Talanta 99 (2012) 175–179

Contents lists available at SciVerse ScienceDirect

Talanta

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

Detection of addition of barley to coffee using near infrared spectroscopy and chemometric techniques

Heshmatollah Ebrahimi-Najafabadi [a,b], Riccardo Leardi [a,*], Paolo Oliveri [a], Maria Chiara Casolino [a], Mehdi Jalali-Heravi [b], Silvia Lanteri [a]

[a] Department of Pharmacy, University of Genova, Via Brigata Salerno 13, I-16147 Genova, Italy
[b] Department of Chemistry, Sharif University of Technology, P.O. Box 11155-9516 Tehran, Iran

ARTICLE INFO

Article history:
Received 9 February 2012
Received in revised form
9 May 2012
Accepted 18 May 2012
Available online 25 May 2012

Keywords:
Coffee
Barley
Near infrared (NIR) spectroscopy
D-optimal design
Partial least squares (PLS) regression

ABSTRACT

The current study presents an application of near infrared spectroscopy for identification and quantification of the fraudulent addition of barley in roasted and ground coffee samples. Nine different types of coffee including pure *Arabica*, *Robusta* and mixtures of them at different roasting degrees were blended with four types of barley. The blending degrees were between 2 and 20 wt% of barley. D-optimal design was applied to select 100 and 30 experiments to be used as calibration and test set, respectively. Partial least squares regression (PLS) was employed to build the models aimed at predicting the amounts of barley in coffee samples. In order to obtain simplified models, taking into account only informative regions of the spectral profiles, a genetic algorithm (GA) was applied. A completely independent external set was also used to test the model performances. The models showed excellent predictive ability with root mean square errors (RMSE) for the test and external set equal to 1.4% w/w and 0.8% w/w, respectively.

1. Introduction

Coffee is one of the three most widely traded foodstuffs and the second largest commodity industry worldwide [1]. There are two varieties of coffee with economic importance: *Arabica* and *Robusta* [2]. Coffee *Arabica* is generally more appreciated for its organoleptic features and, thus, it is the most expensive [3]. Assurance of quality of roasted coffees has attracted widespread attention for controlling and preventing coffee adulteration, also given the great difference in the final sale price [4]. The principal adulterants of coffee include roasted and unroasted coffee husks, twigs, barley, chicory, malt, starch, corn, maltodextrins, glucose sirups, and caramelized sugar [5]. As the simple visual inspection is not an appropriate method for differentiating between the genuine coffee samples and the fraudulent ones, a number of analytical strategies have been developed. El-Abassy et al. [6] applied micro Raman spectroscopy combined with chemometric methods to discriminate between green *Arabica* and *Robusta* coffees based on chlorogenic acid and lipid contents. Hecimovic et al. [7] employed UV–vis spectroscopy and HPLC analysis to determine the polyphenolic compounds and caffeine content of four different types of coffees. Gonzalez et al. [8] used the tocopherol and triglyceride content of roasted and green coffees as features for discriminating between *Arabica* and *Robusta* varieties; principal component analysis (PCA) and linear discriminant analysis (LDA) were employed as pattern recognition tools. Martin et al. [2] applied a relatively similar approach, based on fatty acid profiles as discriminant parameters for coffee variety differentiation. Digital image processing was carried out by Sano et al. [9] to quantify the amounts of brown sugar, coffee husk, maize, and soybean added to coffee *Arabica*. Near infrared spectroscopy combined with multivariate calibration methods was used by Pizarro et al. to quantify the content of *Robusta* variety in roasted coffee mixtures [4]. Jham et al. [10] investigated the potential of tocopherols determined by HPLC analysis as markers to detect coffee adulteration by corn. The feasibility of detection of coffee adulteration with roasted barley, based on volatile compound profiles was studied by Oliveira et al. [11]; solid phase microextraction (SPME) coupled with GC–MS analysis was carried out as analytical tool and chemometric methods were used for data processing. Nogueira and Lago [12] proposed a method based on acid hydrolysis of xylan and starch and consequent electrophoretic separation for identification of adulteration in processed coffee with cereals and coffee husks.

