

Characterization of Cherry Juices by Pattern Recognition Analysis

BELGIN IZGI

Department of Chemistry, Science and Art Faculty

Uludag University, 16059 Bursa, Turkey

Fax: (90)(224)2941899; Tel: (90)(224)2941728

E-mail: belgin@uludag.edu.tr

Provenance analysis has become important role for controlling process of public health and environment. Multivariate data analysis is used most commonly to give information about similarities, relationships, origin of the species from different sources. In this study, eight different cherry juices available in the market in Turkey were analyzed for their metal content, total sugar, titratable acidity and pH. The results were examined by pattern recognition (PARC) techniques. It was showed that samples could be clustered into three groups including sub-groups. By this way, some of the characteristics of the cherries like geographical origin, the production processes and environmental parameters were explained.

Key Words: Pattern recognition techniques, Characterization, FAAS, Cherry juice.

INTRODUCTION

Nowadays, characterization or provenance analysis has become important role for public health and environmental controlling process. Due to the population explosion, technological developments and climate changes in the world, first applications of characterization are started on the sources of the water reservoirs and water quality¹. For example, information on the abundance of trace and rare-earth elements is very useful for petrography investigations of igneous rock systems. To obtain the precise and accurate analytical data is extremely important for rock classification and in understanding the origin of the rock².

Fruit juice is very important in food pyramid to protect the vitamin and mineral balance of metabolism. Technological improvements make our lives multipart and time is restricted so that eating habits is changing day by day. To receive mineral and nutrients, it is easy and practical to drink fruit juice. However, consuming great volume of fruit juice, as reported in the statistics by American Pediatrics Department, could cause some troubles such as stomach burn, colour and surface changes of teeth, allergy, diabetes, obesity *etc.*³. To overcome this kind of problems, it can be offered to intake

less amount of juice and pasteurization of juices. However, only pasteurization is not enough to deal with. Not only environmental pollution but also the procedures of production, which is starting from the farm to the packed product at the market and processes conditions have effects on the quality of fruits and their products. The story of fruit juice starting from the farm and all the steps of production processes are very important to check the every step to understand the quality. Quality of juices is controlled by Food Codex limits⁴. It is essential to apply standard methods for the analysis of the fruit juice and its additives such as dyes, acidity, metal levels, *etc.* However, it is difficult to calculate and identify of the differences due to the great amount of the data and it is time consuming. Therefore, chemometric methods can be used to cope with these problems. Chemometry can be summarized as a solution of chemical data *via* computer, statistics and some mathematical methods. Matlab, Visual Basic and C++ computer programs can be generally used for this purpose⁵⁻⁸. It is well known that the pattern recognition (PARC) techniques as principal component analysis (PCA) and cluster analysis (CA) have been used in order to classify and identify different classes of flavours in samples such as fruit juices, olive oils, seed oils, tomato sauces *etc.* Good separation among flavours in samples belonging to homogeneous and comparable classes has been obtained by PARC methods⁸.

In this study, eight different cherry juices produced in Marmara, Anatolia and Mediterranean region of Turkey were purchased from local markets. The metals such as Cu, Mn, Zn, Cd, Pb, Na, K, Ca, Mg and Fe, organic compounds that taste to juice like total sugar, titratable acidity and pH level were determined by FAAS, classical titration methods, pH-meter, respectively. The results were examined by PCA, CA. PCA shows the similarities and relationships between the cherry juices and it explains which parameter is the most effective for the characterization. The data sets prepared by PCA and CA were compared to gather information about the similarities of cherry juices, geographical origin and environmental relations.

EXPERIMENTAL

Jenway 3010 model pH-meter calibrated using 4.02 and 7.01 pH buffer solution was used for pH measurements. Metal levels were determined by ATI-UNICAM 929 model flame atomic absorption spectrometer (FAAS). Hollow cathode lamp suitable for each element with deuterium background correction and air/acetylene flame were used during the analysis. The optimum instrumental parameters are given in Table-1.

