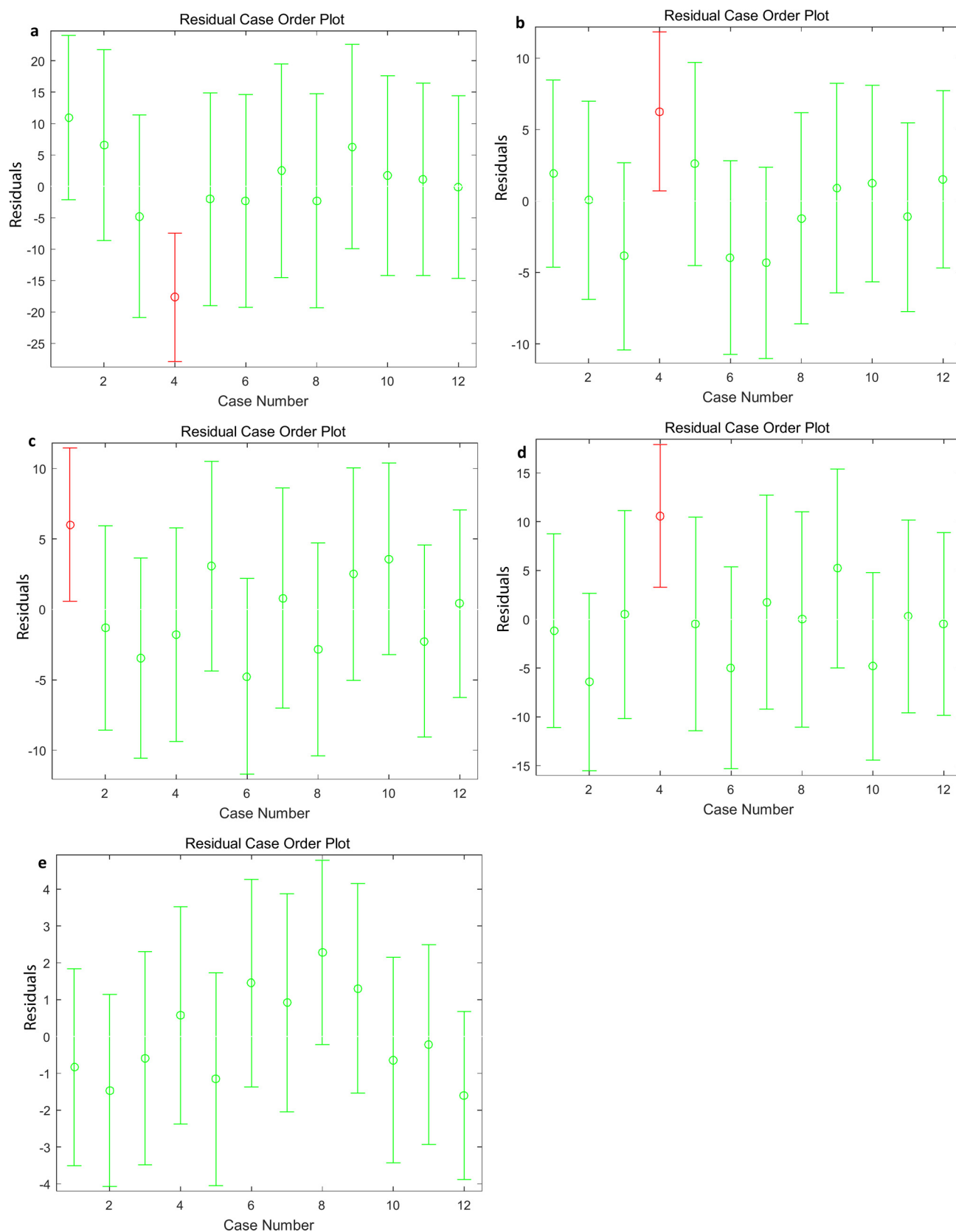


Fig. 7. Residual graph for validation data set. **a** NIR, **b** MIR, and **c** low-level, **d** mid-level, **e** high-level.

Table 5
Evaluation parameters of different modeling on individual data blocks and fused datasets.

Fusion method	pretreatments	Calibration		Validation		Factors	Number of variables
		R ² (CV)	RMSE _{CV}	R ²	RMSEP		
NIR	Baseline + SNV	0.937	6.10	0.896	7.09	8	778
MIR	Baseline + Normalization	0.983	3.17	0.966	4.04	6	3527
NIR(SPA)	Baseline + SNV	0.967	4.41	0.866	8.04	9	19
MIR(SPA)	Baseline + SNV	0.93	6.23	0.97	3.74	9	19
Low -level	Both(Baseline + SG Smoothing)	0.980	3.46	0.975	3.44	10	4305
Mid -level (SPA)	Both(Baseline + SNV)	0.931	6.37	0.923	6.09	5	38
High-level	None	0.973	3.66	0.988	2.86		

The magnitude of this dispersion can be better visualized by the residues graph shown in Fig. 7.

