Analysis of some metals in wine by means of capillary electrophoresis. Application to the differentiation of *Ribeira Sacra* Spanish red wines

M. Núñez, R.M. Peña, C. Herrero and S. García-Martín[*]

Departamento de Química Analítica, Nutrición y Bromatología, Facultad de Ciencias,
Universidad de Santiago de Compostela, Avda. Alfonso X El Sabio s/n, 27002 Lugo, Spain

Abstract. The content of some metals (Na, K, Ca, Mg, Mn and Li) in a variety of red wines produced in different geographical zones of Galicia was determined in a single analysis by capillary electrophoresis. Data were processed using multivariate supervised and non-supervised chemometric techniques involving cluster analysis, principal component analysis (PCA), linear discriminant analysis (LDA), *K* nearest neighbours (KNN) and soft independent modelling of class analogy (SIMCA). The results obtained have allowed for the categorisation of wine samples according to their origin; the wines with *Ribeira Sacra* Certified Brand of Origin (CBO) can be differentiated from the wines of two other CBO from Galicia: *Ribeiro* and *Valdeorras*. The latter two types of wine are made with the same or similar grape varieties; hence, due to their similar organoleptic properties they are possible substitutes for the falsification of *Ribeira Sacra* wines.

Keywords. Mineral content – capillary electrophoresis – wines – multivariate analysis.

Introduction

The classification of product brands and quality of origin is a very active area for the application of chemometric classification procedures [1,2]. The application of pattern recognition techniques for the classification of wines and alcoholic beverages has widely increased in recent years. After the works of Kwan and Kowalski on the classification of wines of *Vitis Vinifera* cv. Pinot Noir from France and the United States [3-5] a number of examples have been reported in literature. Pattern recognition methods have been used in the characterisation and classification of wine and alcoholic beverages: Galician (Spain) wines with metal ions [6-8], Venetian (Italy) wines with classical determinations plus aroma compounds [9], Portuguese wines with free amino acids [10], French red wines with elements, amino acids and aromatic alcohols [11] and German wines with volatile compounds, non-volatile acids and amino acids [12]. Multivariate chemometric techniques were used to classify alcoholic distillates and develop a typification model for Galician liquors, on the basis of data from chromatographic analysis [13].

Minerals seem to be the primary candidates for a fingerprint as they are stable. The metallic composition of wine depends on many factors, some of which are related to the specific production area: grape varieties, soil and climate, culture and wine making practices, yeasts and such. The kind of soil has a fundamental influence, as does the climatology and the geographical situation of the producing area.

The variety of grape is also a very important factor in the content of metallic elements in wine. Other factors that contribute to the metal content are the availability of these elements in soil, fertilising practices and processing conditions. Due to the influence of all these factors a great variability in the metal content in wines from different areas, regions and countries is observed. Determination of typical levels of metal in wines is a very useful tool to differentiate wines from different geographic origins and different Certified Brands of Origin (CBO) as well as to detect adulterations and falsifications of wines.

Atomic absorption spectrometry is the classic technique for the determination of the mineral content in wine samples. Atomic absorption, although adequate for parts per million (flame) and ppb determination (furnace), lacks the ability to carry out simultaneous determinations. In most cases the mineral content of foods is determined by ion chromatography, atomic absorption spectrometry or inductively coupled plasma combined with atomic emission spectrometry. As an alternative technique, capillary electrophoresis can be readily applied to the analysis of both anions and cations. Capillary electrophoresis has important advantages with respect to other commonly used techniques: rapid analysis, versatility, determination of several analytes in a single analysis and good detection limits. Because of these advantages, several research reports have described the determination of cationic mineral elements – potassium, sodium, calcium and magnesium – in a variety of food matrices [14, 15].

The production of quality wines has special economic importance in certain vitivinicole regions of Galicia (Northwestern Spain). To achieve a Certified Brand of Origin (CBO) wines must be elaborated with grapes grown in vineyards located in a controlled geographical area employing only the varieties authorized by the CBO Council. Wines are subjected to certain rules related to the alcoholic content and the acidity. Finally, wines are judged by their organoleptic properties by a expert taste panel. The (CBO) wines from *Ribeira Sacra,* a restricted wine-growing region of Galicia, are typically young red wines produced with a high percentage (more than 70 %) of the Mencía grape variety. Brancellao and Caíño are other traditional Galician grape varieties that may be used in a minor proportion. The same or similar grape varieties are also employed in the elaboration of wines in some nearby geographical areas such as Ribeiro CBO and Valdeorras CBO.

