

HPLC Determination and FT-MIR Prediction of Sugars from Juices of Different Apple Cultivars during Fruit Development

Andruța E. MUREȘAN¹, Sevastița MUSTE^{1*}, Romina A. VLAIC¹, Otilia BOBIȘ², Crina MUREȘAN¹, Carmen SOCACIU³, Vlad MUREȘAN¹

¹University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Food Engineering Department, 64 Calea Florești, 400509, Cluj-Napoca, Romania; andruta.muresan@usamvcluj.ro; sevastita.muste@usamvcluj.ro (*corresponding author); romina.vlaic@usamvcluj.ro;

crina.muresan@usamvcluj.ro; vlad.muresan@usamvcluj.ro

²University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Department of Apiculture and Sericulture, Laboratory of Bee Products Control, 3-5 Calea Mănăștur, 400372, Cluj-Napoca, Romania; obobis@usamvcluj.ro

³University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Food Science Department, 3-5 Calea Mănăștur, 400372, Cluj-Napoca, Romania; carmen.socaciu@usamvcluj.ro

Abstract

Individual sugars were analyzed by high performance liquid chromatography (HPLC) in samples of apple juices obtained from the fruits of 'Jonathan', 'Starkrimson' and 'Golden Delicious' cultivars. Samples were harvested from the inside and the periphery of the crown, at different periods during fruits growth, from 7 to 144 days after full bloom (DAFB). Values from 0.42 to 14.33%, 0.29 to 4.06% and 0 to 4.28% were determined for fructose, glucose and sucrose, respectively. The values of fructose and glucose have increased significantly ($p < 0.05$), starting with the seventh DAFB, regardless of the studied cultivar, while sucrose increased slowly at the beginning and then faster starting 65 DAFB. Fourier transform mid-infrared (FT-MIR) analysis confirmed the differences between juice samples, the region 900-1500 cm^{-1} being the most specific to sugars signals. FT-MIR coupled to partial least squares (PLS) calibration models for predicting individual sugars of apple juices were developed. The optimal regions and pre-treatments of the spectra were 900-1500 cm^{-1} and Savitzky Golay first derivative (d1) for fructose, 900-1200 cm^{-1} and d2 for glucose and 900-1200 cm^{-1} and standard normal variate for sucrose. In cross-validation, the PLS calibration models showed very good performance for fructose ($R_{\text{cv}}^2 = 0.95$; standard error of cross-validation (SECV) = 0.907) and acceptable for glucose ($R_{\text{cv}}^2 = 0.85$; SECV = 0.424), while for sucrose showed only satisfactory performance ($R_{\text{cv}}^2 = 0.75$; SECV = 0.561). For practical relevance, the FT-MIR predicted values were compared against the HPLC determined reference values in external validations tests. The best results were achieved for fructose ($R_p^2 = 0.94$; RPD = 4.9), while glucose ($R_p^2 = 0.84$; RPD = 2.61) and sucrose ($R_p^2 = 0.7$; RPD = 2.08) models reached satisfactory values.

Keywords: apple sugars, chemometrics, fruit development, mid-infrared, rapid method

Introduction

The apple is considered one of the most agreed fruit for both children and adults. It is an important part of the human diet being a rich source of monosaccharides and biologically active compounds (Fuleki *et al.*, 1994; Miller *et al.*, 1997; Zhang *et al.*, 2010; Zheng *et al.*, 2012). Sugars represent more than 98% of the total soluble solids of apple juice (Kelly *et al.*, 2005). Soluble carbohydrates reach significant concentrations in mature apple fruit (fructose 5-15% (w/w), glucose 1-4% (w/w), sucrose 1-5% (w/w)). Carbohydrates provide energy to the fruit and determine fruit's edible quality (Borsanie *et al.*, 2009). Soluble sugars are known for their effect on plant growth and development (Rolland *et al.*, 2006; Ruan *et al.*, 2010). In fleshy fruits, soluble sugar

accumulation during fruit development determines the sweetness at harvest. Compared to other plants, apples have a unique metabolism when it comes to sugar accumulation, because more than 80% of the total carbon flux is going through fructose, so fructose reaches higher levels. Although there are reports on the accumulation of apple carbohydrates and the changes caused by enzymes during growth, the accumulation of sugars remain unclear, being regulated at the level of gene expression (Chourey *et al.*, 1985; Mingjun *et al.*, 2012; Yamaki *et al.*, 1992; Zhang *et al.*, 2010).

It is known that instrumental techniques for carbohydrates determination are expensive and require considerable work. Recently, Fourier transform mid-infrared (FT-MIR) has become a widely accepted method for determining the chemical

composition of food, being quick, practical and environmental friendly. Absorption bands in the MIR region are characteristic linkages and functional groups of a molecule (Kelly *et al.*, 2005). These absorption bands characterize a molecule state, including stretching or bending, so that the whole spectrum is a fingerprint of a specific compound or product. A major advantage of MIR spectroscopy is its sensitiveness to the precise chemical composition of analyzed samples. Near-infrared (NIR) spectroscopy is usually used to identify a class of compounds, and not the individual components (Ścibisz *et al.*, 2011). NIR coupled to chemometrics provided good results in predicting apple fruit quality, including polyphenols content, total level of sugars (Pissard *et al.*, 2013), consistency, color, antioxidant capacity (Giovannelli *et al.*, 2014), or chlorophyll and carotenoids (Beghi *et al.*, 2013). MIR spectroscopy coupled to chemometrics was successfully applied to predict the chemical composition (sugars, organic acids, acidity, dry matter etc.) of different food products, mainly liquids, such as mango juice (Duarte *et al.*, 2002), apple juice (Irudayaraj and Tewari, 2003; Leopold, 2010), or even whole apricot fruit (Bureau *et al.*, 2009) or tomato (Ścibisz *et al.*, 2011). Andrade *et al.* (2003) showed that FT-MIR was able to differentiate “pure” apple drinks from those with added flavors. MIR spectral data was treated using various chemometric techniques to determine the best combination for optimal classification of the beverages. MIR technique was also used to determine apple juice authenticity, together with linear discriminant analysis, principal component (PCR) and partial least squares (PLS) regressions (Sivakesava *et al.*, 2001). FT-MIR spectroscopy has been widely used for the analysis of must and wine. It also became the alternative method for analyzing sugar from foods and beverages such as sugar cane juice, mango juice, fruit juices and refreshments. On the other hand, the FT-MIR was used for authentication and detection of various products obtained from fruits that have suffered deterioration process, as well as classifying wine or honey according to their origin (Bureau *et al.*, 2009).

