ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Postharvest Biology and Technology 29 (2003) 300–307

Postharvest Biology and Technology

www.elsevier.com/locate/postharvbio

Dry matter determination in 'Hass' avocado by NIR spectroscopy

C.J. Clark [a,*], V.A. McGlone [a], C. Requejo [b], A. White [b], A.B. Woolf [b]

[a] Bioengineering, HortResearch, Private Bag 3123, Hamilton, New Zealand
[b] Postharvest and Food Science, HortResearch, PO Box 92 169, Auckland, New Zealand

Received 9 August 2002; accepted 1 March 2003

Abstract

'Hass' avocado harvested at four different times during a growing season, was analysed by both reflectance and interactance NIR spectroscopy to establish its reliability for the non-destructive determination of fruit dry matter (DM). Mean DM increased from 27.7 to 36.8% during the course of the study. Relationships between spectral wavelengths and DM were evaluated by application of chemometric techniques to partial least squares (PLS) and multiple linear regression (MLR) models, using calibration and validation statistics of predictions to compare their efficacy. The interactance mode was a superior predictor of fruit DM compared with reflectance. A PLS model applied to interactance data from fruit from all harvests combined ($n = 239$) gave a goodness-of-fit (R_p^2) between predicted and actual DM measurements of 0.88, and an error of prediction (RMSEP) of 1.8% DM. MLR models incorporating four wavelengths returned equivalent validation statistics. Three of these wavelengths, in the vicinity of 900–920 nm, are consistent with the C–H absorbance band, leaving only a minor role for water-related absorbances. These observations suggest on-line commercial NIR systems capable of operating in an interactance mode have the potential to usefully grade avocado on the basis of their DM content, to improve taste and oil content.

Keywords: Fruit; Avocado; NIR; Non-destructive analysis; Maturity; Dry matter

1. Introduction

Use of near infra-red spectroscopy (NIRS) in horticulture is receiving extensive research effort, and commercial applications are in the early stage of implementation (Abbott et al., 1997; Armstrong, 2000; Hey, 2001; Malagoli, 2002). The ability to rapidly 'scan' fruit on-line, and then sort it, means that if a given characteristic can be accurately measured, fruit can be segregated into distinct classes and either handled or marketed in a different manner.

Commercially, avocado (*Persea americana* Mill.) maturity is measured by dry matter (DM) analysis or the more laborious determination of oil content, both of which are highly correlated (Lewis, 1978; Lee et al., 1983). A means of non-destructively measuring DM could thus be useful

* Corresponding author. Tel.: +64-7-858-4797; fax: +64-7-858-4705.
E-mail address: cclark@hortresearch.co.nz (C.J. Clark).

0925-5214/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0925-5214(03)00046-2

for a range of applications such as rapid and numerous measurement in the field, or grading of fruit during packing so that fruit were of similar DM levels. This may have implications for both consumer acceptability—higher DM fruit is likely to be of better taste—and for improving storage potential or ripe fruit quality.

Of the non-destructive technologies currently being developed for determining internal fruit quality on-line, nuclear magnetic resonance (NMR) and NIRS are leading candidates. Using a variety of NMR methods, Chen et al. (1993) were able to demonstrate a useful relationship between DM and the ratio of the oil/water resonance intensities in intact avocado that had desirable features for high-speed sorting. Their follow-up study (Kim et al., 1999) indicated that this relationship was conserved at conveyor speeds up to 25 cm s^{-1}. Despite the promise of studies such as these, introduction of NMR grading systems for fruit is still some time away, while commercial NIR systems are becoming more prevalent. However, apart from a preliminary study by Schmilovitch et al. (2001), there appears to have been minimal effort to investigate DM relationships in avocado using NIRS.

To determine the feasibility and accuracy of measuring DM in avocado by this approach, we now report the results of NIR measurements using both reflectance and interactance modes obtained from avocado harvested at discrete intervals during a growing season.

2. Material and methods

2.1. Fruit

Mature ‘Hass’ avocados were harvested from three commercial orchards on four occasions during a growing season (23 October 2001, 10 December, 30 January 2002, and 27 March). The three properties were located in close proximity to each other in the Bay of Plenty (North Island, NZ), one of the main avocado production areas in New Zealand. On each occasion, 80 fruit were harvested from each orchard and transported immediately to the laboratory. Fruit were maintained at 20 °C prior to analysis and measurements commenced between 40 and 48 h after harvest.