Despite of the relative success achieved by many of these approaches for determining coffee authenticity [13–18], it is important to consider that they are, in many cases, expensive, complex and/or time consuming. For this reason, a fast, reliable

* Corresponding author. Tel.: +39 010 3532636; fax: +39 010 3532684.
E-mail address: riclea@difar.unige.it (R. Leardi).

0039-9140/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.05.036

and low-cost technique with easy implementation for routine analysis represents a very attractive alternative for adulteration and varietal identification purposes [19–23].

Over the last decades, the application of near infrared spectroscopy (NIRS) as a fast and non-destructive technique for the authentication of food samples has become widespread thanks to the advances in chemometrics. Furthermore, it allows to directly analyze solid samples without any complex physical/chemical pre-treatment. Thus, several studies concerning NIR applications in food quality and authentication assessment have been reported [24–28].

In many studies on coffee adulterations, usually, one or two types of coffee and adulterants have been involved. So, the models obtained are poorly representative and are just applicable for these specific samples. In order to obtain a more representative and, thus, widely applicable model, it is advisable to collect and use a wider variety of coffee and adulterants. Taking into account that the exploration of all the possible combinations of different varieties, blends and roasting degrees would be impractical, it is worth finding the minimum number of samples that is maximally representative of all the variability factors that characterize all the possible combinations. The optimal design techniques select highly representative subsets according to particular criteria. The usual approach is to specify a model, to determine the region of interest, to select the number of runs to be made, to specify the optimality criterion and, finally, to find the subset of designed points from the whole set of candidate points [29]. D-optimality is the criterion most widely applied for such a purpose [30,31].

The objective of the present study was to determine the amount of barley added in coffee samples based on NIR spectral information. In order to obtain a widely applicable model, nine types of commercial coffee samples – chosen as to extensively explore the variability of coffee present on the Italian market – and four types of barley samples, with different roasting degrees, were used for investigation. The concentrations of barley were changed from 2 to 20% (w/w) at 10 levels. The lowest limit is surely lower than the sensorial limit of detection. Also the resolution step (2%) is lower than human sensorial capability for distinguishing between close quantities. By taking into account all the combinations, 360 mixtures (9 coffees $\times$ 4 barleys $\times$ 10 concentrations) should have been prepared. A D-optimal design was therefore applied to reduce the number of experiments maintaining the representativeness. The prediction ability of PLS models, either on the full spectra or after variable selection by means of genetic algorithms (GA) [32–35], was evaluated on a test sample set. All the models obtained were additionally tested for their prediction ability with ten independent mixtures, prepared with one coffee and one barley which were not used for preparing the training mixtures.

2. Materials and methods

2.1. Coffee and barley samples

Nine coffee bean varieties and four barley samples were obtained from specialized markets. The coffee samples were selected as to represent the most common types of coffees available on the Italian market, including both *Arabica* and *Robusta* as well as their mixtures at different roasting degrees.

The coffee pure specimens were labeled with upper-case characters from *A* to *I*, while the barley pure samples were marked by lower-case characters from *a* to *d*.

2.2. Apparatus and procedure

Spectral profiles of powder samples were recorded in the reflection mode in the range 4,000–10,000 cm^{-1} with a resolution of 4 cm^{-1}, by an FT-near infrared spectrophotometer based on a polarization interferometer (Buchi NIRFlex N-500). Before analysis, coffee and barley toasted beans were ground with an electric grinder for about 60 s and, afterward, passed through a 0.3 mm sieve. Mixtures at different concentrations were prepared by separately weighing and accurately mixing the finely ground pure powders. Spectra of two grams of samples were recorded at a temperature of 20 ± 1 °C, in a cylindrical optical-glass cell (Hellma, Müllheim, Germany). Each spectrum recorded was the average of 32 successive scans. In addition, three acquisitions were performed, for each experiment, by manual rotation of the cell. In order to minimize the effect of uncontrollable factors, all the experiments were carried out in a random order.