The study of sugars accumulation dynamics for apple fruit harvested at different intervals (from 7 to 144 days after full bloom) (DAFB) will open new ways to valorise the fruits resulting from physiological fall. Moreover, a direct FT-MIR analysis method will bring significant contributions, the knowledge of individual sugars concentrations during apple fruit growth being an essential tool for selecting the correct valorisations conduit (direct consumption, juice production, fermentation, extraction of antioxidants, pigments etc.). Furthermore, a direct method for determining the sugars concentration of different quality of apple juice will be useful in future genetics and breeding research for monitoring the sugar content of the fruits obtained from new hybrids during growth. Consequently, the current study aimed to determine the individual sugars from juices of ‘Jonathan’, ‘Starkrimson’ and ‘Golden Delicious’ apple cultivars during fruit development by high performance liquid chromatography (HPLC) and based on the same set of samples to develop FT-MIR calibration models.

Materials and methods

Biological materials

Studied apple cultivars were ‘Jonathan’, ‘Starkrimson’ and ‘Golden Delicious’. Samples were collected during fruit development, from the same orchard, from trees having the same

exposure in the orchard. Sampling was done the same for all three apple cultivars. Fruit were collected from the inside and from the periphery of the crown, at different times of growth: 7, 15, 35, 65, 107 and 144 DAFB. Consequently, a number of 36 different samples (3 varieties x 6 harvesting times x 2 crown positions) were harvested in two replicates each. For each harvesting time, the apples were individually vacuum packed after sampling then frozen immediately and stored at -18 °C. At the end of the harvesting period, all samples were unfrozen, peeled and the pulp was pressed. The correspondent apple juice obtained by pressing the pulp of each sample was stored at -18 °C in two separate containers until further analysis (e.g. HPLC and FT-MIR analyses).

Determination of individual sugars by HPLC

Determination of sugars in apple juice obtained from studied samples was performed according to Bonta *et al.* (2008) with some modifications. After filtration through a 0.45 µm Millipore filter, the individual sugar content of each sample was determined by HPLC coupled with a refraction index detector. Identification of sugars was based on the retention times of standards. Quantification was based on calibration curves obtained for each sugar, after injection of known concentrations of a standard. The equipment used was a Shimadzu HPLC system consisting of: controller, auto-sampler, degasser, pump, IR detector. Separation was performed on a modified Amino collum – Alltima Amino 100A, 5 µm, 250 × 4.6 mm. The mobile phase flow rate was 1.3 ml/min. The column temperature was set to 30 °C, the sample injection volume being 20 µl. It has been worked on isocratic conditions using as mobile phase acetonitrile /water (75/25; v/v). Each determination was performed in two repetitions, the standard error of the laboratory method being 0.0471, 0.0143 and 0.0051 for fructose, glucose and sucrose, respectively.

FT-MIR analysis

FT-MIR spectra of apple juice obtained from the studied samples were recorded using a Shimadzu IR Prestige-21 equipment. Each spectrum was recorded from 650–4000 cm⁻¹, in three repetitions, consisting of an average of 128 separate scans. Treatments applied to experimental data and mathematical calibration models were made using The Unscrambler v.9.7 chemometric software.

Statistical analysis

In order to show the effect of cultivar, harvesting crown position and time of harvest on apple juice fructose, glucose and sucrose concentrations, three-way ANOVA General Linear Model, as well as one-way ANOVA and Tukey’s comparison statistical tests (significance level $\alpha = 0.05$) were performed on Minitab 16.1.0.

PLS regression was performed to study the predictive ability of the calibration models. The models were validated using the full cross validation technique, in order to determine the optimal number of latent variables and to detect the outlier samples.

Cross validation (leave one out) means that at a time, one sample is not included in the calibration set and the model is obtained with the rest of the samples, being also used to predict the unused sample; this operation is repeated as many times as the number of samples. External validation sets were used for testing the obtained models. The best models were selected based

Table 1. Changes of individual sugars content of three apple cultivars during fruit development

