DE GRUYTER
OPEN

Determination of adulterants in adulterant-fruit spirit blends using excitation-emission matrix fluorescence spectroscopy

Michaela Jakubíková, Jana Sádecká, Pavel Májek

Institute of Analytical Chemistry, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, SK-812 37 Bratislava, Slovak Republic
michaela.jakubikova@stuba.sk

Abstract: This study introduces a reliable method to detect adulteration of spirit drinks. Excitation-emission matrix (EEM) fluorescence in combination with parallel factor analysis (PARAFAC) and partial least squares (PLS) regression was used to determine the content of water and ethanol in adulterated fruit spirit samples. EEM fluorescence spectra recorded in the emission wavelength range of 315—450 nm and in the excitation wavelength range of 240—305 nm were used for PARAFAC. The model created using PARAFAC-PLS was able to predict the water and ethanol level in adulterated apple spirit with the root mean square error of prediction (RMSEP) values of 1.9 % and 1.8 %, respectively. Regarding adulterated plum spirit, the RMSEP values of 0.7 % and 3.5 % were obtained for water and ethanol, respectively. The aim of this work was to determine whether EEM-PARAFAC can be used to distinguish between plum and apple spirit. Better results were obtained for apple spirit and the method is useful also for water-apple spirit blends.

Keywords: excitation-emission matrix fluorescence spectroscopy, plum spirit, apple spirit multivariate analysis, parallel factor analysis, partial least squares

Introduction

Council regulation No 110/2008 of the EU parliament defines the designation of origin of spirit drinks. Commission Regulation (EC) No 716/2013 lays down detailed rules and methods of analysis of spirit drinks are given in Regulation No 2870/2000. Therefore it is very important to develop new methods for the determination of adulteration in plum and apple spirit drinks. A great variety of analytes and matrices have been investigated using the excitation-emission matrix (EEM) fluorescence spectra, with special attention to: (1) contaminants and natural constituents in environmental samples; and, (2) pharmaceuticals and metabolites in biological samples, such as serum and urine. Due to the possibility of analytical determinations in the presence of non-modelled interferents and the identification of the analyte of interest, calibrations based on scores of parallel factor analysis (PARAFAC) decomposition of EEM fluorescence spectra are becoming increasingly important in routine analysis (Ortiz et al., 2006). Recently, a combination of fluorescence spectroscopy and multivariate methods has been used for the quantification of ethanol in ethanol-petrol and biodiesel in biodiesel-diesel blends. Calibration models were made using a combination of synchronous fluorescence spectroscopy with principal component regression and partial least square (PLS), and EEM fluorescence spectroscopy with N-way partial least square (N-PLS) and unfolded-PLS. All four calibration models were highly robust, the errors in the predictions were found to be lower than 2 % (Kumar and Mishra, 2012). Furthermore, calibration models for the determination of adulterants in brandy were designed using the combination of EEM fluorescence, PARAFAC and PLS. The models were useful for the detection of adulterants (ethanol, methanol and water) in adulterated brandy at levels as low as 0.24 % (Markechová et al., 2014).

The aim of the present study was to build PARAFAC-PLS models based on EEMs and to study their ability to predict the amount of adulterant (ethanol or water) in fruit spirits.

Experimental

Samples

Five samples of one brand of apple spirit and five samples of one brand of plum spirit were included in the study. Samples were stored in the dark at room temperature and analyzed without any pretreatment. The company that supplied the brand apple and plum spirit provided the content and geographical indication in agreement with Reg.110/2008.

To predict the content of ethanol or water in adulterated fruit spirit samples, samples with different

Acta Chimica Slovaca, Vol. 8, No. 1, 2015, pp. 52—58, DOI: 10.1515/acs-2015-0010

relative portions of ethanol or water (in % v/v) in the fruit spirit were prepared by adding appropriate volumes of ethanol or water to the fruit spirit. Two calibration sets were prepared as follows. Random selective apple spirit samples were used for preparing 41 calibration samples containing an adulterant (ethanol or water). Relative adulterant fraction (in % v/v) in the calibration samples varied from 0 % to 20 % and the interval was 0.5 %. Two prediction sets were prepared in a similar manner. Relative adulterant fraction (in % v/v) in the prediction samples varied from 0.25 % to 19.25 % (interval of 1 %). Thus, 20 prediction samples were prepared for each adulterant. The same procedure was repeated with the plum spirits to prepare 41 calibration and 20 prediction samples containing an adulterant (ethanol or water). Ethanol, for HPLC, gradient grade, ≥99.8 % (Sigma-Aldrich, UK), and deionized and double distilled water with the resistivity of 18 MΩ cm were used to prepare all solutions.

