Classification of vinegar by raw materials using near infrared spectroscopy

M.J. Sáiz-Abajo, J.M. González-Sáiz and C. Pizarro

Department of Chemistry, University of La Rioja, C/ Madre de Dios 51, 26006 Logroño, La Rioja, Spain

Introduction

Nowadays, the commercialisation of fraudulent mixtures of vinegar is quite a usual practice. The most common fraud in the production and commercialisation of this product implies blending different proportions of wine vinegar and vinegar of other origins. Indeed, these fraudulent mixtures account for a significant part of the total production in the European Union. The quality control techniques currently applied cannot detect these frauds within reasonable limits.

Near-infrared (NIR) spectroscopy has gained wide acceptance in the field of food chemistry mainly due to its suitability for recording the spectra of solid and liquid samples without any pre-treatment and in a non-destructive way.

Experimental

Instrumentation and software

NIR spectra were collected for the 1100-2500 nm range with a FOSS NIRSystem 5000 liquid analyser spectrometer (Foss NIRSystems, Silver Spring, MD20904, USA) equipped with a flow cell. The instrument is controlled by a compatible PC using Vision 2.22 (Foss NIRSystems,Silver Spring, MD20904, USA) for the acquisition of data.

First and second derivatives, PCA and SIMCA models were performed with Unscrambler v7.6[1] and PARVUS package[2]. The figures were produced in MATLAB 6.1[3].

Samples

Fifty-two samples of commercial vinegar of different origins were taken from the industry and supermarkets in Spain to gather a large and representative group of samples.

Results and discussion

Principal component analysis

The Spanish regulations establish eight denominations for vinegar: wine vinegar, cider vinegar, alcohol vinegar, cereal vinegar, malt vinegar, honey vinegar and whey vinegar[4]. All of them must be genuine products and products derived from blends of final products or from the fermentation of mixtures of raw materials are illegal.

The application of spectroscopic techniques in the study of the origin and differentiation of food and drink products has developed considerably in recent years[5, 6] joined to the continuous assessment of the quality and genuineness of the products[7, 8]. The advantage of these techniques lies

Near Infrared Spectroscopy: Proceedings of the 11th International Conference.

www.nirpublications.com

in the almost complete lack of sample preparation required, which makes them especially rapid to apply.

The main purpose of this study was to develop a method based on near-infrared spectroscopy to characterise and classify vinegar of different origins in order to identify the characteristics of a sample from its NIR spectra and develop a fast method to detect possible frauds within the vinegar industry.

Classification procedures, often called pattern recognition methods, are useful for discerning classification rules that characterise groups of samples. The rules indicate the particular sample measurements (often a small subset of the original measurements) that are most useful in determining group identity. Sample measurements may be used to classify samples according to their origin, to distinguish normal and abnormal products or to detect adulteration.

The spectral data were taken from the Vision software and treated with Unscrambler. Two segments of the spectrum were removed: the first one, from 1880 to 2080 nm, due to the saturation of the spectrum caused by the strong combination band of –OH from water (1950 nm), and the second one, from 2300 to 2500 nm because of low value of the signal/noise ratio. Three replicas of each sample were taken and their mean value was calculated. These values were then centered and divided by its standard deviation. The original NIR spectra are shown in Figure 1.

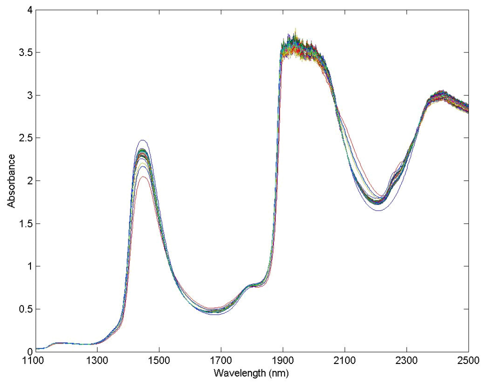
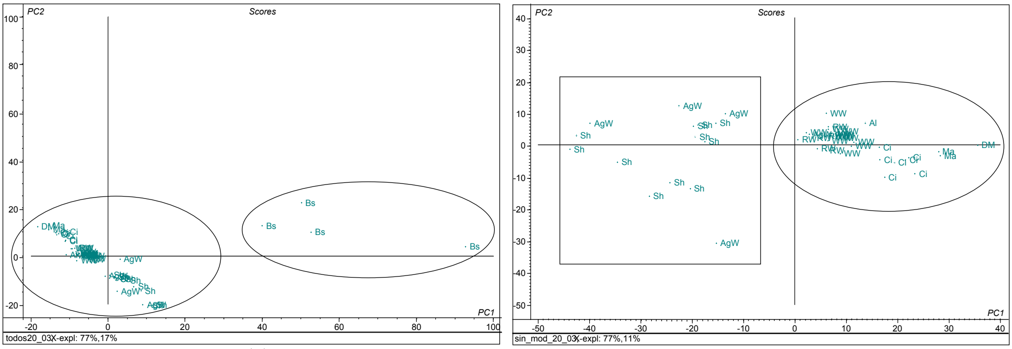


Figure 1. Original NIR spectra of the samples from different origins.