2.3. Experimental design for calibration and validation

The selection of a subset representative of all the possible combinations of coffee and barley samples at ten different concentrations (from 2% to 20% w/w) represented a crucial step prior to carrying out a suitable and significant study, since the total number of combinations was considerably large (360 experiments). A total number of 100 experiments with the maximum representativeness was chosen, based on D-optimal design, to be used in the study as the calibration set. The spectra of all the pure coffee samples as well were recorded and included in the calibration set, which was therefore formed by a total of 109 samples. Four mixtures were prepared and analyzed three times, in order to assess experimental variability, thus the calibration data matrix was formed by 117 spectra. Thirty experiments were also selected as the test set, among the remaining candidate experiments, by applying a subsequent D-optimal design. The experimental matrix is reported in Table 1. An additional evaluation, with a completely external set, was also performed using a new type of coffee and a new type of barley, at ten different concentrations. The new pure coffee was also included in the external set, which was thus composed by 11 spectra.

2.4. Multivariate calibration and variable selection

PLS regression analysis was performed in order to obtain a quantitative model for the prediction of barley amount based on spectral information. PLS is a latent variable-based method, particularly useful when dealing with noisy and collinear data [36]. The latent variables (LVs) are orthogonal directions in the space of the predictors characterized by the maximum covariance with the selected response variable. The optimal complexity – i.e., the number of LVs to be used for building the models – was estimated by a cross-validation procedure, with five deletion groups. In more detail, the number of LVs that provided the minimum error – evaluated as root mean square error in cross-validation, RMSECV – was selected as the optimal complexity.

Column autoscaling was applied on the spectral data.

In the case for spectral data the main goal of variable selection is the elimination of noise, together with the possibility of obtaining models with a reduced complexity.

The selection of variables for multivariate calibration can be considered as an optimization problem. GAs applied to PLS have been shown to be very efficient optimization procedures. They have been applied to many spectral data sets and are shown to provide better results than full-spectrum approaches [37] The major concern with using GAs is the problem of over-fitting. This problem has been addressed using a randomization test [38]. For full details of the algorithm applied, the reader is referred to [39].

Data processing has been performed by programs developed by the authors under MATLAB environment (The MathWorks, Inc., Natick, MA).

Table 1
Design matrix of calibration and test sets. Upper-case characters from *A* to *I* indicate coffee samples while lower-case characters from *a* to *d* indicate barley samples.

