Multi-spectral assessment of ingredients and physical properties of apricot

Ferenc Firtha, Zoltán Gillay, Eszter Vozáry, Tímea Kasza, Anikó Lambert-Meretei and József Felföldi

Corvinus University of Budapest, Department of Physics and Control, 14–16. Somlói str, Budapest, H-1118, Hungary

Introduction

Some quality-related internal parameters of apricot can be predicted by the reflected NIR spectrum. According to the recent publications,[1,2] using 800–2500 nm range, the soluble solids content (SSC) and the titratable acidity (TA) can be predicted properly, but other quality traits, like malic and citric acid, individual sugars, ethylene production and firmness were not satisfactorily modeled.

Non-destructive, non-contact and fast optical measurement methods, such as hyper- or multi-spectral imaging are in more and more demand for on-line industrial quality control tasks. These methods combine the advantages of spectroscopy and conventional image processing, and obtain the spatial distribution of spectral properties on non-homogeneous surfaces, but tend to have much less spectral information in usually noisy environments.

For testing the feasibility of multispectral industrial application, the internal-, rheological and optical properties of apricot cultivars were investigated. Samples of three cultivars, three ripening states and three further categories by storage time were measured, with 20 samples in each group. All the measurements were taken on both blushed and un-blushed sides of the fruit.

The optical properties were measured and checked in different spectral ranges with different instrumentation. Mechanical properties of the samples were measured by a dynamic method, with two impact methods and an acoustic response system. The chemical properties were measured after all non-destructive methods mentioned above had been applied.

Materials and methods

In the experiment described below, samples of Bergeron, Bergarouge and Zebra apricot cultivars were tested with grouping in three ripeness categories (1. immature, 2. ripened for processing, 3. ripened for consumption). The second category was stored for one week (4. category) and two weeks (5. category).

The mass (m) and the three perpendicular diameters (d_1, d_2, d_1) were recorded. The following optical, mechanical and chemical parameters were inspected:

Optical:

- **RGB Imaging System** (RGB images of both side using diffuse illumination):
  average and variance of XYZ colour components of segmented areas
- **Pigment Analyzer** (400–1090 nm range, 3.25 nm resolution):
  Normalised Difference Vegetation Index (NDVI), Normalised Anthocyanin Index (NAI)
- **ColorLite sph850 spectrophotometer** (400–700 nm range, 10 nm resolution):
  reflected spectra, CIE Lab, Luv and XYZ coordinates
- **PCM Spectralyzer 10–25** (1000–2500 nm range, 2 nm resolution):
  the average reflectance on a 25 mm diameter area
- **NIR Multispectral Imaging system** (12 images at 1000–1550 nm):
  average and variance of intensity values on segmented areas

Mechanical:

- **Acoustic resonance method**:
  measured: resonance frequency (f, Hz) and width of the resonance peek
  calculated: $\mathbf{s_1} = f^2 * m^{2/3}$ and $\mathbf{s_2} = f^2 * d_1^2$ acoustic stiffness coefficients
- **Impact method**:
  measured: time of deceleration of the impact hammer (**dT**, ms)
  calculated: $\mathbf{D} = 1/dT^2$ impact stiffness coefficient
- **Sinclair Internal Quality tester**:
  measured: Sinclair firmness coefficient on 1–100 scale (**IQ**)

Destructive:

- **pH** of the apricot flesh (Vaiseshika pH-conductivity TDS+DO meter and inserted flesh probe)
- **SSC (Brix)** of the apricot juice (Atago digital refractometer PAL-1)
- **Sugar-content** (fructose, glucose, saccharose, xylose, raffinose) (HPLC measurement)
- **Free acid content** of samples (titration)

The correlation between **NIR spectrophotometer measurement** and different quality traits was investigated by partial least squares (PLS) linear regression method. Since randomly selecting two thirds of the whole dataset as calibration subset did not result in reliable prediction, finally the first (immature) and third (ready for consumption) groups were used as the calibration set, having a wide distribution of internal properties. The optimal number of latent variables was determined on the base of the minimal value of *RMSEV* and the maximal value of relative performance determinant (*RPD*). The over-fitting of regression was checked by the Beta-vector.

The **multispectral images** were segmented, then the average spectra of the investigated area were compared with spectrophotometer measurements. The prediction model was the PLS method. Multi-linear regression (MLR) was used to determine the significant wavelengths of given internal properties. All of the statistical algorithms were developed in Mathcad software package (version 14.0, MathSoft, USA).

