ELSEVIER

Spectrochimica Acta Part B 60 (2005) 717–724

SPECTROCHIMICA
ACTA
PART B

www.elsevier.com/locate/sab

Pattern recognition applied to mineral characterization of Brazilian coffees and sugar-cane spirits☆

Andréa P. Fernandes[a], Mirian C. Santos[a], Sherlan G. Lemos[a,b], Márcia M.C. Ferreira[c], Ana Rita A. Nogueira[b], Joaquim A. Nóbrega[a,*]

[a] *Grupo de Análise Instrumental Aplicada, Departamento de Química, Universidade Federal de São Carlos, Caixa Postal 676, 13560-970, São Carlos, SP, Brazil*
[b] *Embrapa Pecuária Sudeste, São Carlos, SP, Brazil*
[c] *Instituto de Química, Universidade Estadual de Campinas, Campinas, SP, Brazil*

Received 8 September 2004; accepted 13 January 2005
Available online 22 April 2005

Abstract

Aluminium, Ca, Cu, Fe, K, Mg, Mn, Na, Pb, S, Se, Si, Sn, Sr, and Zn were determined in coffee and sugar-cane spirit (cachaça) samples by axial viewing inductively coupled plasma optical emission spectrometry (ICP OES). Pattern recognition techniques such as principal component analysis and cluster analysis were applied to data sets in order to characterize samples with relation to their geographical origin and production mode (industrial or homemade and organically or conventionally produced). Attempts to correlate metal ion content with the geographical origin of coffee and the production mode (organic or conventional) of cachaça were not successful. Some differentiation was suggested for the geographical origin of cachaça of three regions (Northeast, Central, and South), and for coffee samples, related to the production mode. Clear separations were only obtained for differentiation between industrial and homemade cachaças, and between instant soluble and roasted coffees.

Keywords: Coffee; Sugar-cane spirit; Chemometrics; Pattern recognition

1. Introduction

Globalization has caused a revolution in the consumption habits all over the world. As well as in the internal and external markets, the improvement of the products quality, the decisive aspect of commercial barriers, and expansion of markets accentuate the pressures. The quality certification, based on standards, patterns, and technical specifications, will be the pre-requirement of any product. The search for superior levels of quality, time, and competitiveness is a constant concern of economic agencies, and in the agribusiness sector, it could not be different. The national producer has to be attentive on not losing market to other countries.

Among the representative products of the Brazilian agribusiness, two are prominent in a differentiated way: coffee and sugar-cane spirit, denominated Brazilian cachaça. The former for still highlighting Brazil as a major world producer, consumer, and exporter, and the latter, for waking up recently the interest on the increase of its export front to the growing search by external markets.

Coffee is the world most popular beverage after water, with over 400 billion cups consumed annually [1]. It is one of the most important agricultural products in the international trade, putting into motion approximately US$35 billion per year and being supplanted only by petroleum [2]. In 2004, the coffee industries estimate total sales–internal market and exports–of approximately US$1.5

☆ This paper was presented at the 8th Rio Symposium on Atomic Spectrometry, held in Paraty, RJ, Brazil, 1–6 August 2004, and is published in the special issue of *Spectrochimica Acta Part B*, dedicated to that conference.

* Corresponding author. Tel.: +55 16 3351 8088; fax: +55 16 3351 8350.
E-mail address: djan@terra.com.br (J.A. Nóbrega).

0584-8547/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.sab.2005.02.013

billion [3]. In the case of cachaça, official and exclusive denomination of the spirit distilled of sugar cane produced in Brazil, it comes conquering new markets. With a production of 1.5 billion liters/year, it generates profits of US$700 million and an expected growth of 27% in the exports in this year, with the European countries being the largest importers [4].

Therefore, it is easy to understand the huge relevance of the availability of suitable analytical methods to characterize these products of great consumption for millions of people worldwide. Additionally, the determination of geographical origin of commodities and food products is becoming an increasingly active research area, focused on both geographical authenticity and adulteration of foods [5]. Chemical analyses in conjunction with pattern recognition techniques provide interesting tools for the study of the quality and origin of food products [6].

The objective of this study was to evaluate the feasibility of using a multielement analysis combined with pattern recognition tools, in order to contribute to the development of the Brazilian agribusiness, supplying important information regarding the identity and quality of the national product. Thus, 48 samples of coffee and 156 samples of cachaça from different geographical sources were selected and analyzed with relation to their mineral content. The determination of Al, Ca, Cu, Fe, K, Mg, Mn, Na, Pb, S, Se, Si, Sn, Sr, and Zn was carried out by axial viewing inductively coupled plasma optical emission spectrometry (ICP OES). Afterwards, pattern recognition techniques were applied to characterize samples with relation to the geographical origin and production mode (industrial or homemade and organically or conventionally produced).