The residual graph (Fig. 7) presents the difference between the true values and the values obtained from the prediction model. As can be seen, the five residual plots show a random distribution of data, without a tendency to increase or decrease the predicted values as a function of concentration, indicating a synchronous distribution of data. However, a greater dispersion of data for the mid-level data fusion is identified. This greater dispersion is in accordance with the values of RMSEP and RMSE_{CV} (shown in Fig. 3).

Finally, the values obtained from the PLS prediction models for NIR, MIR, low-level data fusion, mid-level data fusion, high-level data fusion were evaluated with *t*-test, together with the true values; the *p*-value obtained was 0.975 (NIR), 0.824 (MIR), 0.884 (low-level), 0.718 (mid-level) and 0.656 (high-level). As a result, no differences were found between the true values and those obtained by PLS, in 95% level of confidence, indicating that the approaches can be used for the prediction of adulteration and mixtures of rapeseed oil in olive oil, even if mid-level data fusion have presented a greater dispersion, as mentioned earlier.

The PLS results of NIR, MIR, feature extraction (SPA) of NIR and MIR and the three data fusion strategies are summarized in the Table 5. The low-level data fusion retained all spectral information, so the results of low-level data fusion were better than the NIR and MIR. SPA extracted the characteristic variables of NIR and MIR, which reduced the number of variables, but lost some useful information, so the PLS model did not improve the results. Thus affected the mid-level data fusion results, and indicated that the results of mid-level data fusion are closely related to the method of feature extraction. The high-level data fusion is established on the basis of the full spectrum, this is because the results from the full spectrum are better than feature extraction. The high-level data fusion improved the predicted accuracy and enhanced the stability of the model, which could be used for quantitative analysis.

4. Conclusions

This study compared three data fusion strategies in the quantification analysis of adulterated olive oil, and all three data fusion strategies could be used for the prediction of adulteration and mixtures of rapeseed oil in olive oil. The RMSEP of low-level (3.44) and high-level (2.86) data fusion were better than mid-level (6.09) and the validation coefficient of determination R² of low-level data fusion (0.975) and high-level data fusion (0.988) were better than the NIR (0.896) and MIR (0.966) individually. The three data fusion strategies can detect the extra-virgin olive oil adulteration with rapeseed oil: Low-level data fusion contained all spectral information from NIR and MIR and could improve model prediction accuracy; For mid-level data fusion, we selected SPA used for extracting feature variables. This reduced the number of variables, but it did not improve the results; therefore, the result of mid-level data fusion can be influenced by the feature extraction method. The high-level data fusion adopted the binary linear regression model, which was established based on the predicted values

from the NIR and MIR calibration sets; the weights of NIR and MIR and the intercept were obtained and used to calculate the high-level fusion predicted values of the validation set. High-level data fusion improved the R² and the RMSE, and can be used as a reliable tool for quantitative analysis.