Sugar	Culti- var	Crown position	Days after full bloom					
			T1 - 7 days ⁶	T2 - 15 days	T3 - 35 days	T4 - 65 days	T5 - 107 days	T6 - 144 days
Content (g/100 g juice) ⁷								
Fructose	J ¹	IC ³	0.5831 ^{EaA} ± 0.01	0.6072 ^{EaB} ± 0.02	2.1984 ^{DaA} ± 0.01	5.3155 ^{CaA} ± 0.06	9.2764 ^{BaA} ± 0.08	14.3371 ^{AaA} ± 0.17
		PC ⁵	0.5476 ^{EbA} ± 0.01	0.8455 ^{EaA} ± 0.02	1.9889 ^{DabB} ± 0.03	5.5385 ^{CbA} ± 0.10	8.6919 ^{BbB} ± 0.15	12.1343 ^{AbB} ± 0.19
	GD ²	IC	0.6168 ^{EbB} ± 0.01	0.9160 ^{EbA} ± 0.02	2.0886 ^{DcA} ± 0.02	3.4863 ^{CbB} ± 0.02	6.1224 ^{BbB} ± 0.09	9.6471 ^{AbA} ± 0.21
		PC	0.4741 ^{EaA} ± 0.01	0.6769 ^{EaA} ± 0.02	2.0194 ^{DbbB} ± 0.03	6.0804 ^{CbA} ± 0.12	8.2688 ^{BbA} ± 0.12	10.3738 ^{AcA} ± 0.04
	S ³	IC	0.4220 ^{EaA} ± 0.01	0.8010 ^{EaA} ± 0.02	1.9980 ^{DbaA} ± 0.00	4.0390 ^{CcB} ± 0.06	6.9180 ^{BbB} ± 0.08	11.2070 ^{AcA} ± 0.29
		PC	0.6360 ^{EbB} ± 0.01	0.8460 ^{EbB} ± 0.02	1.8790 ^{DbaA} ± 0.03	5.5010 ^{CaA} ± 0.01	8.0900 ^{BaA} ± 0.14	9.8690 ^{AcB} ± 0.18
Glucose	J	IC	0.4054 ^{EaA} ± 0.01	0.4054 ^{EaA} ± 0.01	1.3257 ^{DabB} ± 0.03	1.7711 ^{CbB} ± 0.03	2.0509 ^{BbB} ± 0.04	4.0683 ^{AaA} ± 0.04
		PC	0.2951 ^{EbB} ± 0.01	0.4356 ^{EbA} ± 0.01	1.9889 ^{CaA} ± 0.01	1.9505 ^{CaA} ± 0.03	2.6049 ^{BaA} ± 0.02	3.7027 ^{AaB} ± 0.05
	GD	IC	0.3681 ^{FaB} ± 0.00	0.6068 ^{EbA} ± 0.01	1.4077 ^{DcA} ± 0.01	1.5521 ^{CaA} ± 0.01	0.9146 ^{BaA} ± 0.02	2.0964 ^{AbB} ± 0.01
		PC	0.2915 ^{FaA} ± 0.00	0.3953 ^{EaA} ± 0.01	1.4200 ^{DcB} ± 0.00	1.9766 ^{CcA} ± 0.04	1.8861 ^{BbB} ± 0.06	2.7857 ^{AaA} ± 0.02
	S	IC	0.3200 ^{FbA} ± 0.00	0.5301 ^{EaA} ± 0.01	1.3699 ^{CcA} ± 0.02	2.0507 ^{BcB} ± 0.04	2.9560 ^{DcB} ± 0.06	3.3553 ^{AcB} ± 0.05
		PC	0.3911 ^{DbbB} ± 0.01	0.5631 ^{DcB} ± 0.01	1.1708 ^{CbA} ± 0.00	2.0330 ^{BaA} ± 0.01	2.4676 ^{BcA} ± 0.02	3.6455 ^{AbB} ± 0.01
Sucrose	J	IC	0.0000 ^{DcA} ± 0.00	0.0000 ^{DbaA} ± 0.00	0.0036 ^{DabB} ± 0.00	0.5127 ^{CcB} ± 0.01	2.1910 ^{BaA} ± 0.01	4.2869 ^{AaA} ± 0.01
		PC	0.0000 ^{DcA} ± 0.00	0.0170 ^{DbaA} ± 0.00	0.0058 ^{DcB} ± 0.01	0.6520 ^{CbA} ± 0.01	1.3206 ^{BbB} ± 0.03	2.7887 ^{AaB} ± 0.02
	GD	IC	0.0000 ^{DcA} ± 0.00	0.0032 ^{DbaA} ± 0.00	0.0028 ^{DcA} ± 0.00	0.3419 ^{CcB} ± 0.00	0.8817 ^{BcB} ± 0.01	3.2694 ^{AcA} ± 0.04
		PC	0.0000 ^{DcA} ± 0.00	0.0000 ^{DbaA} ± 0.00	0.0086 ^{DcA} ± 0.00	0.8541 ^{CcA} ± 0.01	1.9047 ^{AcB} ± 0.00	2.5878 ^{BcA} ± 0.04
	S	IC	0.0000 ^{DcA} ± 0.00	0.0000 ^{DbaA} ± 0.00	0.0000 ^{DcA} ± 0.00	0.2888 ^{CbB} ± 0.00	0.6445 ^{BbB} ± 0.01	2.0860 ^{AbB} ± 0.01
		PC	0.0000 ^{DcA} ± 0.00	0.0000 ^{DbaA} ± 0.00	0.0000 ^{DcA} ± 0.00	0.4730 ^{CcA} ± 0.01	1.1917 ^{BaA} ± 0.01	1.2521 ^{AbA} ± 0.01

¹Jonathan; ²Golden Delicious; ³Starkrimson; ⁴Inside of the three crown; ⁵Periphery of the crown; ⁶T1–T6 – Harvesting time during fruit growth; ⁷Capital identical letters for each cultivar and each position in the crown denote no significant differences ($p > 0.05$); Lowercase identical letters for each position and each time denote no significant differences ($p > 0.05$); Italic capital identical letters for each cultivar and each time denote no significant differences ($p > 0.05$)

on the highest coefficients of determination for calibration and cross-validation (R_{cal}^2 , R_{val}^2) or external validation (R_p^2), and the smallest standard error of calibration (SEC), standard error of cross validation (SECV) or standard error of prediction (SEP). Ratio of performance of deviation (RPD) was also used. RPD is the ratio between standard deviation (SD) and SEP, indicating the model's ability to predict future results. Cozzolino *et al.* (2005) mentioned that models with RPD value higher than three can be used only in current standard screening analysis, while the models having RPDs higher than five are considered suitable for quality control products.