No.	Coffee	Barley	Conc[a]	No.	Coffee	Barley	Conc.	No.	Coffee	Barley	Conc.
1	*E*	*d*	4.09	50	*G*	*c*	2.01	99	*F*	*c*	15.98
2	*G*	*c*	13.95	51	*H*	*a*	10.00	100	*D*	*b*	6.02
3	*B*	*c*	8.02	52	*H*	*c*	16.00	101	*E*	*c*	8.00
4	*H*	*a*	11.89	53	*A*	*a*	6.02	102	*B*	*a*	12.01
5	*A*	*b*	14.02	54**	*I*	*c*	6.02	103	*F*	*d*	6.04
6*	*D*	*a*	3.99	55	*I*	*b*	10.00	104	*G*	*a*	3.98
7	*E*	*b*	5.98	56	*I*	*c*	2.00	105	*A*	–	0.00
8	*C*	*a*	2.06	57	*F*	*a*	20.04	106	*F*	*b*	9.99
9	*B*	–	0.00	58	*G*	*d*	16.01	107	*A*	*d*	20.00
10	*G*	*c*	10.13	59***	*F*	*c*	13.99	108	*G*	*d*	20.00
11	*E*	*b*	17.94	60	*B*	*d*	10.00	109	*H*	*a*	17.98
12	*I*	*d*	17.94	61	*A*	*c*	4.00	110	*A*	*d*	18.01
13	*E*	*c*	14.01	62	*F*	–	0.00	111	*I*	–	0.00
14	*F*	*d*	2.00	63	*G*	*b*	11.99	112	*B*	*c*	4.00
15	*H*	*d*	6.02	64	*G*	–	0.00	113	*D*	*c*	17.99
16**	*I*	*c*	6.04	65	*D*	*c*	12.03	114	*E*	*a*	20.00
17	*B*	*b*	8.03	66	*H*	*b*	4.05	115****	*A*	*b*	9.99
18	*B*	*c*	5.98	67	*D*	*c*	20.02	116	*B*	*d*	13.99
19	*A*	*c*	17.99	68	*C*	–	0.00	117****	*A*	*b*	10.03
20	*H*	*d*	2.01	69	*I*	*a*	8.01	1[t]	*B*	*D*	3.99
21	*I*	*b*	2.03	70	*B*	*b*	16.00	2[t]	*F*	*B*	2.02
22	*G*	*b*	18.03	71	*I*	*b*	14.04	3[t]	*G*	*C*	17.98
23	*C*	*a*	8.00	72**	*I*	*c*	6.03	4[t]	*A*	*d*	12.03
24	*G*	*b*	8.22	73	*E*	–	0.00	5[t]	*C*	*c*	2.03
25	*C*	*c*	12.03	74	*D*	*a*	6.00	6[t]	*C*	*d*	14.00
26	*A*	*c*	19.98	75	*G*	*a*	14.03	7[t]	*C*	*d*	20.00
27	*D*	*d*	8.00	76	*C*	*b*	19.99	8[t]	*E*	*d*	12.00
28	*C*	*b*	16.01	77***	*F*	*b*	14.03	9[t]	*F*	*a*	8.00
29	*A*	*a*	16.01	78	*H*	*c*	16.01	10[t]	*H*	*d*	18.00
30	*I*	*a*	19.98	79	*B*	*b*	9.99	11[t]	*E*	*c*	20.00
31	*I*	*d*	16.02	80	*I*	*a*	1.99	12[t]	*F*	*c*	3.99
32	*B*	*d*	19.97	81	*F*	*b*	4.00	13[t]	*I*	*c*	10.00
33*	*D*	*a*	3.98	82	*E*	*d*	7.98	14[t]	*A*	*b*	6.04
34***	*E*	*b*	3.98	83***	*A*	*b*	9.99	15[t]	*I*	*a*	15.99
35	*F*	*c*	14.01	84	*E*	*a*	1.99	16[t]	*B*	*a*	9.98
36	*H*	*d*	8.00	85	*D*	*a*	12.02	17[t]	*E*	*d*	13.98
37	*H*	*b*	20.01	86	*F*	*a*	2.00	18[t]	*H*	*a*	8.00
38	*H*	–	0.00	87*	*D*	*a*	4.03	19[t]	*H*	*b*	12.00
39	*C*	*a*	14.01	88	*C*	*b*	17.99	20[t]	*D*	*b*	20.00
40	*E*	*c*	16.02	89	*D*	–	0.00	21[t]	*B*	*a*	14.05
41	*F*	*c*	11.99	90	*C*	*c*	6.00	22[t]	*E*	*b*	16.03
42	*A*	*d*	7.99	91	*A*	*b*	11.99	23[t]	*G*	*d*	2.02
43	*F*	*d*	4.00	92	*I*	*d*	12.04	24[t]	*G*	*b*	3.99
44	*H*	*b*	14.01	93	*D*	*d*	16.05	25[t]	*D*	*a*	9.99
45	*A*	*a*	9.99	94	*G*	*d*	6.02	26[t]	*A*	*c*	16.00
46	*C*	*d*	12.01	95	*B*	*a*	18.05	27[t]	*D*	*c*	8.02
47	*E*	*a*	2.00	96	*C*	*c*	4.02	28[t]	*I*	*b*	18.01
48	*C*	*d*	10.02	97	*D*	*d*	14.05	29[t]	*H*	*c*	5.99
49	*F*	*a*	18.02	98	*D*	*b*	9.99	30[t]	*G*	*a*	5.98

*, **, *** , **** Indicates groups of replicates.

[a] Weight percentages of barley in coffee (w/w%) (real values of concentration were used for modeling).
[t] Refers to test set.

3. Results and discussion

3.1. Evaluation of different sources of variability

In an ideal system, the main variability is related to the variation of the factor of interest—that is, in the present study, the concentration of barley. However, various external factors – e.g., related to sampling – can affect the global variance of the system.

As a first step, it has been verified that the variability related to sample preparation (mainly grinding of the raw materials, weighing of the coffee and barley powders, mixing) is smaller than the measurement variability (results not reported).

