Analytical Methods

PAPER

View Article Online
View Journal | View Issue

Cite this: *Anal. Methods*, 2016, **8**, 7537

UV-Vis spectral fingerprinting and chemometric method applied to the evaluation of *Camellia sinensis* leaves from different harvests

Elis Daiane Pauli,[a] Roy Edward Bruns[b] and Ieda Spacino Scarminio*[a]

Received 1st September 2016
Accepted 15th September 2016

DOI: 10.1039/c6ay02462a

www.rsc.org/methods

UV-Vis spectral fingerprinting was used to discriminate *Camellia sinensis* leaves of two different harvests and multivariate data analysis was applied to determine the relevant metabolites for separation. First statistical mixture designs of pure ethanol, ethyl acetate, dichloromethane and chloroform solvents as well as their binary, ternary and quaternary mixtures extracted larger varieties and amounts of *C. sinensis* leaf metabolites than would be obtained from classical solvent extractions. UV-Vis spectral fingerprints of crude extracts were subjected to Orthogonal Signal Correction and Partial Least Squares-Discrimination Analysis (OSC-PLS-DA) for classification. The spectra were all correctly identified and classified, showing that the OSC-PLS-DA model possesses a good predictive ability to separate spectral fingerprints of different harvests. VIP score values showed that bands at 272, 410 and 663 nm were responsible for separation. These metabolites were identified by HPLC-DAD as caffeine and pheophytin a. According to the mixture model, the maximum values of relative abundances of both caffeine and pheophytin a can be extracted with pure dichloromethane.

1. Introduction

Tea (*Camellia sinensis* (L.) O. Kuntze) is one of the most widely consumed beverages and an important agricultural product.[1–4] Green, oolong and black are three general classes of tea based on minimum, partial and full fermentation.[5,6] In the manufacturing process of green tea, fresh leaves are manually or mechanically picked and heat processed (steamed or roasted). The oxidation process, commonly known as fermentation, is interrupted, preserving the active components of the leaves. These leaves are then rolled and consequently the enzymes are inhibited and oxidation is prevented.[7] In Brazil, harvesting and full production of *C. sinensis* leaves begins in September. First, the bushes are completely pruned and two weeks after the development of buds, the first leaves are harvested which are called the "first harvest leaves". After about two weeks the "second harvest leaves" are collected and so on until the middle of April, allowing about twelve crops during that period.

Metabolites have a wide variety of physiological and ecological functions in plants. Green tea provides beneficial effects to human health. These effects have been attributed to their complex chemical compositions consisting of phenolic compounds, volatile compounds, amino acids, carbohydrates, proteins, trace elements, *etc.*[8–12] Fresh green tea leaves are very rich in catechins and have antioxidant properties. Antioxidants and their constituent catechins have been reported to prevent cancer and cardiovascular and neurodegenerative diseases.[13] Already caffeine is the main purine alkaloid accounting for about 3 to 6% of the dry weight in *C. sinensis* leaves.[14,15] There are two hypotheses about the importance of caffeine in plants: chemical defense and allopathic function. The first hypothesis states that young leaves possess a higher caffeine content to protect their tissues from mold pathogens and herbivores. The second hypothesis attributes this to an autotoxic effect, *i.e.* when leaves fall to the ground they release germination inhibitors of other plants.[14–16] The consumption of this alkaloid increases alertness, reduces fatigue[17] and has diuretic effects.[18]

Besides the compounds mentioned above, tea leaves contain lipophilic pigment metabolites such as chlorophyll a and b.[19] A study by Lee *et al.* suggested that heat applied during the manufacture of green tea can activate the chlorophyllase enzyme and thus increase the concentration of pheophytin.[20] According to Harpaz-Saad *et al.* this enzyme is responsible for the degradation of chlorophyll.[21] Furthermore, the conversion of chlorophyll into pheophytin may occur during the storage of green vegetables.[22] Even though pheophytin gives green tea a darker color upon processing, studies show that this derivative has antioxidant activities as well as being antigenotoxic.[23,24]

The composition of green tea leaves is very similar to that of the leaves before they are harvested. However, this depends on a variety of factors such as weather, season, culture techniques, type, plant age, and also leaf processing.[25,26] The variation in the

[a]*Laboratório de Quimiometria em Ciências Naturais, Departamento de Química, Universidade Estadual de Londrina, CP 6001, 86051-990, Londrina, PR, Brazil. E-mail: ieda@qui.uel.br; Fax: +55 43 33714286; Tel: +55 43 33714811*

[b]*Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13083-970, Campinas, SP, Brazil*

amounts of metabolites related to harvested *C. sinensis* leaves has also been studied,[27,28] but needs to be further investigated. Plant metabolites are known to be susceptible to environmental and genetic changes and their responses can be evaluated by metabolomic analysis.[29] The metabolome represents a critical aspect of a plant's physiology, growth characteristics and, ultimately, economic value.[30] A more global metabolomics approach is called metabolic fingerprinting that can be used for purposes such as quality control of plants and characterization and classification of medicinal plants.[31]

Studies using faster and simple instrumental techniques such as ultraviolet-visible, infrared and near infrared spectroscopy associated with chemometric methods have been conducted as alternatives to the discriminant analysis and/or classification between caffeinated and decaffeinated coffee,[32] classification and authenticity determination of virgin olive oils from different geographical regions,[33–35] as well as wines,[36] varieties of vinegar,[37] and varieties of tea,[38] among others.

Metabolic fingerprinting of an herbal sample is a characteristic profile of chromatographic or spectroscopic origin, which represents its composition.[39] Compared with conventional analytical approaches, this technique emphasizes the integral characterization of a complex system with a quantitative degree of reliability.[40] It is an untargeted approach and aims at assessing global metabolic profiles without the knowledge of the classes of compounds. It can reveal the variability in the metabolic composition of plants grown under the same conditions, and in combination with multivariate methods can identify the metabolites important to sample discrimination. ^{1}H NMR, HPLC and IR spectroscopy are the most commonly used techniques for fingerprinting with the purpose of sample classification. UV-Vis spectrophotometry is a simple, cheap and easy to handle technique, but is rarely used for the purpose of metabolic fingerprinting.