Results and discussions

HPLC analysis of apple juice sugars obtained from fruits harvested during development

Table 1 shows the carbohydrate concentration for each cultivar, as a function of position and harvesting time. Fructose and glucose concentrations increased significantly ($p < 0.05$), starting seven DAFB, irrespective of studied cultivars. Sucrose content of the fruits increased slowly in early development stages and faster starting from 65 DAFB. A similar increase in fructose and sucrose content was reported in previous studies for other apple cultivars, which concluded that the increase of sucrose occurs at the time of starch degradation (Berüter and Feusi 1997; Hecke *et al.*, 2006; Petkovsek *et al.*, 2007; Teo *et al.*, 2006; Zhang *et al.*, 2010).

Comparing the cultivars as a function of fruit position in the crown for each time separately, one can notice statistically significant differences ($p < 0.05$) from seven DAFB for fructose and glucose and 15 DAFB for sucrose. Carbohydrate content oscillations occurred throughout the fruit development, being almost impossible to rank the cultivars by their carbohydrates accumulation.

Fruit position in the tree crown can play an important role on fructose, glucose and sucrose concentrations. During the current study oscillations were noticed among the harvesting

crown positions, an overall conclusion regarding this experimental factor being impossible to be drawn. However, at the last harvesting time (144 DAFB), fruits from inside of the crown contained a higher content of fructose than the fruits harvested from the periphery of the crown for all the three apple cultivars studied. Regarding glucose and sucrose content at 144 DAFB, 'Jonathan' cultivar had the highest content in fruits harvested from fruits inside of the crown, while 'Starkrimson' and 'Golden Delicious' cultivars had the highest glucose and sucrose content in fruits harvested from the periphery of the crown.

The applicability of FT-MIR and chemometrics for the prediction of sugars from apple juices obtained from fruits harvested during development

Spectra description

A set of 30 standard sugar solutions was developed, the concentrations chosen being between 0-15% for fructose and between 0-5% for glucose and sucrose (Table 2). FT-MIR spectra of the standard solution mixes, respectively, those of the individual standard solutions of fructose (conc. 15%), glucose (conc. 5%) and sucrose (conc. 5%) are shown in Fig. 1A. Spectra showed strong absorption bands, characteristic to water molecules, between 1500-1800 cm^{-1} and between 2800-3700 cm^{-1} , indicating an apparent similarity of all solutions. However, a differentiation of sugars standard solutions can be achieved by analyzing the spectral region characteristic of this class of compounds, between 900-1500 cm^{-1} . For this region, characteristic absorption bands were identified (Fig. 1C) for glucose (specific max at 991, 1031, 1080, 1107, 1151, 1367 cm^{-1}), fructose (specific max at 966, 1063, 1083, 1155, 1254, 1346, 1416 and 1456 cm^{-1}) and sucrose (specific max at 995, 1055, 1113, 1138 cm^{-1}). Carbohydrates characteristic absorption bands in the region 904-1153 cm^{-1} were attributable to C-O s and C-C bonds (stretching modes) whereas those from the region 1199-1474 cm^{-1} were due to the groups O-C-H, C-C-H and C-O-H (bending modes) (Irudayaraj and Tewari, 2003).

Table 2. Concentrations of sugars standard solutions mixes (range 0-15% for fructose, 0-5% for glucose and 0-5% for sucrose)

Name	Glucose % (w/w)	Fructose % (w/w)	Sucrose % (w/w)	Name	Glucose % (w/w)	Fructose % (w/w)	Sucrose % (w/w)
STD1	0.3	0.4	0.2	STD16	3.0	8.5	1.0
STD2	0.5	0.6	0.3	STD17	4.0	9.0	2.0
STD3	1.3	0.8	0.5	STD18	1.0	9.5	3.0
STD4	2.0	1.0	0.7	STD19	2.0	10.0	4.0
STD5	3.0	1.5	0.8	STD20	3.0	11.0	5.0
STD6	3.5	2.0	0.9	STD21	4.0	11.5	5.0
STD7	4.0	3.0	1.0	STD22	5.0	12.0	4.0
STD8	2.5	3.5	1.5	STD23	1.0	13.0	3.0
STD9	4.5	4.0	2.0	STD24	2.0	13.5	2.0
STD10	0.0	5.0	2.5	STD25	3.0	14.0	1.0
STD11	2.0	5.5	3.0	STD26	4.0	14.5	2.0
STD12	3.0	6.0	3.5	STD27	5.0	15.0	5.0
STD13	4.0	6.5	4.0	STD28	5.0	0.0	0.0
STD14	1.0	7.0	4.5	STD29	0.0	15.0	0.0
STD15	2.0	8.0	5.0	STD30	0.0	0.0	5.0