3.2. Multivariate calibration and variable selection

Table 2 shows the root mean square error, the bias and the R^2 values obtained for the PLS models. For the full-spectrum approach, the optimal complexity (estimated by cross-validation) was 12 latent variables.

PLS has a high capability to extract relevant information and to produce a reliable prediction. However, in the last two decades, it has been recognized that an efficient feature selection can be highly beneficial, both to improve the predictive ability of the model and/or to reduce its complexity [37]. For the sake of selecting the proper regions of the spectra, a genetic algorithm (GA) was chosen as the feature selection technique. For the purpose of reducing the search domain, the 1501 original spectral variables were reduced to 188 by sequentially averaging the values of eight contiguous data-points. The GA–PLS algorithm was applied five times on the calibration set. The number of evaluations of each run was set to 200. According to the GA results, 16 variables were selected. The wavenumber intervals corresponding to the variables selected are shown in Fig. 1. Four spectral regions have been selected: the first one is between

Table 2
Statistical performance of PLS and GA–PLS models on the calibration (cross-validation with five deletion groups), test and external sets.

	Calibration set (CV)		Test set		External set	
	RMSE	Bias	RMSE	Bias	RMSE	Bias
PLS	1.23	0.04	1.46	0.53	0.85	−0.58
PLS–GA	1.18	−0.01	1.42	0.21	1.10	−0.95

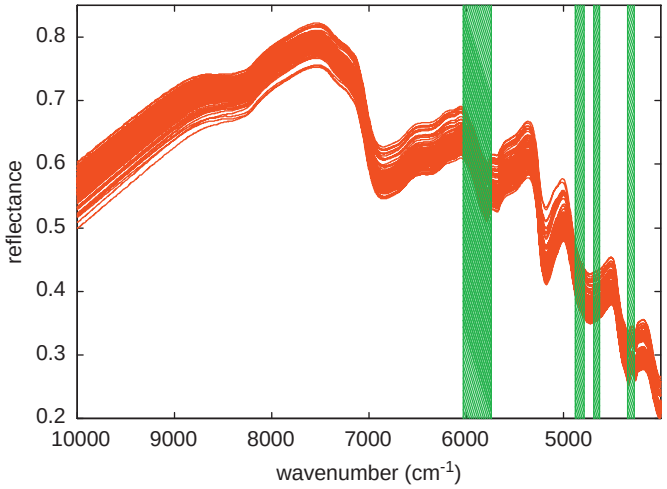

Fig. 1. Spectral regions selected by applying GA.

6032 and 5748 cm^{-1}, the second one includes the 4880–4788 cm^{-1}range, the third one in the 4688–4628 cm^{-1} range, and the fourth includes the narrow range between 4336 and 4276 cm^{-1}. These regions contain 128 wavenumbers in total. Although for the nature of NIRS it is not possible to univocally assign the vibrational transition related to the selected spectral bands, the majority of them are ascribable to the first overtone of N–H, C═O, O–H, C–H, and S–H functional groups of ArOH, H_2O, ROH, CONHR, and RNH_2. The results of PLS applied on the variables selected by GA are presented in Table 2. Eight latent variables were selected for modeling. As it can be seen, the predictive ability of the model is of the same order of that of PLS applied on the whole spectral range. However, the complexity of this model is considerably reduced in comparison to the full-spectrum one. It is worth noticing that PLS outcomes for the three sets (calibration, test and external set) are consistent. This indicates that the models are correctly accounting for valuable spectral information—rather than for chance correlations, which are a frequent cause of overfitting. The parity plots of the two models are shown in Fig. 2. As it can be noticed by plot examination, the models obtained are capable to predict barley concentration with a very satisfactory accuracy, not only in calibration but also on external samples. Residuals are randomly distributed about their mean value, which is satisfactorily close to 0 (low bias), as shown in Fig. 3.