All the reagents used were of analytical grade. Double distilled water was used through all experiments. Standard solutions were prepared from 1000 mg L⁻¹ of stock standard solution (Merck) by appropriate dilution with 0.2 % HNO₃. Calibration curve was prepared by standard addition

TABLE-1
OPTIMUM INSTRUMENTAL CONDITIONS OF FAAS

	Na	K	Ca	Mg	Fe	Cu	Zn	Mn	Cd	Pb
Wavelength (nm)	589.0	766.5	422.7	285.3	248.3	324.8	213.9	279.5	228.8	217.0
Lamp current (mA)	10	10	10	5	10	4	8	8	5	6
Bandpass (nm)	0.2	0.5	0.5	0.5	0.5	0.5	0.5	0.2	0.5	0.5

method from pooled fruit juice prepared by taking standard volume from each juices sample. 0.1 N NaOH was prepared to measure the titratable acidity and standardized by 0.1 N HCl standardized *via* primer standard solution of 0.1 N Na₂CO₃ *via* methyl orange indicator. Thymol blue indicator was prepared at 0.1 % (w/v) in 20 % (v/v) ethyl alcohol solution. Fehling A solution was prepared by dissolving 69.28 g copper(II) sulphate pentahydrate in water together with 1.0 mL of 1 M H₂SO₄ and making up to 1000 mL with water. In addition, Fehling B solution was prepared by dissolving 346 g potassium sodium tartrate tetrahydrate and 100 g sodium hydroxide in water and making up to 1000 mL with water and then filtering through glass wool after standing. Mixed Fehling's solution was prepared by mixing equal volumes of Fehling A and B solutions very slowly in a dry bottle. After stopping the bottle to avoid carbon dioxide absorption, it was stored in the dark. Methylene blue indicator was prepared as 1 % (w/v) with water and filtered before use.

Cherry juice samples produced in different regions, Marmara, Middle Anatolian and Mediterranean part of Turkey were purchased from local markets as 250 mL packed product. All cherry juices were labeled as A, B, C, D, E, F, G and H to protect the firms from speculations. The five samples of last products from packed juices were analyzed for each eight samples.

Analytical procedures: Wet digestion technique was employed for sample preparation. For wet digestion of cherry juices, 40 mL of cherry juice after adding 5 mL of concentrated HNO₃ was boiled until the nitrogen vapour disappeared. After cooling, 5 mL of concentrated H₂O₂ added on to the solution, boiled again nearly for 10 min and made up to 50 mL with water. The last solution was used for all metal analysis. Content of elements in digested samples was calculated from standard addition calibration curve plotted for each element. Acidity of juices is significant that they can source of the quality chances⁹. Titratable acidity and total sugar level were analyzed using standard methods¹⁰. For titratable acidity (TA) analysis, 10 mL of previously filtered cherry juice diluted with water to 80 mL. After the addition of an optimum amount (14-15 drops) of thymol blue indicator,

final solution was titrated by standardized NaOH solution. For total sugar (TS) analysis, 20 mL of cherry juice was boiled with 2 mL 0.1 N HCl. Alkalinity of this solution was adjusted by 10 % NaOH *via* phenolphthalein and then diluted up to 50 mL with water (solution C). Besides, 10 mL of mixed Fehling's solution was boiled after adding 4 drops of methylene blue. This hot solution was titrated with solution C till the blue colour is discharged as described by James¹⁰.

Statistical analysis: Pattern recognition can be defined as the mapping from a given pattern space into a class membership space. Typically, analytes are classified according to a specific property by taking indirect measurements of that property. Chemometric methods are used to relate these indirect instrumental observations to chemical composition. This field includes linear regression, multivariate analysis and the use of artificial neural networks. Factors or principal components are weighed linear combinations of the original variables. Factor analysis methods can be used to determine the number of factors that needed to describe the statistically significant variance within the data set. Principal components can be exploiting for qualitative identification of samples, in pattern recognition or for quantification using principal components regression (PCR) or partial least squares (PLS) algorithms⁷. The principal component analysis decomposes of X matrix into scores, loadings and residual errors matrices. The plots of loadings allow the relationship of variables to be assessed. The plot of scores of PC1 versus PC2 is a general method of classifying samples from their measured properties. Cluster analysis (CA) was performed to classify the samples *via* similarities of their chemical properties.

In this study, analysis was planned using above eight samples with five replicate of each firm in three different regions of Turkey. Each of the samples was analyzed in triplicate, average value evaluated after standardization of each raw result and the data were analyzed using multivariate data analysis such as PCA and CA under Microsoft Excel program. The results were compared with commercially available programs Statistica 99 from Statsoft. The similarity between samples was calculated from Euclidian distance.

