

Introduction

The aim of this work is to study the possible discrimination of Italian olive oil samples according to their region of origin. For this reason, FT-Raman spectroscopy has been used to measure the virgin olive oil samples. Then in a first step, unsupervised techniques as Principal component analysis (PCA) were performed in order to see if it is possible to gather "naturally" samples according to their qualitative attributes. In a second step, supervised techniques as Partial Least Squares Discriminant Analysis (PLS-DA) were also applied. The aim of these supervised classification methods is to develop classification rules. Representative samples of each group (known beforehand) are available, from which the relevant characteristics were known. Based on these data, classification rules are determined.

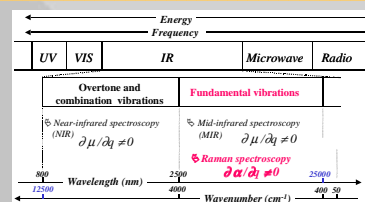
Data

The data set consists on 226 samples of olive oil measured with a FT-Raman spectrometer. The samples are coming from 12 different Italian regions. The number of samples for each region is represented in the map.



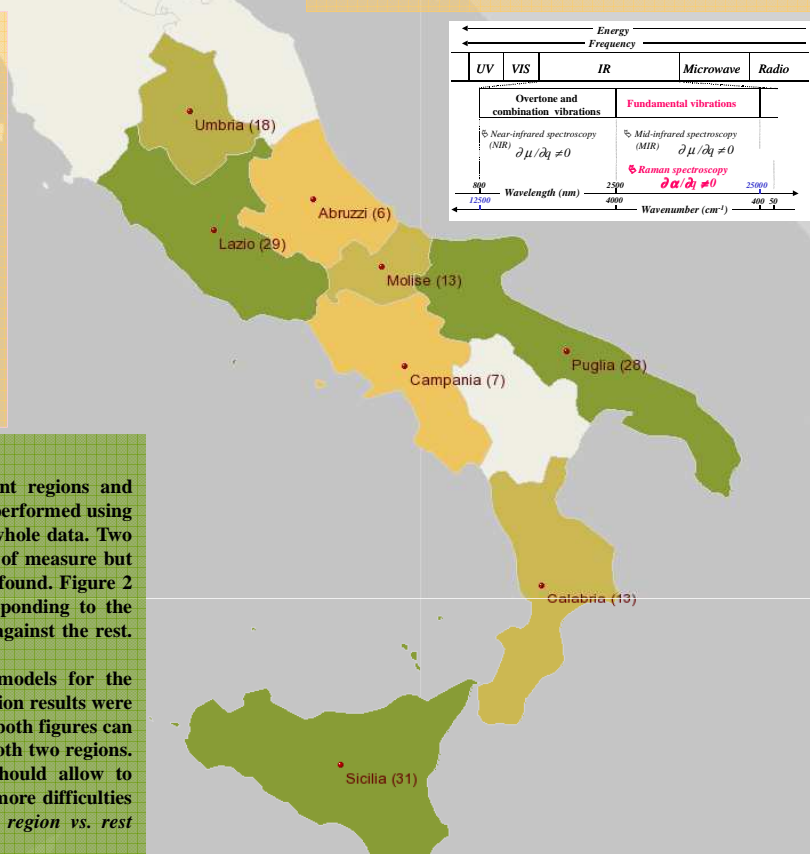
Raman spectroscopy

Raman spectroscopy is a branch of vibrational spectroscopy that provides information about the vibrations of molecules. Figure 1 shows the location of Raman phenomenon in the electromagnetic spectrum. This information is contained within the wavenumber or frequency (or Raman shift) spectrum of scattered intensity. From a chemical point of view, Raman and infrared spectroscopy are based on the vibrational transitions occurring in the ground electronic state of the molecules. Raman scattering arises from the changes in the polarisability or shape of the electron distribution in the molecule as it vibrates; in contrast, infrared absorption requires a change of the intrinsic dipole moment with the molecular vibration.



Instrumentation

FT-Raman spectra are acquired on a Perkin Elmer 2000R FT-Raman spectrometer where the excitation beam is a Nd:YAG laser (yttrium aluminium garnet crystal doped with triply ionised neodymium) with an output in the near infrared at 1064 nm. The laser power is variable from 10 to 857 mW. The software Spectrum for Windows™ is used for the acquisition, spectra manipulation and transformation. The oil samples are presented to the spectrometer in glass tubes of an internal diameter of 12 mm and a length of 75 mm that are introduced into a dedicated thermostatic sample holder made of aluminium to enhance the collection of the scattered light. The sample accessory is maintained at 40°C by a heated water bath and circulation pump. The laser power is set at 600 mW, the resolution is 4cm⁻¹ and the number of co-added scans is 50 for each spectrum.



Results

This study aims to differentiate between the different regions and mainly to discriminate Liguria region from the rest. This is performed using PCA and PLS-DA. Figure 1 shows the PCA results on the whole data. Two main groups become evident, which correspond to the date of measure but for the rest no clear separation between the regions can be found. Figure 2 shows that several clusters are present in the data corresponding to the regions. Ligurian samples have a tendency to discriminate against the rest. Then supervised techniques are needed.

Figures 3-4 show two examples of PLS-DA calibration models for the Liguria region against other regions. In all cases the calibration results were validated using leave-one-out cross-validation (LOOCV). In both figures can be seen that Liguria samples can be discriminated against both two regions. Then, working with *one region vs. one region* models should allow to discriminate between regions. As observed from the figure more difficulties to discriminate will be observed when working using *one region vs. rest* models as more misclassification errors will be found.

Unsupervised results

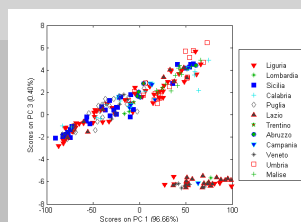


Figure 1: PCA results for the Italian virgin olive oil samples grouped by region after mean center.

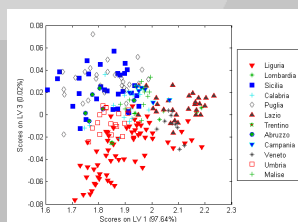
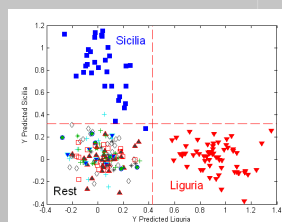
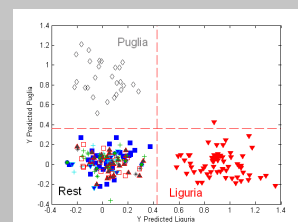


Figure 2: Latent variables from PLS for the Italian virgin olive oil samples grouped by region after SNV and 1st derivative (7/2).

Supervised results



Figures 3-4: Two examples of PLS-DA results of the calibration set (LOOCV) for the Ligurian olive oil samples against Sicilia and Puglia after SNV and 1st derivative (7/2).



This poster has been presented during the 4th Euro Fed Lipid Congress, which took place the 1-4 October 2006 at the University of Madrid (UCM), Spain.

Acknowledgement

This work is funded by the European Commission, under the FP6 Food Quality and Safety Priority, within the framework of the Integrated Project TRACE – 006942 – entitled "Tracing Food Commodities in Europe". The information contained in this poster reflects the authors' views; the European Commission is not liable for any use of the information contained therein. The authors thank the collaborators of the TRACE project for their useful contribution to the elaboration of the sample bank used.