Determination of the arabica/robusta composition of roasted coffee according to their sterolic content

The Analyst

M. S. Valdenebro,[b] M. León-Camacho,[b] F. Pablos,[a] A. G. González[a] and M. J. Martín*[a]

[a] *Department of Analytical Chemistry, University of Seville, 41012 Seville, Spain*
[b] *Instituto de la Grasa, CSIC, 41012 Seville, Spain*

Received 22nd March 1999, Accepted 24th May 1999

A method for the determination of the percentage of the *arabica* coffee in mixtures of roasted coffee is proposed. The sterol content of roasted coffee blends was determined by extracting the coffee oil, saponifying the lipids and the sterols present in the unsaponifiable fraction were separated by thin layer chromatography. Then, they were converted into trimethyl silyl derivatives and analysed by gas chromatography. Twelve sterols were determined in roasted coffee samples which were mixtures of the arabica and robusta classes. Considering the sterols as chemical descriptors, principal component regression was applied. Δ^5avenasterol was found to be a very adequate variable to establish the arabica percentage in roasted coffee blends. The method was applied to the determination of the arabica–robusta composition of commercial roasted coffee samples.

Introduction

The two species of coffee with greatest commercial significance are *Coffea arabica* and *Coffea canephora*, that are known in the trade, respectively, as arabica and robusta, and will be referred to by these names throughout. There are other species like *Coffea excelsa* and *Coffea liberica* but they are of minor interest for the market.[1] Roasted coffee is commercially available as one of the two first classes, arabica and robusta, or blends of them. In the case of mixtures, the percentage of robusta is usually lower than arabica's. Due to its more pronounced and finer flavour[2] the arabica coffees are considered of better quality and consequently they command for higher prices. To take this situation into account, it is important to have appropriate methods to discriminate between the two classes mentioned and to determine the composition of their mixtures. The chemical analysis is a very useful tool to differentiate between these categories. Several studies[3–5] about the chemical composition of coffee beans have been carried out. Some of the parameters used to characterise the arabica and robusta coffees are 16-O-methylcafestol,[6,7] the metal content[8,9] the volatile components.[3,10] The lipid composition of the coffee seeds has also been analysed.[11] Particularly, the sterolic fraction of the lipids present in the coffee oil is a very interesting approach. Within this realm, there are several reports in the literature.[12–16]

This paper deals with a method to determine the arabica percentage in mixtures of roasted coffee. A study of the sterolic fraction of arabica–robusta roasted coffee blends has been carried out. After the extraction of the coffee oil, the lipids have been saponified and the sterols present in the unsaponifiable fraction have been separated by thin layer chromatography (TLC), converted into trimethyl silyl (TMS) derivatives and analysed by gas chromatography (GC) with flame ionisation detection (FID).

Considering the sterols studied in the blends as chemical descriptors and applying principal component regression (PCR), the composition of mixtures of roasted commercial coffees has been determined.

Experimental

Apparatus

A Hewlett Packard (Palo Alto, CA, USA) 5890 II gas chromatograph equipped with a flame ionisation detector, a fused silica capillary column of 30 m × 0.32 mm coated with a 0.2 μm film of HP-5 stationary phase and a Hewlett Packard 7673 automatic injector was used. The oven was isothermally operated at 265 °C and the injector and detector were held at 280 and 300 °C, respectively. Hydrogen was used as carrier gas at a flow rate of 0.7 ml min^{-1} through the column with a split ratio of 1:80. Air and hydrogen with flow rates of 430 and 30 ml min^{-1}, respectively, were used for the detector, which had an auxiliary flow of 30 ml min^{-1} of nitrogen.