Fluorescence spectroscopy

Fluorescence spectra were recorded using the Perkin-Elmer LS 50 Luminescence Spectrometer equipped with a Xenon lamp. Samples were placed in a 10 × 10 × 45 mm quartz cell. Excitation and emission slits were both set at 5.0 nm. Scan speed was 200 nm/min. EEM fluorescence spectra were collected in the excitation wavelength range of 200—500 nm with an interval of 5 nm and in the emission wavelength range of 250—600 nm with an interval of 5 nm, respectively. EEM fluorescence spectra in the emission wavelength range of 315—450 nm with an interval of 5 nm and in the excitation wavelength range of 240—305 nm with an interval of 5 nm, which gives 28 and 14 data points along the emission and excitation wavelength axes, respectively, were used for PARAFAC. Applying emission wavelengths above the excitation wavelengths prevented the Rayleigh scatter.

Multivariate analysis

Data were exported to ASCII and processed with the Microsoft Office Excel 2010 software, Statistica version 7.0 (StatSoft, USA, 2004), MATLAB Version 7.0 (The MathWorks Inc., USA, 2005) and PLS_Toolbox version 6.0 (Eigenvector Research Inc., USA, 2010). Autoscale preprocessing was performed for calibration and prediction of data sets. Detailed information on PARAFAC, PLS and related parameters have been discussed earlier (Wold et al., 2001; Wise et al., 2006; Smilde et al., 2004). Below, PARAFAC (Bro, 1997), commonly used for modelling of fluorescence excitation—emission data is briefly described. To discuss the PARAFAC model, the fluorescence data arranged in a three-way array X ($I \times J \times K$) where I refers to the samples, J to the emission wavelengths, and K to the excitation wavelengths, were considered. The fluorescence signals X are decomposed into F three-linear components according to the number of fluorophores present in the samples (Eq. 1):

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk} \quad (1)$$

i = 1, ... I; j = 1, ... J; k = 1, ... K

where x_{ijk} is the intensity of the measured light for sample i at the emission wavelength j and excitation wavelength k, F is the number of components (individual fluorophore moieties), a_{if} is the ith score for the fth component and it is related to the concentration of the fth component in the ith sample, b_{if} and c_{kf} are estimates of the emission and excitation spectrum of the fth component (defined as loadings), respectively, and e_{ijk} is the residual containing the variation not captured by the model (Bro, 1997). To determine the number of components, various numerical characteristics of the model can be used e.g. explained variance (%) (Surribas et al., 2006) and core consistency (corcondia) (%) (Divya and Mishra, 2007) — ideally, both values are 100 %. The results of PARAFAC modelling are relative concentrations (score) and spectral profiles (loadings) of components in the samples (Bro 1997). Excitation and emission profiles can be assigned to fluorophores by comparison with recorded spectra of standards or with data reported in literature. Score values can be used to construct the calibration model describing the relationship between the concentration of the fluorophore and the score value (Surribas et al., 2006). In this work, PLS–regression was used. This method identifies the directions of the greatest variability by comparing both score and concentration information with the new axes called factors. PLS require proper selection of the number of factors in order to avoid both overfitting leading to accurate calibration but poor prediction, and underfitting resulting in better robustness but impaired prediction. In this study, the leave-one-out cross-validation was used to select the optimum number of factors consisting in removing one sample at a time from the calibration set and performing the calibration with all other samples. Concentration of the sample removed was then predicted by the obtained model. This step was repeated for each sample considered. The procedure was repeated after fixing a different number of factors. The number of PLS factors included in the model was chosen considering the lowest root mean square error of the cross-validation (RMSECV) (Wold et al., 2001).

Results and discussion

EEM fluorescence spectra

Fig. 1 shows EEM fluorescence spectrum of apple spirit in the form of a contour map, where the contours are plotted by linking the points of equal fluorescence intensity. EEM contour map spreads in the excitation wavelength range from 200 nm to 350 nm and the emission wavelength region from 295 nm to 500 nm with the fluorescence maxima at the excitation wavelengths of 250 nm and 300 nm, and the emission wavelengths of 325 nm and 420 nm. Volatile compounds as 2-phenylethanol, guaiacol, 4-vinylanisole, 4-methylguaiacol, methyleugenol, 4-ethylguaiacol, eugenol, 4-ethylphenol, 4-vinylguaiacol identified in apple distillates (Genovese et al., 2004; Ledauphin et al., 2003; Ledauphin et al., 2004) can contribute to the observed fluorescence (Song et al., 2008; Zhan et al., 2008; Sádecká et al., 2015). Many substituted phenols or anisols show similar fluorescence properties, for example eugenol exhibit bands with the respective maxima at λ_{ex} = 280 nm and at λ_{em} = 320 nm. The band at $\lambda_{ex}/\lambda_{em}$ = 300/425 nm can be related to coumarins (Tóthová et al., 2009).