2. Experimental

2.1. Instrumentation

An inductively coupled plasma optical emission spectrometer with axially viewed configuration (VISTA AX, Varian, Mulgrave, Australia) was used for Al, Ca, Cu, Fe, K, Mg, Mn, Na, Pb, S, Se, Si, Sn, Sr, and Zn determinations. This equipment involves a simultaneous charge coupled device (CCD) detector that allows readings from 167 to 785 nm. The pre-optical system was purged with Ar in order to enable readings below 190 nm. The polychromator was thermostatized at 34 °C and purged with argon. End-on-gas interface was used to minimize interferences caused by auto-absorption processes. A V-groove nebulizer and a Sturman-Masters spray chamber were employed for sample introduction. The operational parameters are described in Table 1.

A microwave oven, model Ethos 1600 (Milestone, Sorisole, Italy), equipped with 120 ml Teflon PFA vessels was used for digestion of coffee samples. A centrifuge, model HermLe Z 200 A (Labnet, Berlin, Germany), was used for coffee samples extraction.

Table 1
ICP OES operating conditions

RF generator	40 MHz
Power	1.3 kW
Plasma flow	15 l min^{-1}
Auxiliary flow	1.5 l min^{-1}
Nebulizer flow	0.7 l min^{-1}
Replicates	3
Injector tube diameter	2.4 mm
Spray Chamber	Sturman Masters
Nebulizer	V-groove
Emission lines (nm)	Al I 396.152, Ca II 396.847, Cu II 213.598, Fe II 238.204, K I 766.491, K 769.897, Mg II 280.270, Mn II 257.610, Na I 589.592, Pb II 220.353, S I 181.972, Se I 196.026, Si I 251.611, Sn II 189.927, Sr II 421.552, Zn II 202.548

2.2. Reagents and samples

All solutions were prepared using analytical-grade reagents and deionized water (Milli-Q water, 18 MΩ cm, Millipore, Bedford, MA, USA). All glassware and polypropylene flasks were washed with neutron soap, soaked in 10% v/v nitric acid (Merck, Darmstadt, Germany) and rinsed with deionized water prior to use. Reference solutions were prepared by diluting stock solutions containing 1000 μg ml^{-1} of each element (Tec-Lab, Hexis, São Paulo, SP, Brazil) with distilled-deionized water in 8% v/v ethanol.

The data set used in the present study comprised 156 Brazilian cachaça samples and 48 coffee samples of different Brazilian regions. Acquisition of samples was accomplished in Brazilian markets based on the following characteristics: homemade and industrial for cachaça samples, roasted and instant soluble for coffee samples, and organically or conventionally produced for both types of samples.

2.3. Sample preparation

Cachaça samples were diluted in water (1:5 v/v), acidified to 0.014 mol l^{-1} with HNO_3. Adopting this dilution avoided problems of plasma extinction.

For the extraction of coffee samples, a mass of approximately 1.0 g was mixed with 10 ml boiling water and, after 15 min, centrifuged at 4000 rpm for 6 min and the supernatant was collected, acidified to 0.014 mol l^{-1} with HNO_3.

For the digestion procedure, a mass of about 0.25 g of coffee samples was directly weighed in a PTFE digestion vessel, 1 ml of H_2O_2 conc. and 3 ml of HNO_3 conc. were added, and the vessels were placed on the turntable. The microwave oven was operated according to the parameters listed in Table 2. The digestates were diluted to 25 ml with

Table 2
Microwave-assisted acid digestion: experimental parameters

Step	Power (W)	Time (min)
1	250	2
2	0	2
3	350	5
4	500	5
5	700	5

deionized water and the analytes were determined by ICP OES.

2.4. Data analysis

Fifteen metals, Al, Ca, Cu, Fe, K, Mg, Mn, Na, Pb, S, Se, Si, Sn, Sr, and Zn, were selected as chemical descriptors. In the coffee extracts or digestates, Sn did not present any analytical signal and could not be determined. The same was observed for Mn in cachaça samples. Each coffee or cachaça sample was considered as an assembly of 14 variables (i.e., the contents of the metals), which constitute their chemical descriptors. For coffee samples, a data matrix of 14 columns and 96 rows was built, each duplicate analysis being considered as one sample. Another data matrix of 156 rows and 14 columns was built for cachaça samples. These matrixes were used in the chemometric calculations.

Data analysis was performed using the software Pirouette 2.7 (Infometrix, Seattle, WA). Pattern recognition methods were applied to the data sets according to the following discussion.

2.4.1. Data preprocessing

Proper data preparation is integral to a successful multivariate analysis. It is often necessary to adjust a data set before running a multivariate algorithm. When one variable's magnitude is much larger than others are, this variable alone may dominate subsequent computations based on variance or distance. In other words, the elements that are in higher concentration would be the most significant to the model. The data preprocessing used in this study was the autoscaling, in which each variable (each column of the data matrix) is mean centered and then divided by its standard deviation.

