Effect of cultivar and season on the robustness of PLS models for soluble solid content prediction in apricots using FT-NIRS

İbrahim Sani Özdemir¹ · Sylvie Bureau² · Bülent Öztürk³ · Ferda Seyhan¹ · Hatice Aksoy¹

Revised: 12 October 2018 / Accepted: 6 November 2018 / Published online: 10 December 2018
© Association of Food Scientists & Technologists (India) 2018

Abstract FT-NIR models were developed for the non-destructive prediction of soluble solid content (SSC), titratable acidity (TA), firmness and weight of two commercially important apricot cultivars, “Hacıhaliloğlu” and “Kabaası” from Turkey. The models constructed for SSC prediction gave good results. We could also establish a model which can be used for rough estimation of the apricot weight. However, it could not be possible to predict accurately TA and firmness of the apricots with FT-NIR spectroscopy. The study was further extended over 3 years for the SSC prediction. Validation of the both mono and multi-cultivar models showed that model performances may exhibit important variations across different harvest seasons. The robustness of the models was improved when the data of two or three seasons were used. It was concluded that in order to develop reliable SSC prediction models for apricots the spectral data should be collected over several harvest seasons.

Keywords *Prunus armeniaca* L. · FT-NIR · PLS-R · Soluble solid content

Introduction

Turkey is the world’s leading sulphured-dried apricot producer with its 75% global market share (www.faostat.fao.org). Approximately 51% of the fresh apricots of the whole country are produced in Malatya province located in the eastern Turkey, 90% of which is used for sulphured-dried apricot production (Akin et al. 2008).

As a common practice in Malatya province, all the apricots on a tree are harvested on the same day in order to save from the labour cost. Since not all fruit on a tree ripen at the same pace, this practice results in apricot batches of extreme heterogeneity in terms of physical aspects and biochemical composition. This is a serious concern for sulphured-dried apricot producers. Because, apart from process related parameters such as temperature, exposure time and SO₂ concentration, sulphur absorption of the apricots is also affected by the physical and biochemical properties of the fruit such as size, skin thickness, fruit integrity and the biochemical composition (Joslyn and Braverman 1954; Gökçe 1966; Rosselló et al. 1993). It has been reported that SO₂ makes adducts by reversible action with reducing sugars, acetaldehyde, quinones, ketoacids, anthocyanins, and with cysteine residues in proteins (Adachi et al. 1979; Guido 2016). Since the concentration of these compounds exhibit important differences between maturity stages and cultivars, processing of apricots of mixed maturity and/or cultivar may result in important variations in the final SO₂ content of the dried apricots (Gökçe 1966). It was reported that mature apricots with high soluble solid content absorb more SO₂ than immature fruit (Barrett 2002). Türkylmaz et al. (2014) reported that Hacıhaliloğlu cultivar absorbs more SO₂ than Kabaası under identical process conditions and authors attributed this difference also to high soluble solid content of the

✉ İbrahim Sani Özdemir
ibrahim.ozdemir@tubitak.gov.tr;
ibrahim.ozdemir@yahoo.com

¹ TUBITAK Marmara Research Center, Food Institute,
P.O. Box 21, 41470 Gebze, Kocaeli, Turkey

² UMR-A-408, SQPOV INRA, Domaine St Paul, Site
Agroparc, 84914 Montfavet, Avignon Cedex 9, France

³ Apricot Research Institute, 44110 Malatya, Turkey

Hacıhaliloğlu cultivar. Therefore, in order to produce sulphured-dried apricots with uniform SO₂ content it is crucial to sort apricots according to their maturity stage before sulphuration and drying process.

In apricot, sugar content increases constantly during the course of fruit development and ripening. Therefore, among various quality attributes of apricot fruit, soluble solid content (SSC) is the one which correlates the most with the fruit maturity stage (Crisosto 1994; Rossier et al. 2001). In general, SSC of fruits is measured with a refractometer which quantifies the soluble solids content by comparing the refractive index of the fruit juice with that of a sucrose solution and expressed in °Brix (1 °Brix equals to 1 g of sucrose in 100 g of solution). The main drawback of this method is it does not measure the soluble solids other than sugars, and thus, it gives only an estimate of the soluble solid content of the fruit. However, since 85% of the soluble solids in a fruit consist of sugars, the associated error is usually neglected in measuring the SSC of fruits with a refractometer (Magwaza and Opara 2015). In addition, refractometers are easy-to-use, affordable and hand-held instruments which make them widely used by the fruit producers for quality control.