Principal component analysis (PCA) attempts to differentiate the origin of each vinegar sample. Modena and balsamic vinegar samples (Bs) are unique due to their particular method of elaboration[9] (the grape juice is partially fermented and then concentrated by direct exposure to heat). Figure 2a shows the principal components plot of all the samples, and a clear differentiation from the rest of the vinegars can be observed.

When balsamic vinegar was removed from the calculations, two groups of samples with very diverse characteristics could be plotted. The first one corresponded to Sherry vinegars (Sh) from the south of Spain and some aged vinegars (AgW), whereas the rest of the vinegar samples were included in the second group (Figure 2b). Balsamic vinegars and Sherry and aged vinegars are easily differentiated from the rest of vinegar samples with a simple principal component analysis.



(a) (b)

Figure 2. PCA with all vinegar samples (a) and without balsamic vinegar samples (b).

In a third stage, Sherry and aged vinegar samples were removed from the calculations and PCA was applied again to the data. In this case, four or even five groups were identified (Figure 3): a first large group of white (WW) and red (RW) wine vinegar, a second group of cider vinegar (Ci), and finally a group of malt vinegar (Ma) and distilled malt vinegar (DM). The latter can be considered as a new group on its own due to the distance from the rest of the malt vinegar samples. Finally, alcohol vinegar (Al) appeared isolated from the rest of the samples. The area between this vinegar and wine vinegar is particularly relevant given that the most common fraud in the vinegar industry is to sell a blend of both types of vinegar as it was genuine wine vinegar.

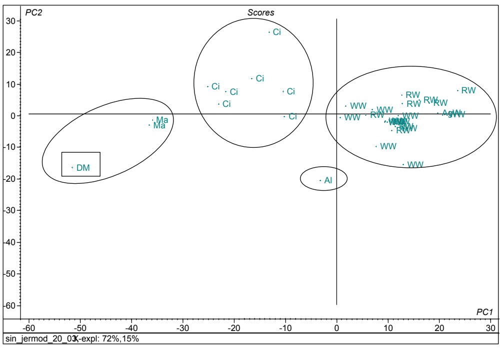


Figure 3. PCA without balsamic, aged and Sherry vinegar samples.

Class modelling

The construction of models is very important as it enables a one to one relationship between the measurements carried out on the objects[10] and concepts as complicated as detection of fraudulent mixtures or different raw materials as in our case.

Modelling is the collection of data in order to develop a formula that describes the behaviour of a system. Once a formula has been obtained, the behaviour of a system can be displayed graphically to aid in its understanding or to select an optimum combination of conditions.

SIMCA(Soft Independent Modelling of Class Analogy) is the modelling technique developed from the modelling properties of the principal component technique which has been applied successfully to solve common pattern recognition problems[11] considering each class separately.

Sensitivity is the proportion of objects belonging to the category which are correctly identified by the mathematical model and specificity is the proportion of objects foreign to the category which are classified as foreign. The sensibility of the models calculated was always above 90%, but in most of the cases it reached 100%. For the specificity, the values are shown individually for each of the remaining classes. It can be seen how wine vinegar model is completely specific compared to the others (Table 1).

Table 1. Performance of the SIMCA models. Specificity: 1, wine vinegar; 2, alcohol vinegar; 3, cider vinegar; 4, malt vinegar.

Model	Sensitivity	Specificity1	Specificity2	Specificity3	Specificity4
Wine vinegar	90	***	100	100	100
Alcohol vinegar	100	60	***	100	66
Cider vinegar	100	90	100	***	100
Malt vinegar	100	100	100	17	***

Figure 4 shows a Coomans plot of the SIMCA model used. The model built for wine vinegar did not admit any other vinegar from other origins so its specificity is 100% for each one of the other categories. It encloses all wine vinegar samples except one of them so its sensitivity is 90%. On the other hand, cider vinegar model admitted one wine vinegar sample so its specificity is 90% while for the rest of categories is 100%. Sensibility is 100% because the whole amount of cider vinegar samples are inside the area of the model. The rest of the vinegar samples (malt vinegar and alcohol vinegar) appear in the upper-right quadrant and are no enclosed by the wine and cider vinegar models.