2.4.2. Pattern recognition tools

Pattern recognition tools applied to data sets included principal component analysis (PCA) as a visualization method, and hierarchical cluster analysis (HCA) as an unsupervised learning method [7]. In PCA the data matrix is decomposed into scores and loadings matrices. The scores vectors describe the relationship between the samples in the model subspace and the loadings vectors describe the importance of each descriptor within the model. It can represent graphically intersample and intervariable relationships and provides a way to reduce the dimensionality of the data. Similarly to PCA, clustering of samples reveals similarities among the samples while clustering of variables pinpoints intervariable relationships [8]. The primary purpose of clustering techniques is to present the data in an approach that demonstrates the grouping in a multidimensional space in such a way that all objects in a single group have some natural relation to one another, and the objects from different groups are somewhat different from each other [9]. The HCA results presented in the form of a dendrogram facilitate the visual recognition of such groups [10].

The measurement of the similarity is based on the squared Euclidean distance. The clustering method used was the Ward's method, which considers in each step the heterogeneity or deviance (sum of the squares of the distance of an object from the baricenter of the cluster) of every possible cluster that can be created by linking two existing clusters [11]. PCA results were validated by using the cross-validation method. As the used methods are non-parametric, they do not make suppositions regarding the statistical distribution behind the data set, and then any evaluation of the normal distribution of the data (Gaussian) is necessary [12].

3. Results and discussion

3.1. Coffee analysis

Although the class of the samples was a priori known, a preliminary study based on an unsupervised pattern recognition method was applied to observe the structure of the data sets. Thus, initially an HCA was applied on the digested coffee samples data set to observe any natural grouping feature. The resulting dendrogram is shown in Fig. 1. Two separated clusters appear, a bigger one containing the roasted coffee samples and a smaller, at the bottom with dots, containing the six instant soluble coffee samples analyzed.

In order to assess this statement and visualize the data structure, a PCA analysis was carried out and the results depicted in Fig. 2. PCA demonstrated that a small number of variables dominate the total data variability, as the three first principal components (PCs) accounted for 76% of the total variability. The first component is responsible for 37%, the second accounts for 30%, and the third explains 9% of the total information. A distinct visual clustering appears when the data were displayed with respect to the first two principal components, which was not surprising since the first principal component accounts for the maximum possible one-dimensional projection of the total variation of the individual data points.

Instant soluble coffees are located at positive scores (1) of PC2, well separated from roasted coffee samples with null and negative PC2 scores (2). Descriptors such as K, Mg, Na, and S were the features with highest positive

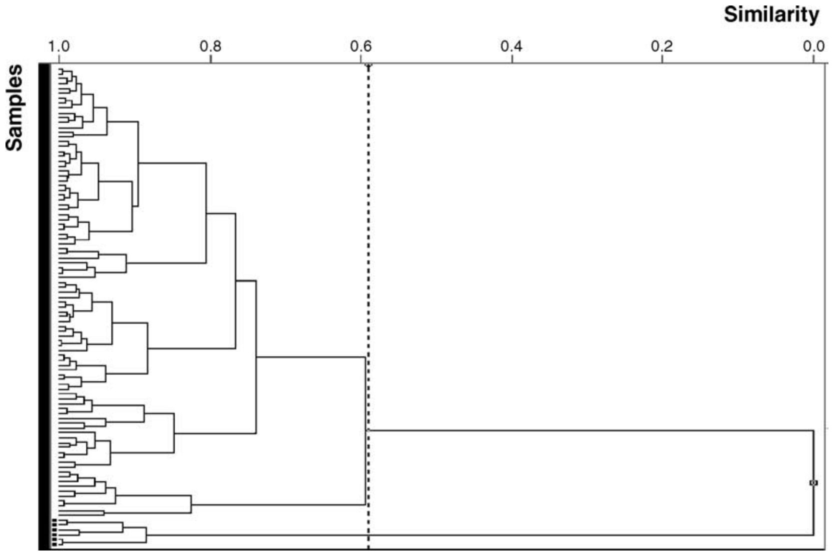

Fig. 1. HCA analysis applied to the differentiation between roasted and instant soluble coffee. Dots at the bottom show the instant soluble coffee group. Data obtained from the digested coffee samples.

loadings in PC2, while Cu and Zn presented the highest negative loadings in this PC. Thus, instant soluble coffees are generally well characterized by relatively high values of K, Mg, Na, and S and low values of Cu and Zn. The reciprocal is true for the roasted coffee group at the bottom of the plot. Both types of coffee present mean contents of the elements with null loadings in PC2, i.e., possess similar amounts of the elements Al, Ca, Fe, Mn, Si, and Sr. According to their proximity K, Mg, Na, and S gave the same kind of information. The same can be mentioned about the descriptors Cu and Zn. The three most discriminating variables were K, Mg, and Zn. The results indicated that by using only the descriptors K, Mg, and Zn, the coffee samples could be classified in their correspondent industrialization type, roasted, or instant soluble coffee.