Reagents and solutions

Cholesterol, campesterol, stigmasterol and β-sitosterol were obtained from Fluka (Buchs, Switzerland). Campestanol, Δ^7campesterol, chlerosterol, sitostanol, Δ^5avenasterol, $\Delta^{5,24}$stigmastadienol, Δ^7stigmastenol and Δ^7avenasterol were obtained and purified from a mixture of lard and olive, rapeseed and sunflower oils. 5α-Cholestane-3β-ol (Fluka) in a 0.2% (m/v) solution of isopropyl ether was used as an internal standard in the gas chromatography determinations. Acetone (Panreac, Barcelona, Spain), chloroform (Riedel-de Haën, Seelze, Germany), diethyl ether (Romil, Cambridge, UK), isopropyl ether (Fluka), ethanol (Riedel-de Haën), *n*-hexane (Romil), anhydrous sodium sulfate (Fluka) and potassium hydroxide (Merck, Darmstadt, Germany) used were of analytical reagent grade. A mixture (9:3:1,v/v/v) of anhydrous pyridine (Fluka), hexamethyldisilazane (Fluka) and trimethylchlorosilane (Fluka) was used as derivatization reagent. A 0.2% (m/v) solution of 2',7'-dichlorofluorescein (Fluka) in ethanol was used to display the bands in the TLC separations.

Samples

Arabica and robusta green coffee samples were supplied by Kraft Jacobs Suchard. They were laboratory roasted, ground

and then stored in polyethylene flasks. Rossell *et al.*[17] pointed out that the temperature treatments do not alter the sterol content of vegetable oil samples. Accordingly, slight differences in roasting from one lot to another should not be considered as a source of spreading with respect to the analytical procedure.

In order to prepare the different arabica–robusta mixtures, two pools of arabica and robusta coffees were prepared by blending nine different lots of each class of coffee. These lots differed both in their origin and harvest year. Thus, for the arabica pool, the origins were: Brazil (2), Nicaragua (2), Honduras, Salvador, Colombia (2) and Guatemala; and for the robusta one: Vietnam, Indonesia, Cameroon (2), Uganda (3) and Ivory Coast (2). The harvest year for the individual lots, according to the information provided by Kraft Jacobs Suchard, ranged between 1992–1995. This randomization procedure may overcome the dependency on origin and crop to some extent. From these two pools, 13 blends were prepared. The *arabica* content varied in the range 100–40% (m/m). A single 100% pure robusta sample was also considered for comparison purposes. Commercial roasted coffee samples were purchased from the market. Before analysis, the samples were dried at 103 °C [ref. (18)] until constant weight to determine their moisture.

Extraction of the coffee oil

The extraction of the coffee oil and its saponification was performed according to the Directive 91/2568/CEE.[19] Thus, about 8.0 g, exactly weighed, of roasted coffee sample, or blend, were extracted with hexane[16,19] in a Soxhlet for 8 h, siphoning six times per hour. The extract was dried over anhydrous sodium sulfate and placed in a 250 ml round bottom flask. Using a vacuum rotary evaporator the solvent was evaporated and the residue was dried at 105 °C to obtain the exact amount of coffee oil. Then, 1.5 ml of internal standard solution was added.

Saponification of the lipids

The coffee oil obtained was treated with 50 ml of 2 M solution of potassium hydroxide in ethanol–water (80 : 20, v/v) and this mixture refluxed, with constant stirring, for 30 min. This way, the esters present in the oil were transformed into potassium salts that are soluble in water. The sterols did not react and could be extracted into non-polar solvents like ethyl ether. Next, 50 ml of water was added and extractions with 3 portions of 80, 70 and 60 ml of diethyl ether were carried out. The organic extract was separated and washed with 5 portions of 100 ml of water. Then, it was dried over anhydrous sodium sulfate, filtered into a 250 ml round bottom flask and concentrated on a rotary evaporator under reduced pressure at room temperature to distil the diethyl ether. At this step, the unsaponifiable fraction was dissolved into chloroform to get a solution of *ca*. 5% (m/v).

Separation of the sterolic fraction from the unsaponifiable one by TLC

The separation of the sterolic fraction from the unsaponifiable one by TLC presents a minor modification with respect to the CEE method.[19] The eluent used for the TLC separation was a mixture hexane–diethyl ether–acetone (60 : 40 : 1, v/v/v) instead of benzene–acetone (95 : 5 v/v) or hexane–ethyl ether (65 : 35 v/v). This modification was recently proposed and established.[20]

Accordingly, the solution obtained in the previous section was spotted on a potassium hydroxide-impregnated silica gel TLC plate. 5 μl of the internal standard solution was also spotted with the solution of the sample. The plate was developed using hexane–diethyl ether–acetone (60 : 40 : 1, v/v/v). Once dried, it was sprayed with the 2′,7′-dichlorofluorescein solution and the pink band of the sterols can be observed under UV light together with the spot of the internal standard. After identifying it, the band was scraped off and the sterols were dissolved with a portion of 10 ml of trichloromethane and two portions of 10 ml of diethyl ether. The solution obtained was filtered through a paper filter, the solvent was evaporated under reduced pressure and the sterolic fraction was dried in an oven at 105 °C.