It was previously reported that soluble solid content of apricots could be effectively estimated by Fourier Transform Near Infrared (FT-NIR) spectroscopy based Partial Least Squares Regression (PLS-R) models with acceptable accuracy (Carlini et al. 2000; Bureau et al. 2009; Camps and Christen 2009; Berardinelli et al. 2010; Buyukcan and Kavdir 2016). In this respect, as a non-destructive and rapid analysis method, FT-NIR spectroscopy might serve as a solution for sorting apricot fruits on the basis of soluble solid content. However, in order to develop effective grading lines for practical applications, developed models should be robust enough to predict SSC of the fruit of different cultivars, maturity stages and seasons (Nicolai et al. 2007). The results of the studies carried out on peaches (Peiris et al. 1998), apples (Peirs et al. 2003; Bobelyn et al. 2010), mandarins (Guthrie et al. 2005), avocado (Wedding et al. 2013), Japanese plums (Louw and Theron 2010) and mango (Rungpichayapichet et al. 2016) showed that the prediction accuracy of the FT-NIR models can exhibit considerable seasonal variations. In all these quoted studies improvement in model robustness was observed when spectral data of several seasons were incorporated in the calibration model. It was also reported that the extent of the seasonal variations depend on the cultivar. For apples (Bobelyn et al. 2010) it was found that mono-cultivar models constructed with the data of several seasons resulted in more robust models than multi-cultivar models whereas for Japanese plums (Louw and Theron

2010) multi-cultivar models for SSC resulted in equal or better precision than mono-cultivar models.

Therefore, the major goal of the present study was to assess the potential of FT-NIR spectroscopy in predicting SSC of two commercially important apricot cultivars, Hacıhaliloğlu and Kabaası, over three successive seasons with an ultimate aim of providing data for the future applications of commercial grading systems. Besides, in the first year of the study, apart from SSC, we also investigated the potential of using FT-NIR spectroscopy for the rapid non-destructive prediction of titratable acidity (TA), firmness and weight of the both cultivars.

Materials and methods

Apricot samples

The study was carried out on two apricot cultivars, namely, Hacıhaliloğlu and Kabaası which are the most prevalent cultivars used for sulphured dried apricots in Malatya province (Akin et al. 2008). In order to obtain robust prediction models which can be used in real life situation, the biological variance was maximized by collecting apricot samples from different orchards in Malatya province. Samples were harvested manually from the beginning of green mature stage to the overripe stage at 2 or 3 different campaigns with 3–5 days intervals depending on the cultivar and harvesting season.

In 2014, 118 fruits of ‘Hacıhaliloğlu’ and 120 fruits of ‘Kabaası’ cultivar consisting of different maturity stages were collected. After harvest, apricot samples were transported to laboratory where FT-NIR spectra were acquired immediately. Afterwards weight and firmness of the samples were measured. Then the samples were chopped into small cubes and frozen with liquid nitrogen and stored at – 86 °C. Samples were then thawed at 4 °C before analysing SSC and TA of the fruits. In 2015 and 2016, only the SSC of the fruits were measured after spectral acquisition. In 2015, 160 and 146, and in 2016, 85 and 90, fruits of Hacıhaliloğlu and Kabaası cultivars were collected, respectively.

Spectral measurements

The FT-NIR spectra of the fresh apricot samples were acquired by a FT-NIR spectrometer (MPA, Bruker Optics) in diffuse reflectance mode from 12,000 (833 nm) to 4000 (2500 nm) cm^{–1} with a resolution of 8 cm^{–1}. The spectra were the average of 32 scans. One spectrum from each fruit was acquired on the un-blushed side of the fruit.

Quality parameters

The SSC of the fruits were measured according to AOAC-932.12 (1998) standard method by placing small amount of puree on a hand-held refractometer (Atago, Tokyo, Japan). Fruit firmness was determined as described by Bureau et al. (2009) by measuring the maximum force (N) required to compress the fruit by 3% of fruit height with a texture analyser (TA-Xt, Stable Microsystems, UK) equipped with 50 kg load cell. For the TA measurement, according to AOAC-942.15 (1995) standard method, apricot samples were grounded into puree with a ball grinder (Retsch) and 10 g of aliquot was titrated against 0.1 M NaOH to a pH end-point of 8.2 manually by adding NaOH dropwise and was expressed as % malic acid.

Multivariate statistical analyses

The spectral data was first subjected to principal component analysis (PCA) by using the Unscrambler-X software (version 10.3, CAMO, Oslo, Norway) in order to assess whether the spectral data tend to cluster on the basis of cultivar, range of soluble solid content and harvest season. The spectral range between 12,000 and 4000 cm^{-1} was used for the analysis. The raw spectra were processed by standard normal variate (SNV) method in order to eliminate any non-informative variance due to additive and multiplicative scattering effect.

Spectral data and reference data were then used to establish calibration models to predict weight, TA, firmness and SSC of apricots. The PLS-R models were established by using the OPUS software (version 7.2, Bruker Optik GmbH, Ettlingen, Germany). The “optimize” function of the software enabled us to select the best combination of model parameters [informative wavelength regions, number of latent variables (LVs) and pre-treatment method] by comparing the root mean square error of cross validation (RMSECV) from the list of proposed models established by the software by using multitude of commonly used pre-processing methods and various spectral regions. In order to better compare the effect of cultivar and seasonal variations on the SSC model performances, same model parameters (spectral region of 1063–1334, 1639–1836, 2173–2355 nm; 7 LVs) and pre-treatment method (SNV) were applied to spectral data sets which contain either the data of individual season and cultivar or their dual or threesome combinations.