RESULTS AND DISCUSSION

Metals content of cherry juices: Amount of elements in juices were calculated by standard addition calibration curve. The equation of calibration curves of analysis with regression coefficient (R^2) could be given in the range of 0.9949-0.9998. Also the limit of detection (LOD, mg L^{-1}) of analysis could be given in as 0.484, 1.183, 0.757, 3.72, 0.189, 0.554, 0.046, 0.113, 0.236 for Na, K, Mg, Ca, Mn, Fe, Zn, Cu and Cd. The lead level of the samples was not used for calculations due to the lead level of all samples was below the LOD (0.1 mg L^{-1}) by quartz tube.

Statistical analysis: The results were firstly standardized by using the formula $SD = [(VS-MVVC)/STDVC]$ where SD is the standardized data, VS is the value of each sample, MVVC is the mean value of the each variable column and STDVC is the standard deviation of the variable column. The standardized results were examined using Excel by PCA and CA with variables such as pH, titratable acidity, total sugar and metal content. Eigen values of PCs were shown in Table-2. The scores and loadings plots on PC2 vs. PC1 are shown in Figs. 1 and 3. The samples studied are clustered into three different groups. The samples in cluster I are from Central Anatolian region, the samples in cluster II and III represent region of Mediterranean and Marmara together.

TABLE-2
EIGEN VALUES AND VARIANCES OF PRINCIPLE COMPONENTS

Number of PC	Eigen value	Variance (%)	Cumulative variance (%)
1	130.6645200	27.919770010	27.91977
2	101.4637100	21.680280790	49.60005
3	77.8671690	16.638283920	66.23833
4	75.3956600	16.110183790	82.34852
5	25.3716060	5.421283231	87.76980
6	20.5712530	4.395566884	92.16537
7	11.1441270	2.381223750	94.54659
8	9.0690568	1.937832658	96.48443
9	5.6531606	1.207940297	97.69237
10	5.2491723	1.121618009	98.81398
11	3.5004124	0.747951358	99.56193
12	2.0501456	0.438065301	100.00000

The similarities of the samples were investigated by complete linkage cluster analysis *via* Euclidian distance, as seen in Fig. 3. Pearson's correlation coefficient is widely used for statistics.

Principal component analysis (PCA): The data obtained by measuring the concentration of nine elements, total sugar, titratable acidity and pH in 40 cherry juice samples from three regions of Turkey. It can be get three PCs explained 66.24 % of the relevant information from the data out of PC1 and PC2. The average of eigen values (Table-2) is 39 when 12 PCs were taken out. Eigen values of three components are greater than the average. The scores and loadings plots on PC2 against PC1 are shown in Figs. 1 and 3. The samples under investigation are clustered into three different groups by selected chemical contents of cherry juices. In addition, characterization parameters of cherry juices will be getting clear by distribution of variables of these groups.