In PC1, one can see a third group (3) formed by two roasted coffee samples (samples 9 and 10), which were considered as outliers. These samples were identified later as low quality products, which contained a great amount of peels, corn, and other strange materials. These samples exhibited high concentrations of Al, Fe, and Si, and therefore, have the highest positive scores in PC1.

A new classification of the coffee samples was assessed by PCA. Now, the objective was the evaluation of the mineral profiles of products originating from conventional and organic production modes. This production procedure is now in evidence for being based on ecological sustainable practices and has been attracting a great consumer market in spite of a superior mean price compared to that for conventionally produced coffees.

Fig. 3 shows the scores and loadings scatter plots of PC1 versus PC3, which provided the best visualization of the separation between the two groups of samples. The three first principal components (PCs) accounted for 70% of the total data variability. Two groups may be suggested based on the distribution of samples along PC1, which explains

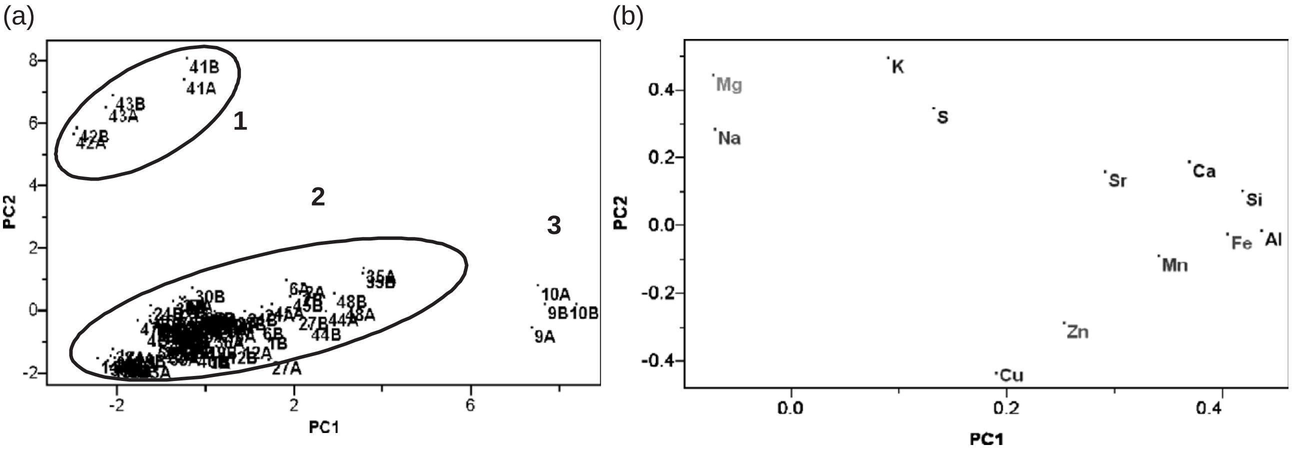

Fig. 2. PC1 versus PC2 (a) scores and (b) loadings plots of coffee samples showing differentiation between roasted and instant soluble coffee. Group legend: (1) instant soluble, (2) roasted and (3) outliers.

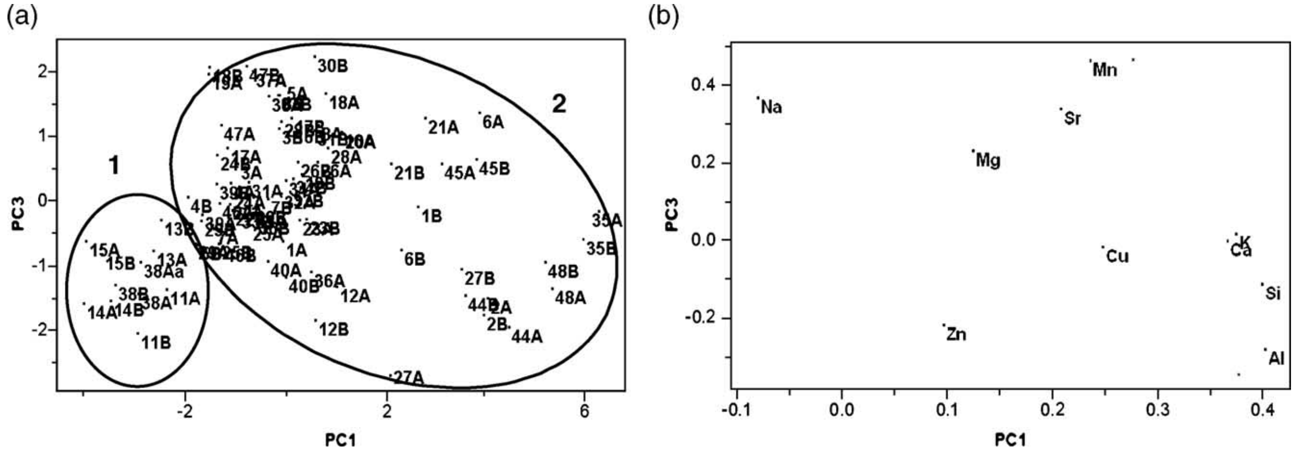

Fig. 3. PC1 versus PC3 (a) scores and (b) loadings plots of coffee samples showing differentiation between production modes (organically or conventionally produced). Group legend: (1) organically produced, (2) conventionally produced.