The purpose of this study is to differentiate between red wines with *Ribeira Sacra* CBO from other Galician wines produced in two nearby CBO such as Ribeiro and Valdeorras that can be used as substrate for falsification of genuine Ribeira Sacra wines due their similar organoleptic properties. The basis for this differentiation is the content of some selected metals (Na, K, Ca, Mg, Mn and Li) determined by capillary electrophoresis with a single analysis.

Experimental

Wine samples

Twenty-five samples of Galician red wines were analysed. To be sure about the geographic origin of the wines and the grape varieties which have been utilized in making it, the wine samples of the production areas were collected as follows. Fifteen samples with *Ribeira Sacra,* CBO representing the main producers of this region, were provided by the Certification of Origin Council in the production area. The rest were Galician red wines of two origins: five samples from Ribeiro CBO and five samples from Valdeorras CBO were collected directly by the authors. All of the wines came from unsuspicious origin and were made with the traditional varieties for these Spanish producing regions during the period 1994-1995. For differentiating purposes, *Ribeira Sacra* wines were considered class RS, while *Ribeiro* and *Valdeorras* wines as class N-RS. Samples were stored at 3-4 °C before analysis.

Analytical procedure

The content of Na, K, Ca, Mg, Mn and Li was determined by capillary electrophoresis. A simple dilution (1:10) of the sample with ultra pure water provided by a Milli-Q water purification system (Millipore Bedford, MA, USA) was necessary prior to injection in the capillary electrophoretic device. Filtration is not really needed. However, if dregs are present in the wines, the clean-up procedure consists of microfiltration through a 0.45 µm Millex HA filter (Millipore Co.). A Waters Quanta 4000 electrophoretic
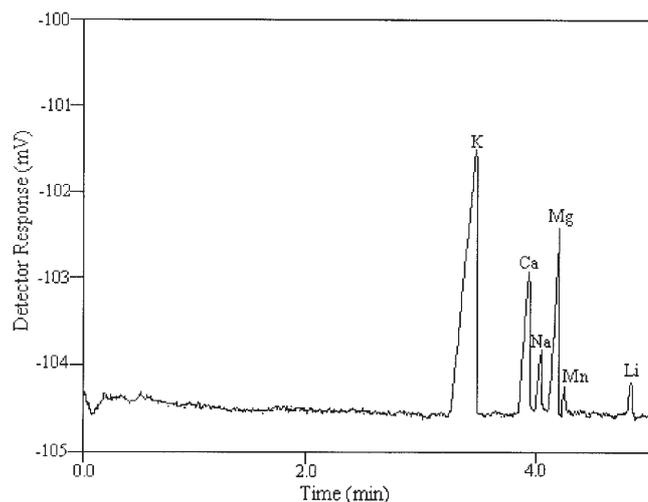
Figure 1. Electropherogram of a sample corresponding to *Ribeira Sacra* CBO wine.

system equipped with a fused silica capillary, 60 cm × 75 µm (Composite metal services Ltd.) was used. The running buffer was 5 mM Cia-Pack UV-cat-1 (Waters Chromatography Milford, MA Reagent WAT 054780), 6.5 mM α-hidroxyisobutiric acid (HIBA) (Fluka Chemica, Buchs, Switzerland, Ref.: 55390) and 2 mM 18-crown-6-ether (Merck Darmstadt, Germany, Ref.: 8.11684.0500) at a final pH 4.5. The applied voltage was 22 kV that provided a current intensity of 4.5 ± 1 µA. Samples were injected using hydrostatic injection for 30 seconds and indirect UV detection at 214 nm was used. The developed method shows good precision in migration time, and peak area, the values of the percentage of relative standard deviation (% RSD) ranged 0.09 to 0.209 for the migration time and 2.5 to 3.4 for the peak area. The linearity of the method was proven up to 150 µg·L^{-1} for Li and to 300 mg·L^{-1} for the other cations. The calibration lines obtained, representing peak area *vs.* concentration for each analyte, presented a correlation coefficient of 0.999 or better in all cases. The obtained detection limits were adequate to determine the six considered cations in the concentration usually present in wines. The results of the recovery were satisfactory for all the elements studied with values comprised between 92 and 102 %. The optimization of the complete procedure, its analytical characteristics and the comparison of the results obtained with those provided by atomic spectroscopy have been described in detail in a previous work [16]. Under the conditions mentioned above, an example of the electropherogram obtained for a wine sample from CBO Ribeira Sacra is presented in figure 1.