The calibration and validation models for weight, TA and firmness were developed by using the data set of 2014. Both mono and multi-cultivar models were established. In order to carry out a validation test, the data set was randomly split into two subsets with two-thirds of the fruits used for calibration and remaining one-third used for

validation. Regarding SSC, first of all, mono and multi-cultivar calibration models were established for each of the three harvest seasons. These models were validated first by internal test set corresponding to the one-third of the data and then by an external data set belonging to another season. Then, the data of two different seasons were combined and established models were validated by the data of the remaining season. Finally, the data of all three seasons were pooled to establish global mono and multi-cultivar models which were validated internally by using the one-third of the whole data set. The model performances were assessed on the basis of the coefficient of determination of the calibration (R_c^2) and prediction (R_p^2) models, root mean squared errors of calibration (RMSEC) and prediction (RMSEP) and residual prediction deviation (RPD) (Nicolăi et al. 2007).

Results and discussion

A representative set of FT-NIR spectra of apricot samples from the present study is given in Fig. 1. With no surprise the spectral profiles of Turkish apricot cultivars used in this study are similar to those previously reported for French (Bureau et al. 2009), Italian (Berardinelli et al. 2010) and Swiss (Camps and Christen 2009) apricot cultivars. FT-NIR spectra of the apricots, and fruit in general, are dominated by the absorbance bands of water due to OH stretching overtones at 970, 1190, 1450 and 1940 (Polesello et al. 1983). The peak at 1790 nm corresponds to C–H overtone region and was reported to be sugar related (Louw and Theron 2010).

In the present study, in order to establish robust calibration models we worked on a large data set comprising fruits of two cultivars with wide range of maturity stages which were harvested over three successive seasons. Therefore, as a first attempt we carried out PCA on the whole spectral data set ($n = 718$) in order to check whether apricots tend to group or not on the basis of SSC (as an indicator of maturity stage), season and the type of cultivar. The scores and correlation loadings plots of the PCA are given in Fig. 2a–d. As can be seen from these figures, apricots can be discriminated on the basis of SSC (Fig. 2a) and cultivars (Fig. 2c). However, apricots of 2014, 2015 and 2016 seasons completely overlapped on the basis of their FT-NIR profile (Fig. 2b). Apricot cultivars (Fig. 2c) were clustered mainly along the first (PC1) and second principal component (PC2) axis which accounts for the 87% of the total variance. On the other hand, discrimination according to SSC occurred mainly along the third principal component (PC3) which explains only the 8% of the total spectral variance. It should be emphasised that in PCA the PCs which explain small portions of the total

Fig. 1 Average FT-NIR spectra acquired from “Kabaası” (n = 367) and “Hacıhaliloğlu” (n = 351) apricots

variance are generally overlooked. But as seen in this study, the PC3 explains only 8% of the total variance and still contains critical information related to soluble solid content variation in apricots which should not be omitted. As can be seen from Fig. 2d, 963, 1062, 1390, 1890 and 2300 nm are the major wavelengths which correlate with PC3. Among these wavelengths, the first four of them are known to be related to OH stretching overtones from water (Polesello et al. 1983) and the one at 2300 falls within the combinations region associated with the C–H and C–H-combinations grouping (Louw and Theron 2010).

In the first year of the study, we mainly focused on evaluating the potential of using FT-NIR spectroscopy technique for predicting the SSC, TA, weight and firmness of the apricot cultivars. The descriptive statistics results for the reference data for each quality trait used for PLS-R models are given in Table 1. In order to validate the calibration models, the whole data set was randomly split into two subsets with two-thirds of the fruits used for calibration and remaining one-third of the fruits were used for validation. We established both mono and multi-cultivar prediction models, in order to better evaluate the effect of biological variance arising from the type of cultivar on the prediction ability of the models. As it can be seen from Table 2, among all the quality traits tested the highest correlation was obtained between the FT-NIR spectra and the measured SSC values for mono and multi-cultivar models. The prediction ability of the models was largely influenced by the cultivar. The mono cultivar SSC models established for ‘Hacıhaliloğlu’ apricots gave better results ($R_p^2 = 0.913$, RMSEP = 1.43, RPD = 3.39) than that of ‘Kabaası’ ($R_p^2 = 0.827$, RMSEP = 1.85, RPD = 2.42). The multi-cultivar model presented $R_p^2 = 0.874$ and RMSEP = 1.72 values which fell between the values for two mono-

cultivar models. These results were better than those previously reported for French (Bureau et al. 2009), Italian (Berardinelli et al. 2010), Turkish (Buyukcan and Kavdir 2016) and Swiss (Camps and Christen 2009) apricots for which the reported model parameters were $R_p^2 = 0.85$ and RMSEP = 0.98, $R_p^2 = 0.72$ and RMSEP = 0.80, $R_p^2 = 0.75$ and RMSEP = 1.51 and $R_p^2 = 0.81$, RMSECV = 1.10, respectively. This difference between the prediction results can be attributed in part to the differences in the range of SSC of the apricots used. Indeed, the SSC range of the apricots used in our study (14.2–35.2 °Brix) was considerably wider than that of above quoted studies which were in the range of 7.3–21.7% Brix, 8–18 °Brix, 9–25 °Brix and 8.7–16.3 °Brix, respectively.

Regarding TA, the prediction ability of the mono-cultivar models were very poor with $R_p^2 = 0.292$, RMSEP = 1.01, RPD = 1.19 for ‘Hacıhaliloğlu’ and $R_p^2 = 0.349$, RMSEP = 0.85, RPD = 1.34 for ‘Kabaası’ cultivar. However, significant improvement was achieved in model performance when the data for both cultivar was pooled which resulted in $R_p^2 = 0.701$, RMSEP = 0.99, RPD = 1.84 in agreement with the results of the multi-cultivar model developed for the French apricots (Bureau et al. 2009).

