Raman spectroscopy and multivariate classification as a tool for different ketchup samples discrimination

S. Kolašinac[1], I. Pećinar[1], Z. Dajić-Stevanović[1]
[1] *Faculty of Agriculture, Departman of Agricultural Botany, Belgrade, Serbia*
e-mail:stefan.kolasinac@agrif.bg.ac.rs

Introduction:

Ketchup derives from thermally processed tomatoes and represents the main source of lycopene and beta-carotene (Schulz et al., 2006). There are a lot of varieties that can be used as a raw material for producing ketchup. However, depending on which variety of tomato was used in production, there are differences in physico-chemical, organoleptic and chemical quality as well as nutritional value (Desai, 2009).

Aim:

The main aim of this paper is to test PCA-LDA (Principal Component Analysis – Linear Discriminant Analysis) coupled with Raman spectroscopy (RS) for discrimination of two ketchups commercially available at the local markets in Serbia. Raman spectra were recorded at two different wavelengths (785 and 532 nm) and 30 spectra per sample are obtained.

Material and methods:

The data were divided into the training (3/4 of samples) and validation (1/4 of samples) data. Two types of pre-processing methods that were applied and results of discrimination are represented in table 1. Pre-processing analysis of the spectra was performed using the software The Unscrambler X version 10.4 (Camo Software, Oslo, Norway) while supervised classification models were performed using the Python software.

Results and discussion:

Obtained results (table 1 and figure 1) showed that second-order derivatives did not improve discrimination power. On the other hand, laser at 785 nm provided a better classification of samples which can be related to the fact that 785 nm laser reveals bands that are masked by the high fluorescence background seen when using the 532 nm laser and consequently gives wider chemical information about the sample (Haraa et al., 2018).

References:

[1] R. Haraa, M. Ishigakib, Y. Kitahamab, Y. Ozakib, T. Genkawaa. Food Chem. 258, 308 (2018)
[2] H. Schulz, W. Schütze and M. Baranska. Acta Hort., 712, 603 (2016).
[3] Desai, S. Doctoral Theses. (2009).

Funding information:
This work was supported by Ministry of Education, Science and Technological Development of Republic of Serbia, the contract No. 451-03-68/2020-14/200116 and the EthnoHERBS-H2020-MSCA-RISE-2018 project.

Table 1. Classification results of training and test sets of PCA-LDA

Excitation wavelength			785 nm		532 nm	
Pre-processing methods			Smoothing + baseline correction + normalization + principal component analysis	Smoothing + baseline correction + normalization + 2nd order derivative + principal component analysis	Smoothing + baseline correction + normalization + principal component analysis	Smoothing + baseline correction + normalization + 2nd order derivative + principal component analysis
Correct classified (%)	Training data	Sample 1	95.45	95.45	95.45	90.91
		Sample 2	100.00	95.45	95.45	95.45
	Test data	Sample 1	87.50	87.50	62.50	37.50
		Sample 2	100.00	87.50	37.50	37.50

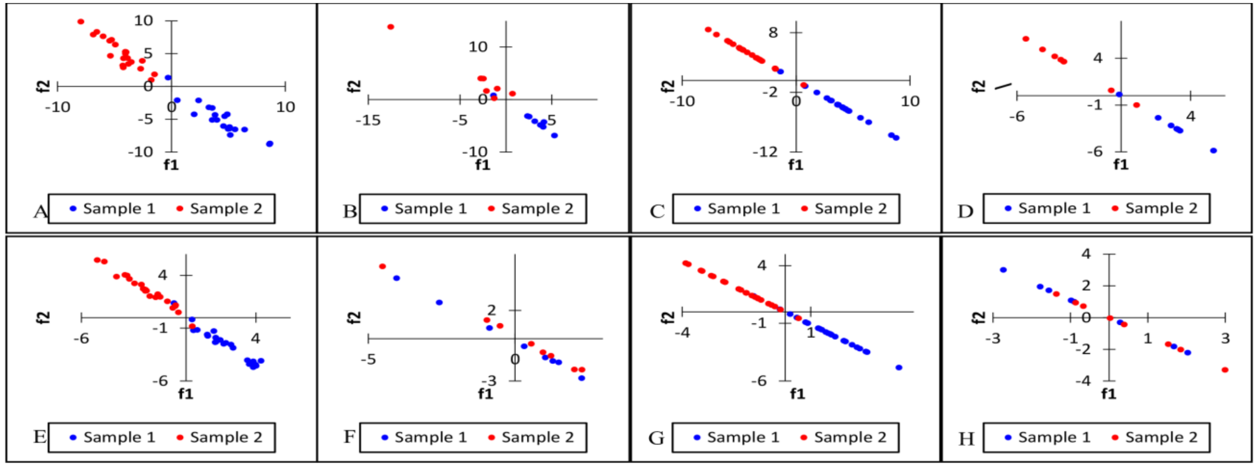

Figure 1. PCA-LDA discrimination plots: A (training data) and B (test data) illustrate discrimination between two samples using RS equipped with a 785 laser and pre-processing without second order derivative; C (training data) and D (test data) illustrate discrimination between two samples using RS equipped with a 785 nm laser and pre-processing with second order derivative; E (training data) and F (test data) illustrate discrimination between two samples using RS equipped with a 532 nm laser and pre-processing without second order derivative; G (training data) and H (test data)) illustrate discrimination between two samples using RS equipped with a 532 nm laser and pre-processing with second order derivative Discriminant scores (f1, f2) determine the sample's membership in appropriate class.