Data analysis

The data analysis was performed in three steps.

1. Firstly, a preliminary data analysis by cluster and principal components analysis was performed. Cluster analysis describes the nearness between wine samples (objects). A

similarity matrix was built using the autoscaled data. The elements of this matrix were the squared Euclidean distances of one object from the rest. To obtain clusters, the Ward method was used. In each step this agglomerative method considers the heterogeneity or deviance (the sum of the square of the distance of an object from the barycenter of the cluster) of every possible cluster that can be created by linking two existing clusters [17]. Principal component analysis was used to provide a data structure study in a reduced-dimension retaining the maximum amount of variability present in the data. The principal components or eigenvectors are orthogonal and they are a linear combination of the original variables [18].

2. Secondly, classification techniques LDA [19], KNN [20] and SIMCA [21] were applied to complete data set with all samples as training set, divided in a two category arrangement: *Ribeira Sacra* wines (class RS) and *Non-Ribeira Sacra* wines (class N-RS).

3. In step three, validation of the classification systems developed in step 2 was performed by means of a cross-validation method with 25 % of samples as the test set. The same process was repeated four times with different compositions of both training and test sets, to ensure that all samples were included in the test set at least once. The performance of the classification model for the three classification procedures has been studied in terms of recognition and prediction abilities. The recognition ability is characterized by the percentage of the members of the training set that are correctly classified. The prediction ability is determined by the percentage of the members of the test set correctly classified by using the classification rules developed during the training step.

Pattern recognition analysis was performed by means of the statistical software packages Statgraphics [22] and Parvus [23].

Results and discussion

The results of the six selected metals for the *Ribeira Sacra*, *Ribeiro* and *Valdeorras* CBO wines are presented in table I. The levels obtained in Galician wines were comparable to those reported by other authors in wine samples from Galicia [6-8].

The search for natural groupings among the samples is one preliminary way to study the data structure. The results of the cluster analysis are shown as a dendrogram in figure 2. The results obtained showed the presence of wine clusters; the data of metal content in wines contained information useful in achieving a category classification that confirms the origin of the wines. At a similarity level of 0.6, three clusters were found. The first cluster contained all N-RS wines from CBO Valdeorras. The second cluster was made of wines from classes RS and N-RS (in this case Ribeiro CBO wines) brands and the third cluster is formed only by wines of RS class.

Table I. Metal content in Galician red wines (all results are in mg·L^{-1} except Li, which is in μg·L^{-1}).

Element	Mean	SD	Max	Min
Ribeira Sacra Wines				
Li	46	20	85	19
Ca	68.1	9.3	86.1	50.7
Na	10.8	7.2	22.8	1.0
Mg	70.1	4.2	77.3	64.7
Mn	4.4	3.3	9.5	0.6
K	1116	145	1368	923
Ribeiro Wines				
Li	51	10	64	40
Ca	78.2	8.4	92.7	71.0
Na	24.2	24.6	67.7	7.7
Mg	75.5	6.4	85.8	70.6
Mn	0.5	0.1	0.6	0.4
K	1055	136	1202	839
Valdeorras Wines				
Li	46	25	72	20
Ca	27.7	3.0	30.6	22.8
Na	8.0	3.6	12.0	4.0
Mg	21.0	3.5	26.5	16.8
Mn	0.8	0.3	1.2	0.5
K	1702	118	1883	1562
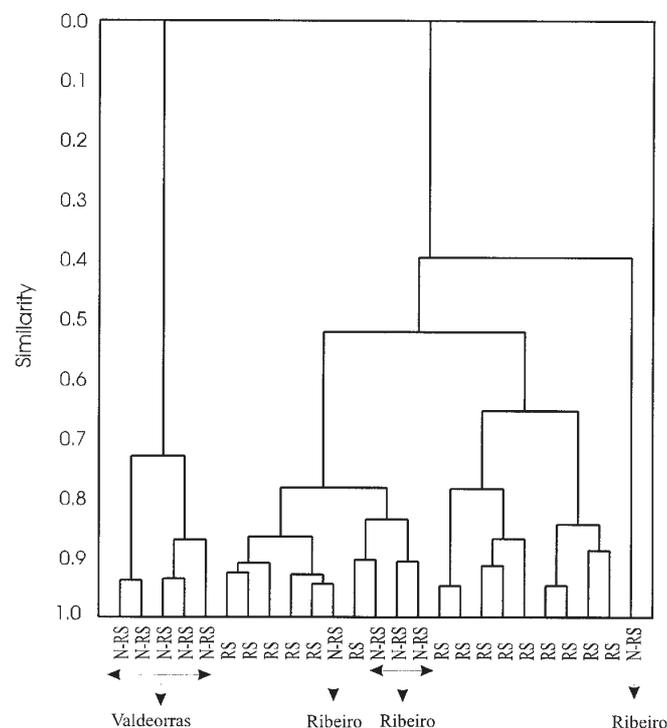
Figure 2. Dendogram of cluster analysis of wines with *Ribeira Sacra* origin (class RS) and Non-*Ribeira Sacra* origin (class N-RS).