Determination of the sterols by GC

The procedure used for the determination of sterols in the extracted coffee oil was an international standard one developed for determining the content and composition of sterols in animal and vegetable fats by gas chromatography.[21] The reliability of the method was established already by an international interlaboratory test including 14 laboratories, organised by the ISO German member body in 1996/97 on samples of olive oil, sunflower seed and oil rapeseed in accordance with ISO 5725 standard.

Accordingly, the sterolic fraction isolated from TLC, as indicated above, was treated with the derivatisation reagent to obtain the trimethylsilyl (TMS) derivatives, which are much more volatile. A volume of 0.05 ml of reagent for each mg of sterol was added. Aliquots of 2 μl of this solution were injected into the gas chromatograph and the ratio of the peak areas of the analyte and internal standard was used as analytical signal. The content of individual sterols was expressed as the percentage of the sterolic fraction obtained.

Albeit the method is sound and well established, a previous study on the accuracy based on recovery assays from spiked samples was done. Four additions of standard mixtures of the studied sterols were spiked on samples of extracted coffee oils, and then, the analytical method was applied. Average recoveries (in %) for each sterol were calculated therefrom and are presented as follows: cholesterol (106.2%), campesterol (98.0%), campestanol (91.2%), stigmasterol (96.1%), Δ^7campesterol (99.0%), chlerosterol (97.9%), β-sitosterol (99.7%), sitostanol (93.9%), Δ^5avenasterol (94.6%), $\Delta^{5,24}$stigmastadienol (98.5%), Δ^7stigmastenol (98.0%) and Δ^7avenasterol (100.9%). These results show that the analytical procedure can be considered accurate according to the AOAC guidelines.[22]

Data analysis

Twelve sterols have been analysed in the roasted coffee mixtures and considered as chemical descriptors. These sterols will be viewed as follows: cholesterol (COL), campesterol (CPR), campestanol (CPN), stigmasterol (STR), Δ^7campesterol (D7C), chlerosterol (CLE), β-sitosterol (BSIT), sitostanol (SIT), Δ^5avenasterol (D5), $\Delta^{5,24}$stigmastadienol (D524), Δ^7stigmastenol (D7S) and Δ^7avenasterol (D7A). Table 1 shows the content of the sterols found in the analysed blends, indicating the percentage of *arabica* in each of the mixtures.

PCR calculations were made for the compositional analysis of the blends, using the statistical package CSS: STATISTICA from Stafsoft™ (Tulsa, OK, USA).

Results and discussion

The analysis of the sterolic content by GC after TLC separation of the unsaponifiable fraction of the oil present in coffee samples has been proved to be very adequate to differentiate between arabica and robusta.[15] Fig. 1 shows the chromatograms corresponding to 100% *arabica*, 100% *robusta* and 50% *arabica* roasted coffee samples, in which the peaks of the TMS derivatives of the sterols can be observed. As can be seen, the profile of the chromatogram is distinct and the major differences appear for peaks 9, 11 and 12, that correspond to D5, D7S and D7A, respectively.

Resolution of arabica–robusta mixtures by PCR

As indicated above, 12 descriptors feature each coffee mixture, which are the contents (in % m/m) of 12 different sterols. Thus

we have 13 mixtures (within 100–40% arabica) that can be arranged as a column vector ***y*** (column 1 of Table 1 except for the last row) and a descriptor data matrix ***X*** of dimension 13 × 12 (from columns 2 to 13 of Table 1). The last row of Table 1 refers to a pure *robusta* sample for the sake of comparison, but cannot be considered in the modelling of blends.