4. Conclusions

The excellent prediction ability obtained by multivariate calibration and indicated by the low values of root mean square errors (RMSE) confirmed that non-destructive NIR measurements can be successfully employed for the detection and quantification of fraudulent addition of roasted barley to roasted coffee. Variable

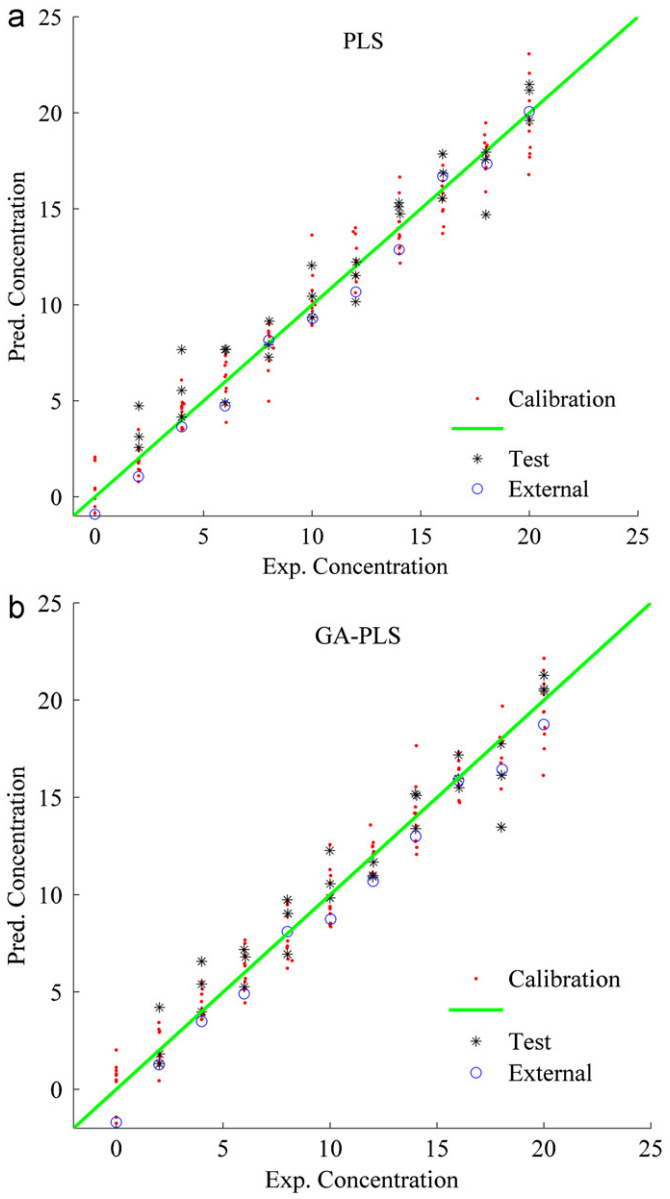

Fig. 2. Experimental vs. predicted values of concentration (% w/w) of barley in coffee samples for (a) PLS model on the whole spectral range and (b) PLS model on spectral bands selected by GA.

selection by using genetic algorithms helped to determine the spectral regions most useful to identify the adulteration of coffee with barley. The methodology allowed to quantify the amount of adulterant up to a level of 2% w/w of barley.

This paper clearly shows that the representativity of the training set is a key point in the success of a calibration model. The achievement of very low prediction errors on a totally external test set (i.e., on mixtures composed by qualities of coffee and barley unknown to the model) has been possible only as a consequence of the fact that the training set was made by taking into account a relatively large number of varieties of coffee and barley. Another key point is the application of D-optimal design for the selection of a subset of adequate size from the very high set of candidate experiments.

The results reported in the present study indicate NIRS to be a promising procedure to be considered in future applications to quantify different adulterants in coffee powder. Indeed, sample collection and analysis should be performed through a number of years, in order to account for variability factors closely related to