most of the variability, around 45%. The second component accounts for 15% and the third explains 10% of the total information. One group, denoted as (1), with negative PC1 scores, composed by the organically produced coffees, and another group (2) with scores located at the positive side of PC1, characterized by the conventional coffees.

The PC1 loadings profile shows negative values for the descriptor Na, indicating that higher amounts of this element may characterize organically produced samples. In PC3, a clear separation does not exist among the groups. However, it can be pointed out that the coffees produced organically can be characterized due to the higher amounts of Al and Zn, indicated by their negative values of scores. However, only by using the descriptor Na the coffee samples can be classified in their correspondent production mode, organically or conventionally.

Attempts to correlate the metal content of coffee samples with their geographical origin have been unsuccessful. This may be related to the fact that industrialized Brazilian coffees are blends of coffees from different regions, due to seasonal effects, which changes the supplies around the country during the year.

It should be pointed out that the same chemometric results were obtained for the extracted coffee data set. It means that further studies with a large number of samples can be carried out using extracts and, consequently, microwave-assisted digestions are unnecessary. This is advantageous to decrease costs and speed-up frequency of analysis.

3.2. *Cachaça analysis*

Similarly to the approach adopted for coffee samples, a preliminary study based on an unsupervised pattern recognition method was applied to observe the structure of the data matrix. Thus, a cluster analysis on the cachaça samples data set was carried out in order to observe any natural grouping feature. Fig. 4 shows the resulting dendrogram, in which two distinct clusters appears, a smaller one at the top (with dots) containing the industrially produced samples and a bigger one, containing the homemade cachaças. It can be seen that there are some misclassified samples indicated by the dots in the center of the biggest cluster, pointing out that the classification between industrially produced and homemade cachaças is not straightforward.

PCA was carried out in order to assess this statement and visualize the data structure and the results are depicted in Fig. 5. Three first principal components accounted for 73% of the total variability. The first component accounts for 40%, the second for 20%, and the third explains 13% of the total information. In addition, a distinct visual clustering appears when the data were displayed with respect to the second and third PCs. Industrially produced cachaças are located at negative PC2 scores (1), well separated from homemade samples, which are located at null and positive PC2 scores (2). Descriptors such as K, Mg, Ca, Cu, and Pb were the features with highest positive PC2 loadings, while Na and Si presented the highest negative loadings in this PC. Thus, homemade cachaças are generally well characterized by relatively high values of K, Mg, Ca, Cu, and Pb, and low values of Na and Si. The reciprocal is true for the industrially produced group. Both types of samples presented mean contents of the other elements with loadings close to zero in PC2.

The PCA indicated a differentiation based on Cu and Pb, because the remaining descriptors are constituted of common macroelements. The great relative amount of copper in the homemade cachaças is caused by the use of copper stills in their production process. According to the Brazilian legislation, the maximum Cu content allowed in cachaça is 10.0 mg l^{-1} [13]. The careless cleaning of the copper apparatus after the distillation process is probably the main cause of high Cu concentration in homemade products [14]. The Pb content may also be related to the lack of proper process control in the homemade production. It should also be mentioned that Pb is a common contaminant found in the production of Cu metal. These cachaças containing measurable amounts of Pb are not mass-

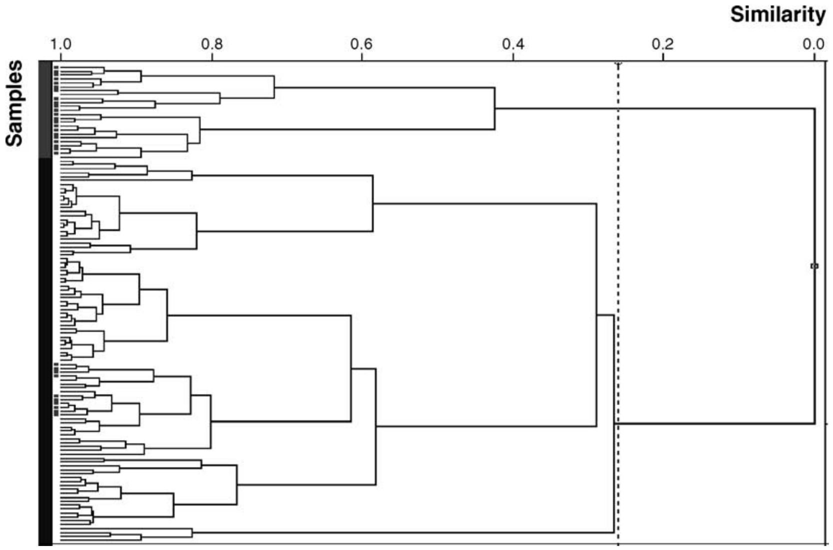

Fig. 4. HCA analysis applied to the differentiation between homemade and industrial cachaça. Dot points at the top and center show the industrial cachaça group.

produced and are mainly consumed near to their production area.