In this study the objective was to evaluate the potential of UV-Vis spectrophotometry fingerprinting associated with Orthogonal Signal Correction and Partial Least Squares-Discrimination Analysis (OSC-PLS-DA) for discriminating *C. sinensis* leaves of two harvests and identify the most important metabolites for the discrimination. In addition, an exploratory analysis was performed to investigate the effects of extraction solvent mixtures on the identified metabolites using a four component simplex-centroid mixture design.

2. Experimental

2.1. Plant materials

Two sets of leaf samples from *Camellia sinensis* (L.) Kuntze were kindly provided by the Agrochá Boa Vista farm (Araucária, PR, Brazil). The first sample set was collected in September 2011 and corresponds to the first harvest after pruning. Eighteen days later, the second set of samples was collected, when the new shoots were ready to be picked. Since the tea leaves contain oxidizing enzymes, they were subjected to steaming prior to oxidation, thus disabling oxidizing enzymes, while retaining the color of chlorophyll as well as the active components contained in tea. A voucher specimen catalogued as FUEL 49288 has been stored in the Herbarium of the Universidade Estadual de Londrina (UEL). The species was identified by A. O. Vieira, Departamento de Biologia, UEL.

2.2. Chemicals and reagents

For plant extraction, all organic solvents were of analytical grade and obtained from F. Maia (São Paulo, Brazil). All crude extracts were filtered with 25 mm PTFE 0.2 μm syringe filters purchased from Chromafil (Macherey-Nagel, Düren, Germany). HPLC grade acetonitrile and methanol were purchased from J. T. Baker (Phillipsburg, USA). HPLC grade water (18.2 MΩ cm) was prepared using a Millipore Milli-Q Gradient purification system (Bedford, USA) and used for the mobile phases.

2.3. Extract preparation

C. sinensis leaves were ground in a blender with the aim of obtaining small fragments for subsequent extraction. The crude extracts were prepared using mixtures of (e) ethanol, (a) ethyl acetate, (d) dichloromethane and (c) chloroform whose proportions were varied according to the simplex-centroid mixture design presented in Table 1. Extraction consisted of four pure solvents, their six binary 1 : 1 mixtures, four ternary 1 : 1 : 1 mixtures and a quaternary 1 : 1 : 1 : 1 mixture. These mixtures were prepared in a random order including three replicates for the quaternary mixture (center point) to estimate experimental error. The selection of each solvent was made considering Snyder's solvent selectivity triangle, since solvents from different groups in the triangle have different selectivity characteristics.[41]

Each crude extract was prepared by adding 6 mL of extraction solvent to 2.0 g of crushed leaves. The mixtures remained in an ultrasound bath for 30 min with the bath water being

Table 1 Mixture design solvent proportions used for extraction (given in mL)

	Solvents			
Extract notation	Ethanol (mL)	Ethyl acetate (mL)	Dichloromethane (mL)	Chloroform (mL)
e	6	0	0	0
a	0	6	0	0
d	0	0	6	0
c	0	0	0	6
ea	3	3	0	0
ed	3	0	3	0
ec	3	0	0	3
ad	0	3	3	0
ac	0	3	0	3
dc	0	0	3	3
ead	2	2	2	0
eac	2	2	0	2
edc	2	0	2	2
adc	0	2	2	2
eadc1	1.5	1.5	1.5	1.5
eadc2	1.5	1.5	1.5	1.5
eadc3	1.5	1.5	1.5	1.5

changed to maintain temperature. This procedure was repeated fourteen more times, for a total of 15 repetitions for each point of the simplex-centroid design. The extracts were left to rest under forced ventilation until they reach a constant weight.

2.4. Spectroscopic fingerprint measurements

UV-Vis spectral fingerprinting was carried out with 20 mg of each crude extract dissolved in 10 mL of respective extraction solvent (Table 1). The absorptions were performed in a 1 cm quartz cuvette, with a Thermo Scientific, Evolution 60S model UV-Vis spectrophotometer, coupled with Vision Lite software and monitored in the 200–800 nm range with a resolution of 1 nm.

2.5. HPLC analysis

For HPLC analysis, 2.00 mg of extract were dissolved in 2.0 mL of methanol and 100 μL of this solution was diluted in 400 μL of initial mobile phase ACN/H_2O (95 : 5, v/v) and filtered through a 0.2 μm polytetrafluoroethylene membrane (PTFE, Chromafil). These analyses were conducted on a Finnigan Surveyor 61607 system coupled with a Finnigan Surveyor PDA Plus photodiode array detector (PDA) and manual sample injector with a 20 μL loop. The metabolites were separated by using a Hilic C_{18} column (150 × 4.6 mm, Kinetex model with 2.6 μm particle size) from Phenomenex equipped with a security guard holder. The flow rate was 0.7 mL min^{-1}. The mobile phase consisted of a combination of solvent A (acetonitrile) and solvent B (water). The gradient elution was as follows: t = 0.00 min, 5% B; t = 1.00 min, 5% B; t = 1.01 at 4.00 min, 19.35% B; t = 4.01 at 7.00 min, 35.50% B; t = 7.01 at 10.00 min, 50% B; t = 10.01 at 15.00 min, 5% B. The photodiode detector continuously recorded absorbance from 190 to 800 nm. The data were processed using ChromQuest 4.2 software.