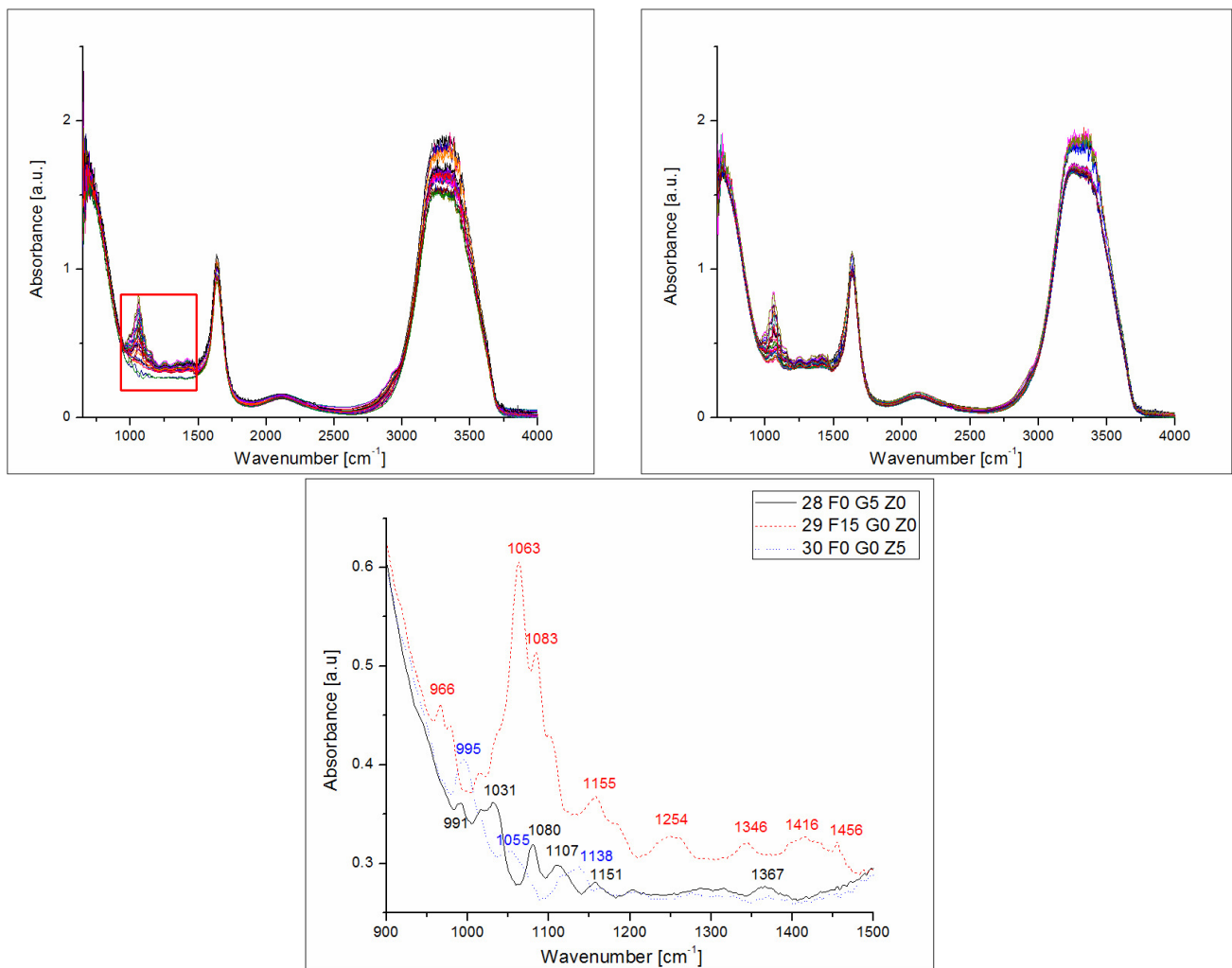


Fig. 1. FT-MIR spectra for: (A) sugars standard solutions mixes; (B) apple juices obtained from fruits of three apple cultivars harvested during development; (C) spectral zone 900-1500 cm⁻¹ characteristic for sugars - standard solution of 5% glucose (STD28 - continuous line) 15% fructose (STD29 - dotted line) and 5% sucrose (STD30 - interrupted line)

FT-MIR spectra (650-4000 cm⁻¹) of the fruit juices obtained from 'Golden Delicious', 'Starkrimson' and 'Jonathan' apple cultivars harvested during their development (Fig. 1B), showed, similar to spectra of the standard sugar solutions discussed above, three major areas: 2800-3700 cm⁻¹, 1500-1800 cm⁻¹ and 900-1500 cm⁻¹. Previous research showed that the first two areas are

attributed to water molecules, apple juice containing a significant quantity of water (dry matter content of the juice was between 4.75-22%) (Bureau *et al.*, 2009; Ścibisz *et al.*, 2011; Wilkerson *et al.*, 2013). In the current study, the characteristic region for sugars (900-1500 cm⁻¹) indicated the accumulation of sugars in the apple fruits harvested during development. Glucose, fructose

Table 3. FT-MIR prediction of sugars in apple juice obtained from three cultivars fruits harvested during their development - descriptive statistics and PLS calibration models performance for selecting optimal spectral region and pre-treatments