2.2. NIR methods

Vis/NIR spectra of intact avocado were acquired in both the reflectance and interactance modes. In each case, a custom-designed laboratory system was utilised that consisted of a wide-band light source, a fruit holder/light collection fixture and a non-scanning polychromatic/diode array spectrometer (Zeiss MMS1-NIR, Germany). For reflectance measurements, the light source was a 50 W, 12 V DC tungsten halogen lamp. For interactance, the illumination source was a ‘Fruit Selector’ instrument with a 650 W projector bulb and condenser lens (Cabin Industrial Co., Japan). The spectrometer had a spectral range of 300–1140 nm with a 3.3 nm sampling interval (256 points per spectrum). A 15-bit ADC electronics unit (FEE-001, tec5, Germany) was used to amplify and digitise the spectral signal. Data acquisition and storage was achieved with a PC running in-house software (NIR Fruit, HortResearch, 1999).

Each reflectance or interactance spectrum was accumulated over 200 ms, or 1.5 s, from five contiguous acquisitions at a single location on the fruit at integration times of 40 or 300 ms, respectively. All spectra were converted to relative transmittance by subtracting the dark current and dividing the measured intensity at each wavelength by the corresponding intensity from a standard reference spectrum. The sensor dark current was measured by collecting the spectrum obtained by covering the fruit holder with a steel ball that prevented light reaching the sensor. The standard reference spectrum, measured after every 20 fruit, was obtained by placing a 70 mm diameter Teflon sphere in the fruit holder, and measuring the intensity of reflected light; or by placing a 50 × 80 × 80 mm Teflon block directly above the fruit holder (for interactance).

2.3. Measurements

Implications concerning the spatial variability of DM in fruits on NIR spectrometric measurements have been reported by Peiris et al. (1999), and physiological gradients in the distribution of DM are known to occur in avocado (Schroeder, 1985). Taking these factors into account, a single position was identified on the fruit where a reflectance and interactance spectrum was collected, and DM measured. The measurement site was in the same plane as the peduncle insertion point, but on the opposite side of the fruit at the equator (Fig. 1).

A plug of tissue, collected perpendicular to the fruit surface, was subsequently removed from the NIR sampling site using a 17.3 mm diameter cork borer. It should be noted that the diameters from which the NIR spectra were collected were 12 and ~24 mm for interactance and reflectance, respectively. After removal of any seed coat, and the skin, disks of plug tissue were cut into slices to facilitate drying and oven-dried (65 °C) to constant weight for determination of DM.

2.4. NIR data analysis

Analysis of the NIR methods involved 20 separate modelling/validation exercises. Each exercise used a different random split of the fruit in a data set into two subsets: a calibration/modelling set ($n = 180$ fruit) and a validation set ($n = 60$ fruit). The minimum standard error in cross validation (RMSEV; 10 subsets) was used to choose the number of latent variables in the model, up to a maximum of 20. Once complete, the calibration model was applied to the validation set, and regression analysis between predictions and known physical measurements was used to judge the predictive performance of the model using the statistics R^2, and the root mean square error of prediction (RMSEP). Repeating the analysis across 20 separate modelling/validation exercises allowed us to gauge the stability of the model against variation in fruit selection.

The principal modelling tool used was the partial least squares (PLS) method (MatLab, MathWorks, USA), as well as multiple linear regression (MLR). Spectral pre-processing options were investigated in conjunction with the modelling. Options included absorbance transformation, second order differentiation, area normalisation, multiple scatter correction and the standard normal variate transformation. Four separate spectral windows identified as: std (500–1050 nm); vis (500–750); NIR1 (750–1050); and NIR2 (800–1000), were studied in factorial combination with the pre-processing options, as were various levels of spectral smoothing based on the Savitsky–Golay (S–G) algorithm. The optimal combination was chosen on the basis of the minimum RMSEV, in conjunction with latent variable determination as described above.