Prediction of weight by FT-NIR spectroscopy has not been reported previously for apricots but encouraging results were reported for nectarines (Pérez-Marín et al. 2009) with $R_p^2 = 0.97$ and RMSEP = 6.76 g and for Japanese plums (Louw and Theron 2010) with $R_p^2 = 0.751$ and RMSEP = 12.8 g. In the present study, we obtained poor results for mono-cultivar models for the weight prediction. However, for the multi-cultivar model the results were acceptable with $R_p^2 = 0.826$, RMSEP = 3.54 g and RPD = 2.84 (Table 2). As stated by Nicolai et al. (2007) the models with an RPD value of 2–2.5 can be used for coarse

Fig. 2 Principal component analysis of the FT-NIR spectra acquired from the fruits of two apricot cultivars, Hacıhaliloğlu (HH) and Kabaaşı (KA) during three successive years 2014, 2015 and 2016. In total 718 fruits were analyzed 351 of which belongs to Hacıhaliloğlu and the remaining 367 belongs to Kabaaşı cultivar. The raw spectra were pretreated with SNV transformation procedure before PCA. The scores were colored

according to **a** soluble solid content (SSC) with a range of 3 °Brix between 14 to 29 °Brix, **b** harvest years 2014 (red), 2015 (orange) and 2016 (green), **c** cultivar type, “Hacıhaliloğlu” (HH, green) and “Kabaaşı” (KA, orange). Correlation loadings values of absorbances **d** on the first (PC1, blue solid line), second (PC2, red dotted line) and third (PC3, green dashed line) principal components (color figure online)

classification purposes such as small/large. As indicated by Louw and Theron (2010), weight is a physical parameter which cannot be measured directly by FT-NIR spectroscopy but fruit water content which is known to be correlated with its weight can be directly measured by FT-NIR spectroscopy and linked indirectly to the prediction of

fruit weight. These results are of practical importance since fruit size which is highly correlated with the weight is an important quality criterion in the marketing standard for fresh apricots (European Commission 2000). Moreover, fruits need to be sorted according to their size before processing such as drying or sulphuration otherwise large

Table 1 Descriptive statistics of the calibration and validation data sets used for the developing PLS models for the prediction apricot quality traits of two apricots cultivar harvested over three successive seasons between 2014–2016

Season	Parameter	Cultivar	Calibration					Validation				
			n	Range	Mean	Median	SD	n	Range	Mean	Median	SD
2014	Weight (g)	Hacıhaliloğlu	80	16.0–40.9	25.7	24.7	5.1	40	18.5–34.2	25.7	25.1	4.3
		Kabaaş1	79	21.0–50.9	38.5	39.3	5.9	39	27.1–47.0	39.2	40.5	5.3
		Pooled	158	16.0–50.9	32.0	31.4	8.4	80	16.1–47.0	32.5	31.7	8.5
	TA (meq/100gFW)	Hacıhaliloğlu	80	4.6–11.4	7.2	7.2	1.5	40	5.0–10.2	7.1	6.9	1.2
		Kabaaş1	79	2.4–7.4	4.6	4.6	1	39	2.6–7.0	4.3	4.0	1.1
		Pooled	158	2.4–11.4	5.9	5.6	1.8	80	2.6–10.2	5.7	5.7	1.8
	Firmness (N)	Hacıhaliloğlu	80	2.3–5.6	4.0	4.0	0.7	40	2.7–5.3	4.0	4.1	0.7
		Kabaaş1	79	2.3–7.7	4.5	4.3	1.1	39	2.4–7.6	4.5	4.3	1.1
		Pooled	158	2.3–7.7	4.2	4.1	1.0	80	2.4–7.6	4.3	4.2	1.0
2015	SSC (°Brix)	Hacıhaliloğlu	80	14.2–30.9	20.1	18.1	4.8	40	14.9–29.9	19.9	17.4	4.9
		Kabaaş1	79	14.4–35.2	22.9	23.2	5.2	39	15.1–30.8	22.3	21.9	4.5
		Pooled	158	14.2–35.2	21.5	20.2	5.1	80	14.4–30.8	21.1	19.7	4.9
	SSC (°Brix)	Hacıhaliloğlu	97	11.8–31.8	21.0	20.5	4.6	49	12.2–31.2	20.8	21.2	4.3
		Kabaaş1	107	12.9–30.5	20.3	20.7	4.1	53	12.9–29.0	20.3	21.2	3.8
		Pooled	204	11.8–31.8	20.5	20.5	4.3	102	12.2–31.2	20.7	21.2	4.2
	SSC (°Brix)	Hacıhaliloğlu	56	13.8–29.9	21.8	21.4	4.0	29	14.9–28.8	20.9	19.8	3.5
		Kabaaş1	59	9.0–24.7	16.7	17.2	3.7	30	9.9–23.8	16.5	16.2	3.3
		Pooled	116	9.0–29.9	19.2	19.2	4.6	58	9.7–28.8	18.8	19.2	4.0