Table 1. Average and standard deviation of quality traits.

		Stiffness					
Cultivar	Ripeness/Storage	Sinclair		Impact		Acoustic	
	Categories	**avg**	std	**avg**	std	**avg**	std
Bergeron	immature (1)	**14.0**	2.5	**0.140**	0.028	**2050099**	475643
	processing (2)	**12.1**	1.3	**0.133**	0.017	**2135610**	370329
	consumption (3)	**9.8**	1.3	**0.092**	0.017	**1572507**	294056
	processing (2)	**12.1**	1.3	**0.133**	0.017	**2135610**	370329
	1 week (4)	**10.3**	7.2	**0.073**	0.020	**1011637**	465146
	2 weeks (5)	**–**	**–**	**0.039**	0.016	**186192**	107327
Bergarouge	immature (1)	**14.9**	2.5	**0.141**	0.027	**2955759**	1117264
	processing (2)	**11.2**	2.6	**0.110**	0.022	**2085061**	847019
	consumption (3)	**8.2**	2.0	**0.073**	0.016	**956134**	315487
	processing (2)	**11.2**	2.6	**0.110**	0.022	**2085061**	847019
	1 week (4)	**13.5**	12.4	**0.056**	0.012	**461506**	202062
	2 weeks (5)	**–**	**–**	**0.034**	0.011	**186482**	131904
Zebra	immature (1)	**21.8**	3.1	**0.228**	0.041	**4262904**	1238783
	processing (2)	**20.3**	4.1	**0.181**	0.030	**3339701**	1033934
	consumption (3)	**13.2**	4.2	**0.123**	0.027	**1609299**	554809
	processing (2)	**20.3**	4.1	**0.181**	0.030	**3339701**	1033934
	1 week (4)	**12.3**	3.1	**0.092**	0.019	**928187**	282781
	2 weeks (5)	**–**	**–**	**0.057**	0.011	**521339**	168552

Results

All the measured optical (for example, Lab, NDVI, NAI) and mechanical (Sinclair-, impact- and acoustic stiffness) parameters were changed monotonically by the ripeness state and by the storage time (Table 1).

The chemical properties, however, behaved irregularly. The **pH** increased during ripening, but during storage, it increased significantly at the first week, then slightly decreased. The SSC (**Brix**) increased during ripening, but was not affected by the storage time. The change of **sugar-content** (fructose, glucose, saccharose, xylose, raffinose) and **titratable acid** (TA) was not clearly associated with the state of ripeness nor by storage time (Table 2).

This paper will focus on the prediction of pH and Brix.

Spectrophotometer data analysis:

The general profile of the absorption spectra was dominated mostly by water absorption bands with bands at 970 nm, 1190 nm, 1450 nm and 1940 nm. There was a definite tendency in the changes of absorption spectra in the function of both the ripeness state and the storage time on all cultivars and on both sides (blushed and un-blushed) of samples.

a) The PLS prediction model of pH gave $LV=3$ as optimal number of factors. At this value, the *RMSEV* was minimum and the *RPD* was maximum. The vector B-coefficients was found to be smooth, showing that there was no overfitting. The $RMSEV=0.24$ value of validation was acce ptable compared to the standard deviation of pH values of sample set. Both the value of $RPD=2.45$ and the value of the coefficient of determination ($R^2=0.58$) showed good correlation between NIR spectra and the pH of apricot (Figure 1).
b) The PLS model of Brix gave $LV=3$. The Beta-vector showed no over-fitting. The value of $RMSEV=1.22$ was promising compared to the standard deviation of Brix of the validation sample set. Despite the rather small $RPD=1.58$, the $R^2=0.58$ showed encouraging correlation.

Multispectral data analysis:

The frames, acquired through different filters, were processed for segmenting areas to be investigated. The average reflection was calculated for each image using 12 bit (0..4095) resolution. The relation between NIR absorption measurements and multispectral reflection data was found to be satisfactory. The changes of reflected spectra were also consistent with both ripeness state and storage time on all cultivars, and on both sides of samples.

a) Using multiplicative scatter correction, with $LV=5$ factors, the $RPD=1.73$ and the small $RMSEV=0.27$ show acceptable values. $R^2=0.24$ is small, because the cultivars have different behaviours on these wavelengths (Figure 2).
b) Building the model for a given cultivar (for example, Bergeron), the results improved. Without multiplicative scatter correction, $LV=2$ factors resulted in $RPD=1.38$ and $RMSEV=0.37$ values. The significant wavelengths were calculated by the MLR method.
c) The model for prediction of Brix resulted in the $LV=6$ as optimal number of latent variables. The value of $RPD=1.20$ was barely bigger than 1, $RMSEV=1.52$ was acceptably small, but the coefficient of determination was almost zero, indicating that using these wavelengths, the

Table 2. Summary of chemical properties.