2.6. Data analysis

2.6.1. Response surface analysis. Response surface analysis has been used to model and optimize the extraction of vegetal materials.[42] A four component polynomial model was adjusted to the experimental data:[43]

$$\hat{y} = \sum_{i}^{4} b_i x_i + \sum_{i} \sum_{<j}^{4} b_{ij} x_i x_j + \sum_{i} \sum_{<j} \sum_{<k}^{4} b_{ijk} x_i x_j x_k + \ldots + b_{1234} x_1 x_2 x_3 x_4 \quad (1)$$

where $\hat{y}$ is the predicted absorbance value of the analyzed metabolite, and x_i is the i^{th} solvent proportion. b_i is the linear coefficient representing the expected response for the pure component i, b_{ij} is the coefficient of the interaction between i and j components, b_{ijk} is the coefficient of the interaction between the i, j and k components and b_{1234} is the coefficient of interaction between all four components. Here linear, quadratic and special cubic models were tested and ANOVA regression results were obtained using Statistica 6.0 (Statistica for Windows 6.0, Statsoft, Tulsa, OK, USA).

2.6.2. Orthogonal Signal Correction for Partial Least Squares-Discrimination Analysis (OSC-PLS-DA). Pre-processing treatments, OSC-PLS-DA model calculations, cross validation and predictions were performed using Matlab R2007a (Mathworks Inc. Natick, MA, USA) functions included in PLS Toolbox 5.8.1 (Eigenvector Research Inc., Wenatchee, WA, USA).

Spectral fingerprints of crude extracts of *C. sinensis* leaves from the first and second harvests were monitored in the 200 to 800 nm region. However the low and high spectral wavelengths contained no information and the analysis region was reduced to the 250–720 nm range. This fingerprint matrix, with 34 (crude extracts, 17 for each harvest) × 471 (wavelengths) dimensions, was called the X matrix. Different pretreatments were tested on the raw spectroscopic data, such as the Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), first and second derivatives (performed according to the Savitzky-Golay method), Orthogonal Signal Correction (OSC) filtering and mean centering, in order to get the most suitable Partial Least Squares-Discriminant Analysis (PLS-DA) model.

PLS-DA[44] is a supervised method used for classification purposes. It is developed from algorithms for Partial Least Squares (PLS) regression employing a set of predictor variables X and a categorical response y of binary variables expressing class membership.[45] PLS-DA estimates in an efficient way the best linear combinations of the independent original X-values, called latent variables (LVs), that correlate with the observed changes of the dependent variable, y.[46] In this case, the better option for spectral pretreatment was the application of the Orthogonal Signal Correction (OSC)[47] filter with the data mean-centered based on the maximum number of samples correctly classified (sensitivity) and misclassified (specificity). OSC removes systematic variations in the X matrix that are not correlated with the class variable, y. OSC-PLS-DA regression was applied to optimally model the class variable y (first harvest extracts categorized in class 0 and second harvest ones in class 1). Furthermore, a threshold value is calculated between the predicted values, and values above this threshold value indicate the sample fit to the modeled variable.[48]

For the OSC-PLS-DA application spectral fingerprints were divided into calibration (26 samples) and validation (eight samples) groups, with the Kennard-Stone algorithm.[49] Both sets contained crude extracts from the first and second harvests. For OSC-PLS-DA model interpretation, the selection of the more relevant variables (wavelengths) having an effect on the separation between the crude extract spectral fingerprints was done using Variables Importance in Projection (VIPs) scores[50] that are a weighted sum of squares of PLS weights for each variable and measure the contribution of each predictor variable to the model.[46] When the VIP value is greater than one its variable is generally considered as being important for the model.[51]

3. Results and discussion

Fig. 1 shows that the distinction between spectra is quite complicated owing to spectral band overlap. OSC-PLS-DA is recommended for attempting to find the best correlation between the X (UV-Vis spectral fingerprints of crude extracts of *C. sinensis* leaves) and y (first or second harvest class) matrices when the variability within groups is greater than the variability

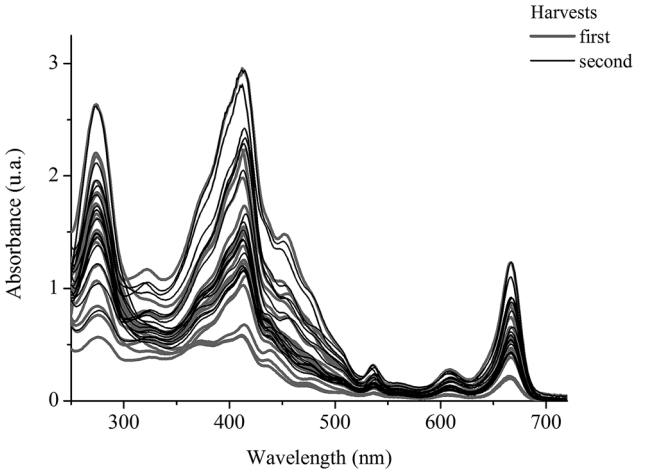

Fig. 1 UV-Visible spectral fingerprinting of crude extracts of *C. sinensis* leaves from first and second harvests obtained for the simplex-centroid mixture design.

among groups. The calibration model was composed of 12 spectra from the first harvest and 14 from the second harvest. The dimensionality of the model was determined by venetian blind cross-validation, based on the lowest RMSECV value. The classification model was developed using two latent variables, containing 94.54% of the variance explained in *X* and 77.27% of the variance in the *y*.

Fig. 2a shows the OSC-PLS-DA results applied to the spectral fingerprints of the crude extracts from the two harvests. The weights plot in Fig. 2b shows the bands that contribute to the discrimination between harvests. The spectral regions which can be considered important for OSC-PLS-DA performance and thus discriminate between classes are observed in the VIP plot of Fig. 2c showing scores with values greater than 1 (ref. 45 and 51) for three bands centered at 272, 410 and 663 nm. Most of the second harvest extracts were discriminated by the positive part of the LV_1. The weights of these three absorptions were positive on LV_1, where the spectral bands of the corresponding metabolites have a higher intensity characterizing the second harvest extracts. The values of the negative weights of LV2 indicate that for leaves from the first harvest the band at 272 nm was more evident. The positive values of the same latent variable demonstrate the influence of the bands at 410 and 663 nm, in the extracts from the second harvest leaves. However the intensities of all three bands were higher for the second harvest extracts. The positive values of the same latent variable demonstrate the influence of the bands at 410 and 663 nm, in the extracts from the second harvest leaves. Therefore, the results above suggest that the difference between the harvests can be attributed to spectral intensity variations of metabolites corresponding to the 272, 410 and 663 nm bands.