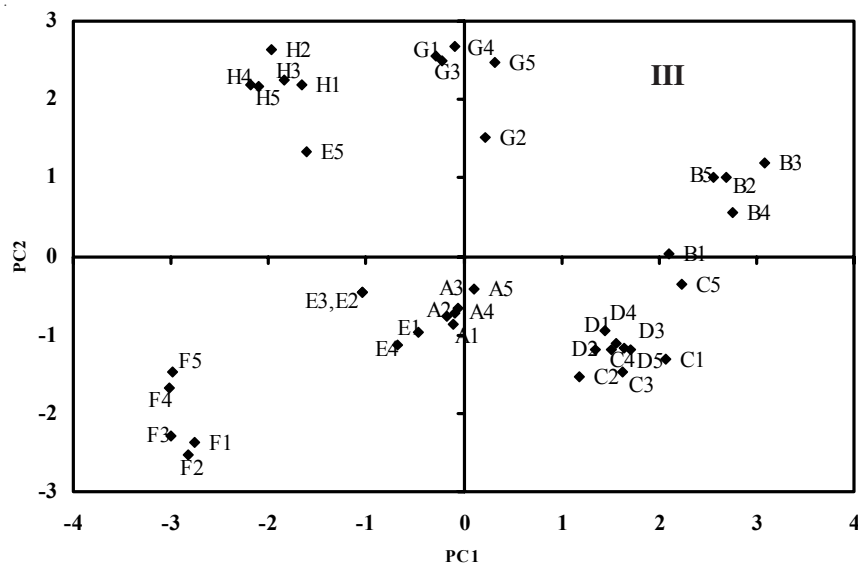


Fig. 1. Distribution of samples on scores plot in PCA

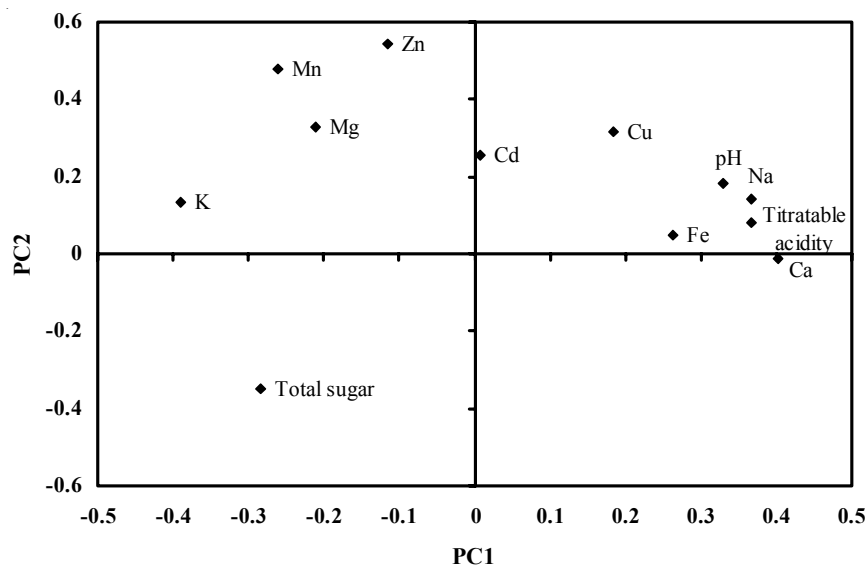


Fig. 2. Distribution of variables on loadings plot in PCA

In-group I, samples are characterized by Mn, Mg, Zn and Cd content. By this way, samples labeled H can be separated from samples labeled G by its Mg and Mn content. From these results, it is obvious that cherry samples in Central Anatolian are taking apart by their Mg, Mn, Zn and Cd content from Mediterranean and Marmara region cherries. The samples

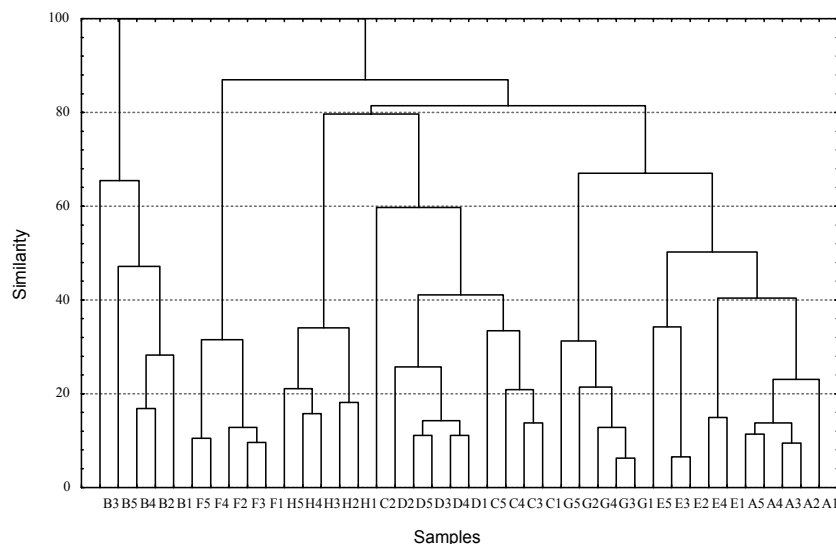


Fig. 3. Dendrogram resulting from complete linkage cluster analysis based on standardized data

in-group **I** are coming from mainly Central Anatolian region. Group **II** is characterized by PC1 and PC2 negative value corresponding for total sugar content. However, it is thought that sugar added in the process. In Food Codex; except grape and pear juice, sugar can be added maximum 15 g L^{-1} above dry mass to arrange the acidity and to improve the taste 150 g L^{-1} of fruit juice and fruit extracts. In addition, it is obviously clear that not allowed adding acid and sugar at the same time due to Food Codex¹¹. From this perspective, the results labeled in-group **II** overlapped with this rule. Group **III** is characterized by positive PC1 values of Cu, pH, Na, titratable acidity, Fe and Ca in the loadings plot. Especially, the samples labeled B can be separated from by Cu, Na from samples labeled C, D. Copper level gives some information about the origin and production quality of cherries. It can be explained that Cu originates from fertilization, some contaminants in the production process or from industrial activity, which is high in Marmara region and the samples; labeled B comes from this part. There are no big differences in samples C and D, so that from these results it can be concluded that mixture of cherries from Mediterranean and Marmara region was used in the production process. Samples in-group **II** and **III** originate from Mediterranean and Marmara region as a mixture. Both groups have only one sample from other region.