EEM contour map of plum spirit (Fig. 2) spreads in the excitation wavelength range from 200 nm to 310 nm and the emission wavelength region from 280 nm to 480 nm with the fluorescence maxima at the excitation wavelengths of 220 nm, 280 nm and 300 nm, and the emission wavelengths of 320 nm and 420 nm. Volatile compounds as 2-phenylethanol, methylphenol, 4-ethylphenol, 4-propenylphenol, eugenol, methyleugenol, 4-vinylanisole and

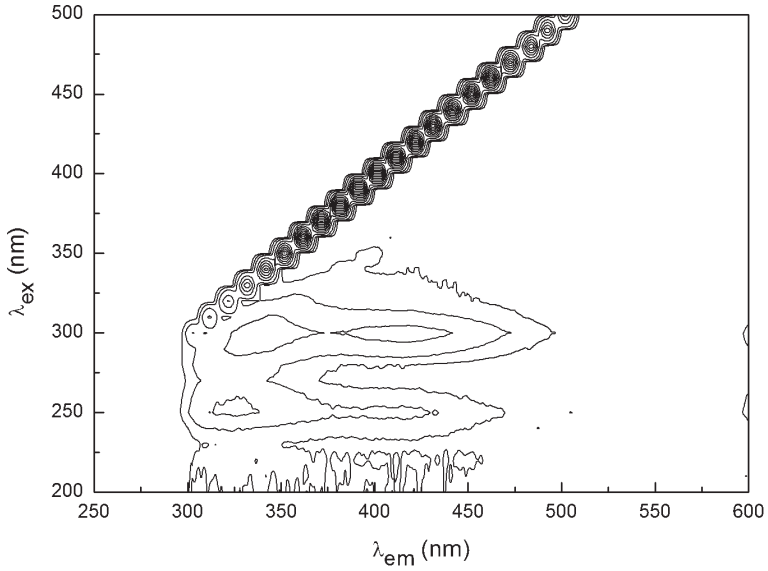

Fig. 1. EEM spectrum of apple spirit.

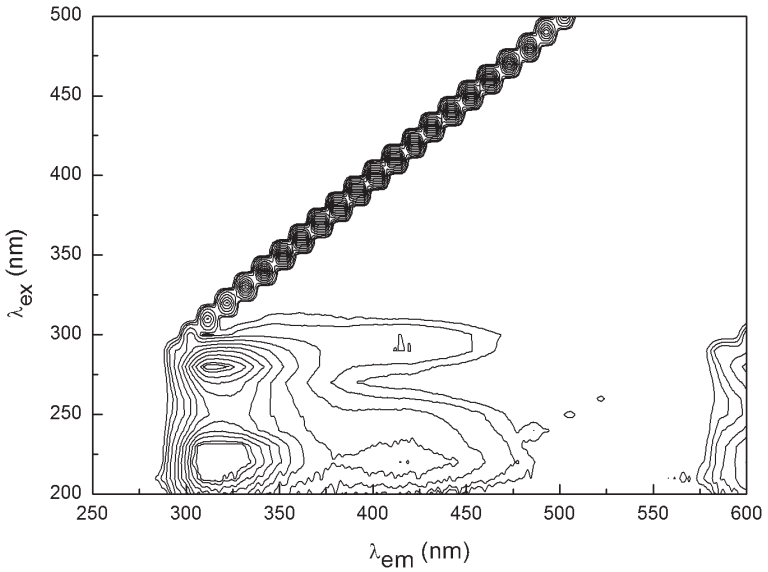

Fig. 2. EEM spectrum of plum spirit.

p-cymene identified in plum distillates (Satora and Tuszyński, 2008; Velíšek et al., 1982; Tešević et al., 2005; Miličević et al., 2012; Coldea et al., 2014) are known fluorophores (Song et al., 2008; Zhan et al., 2008; Sádecká et al., 2015), which can contribute to the observed fluorescence.