On the other hand, in the industrial production process, the stills are made of stainless steel and the whole process is accomplished at a constant flow-rate, generating product standardization.

Additionally, homemade cachaça production is not based on a rigorous quality control in most of the stills spread over the country. The use of Cu stills condensers is recommended because this metal acts as a catalyst, favoring the formation of volatile components in the final product (aromas and bouquets), and improving the quality of cachaças. This artifice is also used in many column distilleries in Brazil.

As indicated by HCA analysis, some misclassified industrially produced samples were found inside the homemade group, located next null values on PC2. This aspect shows that the great majority of samples possess relative homogeneity related to the mineral chemical composition, showing an evolution towards national product standardization.

Further, a new classification of samples was evaluated based on PCA analysis. At this time, the objective was to correlate the mineral profile with the geographical origin through chemometrics. Cachaça samples were collected from 15 different states of the Brazilian federation. Preliminary analyses accomplished in the attempt of classifying the samples according to their origin (15 classes) were not successful. This can be related to the chemical homogeneity of the samples, as already pointed out. However, when the number of classes was reduced to three according to the main geographical regions, a relative separation among the groups was obtained, as it can be seen in Fig. 6. Class number 1 is constituted by samples from the following states: Maranhão, Piauí, Ceará, Alagoas, Pernambuco, Paraíba, and Bahia, denominated as the Northeast region. Class number 2 is constituted by samples from the

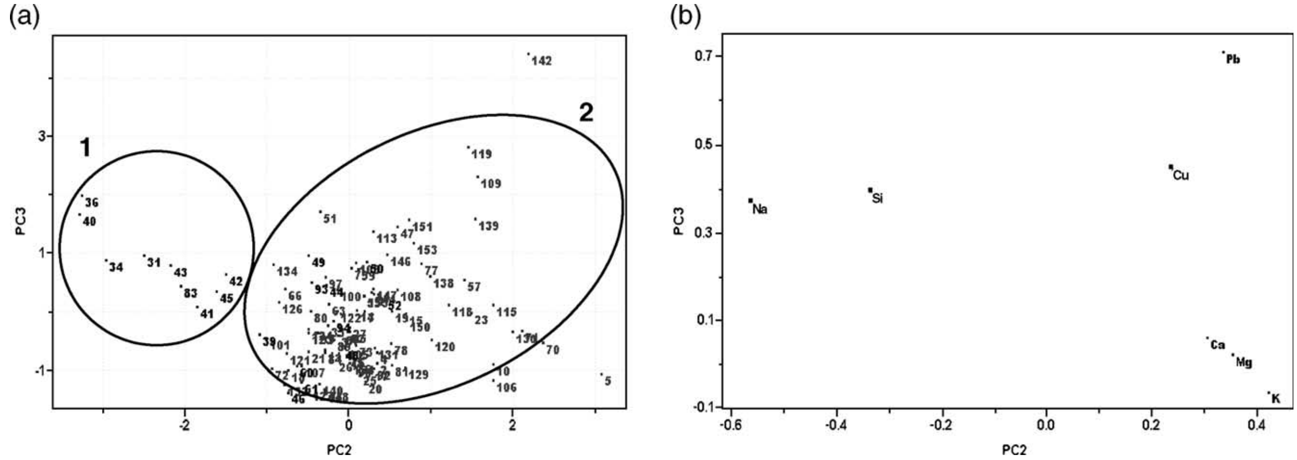

Fig. 5. PC2 versus PC3 (a) scores and (b) loadings plots of cachaça samples showing differentiation between production types (homemade or industrial). Group legend: (1) industrial, (2) homemade.