From loading plots of PC1 and PC2, variables Cd, Ca and Fe have very near to zero level so that they are not very effective in the characterization. Limit levels of some parameters of the samples were summarized in Table-3⁴.

Samples labeled H and G have founded 0.23-1.2 mg kg⁻¹ Mn level than the limits (in drinkable water 0.05 mg kg⁻¹). This can be explained by water and production quality of the firms. This may also give information about the water quality in Central Anatolian Region.

Microbial chances in acidic foods pH < 4.5 such as fruit juices and purees generally do not involve toxicological hazard but they can cause degradation in quality with notable commercial damages⁹. Therefore, titratable acidity is important for packed product to understand its taste quality; it gives information on sugar and acid addition of the process. As it was determined from Pearson's correlation coefficients given in Table-4, there is no big correlation with sugar and acidity.

From loadings plot of PC1 and PC2, it can be said that potassium is not one of the characterization factor for all of the cherry samples due to its location is far from all of the groups seen in Fig. 1. The results of samples labeled A, C (except C5) and D are located very near in each five samples (Fig. 1). This can give reasonable explanation about the control quality of the production process.

Correlation coefficients between samples and variables give also more information about characteristics of the samples. According to correlation coefficient of the samples (Table-3), each sample group has very good correlation within their group. Only coefficient values of the samples labeled E changed in wide range. Generally, it can be concluded that all of the cherry samples are well adjusted among the same group. Samples labeled F, G and H have not correlation with other samples (Fig. 1). According to correlation coefficient of analysis parameters (variables) in Table-4, high degree of correlation existing between Ca, Na, Cu, Mn, Fe, Zn, K and titratable acidity with 'r' coefficient greater than 0.5.

Cluster analysis (CA): The resulting dendrogram is given in Fig. 3. Mainly two clusters are appeared. This result is supported the observations in PCA analysis as mixture results of group **II** and **III**. It is easy to perform the results obtained from PCA and CA with significant PCs instead of original data in cluster analysis¹². A complete linkage cluster analysis was performed on the significant eight PCs scores. The dendrogram result is shown in Fig. 4. Samples labeled B and F are seen as a separate group. Others have some features that are very similar to each other. From these results, it can be extracted that PCA (Fig. 1) can be divided into two main group and two sub-group as well. Samples labeled B and F are harmonious with this information. High similarity was observed between Mediterranean samples [E3 and E2 (3.8 %), F2-F3 (4 %)] and Central Anatolian sample [G1-G3 (4.8 %)].

TABLE-4
PEARSON'S CORRELATIONS COEFFICIENTS
CALCULATED BETWEEN VARIABLES

	Na	K	Mg	Ca	Fe	Cu	Zn	Mn	Cd	T. Acidity	T. Sugar	pH
Na	1											
K	-0.35	1										
Mg	-0.38	0.06	1									
Ca	0.55	-0.30	-0.470	1								
Fe	0.08	-0.01	-0.260	0.41	1							
Cu	0.60	0.07	-0.200	0.23	0.10	1						
Zn	0.07	0.44	0.462	-0.04	0.12	0.30	1					
Mn	-0.19	0.65	0.429	-0.28	0.04	0.22	0.79	1				
Cd	0.43	-0.12	0.036	-0.07	-0.50	0.43	0.19	0.23	1			
T. Acidity	0.27	-0.23	-0.200	0.40	0.70	0.31	0.01	-0.10	-0.25	1		
T. Sugar	-0.24	0.43	-0.430	-0.33	-0.20	-0.24	-0.3	-0.10	-0.03	-0.35	1	
pH	0.25	-0.73	0.236	0.20	0.19	-0.03	0.06	-0.10	0.17	0.32	-0.64	1