When the adulterant content in both the apple and the plum spirit increases from 0.25 to 20 % v/v, the fluorescence intensity of the blends increases; however, the excitation and emission maxima remain the same. When the adulterant content in both the spirits is higher than 25 % v/v, the excitation and emission maxima are shifted towards lower wavelength.

PARAFAC-PLS of apple spirit blends

EEMs fluorescence spectra of apple spirit blends combined with PARAFAC-PLS were used for quantitative analysis (Surribas et al., 2006). EEMs of five apple spirit samples were arranged in a three-way array (samples × number of λ_{em} × number of λ_{ex}) of the dimensions of 5 × 28 × 14. EEMs of each calibration set were arranged in a data matrix of the dimensions of 41 × 28 × 14 (samples × number of λ_{em} × number of λ_{ex}). The three arrays were decomposed using PARAFAC, which resulted in loadings and relative quantities (scores) of components. PARAFAC models were estimated with one to four components and the results were compared. The non-negativity constraint was applied in all three modes (emission, excitation and concentration) since both the fluorescence intensities and the concentrations were positive.

Based on the core consistency, explained variance, split-half validation and visual inspection of the residuals, the PARAFAC models with two components were chosen (Reis et al., 2001). Core consistency >98 % and explained variance >99 % were obtained using the two-component model for all data sets, while the values are almost the same (Table 1).

The loadings decomposed by the two-component PARAFAC models in emission and excitation modes were also identical for all data sets (Table 1). Loading peaks for component 1 were observed at the excitation wavelengths of 250 nm and 300 nm and at the emission wavelength of 420 nm for apple spirit, ethanol- and water-apple spirit blends. Loading peaks for component 2 were observed at the excitation wavelengths of 260 nm and 285 nm and at the emission wavelength of 325 nm. This implies that the fluorescence response recorded for one "fluorophore" in different samples has the same excitation and emission profiles as required by PARAFAC. The two adulterants (ethanol and water) are polar species and their addition to the apple spirit similarly changed the solution environment of fluorophores, which resulted in identical EEM profiles.

The score of component 1 and component 2 depend on the adulterant additions (results not shown). Component 1 was not much modified by the adulteration, however, significant changes in the score of the component 2 were observed. Thus, scores of component 1 and 2 in combination with PLS were used for quantitative analysis. The PLS model was developed in the PLS1 mode (PLS was run for each component separately) using the calibration score matrix with dimensions of 41 × 2 (calibration samples × number of components). PLS models require proper selection of the number of factors in order to avoid overfitting or underfitting. The leave-one-out cross-validation method was used to select the optimum number of factors. Once the optimal number of factors was determined, the final calibration, using all calibration samples with the optimal number of factors, was performed. The resulting calibration characteristics, number of factors (latent variables, LV), total percent of explained variation for score block and concentration block, root mean squares error of calibration (RMSEC), root mean squares error of cross-validation (RMSECV), coefficient of determination of calibration (R^2 Cal) and coefficient of determination of cross-validation (R^2 CV), are listed in Table 2. According to Jerome and Workman (2008), RMSEC, RMSECV and RMSEP values should be as low as possible (close to 0), while R^2 should be as high as possible (close to 1). Thus, the low RMSEC and RMSECV values and high R^2 value (close to 1) confirm the ruggedness of the models (Divya and Mishra, 2007).

Then, the PLS model was used to predict the adulterant concentration in the prediction data set as follows. EEMs of the prediction set were arranged in a data matrix of the dimensions of 20 × 28 × 14 (samples × number of λ_{em} × number of λ_{ex}). PARAFAC model developed above (the loading matrices were kept constant) was used to obtain the new prediction scores. The prediction score matrix with the dimensions of 20 × 2 (prediction samples × number of components) and the PLS model developed before were used to predict the concentrations of the adulterant. The resulting prediction characteristics, root mean square error of prediction (RMSEP) and the coefficient of determination of prediction (R^2 Pred), are given in Table 2. Low errors in the calibration and prediction and high R^2 values (close to 1) indicate good performance of the proposed PLS models for adulterants determination. Better prediction is obtained for apple spirit adulterated with ethanol, showing R^2 Pred of 0.917 and RMSEP of 1.8.

Tab. 1. Excitation and emission maxima, explained variance and core consistency versus apple spirit and adulterant-apple spirit blends for two-component PARAFAC models.