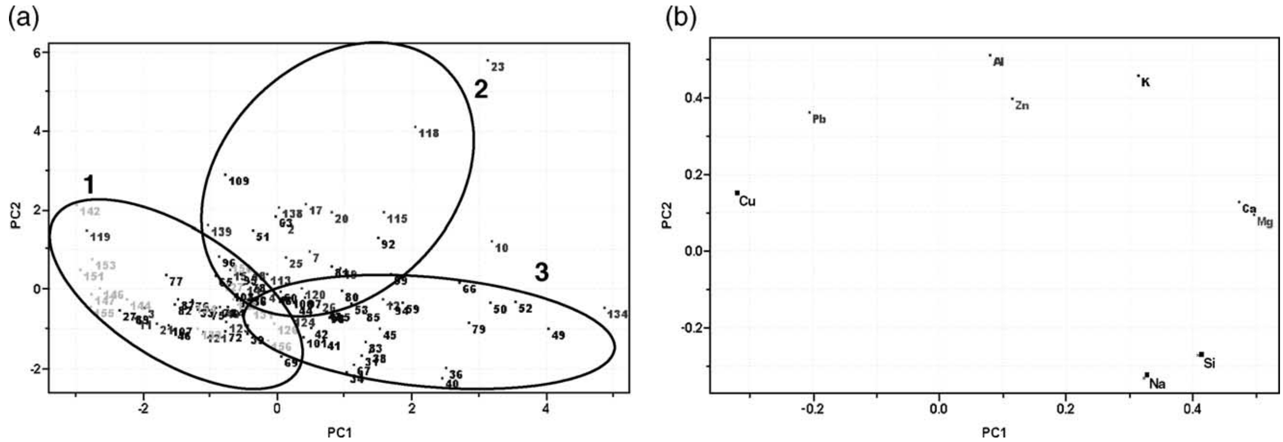

Fig. 6. PC1 versus PC2 (a) scores and (b) loadings plots of cachaça samples showing differentiation by sample origin (Northeast, Central and South). Group legend: (1) Northeast, (2) Central, and (3) South.

following states: Goiás, Minas Gerais, Espírito Santo, and Rio de Janeiro, denominated as the Central region. Finally, class number 3 is constituted by samples from the following states: São Paulo, Paraná, Santa Catarina, and Rio Grande do Sul, denominated as the South region.

In this case, the three first principal components accounted only for 56% of the total variability, showing a modest visual clustering of data. The first component accounts only for 29%, being the second responsible for 17%, and the third explains 10% of the total data information. Three groups may be suggested based on the distribution of samples along PC1. The first group with negative PC1 scores (1), composed by the cachaças from Northeast region; another group with scores located near null values of PC1 (2), characterized by the Central region cachaças, and a third one located at the positive side of PC1, characterized by the cachaças produced in the South region. The second group was better characterized on PC2, due to the major part of the samples that possesses positive scores. PC1 loadings profile shows negative values for the descriptors Cu and Pb and highest positive loadings for the descriptors Na, Si, Ca, Mg, and K. Consequently, Northeast samples may be characterized by relative high values of Cu and Pb, and low values of Na, Si, Ca, Mg, and K. The reciprocal is true for the South group. Central group presents mean contents of these elements and relative higher amounts of Al and Zn with loadings next null in PC1. PC2 scores profile shows the apparent separation of Central group from the others and loadings in that same PC characterize this group as possessing relatively higher amounts of Al, K, and Zn.

The obtained classifications are related to the soil type in which the sugar cane was cultivated, being the Northeast the most separated group, which also has the most different soil compared to those from other Brazilian regions.

A cluster analysis was accomplished to verify if the PCA results were plausible. Fig. 7 illustrates the dendrogram obtained for the geographical origin study, which shows the presence of two groups with no clear separation (similarity indices around 0.2 or 0.3) demonstrating the homogeneity

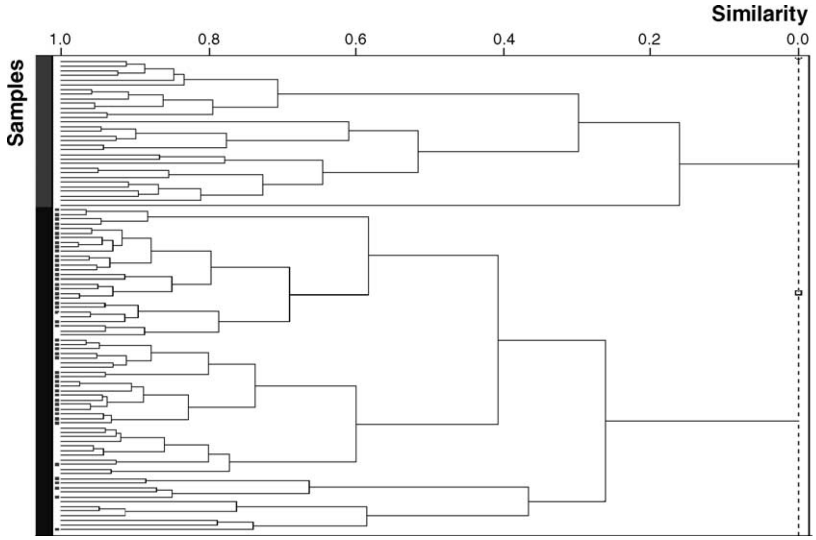

Fig. 7. HCA analysis applied to the differentiation by sample origin (Northeast, Central, and South). Dot points at the center show the northeast group.

presented in PCA analysis. The dots point to the Northeast group samples showing the presence of other misclassified samples in this group as illustrated by PCA analysis.