This OSC-PLS-DA model was then applied for external validation, with a set of 5 spectral fingerprints of the *C. sinensis* leaves from the first harvest and 3 from the second. The prediction performance can be observed in Fig. 3, where the spectral fingerprints of the first harvest are above the threshold value of 0.5550 and the second harvest points are below it. Sensitivity and specificity were determined from data in Fig. 3. Sensitivity is the model's ability to classify the validation samples belonging to a particular class. For the model, the sensitivities were 1.00 for both harvests. Based on these results, the model was able to classify the two different harvests from *C. sinensis* samples. The specificity is related to the incorrect prediction of validation samples of other classes in a particular class. Both harvest classes present specificity equal to 1.00; this result means that no first harvest sample was classified in the second harvest class and *vice versa*. These results show that the OSC-PLS-DA model showed a good ability to discriminate between the spectral fingerprints of the different harvests analyzed.

In the case of *C. sinensis* leaves the absorption band at 272 nm can be associated with caffeine. According to Souto *et al.* absorbance in this region may be attributed to caffeine's $n \rightarrow \pi^*$ transition, in particular, to the chromophore C=O.[32] The two other bands, 410 and 663 nm, probably correspond to pigments present in the *C. sinensis* leaves. Pheophytin is a degradation product of chlorophyll, which is found in tea, and has a Soret band at 410 nm (blue region), followed by other weaker bands at longer wavelengths, called Q bands at around 665 nm. Between the Soret and Q bands, there are minor bands at 505, 535 and 606 nm.[52] These bands are characteristic of the porphyrin ring and occur on promotion of the ligand π electrons to π^* antibonding orbitals.

In order to obtain more information on the metabolites suggested by discrimination in OSC-PLS-DA, the crude extracts were analyzed by high performance liquid chromatography (HPLC) coupled with diode array detection (DAD). Each crude extract was analyzed at the apices of the caffeine and pheophytin peaks, with maximum absorbances at 272 and 410 nm, respectively. According to Milenkovic *et al.* the ratios between the intensities of the absorbance involving the Soret and Q bands (A_{Soret}/A_Q) for pheophytin a and b are about 2.33 and 5.30, respectively.[52] The mean ratio of the intensities of these bands obtained in our experiments was 2.49 ± 0.09 (a.u.) for the crude extracts from the first harvest and 2.45 ± 0.18 (a.u.) from the second, once again suggesting the presence of pheophytin a (Phya).

Paired *t*-tests were applied to the absorption intensity values at 272 nm for caffeine and 410 nm for Phya of all crude extracts. The results showed significant differences at the 95% confidence level between the relative abundances for caffeine and Phya in the two harvests. This was based on the calculated *t* values of 3.51 (caffeine) and 2.43 (Phya) that are higher than the critical value of 2.12. Therefore, this provides statistical evidence that the quantities of these two metabolites are significantly different being higher for extracts of the second harvest leaves.

These results confirm those obtained by spectrophotometry in the UV-Visible region described above. The intensity values at 272 and 410 nm were analyzed as a function of the extraction solvent composition to optimize the caffeine and Phya extracted amounts. According to Ashihara *et al.* caffeine is produced in younger leaves.[14] Just as for the caffeine results, the Phya results showed higher values for extracts from the second harvest. The