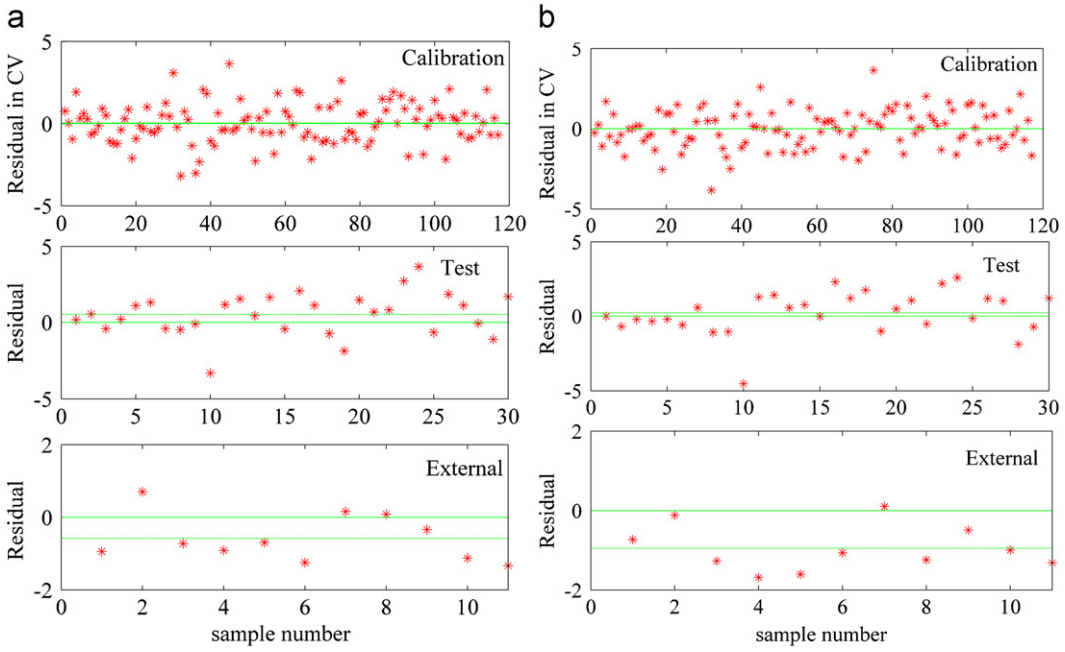

Fig. 3. Residuals vs. sample number of (a) PLS model on the whole spectral range and (b) PLS model on spectral bands selected by GA. Residuals are shown for the 117 calibration spectra, for the 30 test spectra and for the 11 external spectra, respectively. The solid lines represent null residuals while the dashed lines indicate the model bias value.

the harvest and to obtain models characterized by a global applicability.

Acknowledgment

Financial support from the Italian Ministry of University and Research (PRIN 2008, CUP: D31J0000020001) is gratefully acknowledged.

