Postharvest Biology and Technology 58 (2010) 176–184

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Robust prediction models for quality parameters in Japanese plums (*Prunus salicina* L.) using NIR spectroscopy

Esmé D. Louw*, Karen I. Theron

Department of Horticultural Sciences, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

ARTICLE INFO

Article history:
Received 2 March 2010
Accepted 11 July 2010

Keywords:
Firmness
Near infrared (NIR) spectroscopy
Robustness
Prediction models
TA
TSS
Weight

ABSTRACT

Fourier-transformed near infrared (FT-NIR) reflectance spectroscopy was used over a spectral range of 800–2700 nm to develop multivariate prediction models for total soluble solids (TSS), total acidity (TA), sugar-to-acid ratio, firmness and weight in three South African plum cultivars (Pioneer, Laetitia and Angeleno) and a multi-cultivar model. Samples were collected for 7 weeks throughout the ripening period and repeated over two seasons. The validation results had mixed success with TSS (R^2 = 0.817–0.959; RMSEP = 0.453–0.610% Brix), TA (R^2 = 0.608–0.830; RMSEP = 0.110–0.194% malic acid), sugar-to-acid ratio (R^2 = 0.718–0.896; RMSEP = 0.608–1.590), firmness (R^2 = 0.623–0.791; RMSEP = 12.459–22.760 N) and weight (R^2 = 0.577–0.817; RMSEP = 7.700–12.800 g). The cultivar-specific models of 'Pioneer' and 'Laetitia' had a better predictability capacity than the 'Angeleno' model on all parameters. Although the multi-cultivar model for TSS, TA and sugar-to-acid ratio outperformed the single-cultivar models on R^2 values, they had higher prediction errors. The robustness of all the TSS, TA and firmness models is high in terms of seasonality and range.

1. Introduction

Japanese plums (*Prunus salicina* L.) have been bred and cultivated in South Africa for many decades. With more than eight million cartons of 35 different cultivars exported annually, plums are one of the most diverse fruit tree crops traded internationally. As plums have a relatively short shelf life, South Africa is geographically placed at a disadvantage when exporting plums to European markets. As sea freight is the most economical mode of transporting fresh fruit from South Africa, plums are harvested relatively immature and ripen while in transit for up to 42 d. This makes determining and monitoring of quality parameters in the orchard, pack house and delivery points crucial in producing a product that is acceptable to the end user.

Consumer acceptance studies have shown that total soluble solids (TSS) concentration and titratable acidity (TA) are two important quality parameters in plums (Crisosto and Bowerman, 2003; Crisosto and Crisosto, 2005). Fruit firmness has been proven to indicate an acceptable shelf life (Valero et al., 2007) and can be used successfully to determine the optimum harvest date of plums (Guerra and Casquero, 2008). The weight of fruit is not directly associated with fruit quality or consumer acceptance, but can be an indication of water loss and shrivel that have a negative impact on fruit appearance.

Currently all the quality parameters are determined using destructive measures, i.e. paring and crushing to determine firmness and juicing to measure TSS and TA. As this makes it impossible to test every unit of fruit, a statistically determined subset of a batch is tested and the results are taken as representative of the entire batch. Large variability can exist between individual fruit as a result of pre-harvest factors (climatic conditions, e.g. winter chilling, soil type, bearing position of fruit on the tree, age of the tree, irrigation, fertilization schedules, etc.) or postharvest factors (time of harvest, pre-cooling, handling and storage practices, etc.) and with numerous quality parameters to test it can be difficult to accurately assess the quality of the entire batch of fruit. The South African plum industry can benefit from non-destructive technology that rapidly and accurately predicts the quality parameters of individual fruit.

Near infrared (NIR) spectroscopy can possibly serve as a non-invasive technique to determine quality in plums as it is interacts with molecular groups associated with quality parameters such as sugars (C–H group), acids and moisture (O–H group) and scattering from microstructures (Abu-Khalaf and Bennedsen, 2002; Nicolaï et al., 2007) can indirectly indicate physical parameters. Most of the NIR absorption bands associated with these groups are overtone or combination bands of the fundamental absorption bands in the infrared region which are due to vibrational and rotational transitions (Nicolaï et al., 2007). Exposing intact fruit samples to

* Corresponding author. Tel.: +27 21 808 4900; fax: +27 21 808 2121.
E-mail address: esmelouw@gmail.com (E.D. Louw).

0925-5214/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.postharvbio.2010.07.001

NIR spectroscopy will produce an absorption pattern of the chemicals present in the fruit in a rapid and non-destructive way. These spectra can then be manipulated using multivariate data analysis techniques to develop prediction models for each measured variable. Although the initial model building will require reference data based on the traditional destructive methods, a robust model can thereafter be used to predict the quality parameters non-destructively.

When considering a prediction model it is important to take note of the type of validation method that was used. Many different validation methods are available and the choice is often driven by logistics or cost. However, there is no better validation than testing on an entirely independent data set (external validation). NIR spectroscopy prediction models (using different validation methods) have been reported for numerous fruit and vegetable types focusing on TSS, TA and firmness as the predicted quality parameter. Some of these include models for apricots (Bureau et al., 2009; Camps and Christen, 2009), tomatoes (Flores et al., 2009), loquats (Fu et al., 2009), apples (Peirs et al., 2005; Lui et al., 2007; Paz et al., 2009), mango (Schmilovitch et al., 2000), pears (Lui et al., 2008; Cavaco et al., 2009), kiwifruit (McGlone and Kawano, 1998), watermelons (Ito et al., 2002), peaches (Ma et al., 2007), nectarines (Pérez-Marín et al., 2009), prunes (Slaughter et al., 2003) and plums (Onda et al., 1994; Abu-Khalaf and Bennedsen, 2002; Paz et al., 2008).