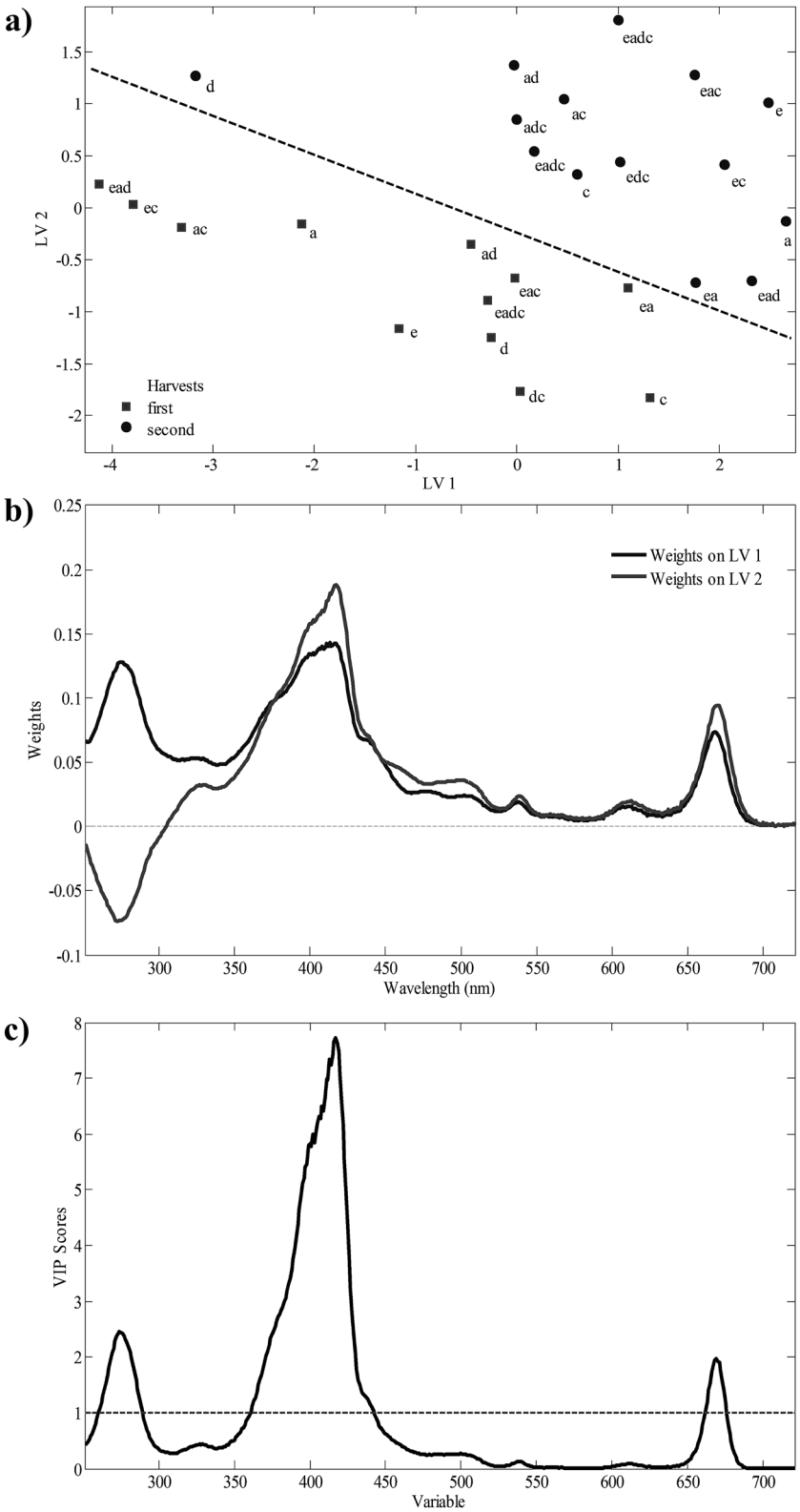

Fig. 2 OSC-PLS-DA results of the spectral fingerprint analysis of crude extracts of *C. sinensis* leaves from first and second harvests were recorded in the 250 to 720 nm region for simplex centroid mixture design extracts: (a) scores plot (LV1 *versus* LV2); (b) LV1 and LV2 weight plot (wavelength); (c) variable importance (VIP) scores plot.

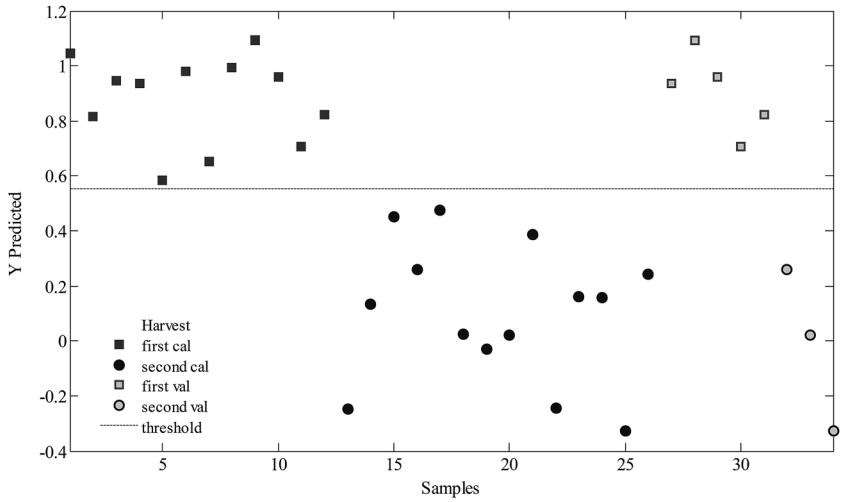

Fig. 3 Results of the calibrations and validation set of the OSC-PLS-DA model of the UV-visible spectral fingerprinting of the *C. sinensis* leaves from first and second harvests.

fact that a second stress occurred in the plants related to leaf cutting leads us to believe that this triggers increased production of caffeine as well as chlorophyll and its derivatives.

Linear, quadratic and special cubic models were fit to the absorption intensities at 272 and 410 nm, for the peaks separated and identified by HPLC-DAD. The quadratic model quality for caffeine and the special cubic one for Phya from the first and second harvests were determined by ANOVA, which showed no evidence of lack of fit at the 95% confidence level. These models are given below:

$$\begin{aligned}\text{Caffeine}_{1^{st}\ \text{harvest}} &= \underset{(\pm0.07)}{\mathbf{0.31e}} + \underset{(\pm0.07)}{\mathbf{0.25a}} + \underset{(\pm0.07)}{\mathbf{0.59d}} + \underset{(\pm0.07)}{\mathbf{0.26c}} + \underset{(\pm0.29)}{0.20\text{ea}} \\ &\quad - \underset{(\pm0.29)}{0.14\text{ed}} + \underset{(\pm0.29)}{0.67\text{ec}} - \underset{(\pm0.29)}{\mathbf{0.77ad}} - \underset{(\pm0.29)}{0.33\text{ac}} - \underset{(\pm0.29)}{0.50\text{dc}} \end{aligned} \quad (2)$$

$$\begin{aligned}\text{Caffeine}_{2^{nd}\ \text{harvest}} &= \underset{(\pm0.07)}{\mathbf{0.59e}} + \underset{(\pm0.07)}{\mathbf{0.28a}} + \underset{(\pm0.07)}{\mathbf{1.11d}} + \underset{(\pm0.07)}{\mathbf{0.30c}} + \underset{(\pm0.31)}{0.62\text{ea}} \\ &\quad - \underset{(\pm0.31)}{0.31\text{ed}} - \underset{(\pm0.31)}{0.13\text{ec}} - \underset{(\pm0.31)}{\mathbf{1.18ad}} - \underset{(\pm0.31)}{0.44\text{ac}} - \underset{(\pm0.31)}{\mathbf{1.44dc}} \end{aligned} \quad (3)$$

$$\begin{aligned}\text{Pheophytina}_{1^{st}\ \text{harvest}} &= \underset{(\pm0.08)}{\mathbf{0.41e}} + \underset{(\pm0.08)}{\mathbf{1.50a}} + \underset{(\pm0.08)}{\mathbf{1.55d}} + \underset{(\pm0.08)}{\mathbf{1.40c}} \\ &\quad - \underset{(\pm0.39)}{0.07\text{ea}} - \underset{(\pm0.39)}{0.74\text{ed}} + \underset{(\pm0.39)}{\mathbf{1.28ec}} + \underset{(\pm0.39)}{0.13\text{ad}} - \underset{(\pm0.39)}{\mathbf{1.33ac}} - \underset{(\pm0.39)}{\mathbf{1.59dc}} \\ &\quad - \underset{(\pm2.50)}{\mathbf{19.99ead}} + \underset{(\pm2.50)}{\mathbf{9.15eac}} + \underset{(\pm2.50)}{4.96\text{edc}} + \underset{(\pm2.50)}{5.47\text{adc}} \end{aligned} \quad (4)$$