Table II. Loadings of the variables in the first three principal components.

	Variable					
	Li	Na	K	Mn	Ca	Mg
Principal component 1	–0.059	–0.168	0.554	–0.170	–0.561	–0.563
Principal component 2	0.608	–0.463	0.015	0.640	–0.049	–0.056
Principal component 3	–0.664	–0.721	–0.092	0.125	0.044	0.111

Principal component analysis was performed in the complete data set. From the loadings of features in the first and second principal components (see Tab. II), potassium, calcium and magnesium are the dominating features in the first principal component (47.42 % of the total variability). This fact can be explained because calcium and magnesium contents in wines will be influenced by the soil conditions, the treatment of the musts with calcium carbonate, the pH, the time and temperature of storage and the use of fining agents. The potassium content in wine is influenced by the variety of grapes, soil and climatic conditions, time of harvest and other variables such as the temperature of fermentation and storage, the pH and the use of ion-exchange resins. Due to the great number of factors influencing the content of this metal in wine, the levels of potassium are different according to the type of wine, their geographical origin, etc. The second principal component is dominated by the contribution of sodium, manganese and lithium. This fact can be chemically explained because lithium and manganese are related to the kind of soil. The sources of sodium in wines, apart from natural concentration, are the addition of sodium salts such a sulphite, sorbate and (in some countries) sulphide. In Spain, sodium chloride may be added to improve clarification. Hence, this principal component can be associated with factors like soil and the kind and quantities of additives used in winemaking. The first three principal components, which account for 83 % of the total variability, were considered to be sufficient for such data. Group identification by PCA achieved interesting results. When a three-dimensional plot of the objects in the space defined by the three principal components was drawn, a natural separation of the objects (wine samples) according to their origin was achieved (Fig. 3). Wines of *Ribeira Sacra* origin form a separate group and wines of Ribeiro and Valdeorras origin appear in two other separated groups. In this factor space, Valdeorras wines make up a well-separated group. This result is consistent with the conclusion obtained by cluster analysis, where all Valdeorras origin wines were grouped in a single cluster. Ribeira Sacra and Ribeiro are the two closest groups as in cluster analysis where some Ribeira Sacra CBO wines were clustered with Ribeiro CBO wines.

In the second step of data analysis, LDA, KNN and SIMCA were applied to the complete data set after autoscaling to eliminate the effect of different size variables. In this step, all the samples are in the training set. LDA achieved a percentage of 93.3 % of correct classification for RS wines

