X Colloquium Chemiometricum Mediterraneum
Menorca, from June 12th to 14th, 2019

DEVELOPMENT OF A METHODOLOGY TO ANALYZE LEAVES FROM *Prunus dulcis* VARIETIES USING NEAR INFRARED SPECTROSCOPY

Sergio Borraz-Martínez [a,c], Ricard Boqué [b], Joan Simó [a,d], Mariàngela Mestre [c], Anna Gras [a]

[a] *Universitat Politècnica de Catalunya, Department of Agri-Food Engineering and Biotechnology, Esteve Terrades 8, 08860 Castelldefels, Spain*
[b] *Universitat Rovira i Virgili, Department of Analytical Chemistry and Organic Chemistry, Campus Sescelades, 43007 Tarragona, Spain*
[c] *Agromillora Iberia S.L.U, Center of Initial Materials, Ctra. BV-2247 km. 3, 08770-Sant Sadurní d'Anoia*
[d] *Fundació Miquel Agustí, Esteve Terrades 8, 08860 Castelldefels, Spain*

sergio.borraz@upc.edu; ricard.boque@urv.cat

Introduction

The appearance of varietal mixtures is an important problem in the nursery plants field. Traditional methods to deal with this problem are based on genetic analysis, but they are expensive and complex. Near-infrared spectroscopy (NIRS) could be a faster and cheaper alternative to traditional methods.

Objective

The aim of this work was to investigate how sampling of vegetal material affects the collection of NIR spectra for building a multivariate discriminant model for *Prunus dulcis* varietal classification.

Points of study

Sources of variation

- Position of the leaves in the trees (young/adult leaves)
- Differences among trees of the same variety
- Differences at varietal level

Processed samples:

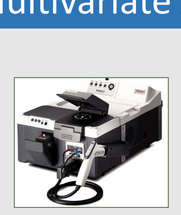
- Fresh leaves (direct measurement)
- Dried leaves (65 ºC for 48h)
- Dried-powdered leaves (65 ºC for 48h + grinding)

Material and Methods

- Three varieties of *Prunus dulcis* (*Avijor*, *Guara*, and *Pentacebas*)
- Principal component analysis (PCA)
- Partial least-squares discriminant analysis (PLS-DA)
- ANOVA simultaneous component analysis (ASCA)

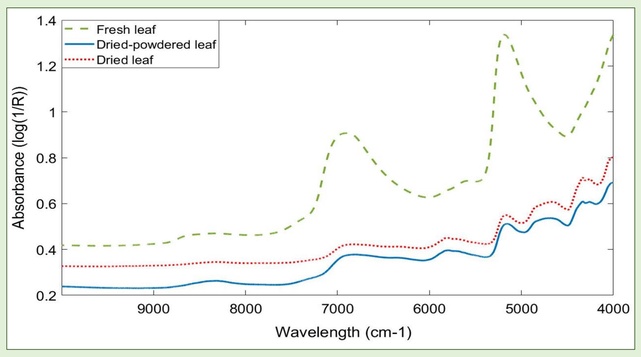
Multivariate Analysis

- Antaris II FT-NIR analyzer (Thermo Scientific, USA)



Experimental Results

Raw spectra



Results of ASCA-ANOVA model. Study of the variance of the factors

Factor	Fresh leaves: Principal components	Fresh leaves: Effect %	Dried leaves: Principal components	Dried leaves: Effect %	Dried-powdered leaves: Principal components	Dried-powdered leaves: Effect %
Variety	2	30.26	2	19.25	2	24.99
Tree	3	1.83	3	4.69	3	1.87
Young / adult	1	19.11	1	24.18	1	6.68
Residual	6	48.80	3	51.88	3	66.46

Variety was the most influential factor, followed by the age of the leaves. The residual term may be related to the physiological state of the leaves (damaged leaves and/or climatologic agents).

PLS-DA model results of the spectra pre-treatment and study of the types of pre-processed samples

Dried-powdered leaves

Real class	Data set	Assigned class: *SNV + Mean center*	Assigned class: *SNV + 1st derivative + Mean center*
Avijor	Cross-validation	87.4 %	99.2 %
	Test set validation	97.5 %	100 %
Guara	Cross-validation	89.9 %	99.2 %
	Test set validation	96.6 %	100 %
Pentacebas	Cross-validation	97.5 %	100 %
	Test set validation	99.2 %	100 %