Sugar	Mean	SD	Range	Spectral region [cm ⁻¹]	Pre-treatment	Calibration			Cross validation ¹		
						F	R _{cal} ²	SEC	R _{cval} ²	SECV	RPD
Fructose	4.584	4.09	0.422-14.337	650-4000		2	0.94	0.943	0.94	1.028	3.98
				900-1600		2	0.95	0.86	0.94	0.943	4.34
				900-1500		2	0.95	0.854	0.95	0.932	4.39
				950-1475		2	0.95	0.859	0.95	0.939	4.36
				900-1200		2	0.95	0.858	0.95	0.936	4.37
				900-1500	SNV	2	0.95	0.839	0.94	0.959	4.26
					d1 (7, 2)	1	0.95	0.84	0.95	0.907	4.51
					d2 (9, 2)	1	0.94	0.965	0.93	1.042	3.93
				Glucose	1.598	1.094	0.291-4.068	650-4000		2	0.84
900-1600		3	0.89					0.362	0.84	0.448	2.44
900-1500		3	0.89					0.361	0.84	0.444	2.46
975-1475		3	0.89					0.359	0.84	0.443	2.47
900-1200		3	0.88					0.368	0.84	0.439	2.49
900-1200	SNV	3	0.88					0.372	0.84	0.448	2.44
	d1 (7, 2)	3	0.89					0.361	0.84	0.438	2.50
	d2 (9, 2)	3	0.90					0.345	0.85	0.424	2.58
Sucrose	0.765	1.108	0-4.286					650-4000		3	0.87
				900-1600		2	0.78	0.509	0.75	0.566	1.96
				900-1500		2	0.79	0.505	0.75	0.565	1.96
				975-1475		2	0.79	0.504	0.75	0.566	1.96
				900-1200		2	0.79	0.514	0.75	0.562	1.97
				900-1200	SNV	1	0.78	0.514	0.75	0.561	1.98
					d1 (7, 2)	1	0.78	0.518	0.74	0.569	1.95
					d2 (9, 2)	1	0.72	0.575	0.69	0.626	1.77

¹Full cross validation; SD – standard deviation; F – number of latent variables; R_{cal}², R_{cval}² – coefficient of determination for calibration and cross validations; SEC – standard error of calibration; SECV – standard error of cross-validation; RPD – ratio of performance of deviation; SNV - Standard Normal Variate; d1 / d2 – first / second derivative

Table 4. External validation sets assessment - descriptive statistics and performance of PLS calibration models for the prediction of sugars in apple juice obtained from three cultivars fruits harvested during their development

Set	Description	N	Calibration						External validation									
			Sugar	Mean	SD	Range	F	R _{cal} ²	SEC	Set	N	Sugar	Mean	SD	Interval	R _p ²	SEP	RPD _p
C1	Randomly selected	24	Fructose ^{a1}	467	419	0.42-14.33	1	0.95	0.84	V1	12	Fructose	441	404	0.54-12.13	0.94	0.82	4.90
			Glucose ^{a2}	1.57	1.04	0.29-4.06	3	0.91	0.31			Glucose	1.64	1.23	0.29-3.70	0.84	0.47	2.61
			Sucrose ^{a3}	0.83	1.20	0-4.28	1	0.78	0.55			Sucrose	0.63	0.91	0-2.78	0.70	0.43	2.08
STD	mix of FG and Z	30	Fructose ^{a1}	7.04	5.09	0-15	2	0.98	0.65	V2	36	Fructose	458	409	0.42-14.33	0.78	0.99	4.13
			Glucose ^{a2}	2.52	1.58	0-5	2	0.95	0.35			Glucose	1.59	1.09	0.29-4.06	0.2	0.51	2.10
			Sucrose ^{a3}	2.41	1.73	0-5	3	0.94	0.39			Sucrose	0.76	1.10	0-4.28	*	1.77	0.62
C2	STD+ Randomly selected	54	Fructose ^{a1}	5.98	4.82	0-15	2	0.96	0.92	V1	12	Fructose	441	404	0.54-12.13	0.88	0.94	4.28
			Glucose ^{a2}	2.09	1.44	0-5	2	0.86	0.53			Glucose	1.64	1.23	0.29-3.70	0.74	0.50	2.42
			Sucrose ^{a3}	1.71	1.70	0-5	3	0.71	0.91			Sucrose	0.63	0.91	0-2.78	0.31	0.73	1.23

N – number of samples for each set; ^{a1}Spectral region 900-1500 cm⁻¹; Pre-treatment: d1 (7, 2); ^{a2}Spectral region 900-1200 cm⁻¹; Pre-treatment: d2 (9, 2); ^{a3}Spectral region 900-1200 cm⁻¹; Pre-treatment: SNV; R_p² – coefficient of determination for external validation; all other abbreviations are similar with Table 3.

and sucrose solutions showed intense and characteristic absorption bands in this area. The bands in the region 1200-1500 cm⁻¹ were assigned to –CH₂ deformation and angular deformation of C-C-H and H-C-O (Bureau *et al.*, 2009). Absorption bands in 950-1200 cm⁻¹ region were explained by stretching peaks of C-C and C-O (Bureau *et al.*, 2009). The differences between juice samples obtained from apples harvested during fruit development were observed, especially in 900-1500 cm⁻¹ region. Baseline shift and multiplicative interferences were corrected using mathematical treatments (standard normal variate (SNV), first and second derivatives).

Optimal spectral region and pre-treatments selection

To find the calibration model with the best predictive power, for each sugar, several spectral regions were tested on the set of samples (N=36) obtained from apples harvested during their development (Table 3). The set contained a wide variety of

samples, showing a high variability in the concentrations of sugars due to different harvesting periods (from seven to 144 DAFB), and also because of different cultivars studied. By obtaining the samples in this manner, we were provided with a set of samples showing a wide range of apple juice concentrations for fructose (0.42-14.33%), glucose (0.29-4.06%) and sucrose (0-4.28%). Moreover, it provided originality by creating spectra databases of apple juice with low concentrations of sugars (genuine / without running dilutions). So far, the studies were based on the analysis of apple juice obtained from fruits harvested at technological and / or consumption maturity (Irudayaraj and Tewari, 2003; Leopold, 2010). Checking the feasibility of the PLS calibration models represented the first step in the development of rapid methods (FT-MIR type) for analyzing apples growth dynamics.