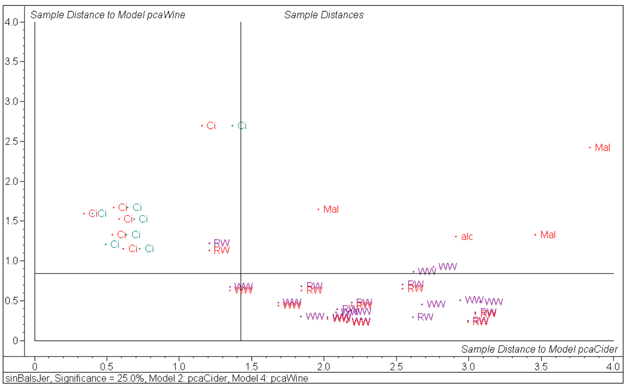


Figure 4. Coomans plot with the squared distance SIMCA for cider and wine vinegar samples.

In Figure 5 appears the SIMCA model built for alcohol and wine vinegar samples. The model developed for alcohol vinegar accepted all the alcohol vinegar samples but also a malt vinegar and a wine vinegar sample. The latter, is susceptible to be a possible blended sample due to its proximity to alcohol vinegar model and required further studies. The model built for wine vinegar did not admit any sample from any other origins, its specificity is 100%, but one of the samples seemed to fit the alcohol vinegar model, its sensitivity is 90%. Cider and malt vinegar samples appear outside of the wine and alcohol vinegar models.

Similar models were calculated for each of the categories and the values of the sensibility and specificity appear in Table 1.

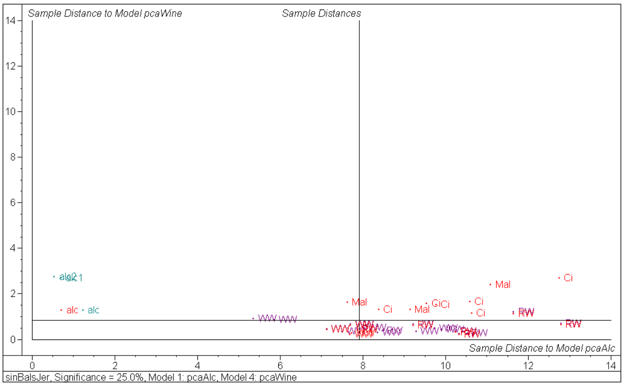


Figure 5. Coomans plot with the squared distance SIMCA for alcohol and wine vinegar samples.

Derivation is a usual treatment in NIR as it allows to correct spectral shifts. Two first-derivative methods (Norris and Savitsky Golay) and a Savitsky Golay's second-derivative method were applied to the data although no improvement of the classification was observed.

Conclusions

NIR spectroscopy, combined with PCA pattern recognition method and SIMCA modelling multivariate technique, can be used to characterise and classify vinegar from different origins in a short period of time and without damaging the samples. This is quite important nowadays that the determination of the authenticity of food products is one of the most crucial issues in the field of food quality control and safety.

References

1. The Unscrambler v7.6, Copyright 1986- 2000 CAMO ASA. Nedre Vollgate 8, N-0158 OSLO, Norway. Phone: (+47) 22 39 63 00 Fax: (+47) 22 39 63 22. E-mail: camo@camo.no Support: support@camo.no, Web: http://www.camo.com.
2. M. Forina, S. Lantieri, C. Armanino, C. Cerrato-Oliveros, V-PARVUS 2003: An Extendable Package of Programs for Data Explorative Analysis, Classification and Regression Analysis, Dipartimento di Chimica e Tecnologie Farmaceutiche ed Alimentari, Genoa, Italy.
3. MATLAB, High-Performance Numeric Computation And Visualization Software, Version 6.1 for Windows, in: The MathWorks, Natick, MA.
4. Real Decreto 2070/1993 de 26 noviembre.

5. W. Tesfaye, M.L. Morales, M.C. García-Parrilla and A.M. Troncoso, *Trends food sci. tech.* **13**, 12 (2002).
6. I.S. Arvanitoyannis, M.N. Katsota, E.P. Psarra, E.H. Soufleros and S. Kallithraka*, Trends food sci. tech.* **10,** 321 (1999).
7. E. Anklam and R. Battaglia, *Trends food sci. tech..* **12,** 197 (2001).
8. A. Garrido-Varo, C. Cobo, J. García-Olmo, M.T. Sánchez-Pineda, R. Alcalá, J.M. Horcas and A.Jiménez, *Near Infrared Spectroscopy: Proceedings of the 9th International Conference*, Ed by A.M.C. Davies and Giangiacomo (Eds). NIR Publications, Chichester, UK (1999).
9. A. del Signore, *J. food eng.* **50,** 77 (2001).
10. M. Blanco and I. Villarroya, *Trends anal. chem.* **21,** 240 (2002).
11. D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte and L. Kaufman, *Data Handling in Science and Technology, volume 2. Chemometrics: a textbook*. Elsevier Science Publishers, Amsterdam, The Netherlands (1988).