Dried leaves

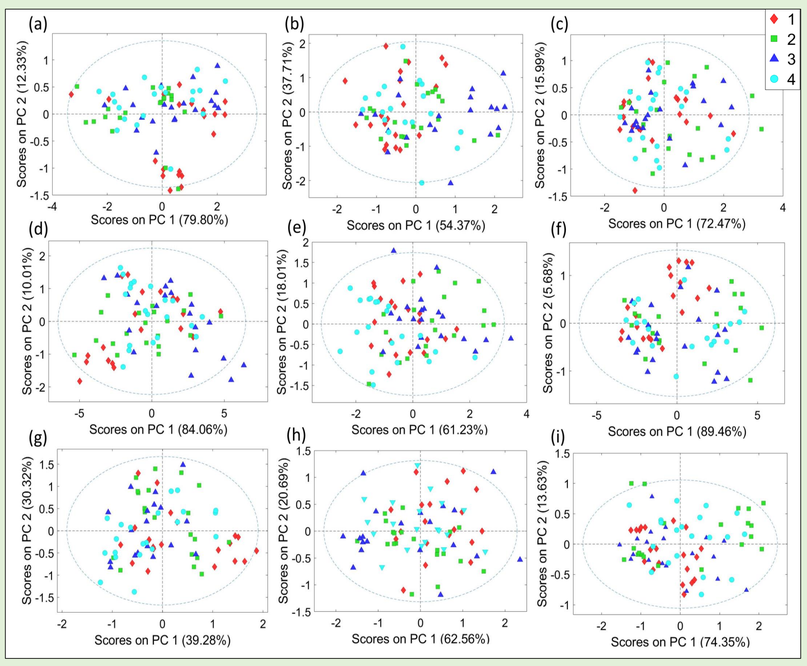
Real class	Data set	Assigned class: *SNV + Mean center*	Assigned class: *SNV + 1st derivative + Mean center*
Avijor	Cross-validation	97.5 %	99.2 %
	Test set validation	95.0 %	98.3 %
Guara	Cross-validation	95.0 %	100 %
	Test set validation	92.5 %	97.5 %
Pentacebas	Cross-validation	97.5 %	99.2 %
	Test set validation	97.5 %	100 %

Fresh leaves

Real class	Data set	Assigned class: *SNV + Mean center*	Assigned class: *SNV + 1st derivative + Mean center*
Avijor	Cross-validation	100 %	97.5 %
	Test set validation	99.2 %	100 %
Guara	Cross-validation	99.2 %	97.5 %
	Test set validation	98.3 %	99.2 %
Pentacebas	Cross-validation	99.2 %	100 %
	Test set validation	99.2 %	99.2 %

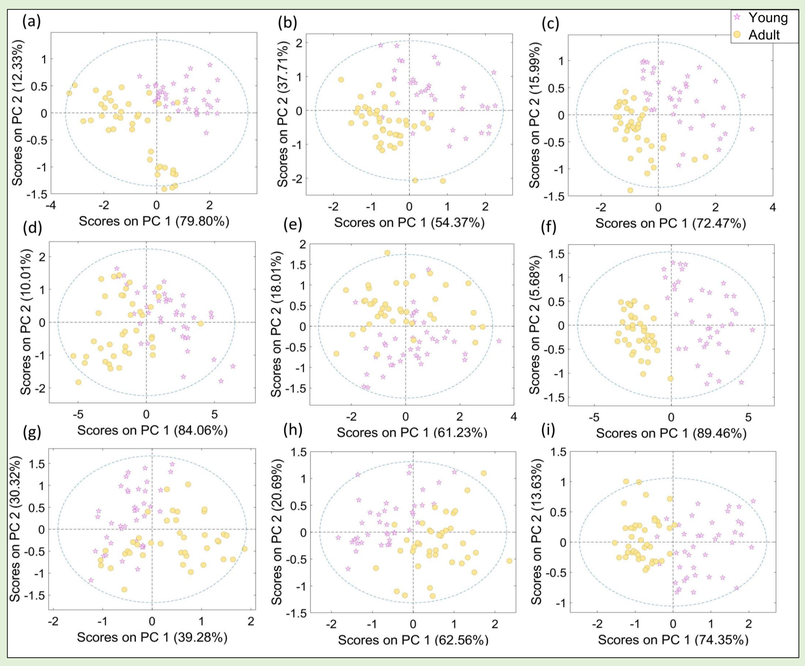
The best results were obtained with dried powdered leaves using SNV + SG first derivative + mean center spectral pre-treatment. Good results were also obtained for fresh leaves, being the easiest and most suitable samples for laboratory or industrial analysis.

Variability between trees from the same variety



No notable differences were detected between trees of the same variety, thus showing homogeneity within varieties.

Variability between leaves from the same variety



Differences were observed at the PCA level between young and adult leaves, which indicated age is important to be considered during the sampling process.

Conclusions

The results indicated that variety was the most important factor for classification. The spectral pre-treatment that provided the best results was a combination of standard normal variate (SNV), Savitzky-Golay first derivative, and mean-centering methods. With regard to the type of processed sample, the highest percentages of correct classifications were obtained with fresh and dried powdered leaves at both the training set and test set validation levels.

Acknowledgment: The authors thank the financial support provided by DI-COF 2017 and project AGL2015-70106-R, AEI/FEDER, UE, and to Thermo Scientific for the temporary loan of the NIR equipment.