For each type of sugar, the important wave numbers range for achieving performant PLS calibration models were proposed

based on regression coefficients developed for the full spectrum (650-4000 cm^{-1}). Wave numbers that are characterized by high absolute regression coefficients have a higher impact on Y variables (reference values), indicating their importance in the selection of calibration models development (Ścibisz et al., 2011). On the other hand, "noise" dominated regions of the regression coefficients must be removed from the model. As shown in Table 3, the highest values for R_{cal}^2 , R_{val}^2 and RPD and the lowest values for SEC and SECV, were obtained when using 900-1500 cm^{-1} spectral region for fructose, 900-1200 cm^{-1} for glucose and 900-1200 cm^{-1} for sucrose. For each type of sugar, these regions were further used to select the optimal pre-treatment of the spectral data (Table 3). SNV pre-treatment, Savitzky Golay first derivative seven smoothing points, polynomial order 2 -d1 (7, 2) and Savitzky Golay second derivative 2-d1(7,2) were used to increase the RPD values. It was concluded that the optimal pre-treatments of the spectra were d1 (7, 2) for fructose, d2 (9, 2) for glucose and SNV for sucrose.

Selecting the optimal spectral region and pre-treatments led to PLS calibration models with very good performance for fructose ($R_{\text{val}}^2 = 0.95$; $\text{SECV} = 0.907$) and acceptable for glucose ($R_{\text{val}}^2 = 0.85$; $\text{SECV} = 0.424$). The best calibration models developed for sucrose showed only satisfactory performance ($R_{\text{val}}^2 = 0.75$; $\text{SECV} = 0.561$), irrespective of regions or pre-treatments used.

External validation

For each type of sugar, the performance of previously best developed PLS models was tested by performing external validation sets. By manual random selection and targeting a representative distribution of all types of samples and a wide range of concentrations, two calibration sets were formed: the set C1 consisted of 24 samples (66.6%), and an external validation set (V1) consisting of 12 samples (33.3%). As shown in Table 4, the mean, SD and the sugars concentrations range of the calibration set ($N = 24$) were very close to those of the validation set ($N = 12$). To verify the feasibility of the calibration models obtained from the 30 mixes of sugars standard solution (STD set; $N = 30$), the V2 validation set ($N = 36$) was used, consisting of all samples of apple juice obtained from fruits harvested during their development. For each type of sugar, the FT-MIR predicted results were compared with the ones determined by the direct HPLC method. Means, SD and concentration ranges of STD and V2 sets are shown in Table 4. To verify the hypothesis that a set containing both the mixes of standard solution and the authentic apple juice would have a higher predictive capacity, it was decided to use the C2 ($N = 54$) calibration set, consisting of the 30 mixes of standard solutions and the C1 set ($N = 24$). Model validation was performed by using the V1 set, consisting of 12 samples of authentic juice, other than the 24 already used in the calibration set (Table 4).

Table 4 shows a synthesis of results for PLS calibration models developed for C1, C2 and STD sets, and their performance to predict V1 and V2 external validation sets. One can observe the high accuracy of PLS calibration models developed for spectra of the standard solution mixes (set STD, Table 4). For each of the three sugars, the models obtained for STD set revealed coefficients of determination very close to unit (0.98 for fructose, 0.95 for glucose and 0.94 for sucrose).

Very good results were obtained for the C1 set used to develop calibration models for fructose and glucose (Table 4), registering an R_{cal}^2 of 0.95 for fructose and 0.91 for glucose. On the other hand, while the results obtained for the STD set had a

high performance in the case of sucrose the ones obtained for C1 and C2 sets were poor (R_{cal}^2 of 0.78 for the C1 set and 0.71 for C2). This can be explained by the high number of samples with zero concentration or very close to it, at the initial sampling times ($T1 \rightarrow T3$, Table 1). Actually, none of the sets of calibration developed for sucrose could be successfully used to predict concentrations of sucrose for the external validation sets (V1-V2, Table 4).

For fructose, the best results attained in predicting the external sets (V1-V2), were obtained when the model was based on the spectra of the authentic apple juice (set C1). Thus, for the fructose, by using the PLS calibration model developed on C1 set, for the V1 validation set, R_p^2 of 0.94 and SEP of 0.826 were obtained (Table 4). Satisfactory performances were obtained (R_p^2 of 0.84 and SEP of 0.472; Table 4) when predicting the glucose concentrations of V1 set and using the C1 set calibration model. For all sugars studied when calibration model was based on the spectra of standard solution there was a drastic reduction of their ability to predict sugar concentrations of apple juice obtained from fruits harvested during development (Table 4, External validation section, V2 set for each sugar). When authentic juice samples were included on the calibration in addition to standard solution samples (C2 set), an improvement of the results was noticed (Table 4, External validation section, for each sugar's V1 set for PLS calibration models developed with the C2 set). These results reinforced the hypothesis that it is desirable to include in the calibration sets similar samples with the ones to be predicted.

Conclusions

The selection of optimal spectral region and pre-treatments resulted in PLS calibration models had very good performance in the case of fructose and acceptable for glucose. The sucrose calibration models showed a satisfactory performance, although different spectral regions and pre-treatments were used. In the case of fructose and glucose, the best results attained for predicting external sets were obtained when the model was based on the spectra of authentic apple juice. None of the calibration sets developed for sucrose could be successfully applied to predict the sucrose concentration for external validation sets.

Acknowledgements

Romina Alina Vlaic and Vlad Mureşan thanks for support from European Social Fund, Human Resources Development Operational Programme, project no. POSDRU/159/1.5/S/132765.