References

[1] C. Pizarro, I. Esteban-Diez, J.M. Gonzalez-Saiz, M. Forina, J. Agric. Food Chem. 55 (2007) 7477–7488.
[2] M.J. Martin, F. Pablos, A.G. Gonzalez, M.S. Valdenebro, M. Leon-Camacho, Talanta 54 (2001) 291–297.
[3] R. Briandet, E.K. Kemsley, R.H. Wilson, J. Agric. Food Chem. 44 (1996) 170–174.
[4] C. Pizarro, I. Esteban-Diez, J.M. Gonzalez-Saiz, Anal. Chim. Acta 585 (2007) 266–276.
[5] J. Prodolliet, M. Bruelhart, M.B. Blanc, V. Leloup, G. Cherix, C.M. Donnelly, R. Viani, J. AOAC Int. 78 (1995) 761–767.
[6] R.M. El-Abassy, P. Donfack, A. Materny, Food Chem. 126 (2011) 1443–1448.
[7] I. Hecimovic, A. Belscak-Cvitanovic, D. Horzic, D. Komes, Food Chem. 129 (2011) 991–1000.
[8] A.G. Gonzalez, F. Pablos, M.J. Martin, M. Leon-Camacho, M.S. Valdenebro, Food Chem. 73 (2001) 93–101.
[9] E.E. Sano, E.D. Assad, S.A.R. Cunha, J. Food Qual. 26 (2003) 123–134.
[10] G.N. Jham, J.K. Winkler, M.A. Berhow, S.F. Vaughn, J. Agric. Food Chem. 55 (2007) 5995–5999.
[11] R.C.S. Oliveira, L.S. Oliveira, A.S. Franca, R. Augusti, J. Food Compos. Anal. 22 (2009) 257–261.
[12] T. Nogueira, C.L. do Lago, J. Sep. Sci. 32 (2009) 3507–3511.
[13] M.S. Valdenebro, M. León-Camacho, F. Pablos, A.G. González, M.J. Martín, Analyst 124 (1999) 999–1002.
[14] M.R. Alves, S. Casal, M.B.P.P. Oliveira, M.A. Ferreira, J. Am. Oil Chem. Soc. 80 (2003) 511–517.
[15] M.J. Martin, F. Pablos, A.G. Gonzalez, Food Chem. 66 (1999) 365–370.
[16] M.J. Martin, F. Pablos, A.G. Gonzalez, Talanta 46 (1998) 1259–1264.
[17] S.K. Schreyer, S.R. Mikkelsen, Sens. Actuators B 71 (2000) 147–153.
[18] S. Casal, M.R. Alves, E. Mendes, M.B.P.P. Oliveira, M.A. Ferreira, J. Agric. Food Chem. 51 (2003) 6495–6501.
[19] J.S. Ribeiro, M.M.C. Ferreira, T.J.G. Salva, Talanta 83 (2011) 1352–1358.
[20] A. Keidel, D.V. Stetten, C. Rodrigues, C. Maguas, P. Hildebrandt, J. Agric. Food Chem. 58 (2010) 11187–11192.
[21] J. Wang, S. Jun, H.C. Bittenbender, L. Gautz, Q.X. Li, J. Food Sci. 74 (2009) 385–391.
[22] I. Esteban-Diez, J.M. Gonzalez-saiz, C. Saenz-Gonzalez, C. Pizarro, Talanta 71 (2007) 221–229.
[23] I. Esteban-Diez, J.M. González-Sáiz, C. Pizarro, Anal. Chim. Acta 514 (2004) 57–67.
[24] E. Ferrari, G. Foca, M. Vognali, L. Tassi, A. Ulrici, Anal. Chim. Acta 701 (2011) 139–151.
[25] P. Oliveri, V. Di Egidio, T. Woodcock, G. Downey, Food Chem. 125 (2011) 1450–1456.
[26] D. Wu, P. Nie, J. Cuello, Y. He, Z. Wang, H. Wu, J. Food Eng. 102 (2011) 278–286.
[27] D. Toher, G. Downey, T.B. Murphy, Chemom. Intell. Lab. Syst. 89 (2007) 102–115.
[28] K. Ruoff, W. Luginbuhl, S. Bogdanov, J.O. Bosset, B. Estermann, T. Ziolko, R. Amado, J. Agric. Food Chem. 54 (2006) 6867–6872.
[29] D.C. Montgomery, Design and Analysis of Experiments, 5th ed., Wiley & Sons Inc., 2001.
[30] T.J. Mitchell, Technometrics 16 (1974) 203–210.
[31] P. Zunin, G.C. Fusella, R. Leardi, R. Boggia, A. Bottino, G. Capannelli, J. Am. Oil Chem. Soc. 88 (2011) 1821–1829.
[32] R. Leardi, A.L. Gonzalez, Chemom. Intell. Lab. Syst. 41 (1998) 195–207.
[33] R. Leardi, R. Boggia, M. Terrile, J. Chemom. 6 (1992) 267–281.
[34] D. Jouan-Rimbaud, D.L. Massart, R. Leardi, O.E. de Noord, Anal. Chem. 67 (1995) 4295–4301.
[35] J. Ghasemi, A. Niazi, R. Leardi, Talanta 59 (2003) 311–317.
[36] S. Rannar, F. Lindgren, P. Geladi, S. Wold, J. Chemom. 8 (1994) 111–125.
[37] R. Leardi, Genetic algorithm-PLS as a tool for wavelength selection in spectral data sets, in: R. Leardi (Ed.), Nature-inspired Methods in Chemometrics: Genetic Algorithms and Artificial Neural Networks, Data Handling in Science and Technology series, vol. 23, Elsevier, Amsterdam, 2003, pp. 169–196.
[38] D. Jouan-Rimbaud, D.L Massart, O.E. de Noord, Chemom. Intell. Lab. Syst. 35 (1996) 213–220.
[39] R. Leardi, Genetic algorithms in feature solution, in: J. Devillers (Ed.), Genetic Algorithms in Molecular Modelling, Academic Press, London, 1996, pp. 67–86.