In multivariate analysis we often have to model a response ***y***, in this case the % arabica of coffee mixtures, as function of several X-variables (here, the selected descriptors). Instead of considering a multiple linear regression to model ***y*** as a function on ***X***, it is better to perform a principal component analysis (PCA) on the ***X*** matrix and use the issued principal component (PC) scores as the basis of fitting a multiple regression model. This approach is called principal component regression (PCR) and has found application in areas such as multivariate calibration, structure–activity relationships and modelling chemical properties of molecules.[23] PCR joins PCA and regression techniques, and it is particularly appropriate to avoid multi-collinearity problems among ***X***-variables.[24] First, PCA is carried out on the autoscaled ***X*** matrix to derive a low dimension representation of the samples through a few PCs (which are orthogonal and cannot exhibit collinearity features). Then the *y*-property can be modelled as a function of the scores of these PCs as follows:

$$y = b_0 + b_1\,(\mathrm{PC}_1\ \text{scores}) + b_2\,(\mathrm{PC}_2\ \text{scores}) + \ldots + b_f\,(\mathrm{PC}_f\ \text{scores}) \quad (1)$$

f being the established dimensionality of the factor space.

Thus, after PCA of our ***X*** autoscaled data matrix, two PCs (*f* = 2) were selected according to the Kaiser criterion,[25] the Malinowski indicator function[26] and the method of cross validation.[27] Kaiser criterion, also called the average eigenvalue criterion is based upon retaining only those PC dimensions whose eigenvalues are above the average eigenvalue. If the data are autoscaled, then any component having an eigenvalue higher than unity is retained. Malinowski devised an indicator function called IND that is very sensitive to identify the true dimensionality of data matrices. IND is an empirical function based on the residual standard deviation of the data matrix and reaches a minimum for the right number of factors. Cross validation aims to identify those dimensions (first *f* PCs) with best predictive ability to rebuild the data matrix once some data groups were deleted. As indicated above, the three methods lead to two underlying factors, that is, a two PC model.

The *y* property (% *arabica*) was then modelled with the corresponding PC scores giving a first linear fit

$$\%\ \text{arabica} = 70 + 19.3\,(\mathrm{PC}_1\ \text{scores}) - 0.8\,(\mathrm{PC}_2\ \text{scores}) \quad (2)$$

with a correlation coefficient of 0.992. However, according to the *t*-test, the coefficient for the PC_2 scores was found without statistical significance and the corresponding term dropped out from the model. The new fit was fairly linear

$$\%\text{arabica} = 70 + 19.3\,(\mathrm{PC}_1\ \text{scores}) \quad (3)$$

with a correlation coefficient of 0.991.

The predicted values were plotted against the actual % arabica in Fig. 2. In order to detect possible outliers and for checking linearity, both analyses of residuals and regression ANOVA were used. A plot of the residuals on normal probability paper shows model adequacy. Residuals are also used to detect outliers. A very straightforward way is to consider as outlier any point whose residual is greater than twice the value of the standard deviation of the regression line.[28] In our case, outliers were absent according to this criterion. The ANOVA of the regression shows that the *F* ratio between the