References

- [1] J.A. Pino, K. Almora, R. Marbot, Volatile components of papaya (*Carica papaya* L., Maradol variety) fruit, *Flavour Fragr. J.* 18 (6) (2003) 492–496.
- [2] B.D. Page, G. Lacroix, Application of solid-phase microextraction to the headspace gas chromatographic analysis of halogenated volatiles in selected foods, *J. Chromatogr.* 648 (1) (1993) 199.
- [3] K.I. Poulli, G.A. Mousdis, C.A. Georgiou, Rapid synchronous fluorescence method for virgin olive oil adulteration assessment, *Food Chem.* 105 (1) (2007) 369–375.
- [4] M.Q. Zou, et al., Rapid authentication of olive oil adulteration by Raman spectroscopy, *J. Agric. Food Chem.* 57 (14) (2009) 6001–6006.
- [5] G. Gurdeniz, B. Ozen, Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data, *Food Chem.* 116 (2) (2009) 519–525.
- [6] L.I. Marcos, et al., Detection of adulterants in olive oil by headspace-mass spectrometry, *J. Chromatogr. A* 945 (1) (2002) 221–230.
- [7] A.H. El-Hamdy, N.K. El-Fizga, Detection of olive oil adulteration by measuring its authenticity factor using reversed-phase high-performance liquid chromatography, *J. Chromatogr. A* 708 (2) (1995) 351–355.
- [8] S. Nicoletta, et al., Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity, *Food Res. Int.* 43 (1) (2010) 369–375.
- [9] S. Sivakesava, J. Irudayaraj, Rapid determination of tetracycline in milk by FT-MIR and FT-NIR spectroscopy, *J. Dairy Sci.* 85 (3) (2002) 487–493.
- [10] L. Wang, et al., Feasibility study of quantifying and discriminating soybean oil adulteration in camellia oils by attenuated total reflectance MIR and fiber optic diffuse reflectance NIR, *Food Chem.* 95 (3) (2006) 529–536.
- [11] O. Jović, Durbin-Watson partial least-squares regression applied to MIR data on adulteration with edible oils of different origins, *Food Chem.* 213 (2016) 791–798.
- [12] S.E. Borrà, et al., Olive oil sensory defects classification with data fusion of instrumental techniques and multivariate analysis (PLS-DA), *Food Chem.* 203 (2016) 314–322.
- [13] W. Sun, et al., Data fusion of near-infrared and mid-infrared spectra for identification of rhubarb, *Spectrochim. Acta A. Mol. Biomol. Spectrosc.* 171 (2017) 72–79.
- [14] U. Ramachandran, et al., Dynamic data fusion for future sensor networks, *ACM Trans. Sens. Netw.* 2 (3) (2006) 404–443.
- [15] G. Pandey, An information theoretic framework for sensor data fusion for robotics applications, *Comptes Rendus Des Séances De La Société De Biologie Et De Ses Filiales* 160 (11) (2014) 2076.
- [16] H. Zhu, O. Basir, A novel fuzzy evidential reasoning paradigm for data fusion with applications in image processing, *Soft comput.* 10 (12) (2006) 1169–1180.
- [17] H. Bao, et al., A fire detection system based on intelligent data fusion technology, *Ind. Instrum. Autom.* 2 (2004) 1096–1101 Vol.2.
- [18] X. Hong, J. Wang, Detection of adulteration in cherry tomato juices based on electronic nose and tongue: comparison of different data fusion approaches, *J. Food Eng.* 126 (4) (2014) 89–97.
- [19] C. Márquez, et al., FT-Raman and NIR spectroscopy data fusion strategy for multivariate qualitative analysis of food fraud, *Talanta* 161 (2016) 80–86.
- [20] A.C. Di, M.P. Callao, I. Ruisánchez, 1H NMR and UV-visible data fusion for determining Sudan dyes in culinary spices, *Talanta* 84 (3) (2011) 829–833.
- [21] A. Biancolillo, R. Bucci, F. Marini, Data-fusion for multiplatform characterization of an Italian craft beer aimed at its authentication, *Anal. Chim. Acta* 820 (2014) 23–31.
- [22] A. Dankowska, A. Domagała, W. Kowalewski, Quantification of *Coffea arabica* and *Coffea canephora* var. *Robusta* concentration in blends by means of synchronous fluorescence and UV-Vis spectroscopies, *Talanta* 172 (2017) 215.
- [23] C. Chen, R. Jafari, N. Kehtarnavaz, A survey of depth and inertial sensor fusion for human action recognition, *Multimed. Tools Appl.* 76 (3) (2017) 1–21.
- [24] B. Muik, et al., Determination of oil and water content in olive pomace using near infrared and Raman spectrometry. A comparative study, *Anal. Bioanal. Chem.* 379

- (1) (2004) 35–41.
- [25] P. Vasuki, C. Aravindan, Improving emotion recognition from speech using sensor fusion techniques, TENCON 2012 - 2012 IEEE Region 10 Conference, (2013).
- [26] X.Z. Song, et al., Several notable problems of wavelength selection in molecular spectroscopy area, Spectrosc. Spectral Anal. (2016).
- [27] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (5) (2016) 772–777.
- [28] N.K. Faber, Multivariate sensitivity for the interpretation of the effect of spectral pretreatment methods on near-infrared calibration model predictions, Anal. Chem. 71 (3) (1999) 557.
- [29] H. Yu, et al., Discrimination between Chinese rice wines of different geographical origins by NIRS and AAS, Eur. Food Res. Technol. 225 (3–4) (2007) 313–320.
- [30] M. Spiteri, et al., Data fusion between high resolution ^1H -NMR and mass spectrometry: a synergetic approach to honey botanical origin characterization, Anal. Bioanal. Chem. 408 (16) (2016) 1–13.
- [31] Y.D. Liu, G.W. Zhang, L.J. Cai, [Analysis of chlorophyll in Gannan navel orange with algorithm of GA and SPA based on hyperspectral], Spectroscopy & Spectral Analysis 32 (32) (2012) 3377–3380.
- [32] B.K. Slinker, S.A. Glantz, Multiple linear regression, Circulation 117 (13) (2008) 1732.