In 2015 and 2016 only soluble solid content (SSC) of the apricots were measured. Prediction models were developed on individual cultivars and their combinations (pooled data). Validation data set was constructed by randomly selecting the one-third of the whole data set in concern

Table 2 PLS model performance parameters (LV: number of latent variables, R_c^2 : coefficient of determination of calibration model, R_p^2 : coefficient of determination of validation model, RMSEC: root mean square error of calibration, RMSEP: root mean square error of prediction, RPD: residual prediction deviation) and informative

spectral regions and pre-processing methods (MSC: multiplicative scatter correction, SNV: standard normal variate, MMN: min–max normalization) used for the calibration and validation of the prediction models for weight, soluble solid content (SSC), titratable acidity (TA) and firmness of the apricots

Parameter	Cultivar	Spectral range	Pre-processing	LV	Calibration			Validation		
					R_c^2	RMSEC	RPD	R_p^2	RMSEP	RPD
Weight (g)	Hacıhaliloğlu	1063–1640, 1835–2175	MSC	5	0.687	2.93	1.79	0.657	2.48	1.74
	Kabaaş1				0.555	4.07	1.50	0.363	4.22	1.32
	Pooled				0.791	3.88	2.19	0.826	3.54	2.48
SSC (°Brix)	Hacıhaliloğlu	1063–1334, 1639–1836, 2173–2355	SNV	7	0.989	0.53	9.45	0.913	1.43	3.39
	Kabaaş1				0.979	0.79	6.90	0.827	1.85	2.42
	Pooled				0.954	1.14	4.64	0.874	1.72	2.81
TA (meq 100 gFW ⁻¹)	Hacıhaliloğlu	1063–1836, 2173–2355	MMN	7	0.848	0.60	2.56	0.292	1.01	1.19
	Kabaaş1				0.810	0.47	2.30	0.349	0.85	1.34
	Pooled				0.838	0.74	2.49	0.701	0.99	1.84
Firmness (N)	Hacıhaliloğlu	1063–2355	SNV	7	0.651	0.47	1.69	0.306	0.59	1.21
	Kabaaş1				0.727	0.61	1.92	0.320	0.90	1.22
	Pooled				0.455	0.73	1.35	0.435	0.73	1.33

The spectra acquired from two cultivars, Hacıhaliloğlu (n = 118) and Kabaaş1 (n = 120), in 2014 were used for the model development. One-third of the data randomly selected for test set validation and the remaining spectra used for the model calibration

differences can be observed in final moisture and sulphur content between small and large sized fruits due to differences in their effective diffusion areas. Therefore, our

results indicate that FT-NIR spectroscopy has the potential of being used as an effective tool for grading the apricots according to their sizes such as small/large.

Table 3 PLS model performance parameters (LV: number of latent variables, R^2 : coefficient of determination, RMSEC: root mean square error of calibration, RMSEP: root mean square error of prediction, RPD: residual prediction deviation) of the calibration and validation models for the soluble solid content (SSC) of the Hacıhaliloğlu (HH) and Kabaası (KA) apricots harvested over three seasons 2014–2016

Year	Cultivar	Validation set	LV	Calibration			Validation		
				R_c^2	RMSEC	RPD	R_p^2	RMSEP	RPD
2014	HH	TSV (1/3)	7	0.989	0.53	9.45	0.913	1.43	3.39
				0.984	0.62	7.92	0.549	3.02	2.03
				0.984	0.63	7.92	0.430	2.86	1.65
	KA	TSV (1/3)	7	0.979	0.79	6.90	0.827	1.85	2.42
				0.972	0.85	6.00	0.776	1.88	2.13
				0.972	0.85	6.00	– 0.09	3.69	1.01
	Pooled	TSV (1/3)	7	0.954	1.14	4.64	0.874	1.72	2.81
				0.947	1.18	4.33	0.567	2.79	1.79
				0.947	1.18	4.33	0.152	4.06	1.37
	HH	TSV (1/3)	7	0.967	0.88	5.48	0.893	1.40	3.19
				0.954	0.99	4.66	0.593	3.06	2.73
				0.954	0.99	4.66	0.832	1.55	2.46
		TSV (1/3)	7	0.952	0.92	4.59	0.843	1.49	2.53
				0.916	1.18	3.46	0.781	2.31	2.55
				0.916	1.18	3.46	0.112	3.32	1.07
	Pooled	TSV (1/3)	7	0.922	1.21	3.58	0.868	1.54	2.84
				0.909	1.30	3.32	0.621	3.10	2.51
				0.909	1.30	3.32	0.586	2.83	1.62
2016	HH	TSV (1/3)	7	0.996	0.27	15.4	0.661	2.00	1.78
				0.975	0.63	6.35	0.533	3.27	2.51
				0.975	0.63	6.35	0.294	3.78	1.89
	KA	TSV (1/3)	7	0.943	0.94	4.19	0.130	3.06	1.07
				0.856	1.39	2.65	0.312	4.09	1.22
				0.856	1.39	2.65	– 0.046	4.07	1.16
	Pooled	TSV (1/3)	7	0.894	1.55	3.06	0.585	2.59	1.58
				0.825	1.88	2.39	0.662	2.93	1.92
				0.825	1.88	2.39	0.510	2.97	1.52