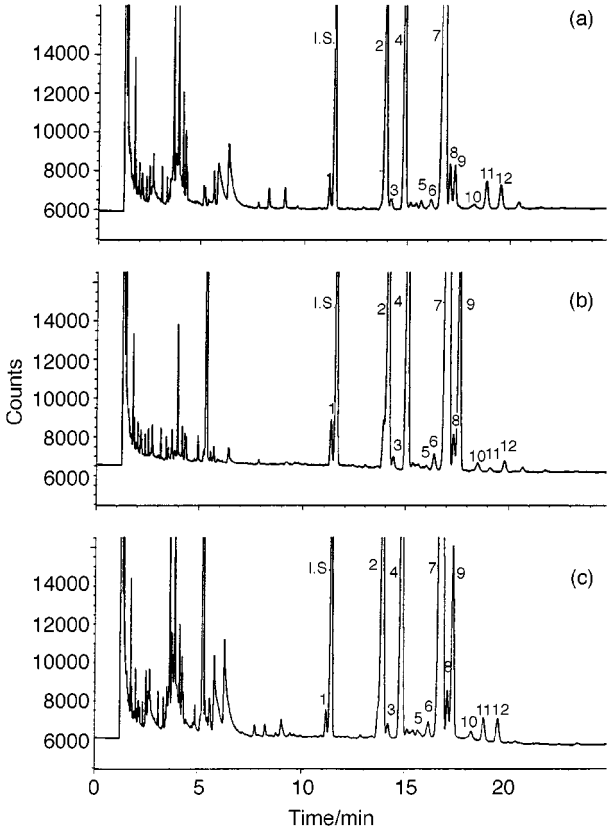

Fig. 1 Chromatograms of the sterolic fraction of roasted coffee samples: (a) 100% arabica, (b) 100% robusta, (c) 50% arabica. (1) Cholesterol (COL), (I.S.) 5α-cholestane-3β-ol, (2) campesterol (CPR), (3) campestanol (CPN), (4) stigmasterol (STR), (5) Δ^7campesterol (D7C), (6) chlerosterol (CLE), (7) β-sitosterol (BSIT), (8) sitostanol (SIT), (9) Δ^5avenasterol (D5), (10) $\Delta^{5,24}$stigmastadienol (D524), (11) Δ^7stigmastenol (D7S) and (12) Δ^7avenasterol (D7A).

Table 1 Composition of the sterolic fraction (%) of roasted coffee blends[a]

% Arabica	COL	CPR	CPN	STR	D7C	CLE	BSIT	SIT	D5	D524	D7S	D7A
100	1.2(1)	15.5(4)	0.73(2)	18.9(7)	0.6(2)	0.87(1)	52.7(2)	2.41(7)	2.84(4)	0.6(2)	2.04(3)	1.74(5)
95	1.2(1)	15.6(4)	0.65(2)	18.0(8)	0.6(4)	0.93(1)	53.0(8)	2.01(9)	3.05(8)	0.4(2)	1.76(9)	1.59(9)
90	0.9(3)	15.8(2)	0.68(1)	18.8(5)	0.5(3)	0.84(4)	52.5(7)	2.02(9)	3.72(8)	0.5(1)	1.75(5)	1.57(5)
85	1.1(2)	15.7(2)	0.67(1)	18.7(3)	0.6(3)	0.96(7)	52.2(2)	1.96(3)	4.27(6)	0.6(2)	1.71(8)	1.56(6)
80	1.1(2)	15.9(1)	0.65(3)	18.7(2)	0.5(5)	0.86(4)	52.3(4)	1.87(7)	4.59(7)	0.5(1)	1.54(2)	1.46(2)
75	1.0(1)	16.2(3)	0.67(3)	19.1(4)	0.6(1)	0.94(8)	51.6(2)	1.89(3)	5.07(6)	0.6(2)	1.49(6)	1.39(8)
70	0.9(2)	16.3(1)	0.56(9)	19.4(1)	0.5(1)	0.86(8)	50.5(4)	1.91(4)	5.70(6)	0.6(1)	1.42(6)	1.38(4)
65	1.0(1)	16.2(1)	0.62(3)	19.0(1)	0.6(2)	0.99(9)	50.2(6)	1.75(6)	6.09(8)	0.8(1)	1.41(9)	1.34(9)
60	0.8(1)	16.4(2)	0.73(9)	19.2(1)	0.7(2)	1.08(9)	49.7(6)	1.85(9)	6.80(4)	0.9(3)	1.40(9)	1.33(7)
55	—	16.6(1)	0.72(9)	19.1(1)	0.6(1)	0.99(8)	50.3(5)	1.83(9)	6.91(5)	0.6(1)	1.20(8)	1.21(3)
50	0.8(1)	16.6(4)	0.61(4)	19.1(4)	0.6(1)	0.95(7)	50.1(8)	1.80(4)	7.54(7)	0.7(3)	1.09(6)	1.14(3)
45	0.8(2)	16.5(1)	0.59(5)	18.8(2)	0.7(2)	0.96(8)	49.1(4)	1.67(8)	7.87(3)	0.8(1)	1.06(7)	1.11(9)
40	1.0(1)	16.7(1)	0.56(2)	19.1(2)	0.5(1)	0.98(6)	49.0(3)	1.53(1)	8.38(3)	0.7(1)	0.84(4)	0.99(2)
0	1.3(1)	17.2(1)	0.51(1)	18.1(2)	0.4(2)	0.84(8)	47.9(1)	1.34(4)	11.07(5)	0.6(1)	0.25(7)	0.55(1)

[a] Average of triplicate determinations. Values between parentheses refer to the error corresponding to the last significant figure.

regression variance and residual variance was above 600, indicating that linearity was fair.