		Chemical			
Cultivar	Ripeness/Storage	Brix, %		pH	
	Categories	**avg**	std	**avg**	std
Bergeron	immature (1)	**14.8**	1.3	**3.46**	0.24
	processing (2)	**15.5**	1.4	**3.92**	0.23
	consumption (3)	**17.1**	1.2	**4.28**	0.22
	processing (2)	**15.5**	1.4	**3.92**	0.23
	1 week (4)	**15.1**	1.3	**5.29**	0.21
	2 weeks (5)	**15.6**	1.5	**5.11**	0.35
Bergarouge	immature (1)	**13.8**	1.5	**3.6**	0.4
	processing (2)	**16.5**	1.0	**4.2**	0.3
	consumption (3)	**17.6**	1.1	**4.4**	0.3
	processing (2)	**16.5**	1.0	**4.2**	0.3
	1 week (4)	**15.5**	0.9	**5.2**	0.4
	2 weeks (5)	**15.7**	1.2	**5.0**	0.5
Zebra	immature (1)	**13.4**	1.2	**2.8**	0.1
	processing (2)	**14.2**	1.0	**2.9**	0.2
	consumption (3)	**14.6**	1.1	**3.6**	0.5
	processing (2)	**14.2**	1.0	**2.9**	0.2
	1 week (4)	**16.1**	1.3	**5.2**	0.2
	2 weeks (5)	**16.8**	1.0	**5.1**	0.5

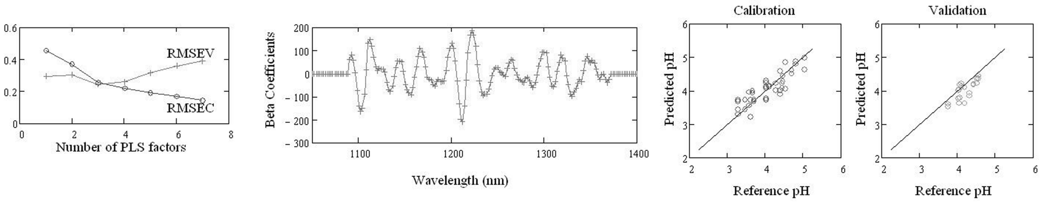

Figure 1. Validation of spectralyzer model.

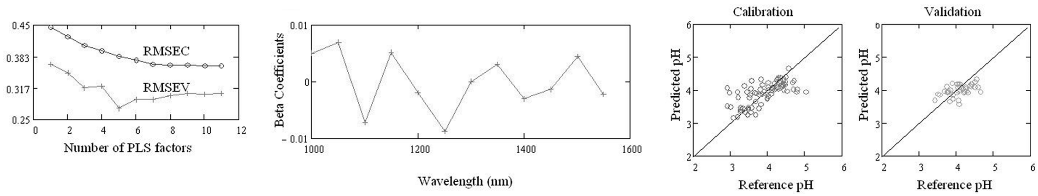

Figure 2. Validation of multispectral model.

soluble solids content can not be predicted. Significant wavelengths for this property must be studied by the hyperspectral method.

Conclusion

The multispectral assessment of ingredients seems to be encouraging, but:

- Sets of samples must be selected for calibration, having wider range of properties.
- All the noise, stray light should be especially excluded at measurements.
- More chemical factors should be measured to explain the irregular changes of sugar and acid components (internal standard addition).
- Significant wavelengths of investigated properties will be studied by the hyperspectral method.
- Image processing algorithm will be developed to segment blushed and un-blushed areas on multispectral images to improve efficiency.

References

1. S. Bureau, D. Ruiz, M. Reich, B. Gouble, D. Bertrand, J. Audergon and M.G.C. Renard, "Rapid and non-destructive analysis of apricot fruit quality using FT-near-infrared spectroscopy", *Food Chem.* **113,** 1323–1328 (2009).
2. C. Camps and D.J. Christen, "Non-destructive assessment of apricot fruit quality using FT-near-infrared spectroscopy", *Food Sci. Technol.* **42,** 1125–1131 (2009).