It should also be pointed out that attempts to correlate metal content of cachaças with their production mode (organically or conventionally produced) were not successful.

PCA and HCA were used here mainly to explore the data and visualize the trends in samples based upon the selected variables. They were not used to make models and predictions (unsupervised pattern recognition methods). After showing that the selected variables are able to fairly discriminate between cachaças according to their geographical origin, the supervised pattern recognition method KNN (Kth Nearest Neighbors) was used to build a classification model [12]. Classification with KNN is based on distance (multivariate Euclidean distance) comparison among samples. The distance between every pair of samples is calculated. The predicted class of a test compound is determined based on the distance of this sample with respect to the closest K samples in the set. Each of the K nearest samples is chosen to "vote" once for its class. The class receiving the highest number of votes is assigned to that sample. Using four nearest neighbors, the predicted classes were compared with the true class memberships giving a success rate of 84.6% (i.e., 84.6% of the cachaça samples were found to be classified correctly).

4. Conclusions

According to the results, pattern recognition analyses were able to differentiate some important features in coffee and cachaça samples based on their metal contents. For coffee samples it was possible the differentiation between instant soluble and roasted coffee, the latter being characterized by greater amounts of Cu and Zn, while the instant soluble presented greater contents of Na, Mg, K, and S. It was also possible for the classification of samples according to the production mode, being Na the most significant chemical descriptor for organically produced samples, which also presented greater amounts of Al and Zn. Separation of samples by their geographical origin was not possible due to the fact that coffees are industrialized as mixtures of different regions (blends), as a result of the variable raw material offer around the country related to the seasonal climatic effects. For cachaça samples, homogeneity among samples was observed, indicating a road towards standardization of the final product used either for internal consumption or for exportation. In spite of that, a clear differentiation was obtained between the industrially produced and homemade samples, with the elements Ca, Mg, K, Cu, and Pb the ones that best described the homemade samples, while Na and Si were the main factors for characterization of the industrially produced samples. The presence of Cu and Pb in measurable amounts indicates the need for a better quality control for some homemade cachaças. A trend was observed towards the classification according to the origin. Three groups, denominated as Northeast, Central, and South, were identified. The first group was the most differentiated due to the higher contents of Cu and Pb and, lower ones of Na, Si, Ca, and Mg. The opposite was observed for samples produced in the South, while samples of the Central region presented mean amounts of those elements and greater concentrations of Al, K, and Zn. The correlation between metal contents and production mode of cachaça (organic or conventional) was not corroborated by experimental data.

Acknowledgments

The authors wish to express their appreciation to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for financial support and for the research scholarships provided.

References

[1] J.M. Esteban-Díez, C. González-Sáiz, Pizarro, An evaluation of orthogonal signal correction methods for the characterization of arabica and robusta coffee varieties by NIRS, Anal. Chim. Acta 514 (2004) 57–67.

[2] E.J. Santos, E. Oliveira, Determination of mineral nutrients and toxic elements in Brazilian soluble coffee by ICP AES, J. Food Compos. Anal. 14 (2001) 523–531.

[3] ABIC. Available on the Internet on www.abic.com.br, August 2004.

[4] SEBRAE. Available on the Internet on www.sebrae.com.br, August 2004.

[5] K.A. Anderson, B.W. Smith, Chemical profiling to differentiate geographic growing origins of coffee, J. Agric. Food Chem. 50 (2002) 2068–2075.

[6] J.R. Piggot, Statistical Procedures in Food Research, Elsevier, London, 1986.

[7] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics, Data Handling in Science and Technology, vol. 20A e B, Elsevier, Amsterdam, 1997.

[8] A. Smoliński, B. Walczak, J. Einax, Hierarchical clustering extended with visual complements for interpretation of environmental data set, Chemometr. Intell. Lab. Syst. 64 (2002) 45–54.

[9] M. Daszykowski, B. Walczak, D.L. Massart, Looking for natural patterns in data. Part 1: density based approach, Chemometr. Intell. Lab. Syst. 56 (2001) 83–92.

[10] Pirouette User Guide, Infometrix.

[11] M. Meloun, J. Militký, M. Forina, Chemometrics for Analytical Chemistry, Ellis Horwood, New York, 1992, pp. 244–269.

[12] M.A. Sharaf, D.L. Illman, B.R. Kowaslki, Chemometrics, Wiley, New York, 1986.

[13] Brazilian Ministry of Health, Complementation on Identification and Quality Standards for Alcoholic Distillates; Portaria No. 685, Brasília 1998 (in Portuguese).

[14] R.F. Nascimento, C.W.B. Bezerra, S.M.B. Furuya, M.S. Schultz, L.R. Polastro, B.S. Lima-Neto, D.W. Franco, Mineral profile of Brazilian cachaças and other international spirits, J. Food Compos. Anal. 12 (1999) 17–25.