However, in spite of these findings, the prediction model was built by using PCs, that is, 'abstract factors' without a proper chemical definition. Our principal aim was to find relationships between the *y* variable and the chemical descriptors rather than PCs. Anyway, after a glance of eqn. 3, PC_1 scores is the predictor variable. If the corresponding PC_1 loadings are considered, one can find that the descriptors with highest contributions (> 0.92) are CPR, BSIT, SIT, D5, D7S and D7A. A further study of correlation shows that all these descriptors are mutually correlated, with correlation coefficients ranging from 0.88 (CPR&SIT) to 1.00 (D7S&D7A). This indicates that all the variables contributing to PC_1 are redundant and consequently, any of them gives the same information as the complete set.

PC_1 is the capital factor for modelling the % *arabica*. Besides, all the descriptors contributing to PC_1 are redundant, and so this factor could be associated to one descriptor alone. Following this deductive way, the next step was the study of the prediction of % *arabica* by using as unique descriptor one of the set (CPR, BSIT, SIT, D5, D7S and D7A). The best fit was achieved with D5:

$$\% \text{ arabica} = 129 - 10.5\ D5$$

with a correlation coefficient of 0.998. In this case, the plot of residuals on normal probability paper was linear and again outliers were absent. The ANOVA of regression leads to a *F* ratio of about 2850 which indicates a very good linearity.

The predicted values were plotted against % arabica in Fig. 3. The standard error of prediction was less than 1.1%.

The result obtained is outstanding: By using the descriptor D5 alone, it is possible to predict the % arabica in coffee mixtures with better reliability than using the first PC. The possible explanation may arise from the fact that the majority of the contributing variables involved in the fundamental factor PC_1, are mutually correlated and give the same information than one of them. The use of D5 as a key descriptor filters the possible noise provoked by the remaining descriptors and gives a good linear fit.

In order to validate the results so obtained in the modelling study, a test set consisting of real coffee samples purchased in the market was considered. All the test samples were certified for the content in arabica or robusta that appeared on the label. The sterolic profile of seven samples 100% arabica, a mixture of 75% arabica and 100% robusta was obtained. From the D5 content, the % arabica was calculated. The results obtained were excellent as can be seen in Table 2.

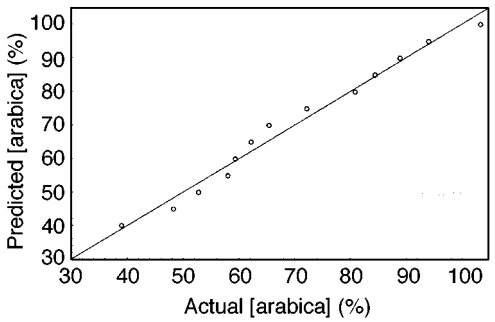

Fig. 2 PCR predicted *vs.* actual of % arabica in arabica–robusta mixtures.

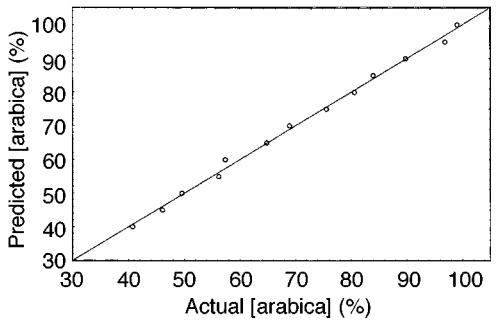

Fig. 3 Δ^5avenasterol predicted *vs.* actual of % arabica in arabica–robusta mixtures.