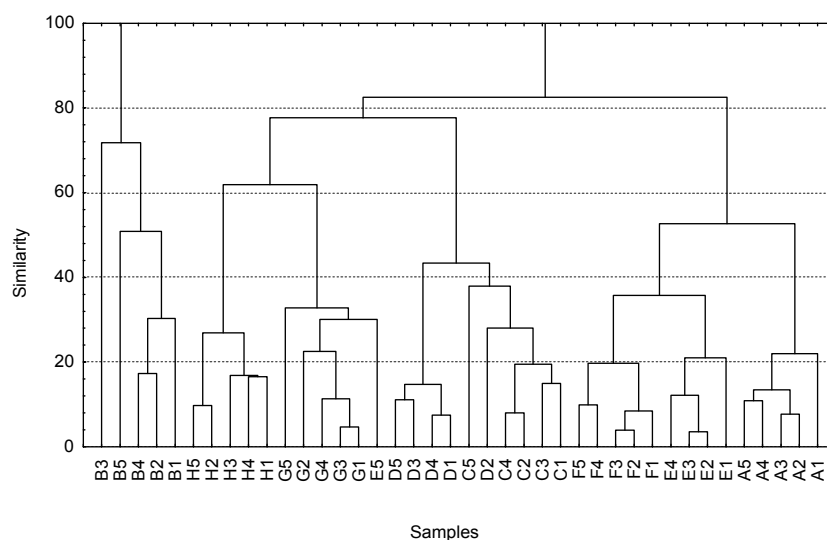


Fig. 4. Dendrogram resulting from complete linkage cluster analysis based on eight PCs scores

Conclusion

The pattern recognition techniques such as the principal component analysis (PCA) and cluster analysis (CA) have been applied successfully in selected cases of food. PCA and CA have demonstrated functional way

for characterization of eight cherry juice samples based on their chemical properties. On the other hand, they give more possibility to check the products in every step of production from the beginning of the farm to packed product of fruit. This study confirms a great potential of chemical content in connection with chemometric methods for food analysis. From the analysis, both PCA and CA can give the same reasonable results. It can be concluded that the pattern recognition technique gives information about origin, geographical distribution and production quality of the food samples. The samples can be grouped with different content of cherry juices distributed in different regions of Turkey.

ACKNOWLEDGEMENTS

The author would like to thank to graduate students Melike Kirmizi and Cüneyt Tan for their co-operations during the laboratory studies and Assoc. Prof. Dr. Cevdet Demir for his valuable scientific discussions.

REFERENCES

1. N. Khalil, M.A. Misdaq, S. Berrazzouk and J. Mania, *Appl. Rad. Isotope*, **56**, 985 (2002).
2. S.V. Panteeva, D.P. Gladkochoub, T.V. Donskaya, V.V. Markova and G.P. Sandimirova, *Spectrochim. Acta*, **58B**, 341 (2003).
3. <http://pediatrics.about.com>
4. <http://www.kkgm.gov.tr/Mevzuat/khk560/Kodeks/Tebbligler/1998-9.htm> (Turkish Food Codex) No. 23172, p. 125 (1997).
5. F. Yan, Z. Liang, C. Jianna, W. Zhengtao, X. Losahan and Z. Zhengxing, *Talanta*, **53**, 1155 (2001).
6. P.P. Mouly, E.M. Gaydou and J. Corsetti, *J. Chromatogr. A*, **844**, 149 (1999).
7. S.K. Schreyer and R.S. Mikkelsen, *Sens. Actuators B*, **71**, 147 (2000).
8. M. Penza, G. Cassano, F. Torterella and G. Zaccoria, *Sens. Actuators B*, **73**, 76 (2001).
9. A. Trifiro, G. Saccani, S. Gherardi, E. Vicini, E. Spotti, M.P. Previdi, M. Ndagijimana, S. Cavalli and C. Reschiotto, *J. Chromatogr. A*, **770**, 243 (1997).
10. C.S. James, *Analytical Chemistry of Foods*, Blackie Academic & Professional, ISBN-07514-0196-X (1996).
11. <http://www.codexalimentarius.net/search/advancedsearch.do> (FAO/WHO, Food Standards, General Standards for Fruit Juices and Nectars, CAC/STAN-247-2005, p. 3).
12. Y. Yücel and C. Demir, *Talanta*, **63**, 451 (2004).