The model validation was carried out either on the internal test set corresponding to the one-third of the whole data (TSV 1/3) or on the data set of another season. Models were developed both for individual cultivars and their combination (pooled data). All the spectral data was pre-treated with standard normal variate (SNV) method and the spectral region of 1063–1334, 1639–1836, 2173–2355 nm were used for model development

Regarding firmness, both mono and multi-cultivar models performed weakly ($R_p^2 = 0.306$ – 0.435 ; RMSEP = 0.59 – 0.90 N) which is in accordance with the results reported for French apricots ($R_p^2 = 0.54$; RMSEP 60.3 kPa) (Bureau et al. 2009). On the other hand, for ‘Tokaloğlu’ cultivar from Turkey better results were reported ($R^2 = 0.80$; RMSECV = 4.68 N) (Buyukcan and Kavdir 2016). It should be noted that fruit firmness was measured by penetration method in the study of Buyukcan and Kavdir (2016) whereas in the present study and in that of Bureau et al. (2009) the apricot firmness was measured by compression method. Therefore, the differences in the methodology used for the firmness measurements in both studies might explain in part the difference in the model performances for the apricot firmness.

Apart from testing the suitability of FT-NIR spectroscopy for the prediction of different quality traits, the

effect of seasonal variations on the mono and multi-cultivar prediction models for SSC was also studied. For this purpose, we established calibration models for apricots harvested over three successive seasons 2014, 2015 and 2016. For the models of individual seasons model validation was made either with an internal test set (TSV) corresponding to the one-third of data [TSV (1/3)] or with the data set of another season. As seen from Table 3 model performances were highly dependent on both cultivar and season. The prediction model which was developed only with the data of 2014 gave very good results for TSV (1/3) validation for ‘Hacıhaliloğlu’ ($R_p^2 = 0.913$, RMSEP = 1.43 , RPD = 3.39). However, for the same cultivar considerably lower R_p^2 and higher RMSEP values were obtained in 2015 ($R_p^2 = 0.893$, RMSEP = 1.40 , RPD = 3.19) and 2016 ($R_p^2 = 0.661$, RMSEP = 2.00 , RPD = 1.78). Seasonal variations in model performances were also observed for

Table 4 PLS model performance parameters (LV: number of latent variables, R^2 : coefficient of determination, RMSEC: root mean square error of calibration, RMSEP: root mean square error of prediction, RPD: residual prediction deviation) of the calibration and validation models for the soluble solid content (SSC) of the Hacıhaliloğlu (HH) and Kabaası (KA) apricots harvested over three seasons 2014–2016

Season	Cultivar	Validation set	LV	Calibration			Validation		
				R_c^2	RMSEC	RPD	R_p^2	RMSEP	RPD
2014	HH	TSV (1/3)	7	0.929	1.28	3.76	0.881	1.58	2.91
2015		2016		0.918	1.36	3.49	0.798	1.71	2.23
	KA	TSV (1/3)	7	0.921	1.33	3.57	0.825	1.83	2.48
		2016		0.914	1.36	3.41	0.129	3.29	1.07
	Pooled	TSV (1/3)	7	0.895	1.53	3.09	0.861	1.68	2.68
		2016		0.890	1.54	3.02	0.667	2.54	1.74
2014	HH	TSV (1/3)	7	0.962	0.91	5.13	0.845	1.71	2.54
2016		2015		0.930	1.20	3.78	0.735	2.32	2.37
	KA	TSV (1/3)	7	0.892	1.85	3.04	0.745	2.50	1.99
		2015		0.874	1.93	2.81	0.431	3.00	1.39
	Pooled	TSV (1/3)	7	0.858	1.89	2.66	0.807	2.14	2.29
		2015		0.853	1.91	2.61	0.759	2.08	2.06
2015	HH	TSV (1/3)	7	0.936	1.13	3.94	0.892	1.36	3.20
2016		2014		0.931	1.14	3.82	0.624	2.94	2.70
	KA	TSV (1/3)	7	0.848	1.71	2.57	0.694	2.24	1.81
		2014		0.811	1.86	2.30	0.713	2.64	2.40
	Pooled	TSV (1/3)	7	0.854	1.71	2.61	0.777	2.01	2.12
		2014		0.830	1.81	2.43	0.836	2.04	2.52
2014	HH	TSV (1/3)	7	0.915	1.34	3.42	0.881	1.53	2.93
2015	KA	TSV (1/3)	7	0.853	1.90	2.61	0.777	2.17	2.14
2016	Pooled	TSV (1/3)	7	0.841	1.88	2.51	0.830	1.90	2.43

The calibration models were built on the different dual combinations of data from different harvest seasons and the model validation was carried out either on the internal test set corresponding to the one-third of the whole data (TSV 1/3) or on the data set of the remaining season. For the global model (last row in the table) the data from three successive harvest seasons were combined and validated with internal test set (TSV 1/3). All the spectral data was pre-treated with standard normal variate (SNV) method and the spectral region of 1063–1334, 1639–1836, 2173–2355 nm were used for model development