$$\begin{aligned}\text{Pheophytina}_{2^{st}\ \text{harvest}} &= \underset{(\pm0.10)}{\mathbf{0.67e}} + \underset{(\pm0.10)}{\mathbf{1.50a}} + \underset{(\pm0.10)}{\mathbf{1.55d}} + \underset{(\pm0.10)}{\mathbf{1.40c}} - \underset{(\pm0.51)}{1.11\text{ea}} \\ &\quad + \underset{(\pm0.51)}{0.78\text{ed}} - \underset{(\pm0.51)}{\mathbf{1.99ec}} - \underset{(\pm0.51)}{\mathbf{4.16ad}} - \underset{(\pm0.51)}{1.26\text{ac}} - \underset{(\pm0.51)}{\mathbf{4.30dc}} + \underset{(\pm3.23)}{4.51\text{ead}} \\ &\quad + \underset{(\pm3.23)}{0.08\text{eac}} - \underset{(\pm3.23)}{4.76\text{edc}} + \underset{(\pm3.23)}{\mathbf{23.10adc}} \end{aligned} \quad (5)$$

where e, a, d and c represent the ethanol, ethyl acetate, dichloromethane and chloroform proportions, respectively. Standard error estimates are presented in parenthesis below their corresponding model coefficients. Significant model coefficients at this level are presented in boldface.

The linear dichloromethane coefficients (d) were highest and significant at the 95% confidence level in the quadratic model for caffeine. By analyzing the other coefficients in eqn (2) and (3) one can observe an antagonistic interaction effect between ethyl acetate and dichloromethane on caffeine extraction for both harvests. However, for crude extracts from the second harvest there was also an antagonistic effect between dichloromethane and chloroform (eqn (3)). Fig. 4a and b show the corresponding response surfaces for the absorption intensity of caffeine in crude extracts of the first and second harvests. The highest absorption intensities are predicted for proportions near the pure dichloromethane vertex (darker region). These results are confirmed by literature studies testing various solvents which show that dichloromethane is more efficient for extracting caffeine.[53–55] Moreover, it is also used as a solvent in the decaffeination of tea.[54]

As for caffeine, the linear dichloromethane coefficients (d) were highest for the Phya special cubic model. The model predictions for Phya from the first harvest (eqn (4)) showed the existence of a binary synergistic effect between ethanol and chloroform as well as a ternary synergistic effect involving ethanol, ethyl acetate and chloroform predicting improved yields compared to the one predicted by only the linear coefficients. However interactions between the ethyl acetate–chloroform, dichloromethane–chloroform and ethanol–ethyl acetate–dichloromethane extraction solvent mixture are antagonistic for Phya intensities. Of the solvent interactions of the crude extracts of the second harvest (eqn (5)) the ternary synergistic interaction involving ethyl acetate–dichloromethane–chloroform improved the extraction over that expected based on only pure solvent effects. Binary mixtures involving ethanol–chloroform, ethyl acetate–dichloromethane and dichloromethane–chloroform had an antagonistic effect on the extraction of Phya, *i.e.* absorption intensity was lower than expected. Fig. 4c and d show the response surface models for the apex absorbance for Phya peaks of crude extracts from the first and second harvests. The highest predicted absorption intensities are obtained by pure dichloromethane and chloroform (darker region). The