This study aims to determine if NIR spectroscopy can be used as a non-destructive alternative for the accurate prediction of quality parameters such as TSS, TA, sugar-to-acid ratio, firmness and weight in three South African plum cultivars harvested at different stages of ripeness over two seasons.

2. Material and methods

2.1. Plum fruit selection

Fruit from three plum (*P. salicina* L.) cultivars grown near Stellenbosch (Western Cape, South Africa) were used in this study. The cultivars selected were 'Pioneer', 'Laetitia' and 'Angeleno'. Eighty fruit were collected weekly over a 7 weeks period starting 3 weeks prior to the expected commercial harvest date and continuing for 3 weeks thereafter. Fruit of similar size and colour were selected from the middle of the canopy approximately 1.5 m from the orchard floor. Non-destructive NIR measurements were taken on the same day as harvest and the destructive measurements were done within 36 h after harvest. Fruit were stored at ambient temperature and not exposed to any postharvest treatments prior to processing. The study was conducted over two plum seasons (2007 and 2008) with total fruit numbers of 1200 for 'Pioneer' (eight harvest weeks in 2008), 1120 for 'Laetitia' and 1040 for 'Angeleno' (six harvest weeks in 2008).

2.2. Non-destructive near infrared spectra collection

FT-NIR spectra were obtained using a multi-purpose analyser (MPA) spectrometer (Bruker Optics, Ettlingen, Germany) fitted with a solid probe fiber optics module containing a high sensitivity, thermoelectrically cooled InGaAs detector with a tungsten lamp as the NIR source. For each plum the probe (5 mm diameter with roughly 100 optic fibers) was directed onto the skin of two opposite sides of the intact fruit and an absorbance spectrum covering a wavelength range of 800–2700 nm (resolution of 8 nm, scanner velocity of 10 kHz) was captured through reflectance geometry. For each spectrum the average of 16 scans with a resolution of 8 nm was used. A white Spectralon tile was used as a 100% reflective background reference.

2.3. Determination of fruit quality parameters (reference data)

The reference data were collected using the conventional destructive methods. Fruit weight was determined in grams using a calibrated balance (GÜSS GS20 FTA, Cape Town, South Africa). Flesh firmness was measured in kilograms on two opposite pared sides of the fruit after exposing the flesh to an electronic penetrometer (GÜSS GS20 FTA, Cape Town, South Africa) fitted with an 11.0 mm tip. All values were converted to newtons by multiplying by 9.81. To determine the total soluble solids (TSS) the fruit were juiced individually using a commercial fruit blender. A drop of juice from each fruit was placed onto a temperature-controlled, digital refractometer (Palette PR-32 ATAGO, Bellevue, USA) which measured the TSS levels in % Brix. Total acid (TA) was expressed as % malic acid by titrating a 10 g aliquot of the individual plum juice with 0.1 M NaOH to a pH end-point of 8.2 using an automated titrator (Metrohm AG 760, Herisau, Switzerland). In cases where the fruit were very small and did not produce enough juice (<10 g) the juice of up to three plums were pooled, measured and given the same TA value. Data from the three cultivars were pooled to create reference data for the multi-cultivar model. The mean and standard deviation values were determined for each quality parameter (Table 1).

2.4. Chemometric data analysis

OPUS version 6.1 (Bruker Optics, Ettlingen, Germany) chemometric software was used to perform all the multivariate calculations. Spectral parameters were selected using the "Optimize" function of the software which checks common wavelength frequency regions in combination with several data pre-processing methods. The software then yields a list of the possible parameter combinations and the resulting RMSECV value and number of latent variables, from this we selected the method that presented the best all-round performance (in terms of frequency region, number of latent variables and error) for each model (Table 1). Only the informative frequency regions for each spectrum were retained from the initial wavelength interval of 800–2700 nm and used in further calculations. The partial least square (PLS) regression method (including mean centering) was applied to the transformed data to create prediction models for each of the quality parameters. Outliers were quantified by deriving a threshold value using the Mahalanobis distance of each calibration spectrum. To construct calibration models with high robustness we combined all availability data for each cultivar (2007 and 2008 seasons) and split it into two equal, unique subsets. One subset was used to build the calibration model and then testing it internally via cross-validation (leaving out 10 samples) to determine the complexity using the number of latent variables (LV's) that presented the lowest RMSECV. The second subset was then used to do an external validation of the calibration model using the complexity as calculated by the cross-validation. To illustrate the robustness in terms of seasonality the data was split into the two seasons (2007 and 2008). Data from one season was used as a calibration set and tested internally via cross-validation to determine complexity. This was followed by an external validation using data from the other season. Robustness in terms of range was illustrated by reducing the sample collection period from the initial 7 weeks (W1–W7) to only 3 weeks (W3–W5) including the week of commercial harvest (W4) and the two flanking weeks. Using less data in this way reduces the range of each of the variables when compared to the initial model. Again each model was tested using a cross-validation and complexity was determined by the lowest RMSECV. External validation was done twice for each model, firstly using a reduced validation set also only containing data from W3–W5 ("reduced validation") and then secondly using the full validation set containing data from W1 to W7 ("full validation"). In all cases the spectra from the three cul-