‘Kabaası’ cultivar. For this cultivar, although we obtained good results for TSV (1/3) validation in 2014 ($R_p^2 = 0.827$, RMSEP = 1.85, RPD = 2.42) and 2015 ($R_p^2 = 0.843$, RMSEP = 1.49, RPD = 2.53), the model performance was remarkably poor in 2016 ($R_p^2 = 0.130$, RMSEP = 3.06, RPD = 1.07). Actually, in 2016, Malatya region received heavy rains for 3 days during the harvest period when the apricots were at turning maturity stage, which possibly affected the physico-chemical quality of the apricots resulting in atypical spectral behaviour. As a result, regardless of the cultivar, single-year models constructed by using the data of 2014 and 2015 were not successful in predicting the SSC of the apricots from 2016 season. Similarly, the calibration models constructed with the data of 2016 was not precise as well in predicting the SSC of the apricots from 2014 and 2015 (Table 3). On the other hand, the results of other single-year models for ‘Kabaası’, ‘Hacıhaliloğlu’ and their combinations were in general successful in predicting the apricots of another year with R_p^2 , RMSEP and RPD values ranging between 0.549–0.781, 3.02–2.31 and 2.03–2.55, respectively. Regarding multi-

year calibration models, R_c^2 and RMSEC values of the models ranged between 0.700–0.917 and 0.97–2.35, respectively (Table 4). The results of the model validation either with TSV (1/3) or external test set showed that except for the calibration model for ‘Kabaası’ cultivar established with the data of 2014–2015 and 2014–2016 and validated with the external data set of 2016 and 2015, respectively, SSC calibration models had prediction ability with R_p^2 , RMSEP and RPD values ranging between 0.624–0.836, 2.94–2.04 and 2.70–2.52. It is worth mentioning that although it contains the atypical data of ‘Kabaası’ from 2016, the global model constructed with the data of three successive years and two cultivars gave very good results ($R_p^2 = 0.830$, RMSEP = 1.90 and RPD = 2.43) (Table 4) which clearly indicates the robustness of the prediction model. Similar results were also reported for some other fruits such as Japanese plums (Louw and Theron 2010), apples (Peirs et al. 2003; Bobelyn et al. 2010), mangos (Rungpichayapichet et al. 2016), avocado (Wedding et al. 2013) and mandarins (Guthrie et al. 2005) it was found that the multi-cultivar

models developed with the data of two or more harvest seasons for the prediction were more robust than the models constructed with the data of single harvest season. It should be noted that in the present study the global model constructed with the data of three successive years gave better results for ‘Hacıhaliloğlu’ ($R_p^2 = 0.881$, RMSEP = 1.53 and RPD = 2.93) than ‘Kabaaşı’ ($R_p^2 = 0.777$, RMSEP = 2.17 and RPD = 2.14) and multi-cultivar model (Table 4). For apples (Bobelyn et al. 2010) it was found that mono-cultivar models constructed with the data of several seasons resulted in more robust models than multi-cultivar models whereas for Japanese plums (Louw and Theron 2010) multi-cultivar models for SSC resulted in equal or better precision than mono-cultivar models.

Conclusion

The results showed that the SSC of the apricots can be accurately predicted by FT-NIR spectroscopy. The SSC prediction model became more robust when the data from several harvest seasons were used for the model development. It was found that the cultivar specific models had higher prediction accuracy than the multi-cultivar models. It was also demonstrated that the apricot weight could also be predicted with acceptable prediction accuracy by using FT-NIR spectroscopy which could allow selecting reliably between small and large apricots. The prediction models developed in this study would be of great interest for the development of grading lines which can be used to sort apricots simultaneously according to their SSC and size. Such systems will be very useful to provide standardized raw material for sulphured-dried apricot processing thus will enable the production of sulphured-dried apricots with standard quality.

Acknowledgements The authors are grateful for the financial support of the Scientific and Technological Research Council of Turkey (TUBITAK research Project Number 112 G 018).