Parameter	Apple spirit	Adulterant	
		Ethanol	water
Excitation wavelength/emission wavelength (nm)			
component 1	250, 300/420	250, 300/420	250, 300/420
component 2	260, 285/325	260, 285/325	260, 285/325
Explained variance (%)			
component 1	61.6	61.5	61.5
component 2	38.4	38.4	38.5
Total	100.0	99.9	100.0
Core consistency (%)	98.6	98.7	98.6

Tab. 2. Calibration, validation and prediction results of the PLS model for adulterant-apple spirit blends.

Parameter	Adulterant	
	Ethanol	water
No LVs	2	2
% of variance score block	89.8	87.0
% of variance concentration block	94.3	90.4
RMSEC	1.4	1.7
RMSECV	1.4	1.8
RMSEP	1.8	1.9
R^2 Cal	0.945	0.918
R^2 CV	0.944	0.917
R^2 Pred	0.917	0.897

No LVs number of latent variables, *RMSEC* root mean square error of calibration, *RMSECV* root mean square error of cross-validation, *RMSEP* root mean square error of prediction, *R^2 Cal* coefficient of determination of calibration, *R^2 CV* coefficient of determination of cross-validation, *R^2 Pred* coefficient of determination of prediction

PARAFAC-PLS of plum spirit blends

EEMs fluorescence spectra of plum spirit blends combined with PARAFAC-PLS were used for quantitative analysis (Surribas et al., 2006). EEMs of five plum spirit samples were arranged in a three-way array (samples × number of λ_{em} × number of λ_{ex}) of the dimensions of 5 × 28 × 14. EEMs of each calibration set were arranged in a data matrix of the dimensions of 41 × 28 × 14 (samples × number of λ_{em} × number of λ_{ex}). The three arrays were decomposed using PARAFAC, which resulted in loadings and scores of components. PARAFAC models were estimated with one to four components and the results were compared. Core consistency >98 % and explained variance >99 % were obtained using the two-component model for all data sets (Table 3).

Loading peaks decomposed by the two-component PARAFAC models in emission and excitation modes were identical for all data sets (Table 3). Similarly to apple spirit, component 1 was only slightly modified by adulteration; however, significant changes were observed for the score of component 2 (results not shown). Therefore, scores of component 1 and 2 in combination with PLS were used for quantitative analysis. The PLS model was again developed in the PLS1 mode using the calibration score matrix with the dimensions of 41 × 2 (calibration samples × number of components). The resulting calibration characteristics are listed in Table 4. The low RMSEC and RMSECV values and high R^2 values (close to 1) confirm the ruggedness of the model for water-plum spirit blends. Regarding ethanol-plum spirit blends, higher values of both RMSEC and RMSECV and low values of R^2 indicate lower accuracy of the calibration model.

Then, the PLS model was used to predict the adulterant concentration in the prediction data set. EEMs of the prediction set were arranged in a data matrix of the dimensions of 20 × 28 × 14 (samples × number of λ_{em} × number of λ_{ex}). The PARAFAC model developed above was used to obtain the new prediction scores. The prediction score matrix with the dimensions of 20 × 2 (prediction samples × number of components) and the PLS model developed before were used to predict the adulterant concentrations. The resulting prediction characteristics are given in Table 4. Low errors in calibration and prediction and high R^2 values (close to 1) indicate good performance of the proposed PLS models for the determination of water in water-plum spirit blends. Regarding ethanol-plum spirit blends, the use of PARAFAC-PLS led to a model with very limited predictive ability showing high RMSECV and RMSEP and low R^2.

Tab. 3. Excitation and emission maxima, explained variance and core consistency versus plum spirit and adulterant-plum spirit blends for two-component PARAFAC models.

Parameter	Plum spirit	Adulterant	
		ethanol	water
Excitation wavelength/emission wavelength (nm)			
component 1	300/420	300/420	300/420
component 2	280, 300/325	280, 300/325	280, 300/325
Explained variance (%)			
component 1	69.1	69.2	69.0
component 2	30.8	30.7	30.9
Total	99.9	99.9	99.9
Core consistency (%)	99.1	98.8	98.7

Tab. 4. Calibration, validation and prediction results of the PLS model for adulterant-plum spirit blends.