(one sample misclassified) and 80.0 % for N-RS wines (two samples misclassified). KNN, that were applied using the inverse square of the Euclidean distance and K = 3, produced a level of hits similar to LDA: 93.3 % for RS and 70.0 % for N-RS wines. For both classification procedures, a high level of correct assignation of wines from Ribeira Sacra with a percentage of success higher than 90 % was achieved. A less successful result was obtained for the Non-Ribeira Sacra wines. This fact indicates that, in these cases, the pattern recognition procedure is more sensitive for Ribeira Sacra wines; the probability of a Ribeira Sacra wine being classified as Non-Ribeira Sacra is low. However, a minor level of hits in classification of Non-Ribeira Sacra suggests that there exists a certain probability that a Non-Ribeira Sacra wine might be considered as genuine Ribeira Sacra. This result coincides with the one obtained by PCA and cluster analysis, where some samples from Non-Ribeira Sacra origin were grouped into genuine Ribeira Sacra wines (see Figs. 1 and 2). On the contrary, SIMCA performed a more sensitive model for N-RS wines, 90.0 % of the objects belonging to the category N-RS were correctly classified as N-RS class, while only 73.3 % of RS wines were assigned to RS class. This fact can be explained by taking into account that SIMCA is a modelling technique. There exists an important difference between modelling (as SIMCA) and classification methods (as LDA and KNN); the last procedures only divide the multidimensional space of the variables in disjoint zones corresponding to each category, and a considered sample is classified in one category or another according to its position in the multivariate space. Modelling procedures involve the construction of an enclosure for each category in the space determined by the six metallic variables.
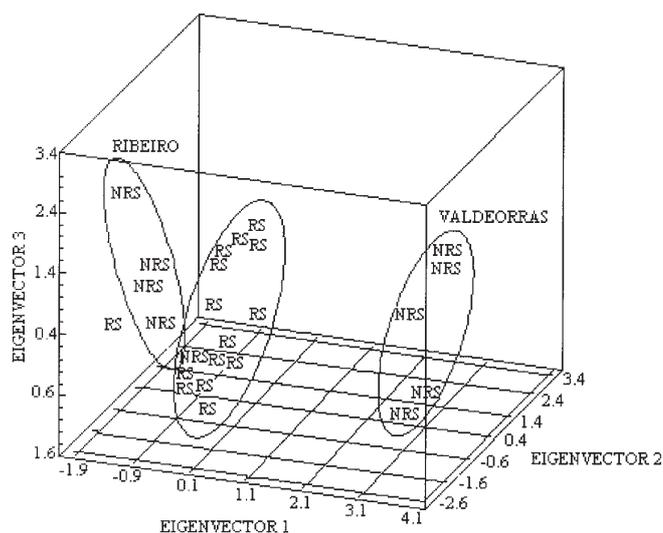
Figure 3. Eigenvector projection of wine samples. RS: Ribeira Sacra origin; N-RS: Non-Ribeira Sacra origin.

Original articles

Table III. Recognition and predicton abilities for LDA, KNN and SIMCA.

Category	% recognition ability	% prediction ability
LDA		
Ribeira Sacra	93	92
Non-*Ribeira Sacra*	78	75
KNN (K=3)		
Ribeira Sacra	94	92
Non-*Ribeira Sacra*	75	75
SIMCA		
Ribeira Sacra	79	75
Non-*Ribeira Sacra*	90	87

In step three, in order to check the reliability of the classification between the RS and N-RS wines obtained in the previous step, the wine samples were divided into a training (or learning) set consisting of 75 % of them and a test (or evaluation) set consisting of the remaining 25 %. Such division allows for a sufficient number of samples in the training set as well as a representative number of wines in the evaluation set. In order to reach a satisfactory test of the recognition and prediction ability of each method and to see how well the samples could be classified quantitatively, the previous procedure of division into training and evaluation sets was repeated four times to obtain four different files with diverse constitutions of the two sets. The three classification procedures were applied to the four mentioned files. The results obtained in terms of percentage of recognition and prediction abilities are presented in table III. All the classification methods achieved a level of correct classification and prediction comparable to those obtained in step 2, confirming the previously indicated differences in sensititivity: LDA and KNN for RS-class, and SIMCA for N-RS class. The adequate agreement between recognition and prediction abilities means that the decision rules do not depend too much on the actual objects in the training set: the solution obtained is stable [20]. Due to the high level of correct recognition and prediction for the wines from Non-Ribeira Sacra origin, it was demonstrated that SIMCA produced a model that can be used to detect RS falsifications, the probability of a N-RS wine being considered genuine RS is small. The combination of the three chemometrical procedures makes for a method to differentiate Ribeira Sacra CBO from the wines of two other CBO from Galicia, Ribeiro and Valdeorras, and to detect falsifications of genuine RS brand wines.

Conclusions

Pattern recognition is able to extract useful information for an amount of data. Information about the content of selected metals in wines from Galicia was used to differentiate between wine samples of different origin. Levels of Ca, Na and Mg in wines can be influenced by regional variations of soil and type and quantity of additives used in the wine elaboration. On the other hand, the content of K is determined by the variety of grapes, soil and climatic conditions, time of harvest, and other variables such as the temperature of fermentation and storage, the pH, the alcoholic degree and the use of ion-exchange resins. The level of Li is related to the soil. The content of this element is very important in regions that have a granite substrate. The level of Mn is related to the soil, grape variety and climatology of an area. PCA and cluster analysis reveals the occurrence of groupings between the analysed samples according to their brand.