References

- Adachi T, Nonogi H, Fuke T et al (1979) On the combination of sulphite with food ingredients (aldehydes, ketones and sugars). II. Zeitschrift für Lebensmittel- und Futtermittelwissenschaft 168:200–205. <https://doi.org/10.1007/BF01123173>
- Akin EB, Karabulut I, Topcu A (2008) Some compositional properties of main Malatya apricot (*Prunus armeniaca* L.) varieties. Food Chem 107:939–948. <https://doi.org/10.1016/j.foodchem.2007.08.052>
- AOAC-932.12 (1998) Solids (soluble) in fruits and fruit products. Official methods of analysis, 16th edn. Association of Official Analytical Chemists, Arlington
- AOAC-942.15 (1995) Acidity (titratable) of fruit products. Official methods of analysis, 16th edn. Association of Official Analytical Chemists, Arlington
- Barrett RB (2002) Characterisation of apricot polyphenoloxidase during fruit development. Thesis (M.App.Sc.). The University of Adelaide
- Berardinelli A, Cevoli C, Silaghi FA et al (2010) FT-NIR spectroscopy for the quality characterization of apricots (*Prunus armeniaca* L.). J Food Sci 75:462–468. <https://doi.org/10.1111/j.1750-3841.2010.01741.x>
- Bobelyn E, Serban A-S, Nicu M et al (2010) Postharvest quality of apple predicted by NIR-spectroscopy: study of the effect of biological variability on spectra and model performance. Postharvest Biol Technol 55:133–143. <https://doi.org/10.1016/j.postharvbio.2009.09.006>
- Bureau S, Ruiz D, Reich M et al (2009) Rapid and non-destructive analysis of apricot fruit quality using FT-near-infrared spectroscopy. Food Chem 113:1323–1328. <https://doi.org/10.1016/j.foodchem.2008.08.066>
- Buyukcan MB, Kavdir I (2016) Prediction of some internal quality parameters of apricot using FT-NIR spectroscopy. J Food Meas Charact. <https://doi.org/10.1007/s11694-016-9434-9>
- Camps C, Christen D (2009) On-tree follow-up of apricot fruit development using a hand-held NIR instrument. J Food Agric Environ 7(2):394–400
- Carlini P, Massantini R, Mencarelli F (2000) Vis-NIR measurement of soluble solids in Cherry and Apricot by PLS regression and wavelength selection. J Agric Food Chem 48:5236–5242. <https://doi.org/10.1021/jf000408f>
- Crisosto CH (1994) Stone fruit maturity indices: a descriptive. Postharvest News Inf 5:65N–68N
- European Commission (2000) Commission Regulation (EC) No 851/2000 of 27 April 2000 laying down the marketing standard for apricots. Off J Eur Communities. 2004R0726 - v.7 of 05.06.2013
- Gökçe K (1966) Malatya kayısılarının kükürtlenmeleri üzerine teknik çalışmalar. Ankara Üniversitesi Ziraat Fakültesi Yayınları 87
- Guido LF (2016) Sulfites in beer: reviewing regulation, analysis and role. Sci Agric 73:189–197
- Guthrie JA, Reid DJ, Walsh KB (2005) Assessment of internal quality attributes of mandarin fruit. 2. NIR calibration model robustness. Aust J Agric Res. <https://doi.org/10.1071/ar04299>
- Joslyn MA, Braverman JBS (1954) The chemistry and technology of the pretreatment and preservation of fruit and vegetable products with sulfur dioxide and sulfites. In: Mrak E, Stewart G (eds) Advances in food research. Academic Press, Boca Raton, pp 97–160
- Louw ED, Theron KI (2010) Robust prediction models for quality parameters in Japanese plums (*Prunus salicina* L.) using NIR spectroscopy. Postharvest Biol Technol 58:176–184. <https://doi.org/10.1016/j.postharvbio.2010.07.001>
- Magwaza LS, Opara UL (2015) Analytical methods for determination of sugars and sweetness of horticultural products—a review. Sci Hortic (Amsterdam) 184:179–192. <https://doi.org/10.1016/j.scienta.2015.01.001>
- Nicolai BM, Beullens K, Bobelyn E et al (2007) Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biol Technol 46:99–118. <https://doi.org/10.1016/j.postharvbio.2007.06.024>
- Peiris KHS, Dull GG, Leffler RG, Kays SJ (1998) Near-infrared spectrometric method for nondestructive determination of soluble solids content of peaches. J Am Soc Hortic Sci 123:898–905
- Peirs A, Tirry J, Verlinden B et al (2003) Effect of biological variability on the robustness of NIR models for soluble solids content of apples. Postharvest Biol Technol 28:269–280. [https://doi.org/10.1016/S0925-5214\(02\)00196-5](https://doi.org/10.1016/S0925-5214(02)00196-5)
- Pérez-Marín D, Sánchez M-T, Paz P et al (2009) Non-destructive determination of quality parameters in nectarines during on-tree

- ripening and postharvest storage. *Postharvest Biol Technol* 52:180–188. <https://doi.org/10.1016/j.postharvbio.2008.10.005>
- Polesello A, Giangiacomo R, Dull GG (1983) Application of near infrared spectrophotometry to the nondestructive analysis of foods: a review of experimental results. *C R C Crit Rev Food Sci Nutr* 18:203–230. <https://doi.org/10.1080/10408398309527363>
- Rosselló C, Cañellas J, Santiesteban I, Mulet A (1993) Simulation of the absorption process of sulphur dioxide in apricots. *LWT Food Sci Technol* 26:322–328. <https://doi.org/10.1006/fstl.1993.1066>
- Rossier J, Darbellay C, Azodanlou R, Villettaz JC (2001) Measurement of organoleptic quality of apricot. In: XII International symposium on apricot culture and decline, vol 701, pp 575–576
- Rungpichayapichet P, Mahayothee B, Nagle M et al (2016) Robust NIRS models for non-destructive prediction of postharvest fruit ripeness and quality in mango. *Postharvest Biol Technol* 111:31–40. <https://doi.org/10.1016/j.postharvbio.2015.07.006>
- Türkyılmaz M, Özkan M, Güzel N (2014) Loss of sulfur dioxide and changes in some chemical properties of Malatya apricots (*Prunus armeniaca* L.) during sulfuring and drying. *J Sci Food Agric* 94:2488–2496
- Wedding BB, Wright C, Grauf S et al (2013) Effects of seasonal variability on FT-NIR prediction of dry matter content for whole Hass avocado fruit. *Postharvest Biol Technol* 75:9–16. <https://doi.org/10.1016/j.postharvbio.2012.04.016>