Parameter	Adulterant	
	ethanol	water
No LVs	2	2
% of variance score block	89.1	84.4
% of variance concentration block	77.8	74.0
RMSEC	3.2	0.71
RMSECV	3.3	0.73
RMSEP	3.5	0.68
R^2 Cal	0.665	0.986
R^2 CV	0.605	0.986
R^2 Pred	0.663	0.997

No LVs number of latent variables, *RMSEC* root mean square error of calibration, *RMSECV* root mean square error of cross-validation, *RMSEP* root mean square error of prediction, *R^2 Cal* coefficient of determination of calibration, *R^2 CV* coefficient of determination of cross-validation, *R^2 Pred* coefficient of determination of prediction

Conclusions

In this work, calibration model for the determination of water and ethanol in plum and apple spirit was carried out. According to the obtained results, especially RMSEP and squared correlation coefficients R^2 of the analysis of calibration and the prediction matrix, it can be concluded that the PARAFAC-PLS method is versatile for simple and rapid determination of water in both water-apple and water-plum spirit blends. The PARAFAC-PLS model is also applicable for the determination of ethanol in adulterated apple spirit. Prediction of the water and ethanol level in adulterated apple spirit with the root mean square error of prediction (RMSEP) values of 1.9 % and 1.8 %, respectively. However, the use of PARAFAC-PLS provided a model with very limited predictive ability for ethanol-plum spirit blends.

Acknowledgement
This research was supported by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences VEGA No 1/0051/13.

References

Bro R (1997) Chemometrics and Intelligent Laboratory Systems 38: 149–171.
Coldea TE, Socaciu C, Moldovan Z, Mudura E (2014) Notulae Botanicae Horti Agrobotanici Cluj-Napoca 42: 530–537.
Divya O, Mishra AK (2007) Analytica Chimica Acta 592: 82–90.
Genovese A, Ugliano M, Pessina R, Gambuti A, Piombino P, Moio L (2004) Italian Journal of Food Science 16: 185–196.
Jerome J, Workman JJ (2008) In: Burns DA and Ciurczak EW (Ed) Handbook of Near Infrared Analysis 3rd ed., (pp. 123–150). CRC Press, Boca Raton.
Kumar K, Mishra AK (2012) Journal of Fluorescence 22: 339–347.
Ledauphin J, Guichard H, Saint-Clair JF, Picoche B, Barillier D (2003) Journal of Agricultural and Food Chemistry 51: 433–442.
Ledauphin J, Saint-Clair JF, Lablanquie O, Guichard H, Founier N, Guichard E, Barillier D (2004) Journal of Agricultural and Food Chemistry 52: 5124–5134.
Markechová D, Májek P, Kleinová A, Sádecká J (2014) Analytical Methods 6: 379–386.
Miličević B, Lukić I, Babić J, Šubarić D, Miličević R, Ačkar D, Miličević D (2012) Technologica Acta 5: 1–7.
Ortiz MC, Sarabia LA, García I, Giménez D, Meléndez E (2006) Analytica Chimica Acta 559: 124–136.
Reis MM, Biloti DN, Ferreira MMC, Pessine FBT, Teixeira GM (2001) Applied Spectroscopy 55: 847–851.
Sádecká J, Uríčková V, Hroboňová K, Májek P (2015) Food Analytical Methods 8: 58–69.
Satora P, Tuszyński T (2008) Journal of the Science of Food and Agriculture 88: 167–174.

Smilde A, Bro R, Geladi P (2004) Multi-way Analysis: Applications in the Chemical Sciences. John Wiley & Sons, New York.
Song LX, Wang HM, Xu P, Yang Y, Zhang ZQ (2008) Chemical and Pharmaceutical Bulletin 56: 468–474.
Surribas A, Amigo JM, Coello J, Montesinos JL, Valero F, Maspoch S (2006) Analytical and Bioanalytical Chemistry 385: 1281–1288.
Tešević V, Nikićević N, Jovanović A, Djoković D, Vujisić L, Vučković I, Bonić M (2005) Food Technology and Biotechnology 43: 367–372.
Tóthová J, Sádecká J, Májek P (2009) Czech Journal of Food Sciences 27: 425–432.
Velíšek J, Pudil F, Davídek J, Kubelka V (1982) Zeitschrift fur Lebensmittel-Untersuchung und Forschung 174: 463–466.
Wise BM, Gallagher NB, Bro R, Shaver JM, Windig W, Koch RS (2006) PLS_Toolboox Version 4.0 for use with Matlab™. Eigenvector Research Inc., Wenatchee.
Wold S, Sjöström M, Eriksson L (2001) Chemometrics and Intelligent Laboratory Systems 58: 109–130.
Zhan H, Jiang ZT, Wang Y, Li R, Dong TS (2008) European Food Research and Technology 227: 1507–1513.