Table 2 Percentage of arabica in commercial coffees

Claimed	100	100	100	100	100	100	100	75	0 (100% Robusta)
Found	104	103	103	95	101	105	100	75	3 (97% Robusta)

Conclusion

The analysis of the sterolic fraction of the oil of roasted coffee blends provides a very useful tool for establishing the percentage of arabica–robusta. D5 is the most adequate chemical descriptor for predicting the arabica–robusta content in coffee samples.

References

1 A. W. Smith, in *Coffee*, ed. R. J. Clarke and R. Macrae, Elsevier, London, 1985, vol. 1, pp. 3–6.
2 R. Briandet, E. K. Kemsley and R. H. Wilson, *J. Agric. Food Chem.*, 1996, **44**, 170.
3 C. P. Bicchi, O. M. Panero, G. M. Pellegrino and A. C. Vanni, *J. Agric. Food Chem.*, 1997, **45**, 4680.
4 M. J. Martín, F. Pablos and A. G. González, *Anal. Chim. Acta*, 1996, **320**, 191.
5 M. Suchánek, H. Filipova, K. Volka, I. Delgadillo and A. N. Davies, *Fresenius J. Anal. Chem.*, 1996, **354**, 327.
6 K. Speer, R. Tewis and A. Montag, *Z. Lebensm. Unters. Forsch.*, 1991, **192**, 451.
7 M. D. Trouche, M. Derbesy and J. Estienne, *Ann. Falsif. Expert. Chim.*, 1997, **90**, 121.
8 M. J. Martín, F. Pablos and A. G. González, *Anal. Chim. Acta*, 1998, **358**, 177.
9 S. J. Haswell and A. D. Walmsley, *J. Anal. At. Spectrom.*, 1998, **13**, 131.
10 A. Murota, *Biosci. Biotechnol. Biochem.*, 1993, **57**, 1043.
11 G. Muratore, M. C. Cataldi-Lupo, F. Fiorenza and C. N. Asmundo, *Ind. Aliment.*, 1998, **37**, 161.
12 E. Tiscornia, M. Centi-Grossi, C. Tassi-Micco and F. Evangelisti, *Riv. Ital. Sostanze Grasse*, 1979, **56**, 283.
13 N. Frega, F. Bocci and G. Lercker, *J. High Resolut. Chromatogr.*, 1994, **17**, 303.
14 G. Lercker, M. F. Caboni, G. Bertacco, E. Turchetto, A. Lucci, R. Bortolomeazzi, E. Pagani, N. Frega and F. Bocci, *Ind. Aliment.*, 1996, **35**, 1186.
15 F. Carrera, M. León-Camacho, F. Pablos and A. G. González, *Anal. Chim. Acta*, 1998, **370**, 131.
16 G. Lercker, N. Frega, F. Bocci and M. T. Rodríguez-Estrada, *Chromatographia*, 1995, **41**, 29.
17 J. B. Rossell, S. P. Kochhar and I. M. Jawad, *Chemical changes in soy oil during high temperature processing. Proceedings of the 2nd ASA Symposium on Soybean Processing*, Antwerp, June 1981, p. 367.
18 International Organization of Standardization ISO 11294, 1994.
19 Directive 91/2568/CEE, Official journal L-248/15, September 1991.
20 F. Carrera, M. León-Camacho, F. Pablos and A. G. González, *Anal. Chim. Acta* 1998, **370**, 131-139
21 ISO/DIS 12228: 1997.
22 AOAC, Peer verified method program, manual on policies and procedures, Arlington, VA, November, 1993.
23 W. P. Gardiner, *Statistical Analysis Methods for Chemists*, The Royal Society of Chemistry, Cambridge, 1997, p. 311.
24 R. Henrion and G. Henrion, *Multivariate Datenanalyse*, Springer-Verlag, Berlin, 1995, p. 129.
25 H. F. Kaiser, *Educ. Psychol. Meas.*, 1966, **20**, 141.
26 E. R. Malinowski, *Factor analysis in Chemistry*, 2nd edn., Wiley, New York, 1991.
27 S. Wold, *Technometrics*, 1978, **20**, 397.
28 A. G. González, M. A. Herrador and A. G. Asuero, *Talanta*, 1999, **48**, 729.

Paper 9/02245G