The content of selected elements in wine can be used to obtain a successful regional classification. LDA, KNN and SIMCA achieved adequate levels of correct classification and prediction, based on stable decision rules, to confirm the authenticity of Ribeira Sacra CBO wines and to distinguish them from similar wines produced in other Galician CBO's such as Ribeiro and Valdeorras. Thus, the potential of pattern recognition techniques in the classification of Galician *Ribeira Sacra* red wines has been demonstrated. However, these results are based on a limited sample set; therefore, more developments with a higher training set that include other possible sources of variation will be necessary in order to obtain an optimised authenticity test that is capable of providing conclusive evidence from a regulatory point of view.

Acknowledgements

References

1. Brown, S. D.; Bear, R. S.; Blank, T. B. *Anal. Chem.* **1992**, *64*, 22-29.
2. Forina, M.; Lanteri, S. In *Chemometrics, Mathematics and Statistics in Chemistry*; Kowalski, B. R. Ed.; Dordrecht, Holland: Riedel Publishing, 1984. pp 305-351.
3. Kwan, W. O.; Kowalski, B. R. *J. Food Sci.* **1978**, *43*, 1320-1323.
4. Kwan, W. O.; Kowalski, B. R. *J. Agric. Food Chem.* **1980**, *28*, 356-359.
5. Kwan, W. O.; Kowalski, B. R.; Schogerboe, R. K. *J. Agric. Food Chem.* **1979**, *27*, 1321-1326.
6. Herrero, C.; Latorre, M. J.; García, C.; Mèdina, B. *J. Agric. Food Chem.* **1994**, *42*, 1451-1455.
7. Latorre, M. J.; Herrero, C.; Mèdina, B. *J. Int. Vigne Vin.* **1992**, *3*, 185-193.
8. Peña, R.; Latorre, M.J.; García, S.; Botana, A; Herrero, C. *J. Sci. Food Agri*c. **1999**, *79*, 2052-2056.
9. Moret, I.; Scarponi, G.; Cescon, P. *J. Agric. Food Chem.* **1994**, *42*, 1143-1153.

10. Vasconcelos, P.; Chaves, H. *J. Agric. Food Chem.* **1989**, *37*, 931-937.

11. Étiévant, P.; Schlich, P.; Symonds, P.; Bouvier, J. C.; Bertrand, A. *J. Sci. Food Agric.* **1988**, *45*, 25-41.

12. Maarse, H.; Slump, P.; Tas, A. C.; Schaefer, J. *Z. Lebensm. Unters Forsch.* **1987**, *184*, 192-203.

13. Cruz, M.; Sáez, J. A.; López-Palacios, J. *Analyst* **1993**, *118*, 801-805.

14. Lindberg, I. *Food Chem.* **1996**, *55*, 73-94.

15. Frazier, R.A.; Ames, J.M.; Nursten, H.E. *Electrophoresis* **1999**, *20*, 3156-3180.

16. Núñez, M.; Peña, R. M.; Herrero, C.; García-Martín, M.S. *J. AOAC. Int.* **2000**, *83(1)*, 183-188.

17. Massart, L.; Kaufman, L. *The interpretation of analytical chemical data by the use of cluster analysis*; New York: John Wiley & Sons, 1983.

18. Joliffe, I.T. *Principal component analysis*; New York: Springer, 1986.

19. Wold, S.; Albano, C.; Dun, W.; Edlund, U.; Esbensen, K.; Geladi, P.; Hellberg, S.; Johanson, E.; Lindberg, W.; Sjöstrom, M. In *Mathematics and statistics in chemistry*; Kowalski, B.R. Ed. Dordrecht, Holland: Riedel Publishing, 1984. pp 17-95.

20. Vandeginste, B.G.; Massart, L.; Buydens, L.M.; De Jong, S.; Lewi, P.J.; Smeyers-Verbeke, J. In *Handbook of chemometrics and qualimetrics: Part B*. Amsterdam, Holland: Elsevier, 1998, Chapter 33.

21. Wold, S. *Pattern Recog.* **1976**, *8,* 127-139.

22. Statgraphics, User's Guide, version 5 STSC, 1991, Rockville, MD.

23. Forina, M.; Leardi, R.; Armanino, C.; Lanteri, S. *Parvus: An extendable package of programs for data exploration, classification and correlation*; Amsterdam, Holland: Elsevier, 1988.