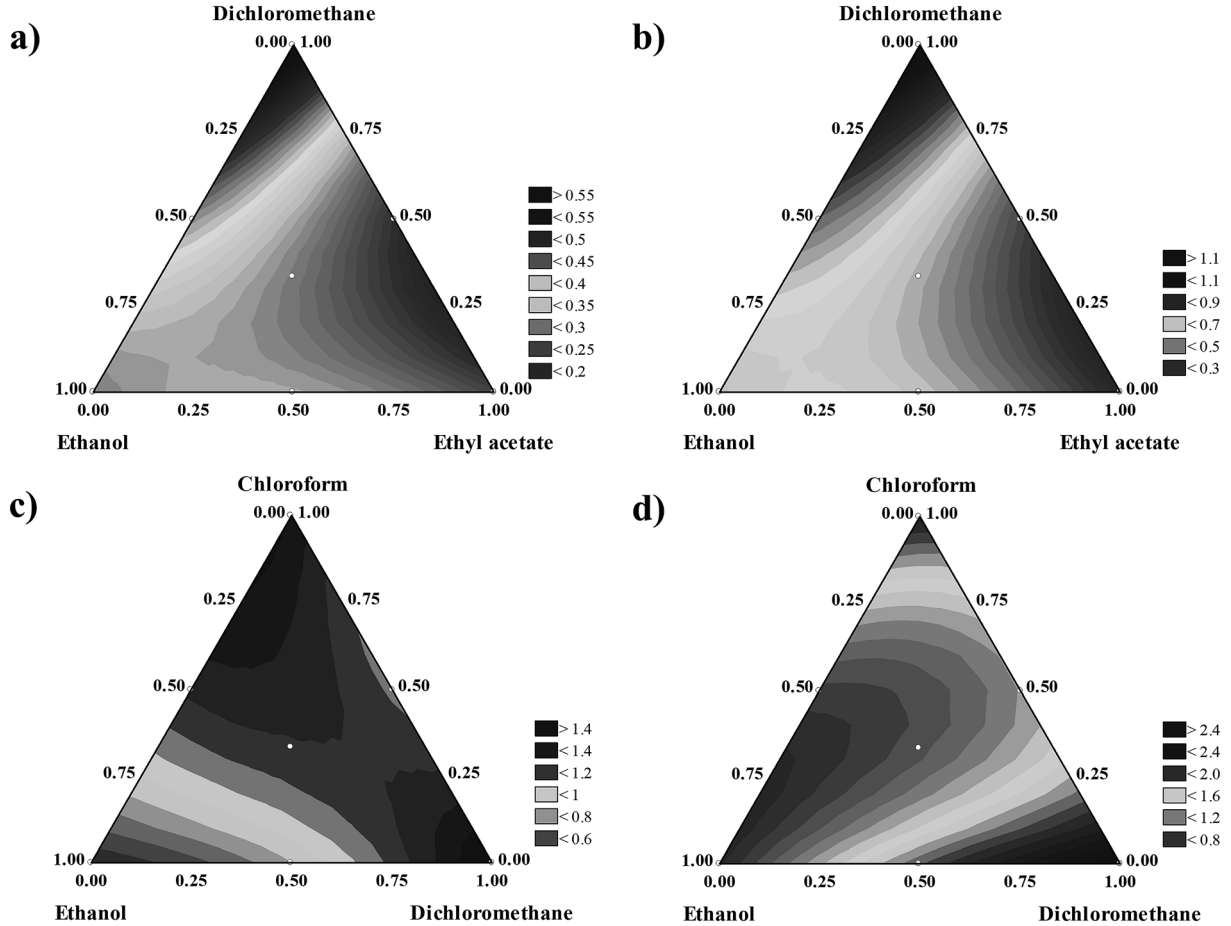

Fig. 4 Response surfaces for the intensity values of caffeine absorption at 272 nm for the (a) first and (b) second harvests and at 410 nm for Phya absorption for the crude extracts of *C. sinensis* leaves of the (c) first and (d) second harvests.

region close to the ethanol vertex (lighter region) indicates lower predicted intensities for both harvests.

4. Conclusion

This work showed the potential of UV-Vis spectral fingerprinting coupled with suitable chemometric tools, in this case OSC-PLS-DA, to discriminate *C. sinensis* leaves of first and second harvests. The advantage of using spectrophotometry is attributed to the ease of handling and quick data acquisition in addition to being inexpensive and simple to use. Thus, it was possible to distinguish the first and second harvest samples, even with large data variations owing to the four solvent simplex-centroid design used to prepare crude extracts. The VIP values from the spectral data indicated that caffeine and Phya were important to distinguish leaves of the two harvests. Pure dichloromethane was the most efficient solvent extractor of both metabolites. These results suggest that the Phya and caffeine may be potential quality markers of *C. sinensis* leaves, since they proved to be sensitive to stress such as pruning. The proposed method is a viable alternative for the evaluation of first and second harvest leaf samples. Future work could investigate spectral fingerprinting combined with the PLS-DA method for other plants and discriminate other factors besides harvests.

Acknowledgements

The Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) and the Fundação Araucária Brazilian granting agencies supported the work done in Brazil. The authors acknowledge and thank Mauricio Takeshi Kitano (Agrochá Boa Vista farm) for his generous donation of the samples. The authors also thank Prof[a] Dr[a] Patrícia Valderrama (Universidade Tecnológica Federal do Paraná-Campo Mourão) for her valuable discussions and suggestions concerning this work.

References

1 D. Kumar, A. Gulati and U. Sharma, *Food Anal. Methods*, 2016, **9**, 1666–1674.
2 V. Sharma, R. Joshi and A. Gulati, *Eur. Food Res. Technol.*, 2011, **232**, 307–317.