References

- Andrade JM, Gomez-Carracedo MP, Fernandez E, Elbergali A, Kubista M, Prada D (2003). Classification of commercial apple beverages using a minimum set of Mid-IR wavenumbers selected by Procrustes rotation. *Analyst* 128:1193-1199.
- Berüter J, Feusi MES (1997). The effect of girdling on carbohydrate partitioning in the growing apple fruit. *J Plant Physiol* 151:277-285.
- Beghi R, Spinardi A, Bodria L, Mignani I, Guidetti R (2013). Apples nutraceutical properties evaluation through a visible and near-infrared portable system. *Food Bioprocess Tech* 6:2547-2554.

- Bonta V, Marghitas LA, Stanciu O, Laslo L, Dezmiorean D, Bobis O (2008). High-performance liquid chromatographic analysis of sugars in Transylvanian honeydew honey. *Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca. Animal Science and Biotechnologies* 65:229-232.
- Borsanie J, Budde CO, Porrini L, Lauxmann MA, Lombardo VA, Murray R (2009). Carbon metabolism of peach fruit after harvest: Changes in enzymes involved in organic acid and sugar level modifications. *J Exp Bot* 60(15):1823-1837.
- Bureau S, Ruiz D, Reich M, Gouble B, Bertrand D, Audergon JM, Renard CMGC (2009). Application of ATR-FTIR for a rapid and simultaneous determination of sugars and organic acids in apricot fruit. *Food Chem* 115:1133-1140.
- Cozzolino D, Murray I, Chree A, Scaife JR (2005). Multivariate determination of free fatty acids and moisture in fish oils by partial least-squares regression and near-infrared spectroscopy. *Lwt-Food Sci Technol* 38:821-828.
- Chourey PS, Beruter J (1985). Sugar accumulation and changes in the activities of related enzymes during development of the apple fruit. *J Plant Physiol* 121:331-341.
- Duarte IF, Barros A, Delgado I, Almeida C, Gil AM (2002). Application of FTIR spectroscopy for quantification of sugars in mango juice as a function of ripening. *J Agr Food Chem* 50(11):3104-3111.
- Fuleki T, Pelayo E, Palabay RB (1994). Sugar composition of varietal juices produced from fresh and stored apples. *J Agr Food Chem* 42:1266-1275.
- Giovanelli G, Sinelli N, Beghi R, Guidetti R, Casiraghi E (2014). NIR spectroscopy for the optimization of postharvest apple management. *Postharvest Biol Tec* 87:13-20.
- Hecke K, Herbinger K, Veberič R, Trobec M, Toplak H, Štampar F (2006). Sugar-, acid- and phenol contents in apple cultivars from organic and integrated fruit cultivation. *Eur J Clin Nutr* 60:1136-1140.
- Irudayaraj J, Tewari J (2003). Simultaneous monitoring of organic acids and sugars in fresh and processed apple juice by Fourier transform infrared attenuated total reflection spectroscopy. *Appl Spectrosc* 57(12):1599-1604.
- Kelly DJF, Gerard D (2005). Detection of sugar adulterants in apple juice using Fourier transform infrared spectroscopy and chemometrics. *J Agr Food Chem* 53(9):3281-3286.
- Leopold L (2010). HPLC and IR/RAMAN markers for oils, juices and cereals authenticity. USAMV Cluj-Napoca, PhD thesis.
- Miller NJ, Rice-Evans CA (1997). The relative contributions of ascorbic acid and phenolic antioxidants to the total antioxidant activity of orange and apple fruit juices and blackcurrant drink. *Food Chem* 60(3):331-337.
- Mingjun L, Engjuan F, Lailiang C (2012). Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. *PLoS One* 7:3-e33055.
- Pissard A, Pierna JAF, Baeten V, Sinnave G, Lognay G, Mouteau A, Dupont P, Alain R, Lateur M (2013). Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near-infrared spectroscopy. *J Sci Food Agr* 93:238-244.
- Petkovsek MM, Stampar F, Veberic R (2007). Parameters of inner quality of the apple scab resistant and susceptible apple cultivars (*Malus domestica* Borkh). *Sci Hortic-Amsterdam* 114:37-44.
- Ruan Y-L, Jin Y, Yang Y-J, Li G-J, Boyer J (2010). Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Mol Plant* 3:942-955.
- Rolland F, Baena-Gonzalez E, Sheen J (2006). Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu Rev Plant Biol* 57:675-709.
- Ścibisz I, Reich M, Bureau S, Gouble B, Causse M, Bertrand D, Renard CMGC (2011). Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. *Food Chem* 125:1390-1397.
- Sivakesava S, Irudayaraj JMK, Korach RL (2001). Detection of adulteration in apple juice using mid infrared spectroscopy. *Appl Eng Agric* 17:815-820.
- Teo G, Suzuki Y, Uratsu SL, Lampinen B, Ormonde N, Hu WK (2006). Silencing leaf sorbitol synthesis alters long-distance partitioning and apple fruit quality. *P Natl Acad Sci Usa* 103:18842-18847.
- Yamaki S, Ino M (1992). Alteration of cellular compartmentation and membrane permeability to sugars in immature and mature apple fruit. *J Am Soc Hortic Sci* 117:951-954.
- Zhang Y, Li P, Cheng L (2010). Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. *Food Chem* 123:1013-1018.
- Zheng HZ, Young-Il K, Shin-Kyo C (2012). A profile of physicochemical and antioxidant changes during fruit growth for the utilisation of unripe apples. *Food Chem* 131:106-110.
- Wilkerson ED, Gordon EA, Diane MB, Sayajon GFG, Santos AM, Rodriguez-Saon LE (2013). Rapid assessment of quality parameters in processing tomatoes using hand-held and benchtop infrared spectrometers and multivariate analysis. *J Agr Food Chem* 61:2088-2095.