3 S. Tsubaki, S. Sakumoto, N. Uemura and J. Azuma, *Food Chem.*, 2013, **138**, 286–290.
4 Y. Zuo, H. Chen and Y. Deng, *Talanta*, 2002, **57**, 307–316.
5 G. M. Hadad, R. A. A. Salam, R. M. Soliman and M. K. Mesbah, *Talanta*, 2012, **101**, 38–44.
6 H. Luo, S. B. Cox, W. Goa, J. Yu, L. Tang and J. S. Wang, *Metabolomics*, 2006, **2**, 235–241.
7 A. A. Shitandi, F. M. Ngure and S. M. Mahungu, *Tea in Health and Disease Prevention*, ed. V. Preedy, Elsevier, London, 2012, pp. 193–206.
8 C. Cabrera, R. Artacho and R. Giménez, *J. Am. Coll. Nutr.*, 2006, **25**, 79–99.
9 H. N. Graham, *Prev. Med.*, 1992, **21**, 334–350.
10 M. Jeszka-Skowron and A. Zgola-Grześkowiak, *Food Anal. Methods*, 2014, **7**, 2033–2041.
11 A. B. Sharangi, *Food Res. Int.*, 2009, **42**, 529–535.
12 L. Y. Wang, K. Wei, Y. W. Jiang, H. Cheng, J. Zhou, W. He and C. C. Zhang, *Eur. Food Res. Technol.*, 2011, **233**, 1049–1055.
13 C. Anesini, G. E. Ferraro and R. Filip, *J. Agric. Food Chem.*, 2008, **56**, 9225–9229.
14 H. Ashihara, H. Sano and A. Crozier, *Phytochemistry*, 2008, **69**, 841–856.
15 P. Mohanpuria, V. Kumar and S. K. Yadav, *Food Sci. Biotechnol.*, 2010, **19**, 275–287.
16 R. J. Maughan and J. Griffin, *J. Hum. Nutr. Diet.*, 2003, **16**, 411–420.
17 A. Smith, *Food Chem. Toxicol.*, 2002, **40**, 1243–1255.
18 J. Nussberger, V. Mooser, G. Maridor, L. Juillerat, B. Waeber and H. R. Brunner, *J. Cardiovasc. Pharmacol.*, 1990, **15**, 685–690.
19 M. Friedman, C. E. Levin, S. H. Choi, S. U. Lee and N. Kozukue, *J. Food Sci.*, 2009, **74**, C406–C412.
20 J. Lee, Y. S. Hwang, I. K. Kang and M. G. Choung, *LWT–Food Sci. Technol.*, 2015, **61**, 201–208.
21 S. Harpaz-Saad, T. Azoulay, T. Arazi, E. Ben-Yaakov, A. Mett, Y. M. Shiboleth, S. Hörtensteiner, D. Gidoni, A. Gal-On, E. E. Goldschmidt and Y. Eyal, *Plant Cell*, 2007, **19**, 1007–1022.
22 S. J. Schwartz and T. V. Lorenzo, *Crit. Rev. Food Sci. Nutr.*, 1990, **29**, 1–17.
23 K. Higashi-Okai, M. Taniguchi and Y. Okai, *J. Ferment. Bioeng.*, 1998, **85**, 555–558.
24 K. Higashi-Okai, M. Taniguchi and Y. Okai, *J. Sci. Food Agric.*, 2000, **80**, 117–120.
25 H. Mukhta and N. Ahmad, *Am. J. Clin. Nutr.*, 2000, **71**, 1698S–1702S.
26 H. H. Sherry Chow and I. A. Hakim, *Pharmacol. Res.*, 2011, **64**, 105–112.
27 S. Ercisli, E. Orhan, O. Ozdemir, M. Sengul and N. Gungor, *Pharm. Biol.*, 2008, **46**, 683–687.
28 L. Yao, N. Caffin, B. D'Arcy, Y. Jiang, J. Shi, R. Singanusong, X. Liu, N. Datta, Y. Kakuda and Y. Xu, *J. Agric. Food Chem.*, 2005, **53**, 6477–6483.
29 J. Zhao, C. Hu, J. Zeng, Y. Zhao, J. Zhang, Y. Chang, L. Li, C. Zhao, X. Lu and G. Xu, *Metabolomics*, 2014, **10**, 805–815.
30 G. G. Harrigan, S. Martino-Catt and K. C. Glenn, *Metabolomics*, 2007, **3**, 259–272.
31 J. Ivanišević, O. P. Thomas, C. Lejeusne, P. Chevaldonné and T. Pérez, *Metabolomics*, 2001, **7**, 289–304.
32 U. T. Souto, M. J. C. Pontes, E. C. Silva, R. V. H. Galvão, M. C. U. Araújo, F. A. C. Sanches, F. A. S. Cunha and M. S. R. Oliveira, *Food Chem.*, 2010, **119**, 368–371.
33 A. Hirri, M. Bassbasi, S. Platikanov, R. Tauler and A. Oussama, *Food Anal. Methods*, 2016, **9**, 974–981.
34 A. Oussama, F. Elabadi, S. Platikanov, F. Kzaiber and R. Tauler, *J. Am. Oil Chem. Soc.*, 2012, **89**, 1807–1812.
35 C. Pizarro, S. Rodríguez-Tecedor, N. Pérez-del-Notario, I. Esteban-Díez and J. M. González-Sáiz, *Food Chem.*, 2013, **138**, 915–922.
36 S. M. Azcarate, M. A. Cantarelli, R. G. Pellerano, E. J. Marchevsky and J. M. Camiña, *J. Food Sci.*, 2013, **78**, C432–C436.
37 F. Liu, Y. He and L. Wang, *Anal. Chim. Acta*, 2008, **615**, 10–17.
38 A. Palacios-Morillo, A. Alcázar, F. Pablos and J. M. Jurado, *Spectrochim. Acta, Part A*, 2013, **103**, 79–83.
39 J. Viaene, M. Goodarzi, B. Dejaegher, C. Tistaert, A. Hoang Le Tuan, N. Nguyen Hoia, M. Chau Van, J. Quetin-Leclercq and Y. Vander Heyden, *Anal. Chim. Acta*, 2015, **877**, 41–50.
40 D. Z. Yang, Y. Q. An, X. L. Jiang and D. Q. Tang, *Talanta*, 2011, **85**, 885–890.
41 L. R. Snyder, P. W. Carr and S. C. Tutan, *J. Chromatogr. A*, 1993, **656**, 537–547.
42 L. M. Z. G. Passari, I. S. Scarminio and R. E. Bruns, *Anal. Chim. Acta*, 2014, **821**, 89–96.
43 B. Barros Neto, I. S. Scarminio and R. E. Bruns, *Statistical Design – Chemometric*, Elsevier, Amsterdam, 2006.
44 M. Barker and W. Rayens, *J. Chemom.*, 2003, **17**, 166–173.
45 S. Wold, M. Sjöström and L. Eriksson, *Chemom. Intell. Lab. Syst.*, 2001, **58**, 109–130.
46 M. Farrés, B. Pinã and R. Tauler, *Metabolomics*, 2015, **11**, 210–224.
47 S. Wold, H. Antti, F. Lindgren and J. Öhman, *Chemom. Intell. Lab. Syst.*, 1998, **44**, 175–185.
48 F. C. G. B. S. Alves and P. Valderrama, *Anal. Methods*, 2015, **7**, 9702–9706.
49 R. W. Kennard and L. A. Stone, Computer aided design of experiments, *Technometrics*, 1969, **11**, 137–148.
50 S. Wold, E. Johansson and M. Cocchi, *3D QSAR in Drug Design: Theory, Methods and Applications*, ed. H. Kubinyi, ESCOM Science Publishers, Leiden, 1993, pp. 523–550.
51 I. G. Chong and C. H. Jun, *Chemom. Intell. Lab. Syst.*, 2005, **78**, 103–112.
52 S. M. Milenković, J. B. Zvezdanović, T. D. Anđelković and D. Z. Marković, *Adv. Technol.*, 2012, **1**, 16–24.
53 T. Atomssa and A. V. Gholap, *Afr. J. Pure Appl. Chem.*, 2011, **5**, 1–8.
54 A. Belay, K. Ture, M. Redi and A. Asfaw, *Food Chem.*, 2008, **108**, 310–315.
55 R. Sharif, S. W. Ahmad, H. Anjum, N. Ramzan and S. R. Malik, *J. Food Process Eng.*, 2013, **37**, 46–52.