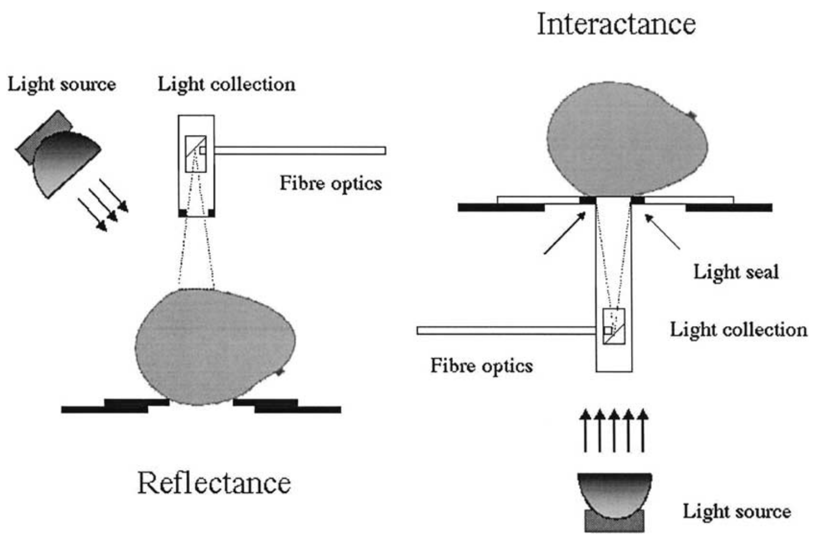

Fig. 1. Schematic of reflectance and interactance measurements.

3. Results

3.1. Fruit measurements

Summary statistics for data sets used in the NIR analysis are shown in Table 1. These were data sets corresponding to each harvest date (H1–H4) and a combined data set (Hall).

Box plots (not shown) of the DM and mass distributions looked normal and harvest date trends indicated DM increased steadily at an average 1.5% per month whilst fruit mass remained nearly constant (Table 1).

3.2. Spectral observations

Interactance spectra were characterised by broad visible light absorbance between 400 and 700 nm, a strong water absorbance around 970 nm, and a small dip around 550 nm—the wavelength of green light (Fig. 2a). Single wavelength correlations with DM were poor with the highest correlation, $r = -0.49$, being at 950 nm (Fig. 2b). There was no significant correlation associated with green light at 550 nm.

Reflectance spectra were somewhat similar in shape to the interactance spectra. The main differences were lower absorbance levels, a relatively weaker water absorbance around 970 nm, and a more pronounced dip at 550 nm (Fig. 2c). Single wavelength correlations with DM were higher for reflectance (Fig. 2d) than the interactance mode, but were still poor with the highest correlation, $r = 0.56$, occurring at 550 nm.

Table 1
Mean physical measurements for 'Hass' avocado by harvest data set

Harvest	DM (%)	Mass (g)
H1	27.7 (2)	–
H2	31.9 (3)	245 (2)
H3	34.7 (3)	234 (2)
H4	36.8 (3)	233 (3)
Hall	32.8 (2)	237 (1)

Standard error (least significant digit) listed in parenthesis. Note, no mass data were collected at H1.

3.3. Prediction of dry matter

Calibration and validation statistics for DM prediction are listed in Table 2. The best spectral pre-processing combination for the interactance mode was the NIR1 window (750–1050 nm), absorbance transformation, 9-point S–G smoothing and multiple scatter correction. The best combination for the reflectance mode was the std window (500–1050 nm), 15-point S–G smoothing and second derivative transformation. High numbers of latent variables were selected, particularly for reflectance models. This suggests the models struggled against spectral noise requiring incorporation of many small spectral features to improve accuracy.

The interactance mode was a superior predictor of fruit DM compared with reflectance (Table 2). With one exception, validation statistics (R_p^2) exceeded 0.83 and RMSEPs were $<1.8\%$ DM, whereas the corresponding reflectance results were <0.75 and $>1.9\%$ DM, respectively. There is an apparent anomaly in the interactance data concerning the H4 results, where predictions were more scattered compared with those at earlier harvests. Indeed, there is a trend of deteriorating predictive performance, as measured by RMSEP value, with increasing harvest date (Table 2). We attribute the low R_p^2 for this harvest to the non-linear relationship that exists between R_p^2 and RMSEP, such that small incremental increases in the RMSEP markedly affect the magnitude of R_p^2 (Weisberg, 1985). We suggest the trend itself is attributable to changes in fruit attributes unrelated to DM, such as fruit texture, that affect optical density and make NIR prediction of DM more difficult later in the season.

For the data set combining all 4 harvests (Hall), the best PLS interactance DM model delivered predictions that correlated well with the actual DM measurements (Fig. 3a). The model, as represented by the β-vector of regression coefficients, showed strong oscillatory behaviour across the spectral range with numerous spikes above 950 nm (Fig. 3b). A persistent spiky sequence in a β-vector regression coefficient display normally indicates over-fitting of the data. This supports the contention that the PLS models are using too

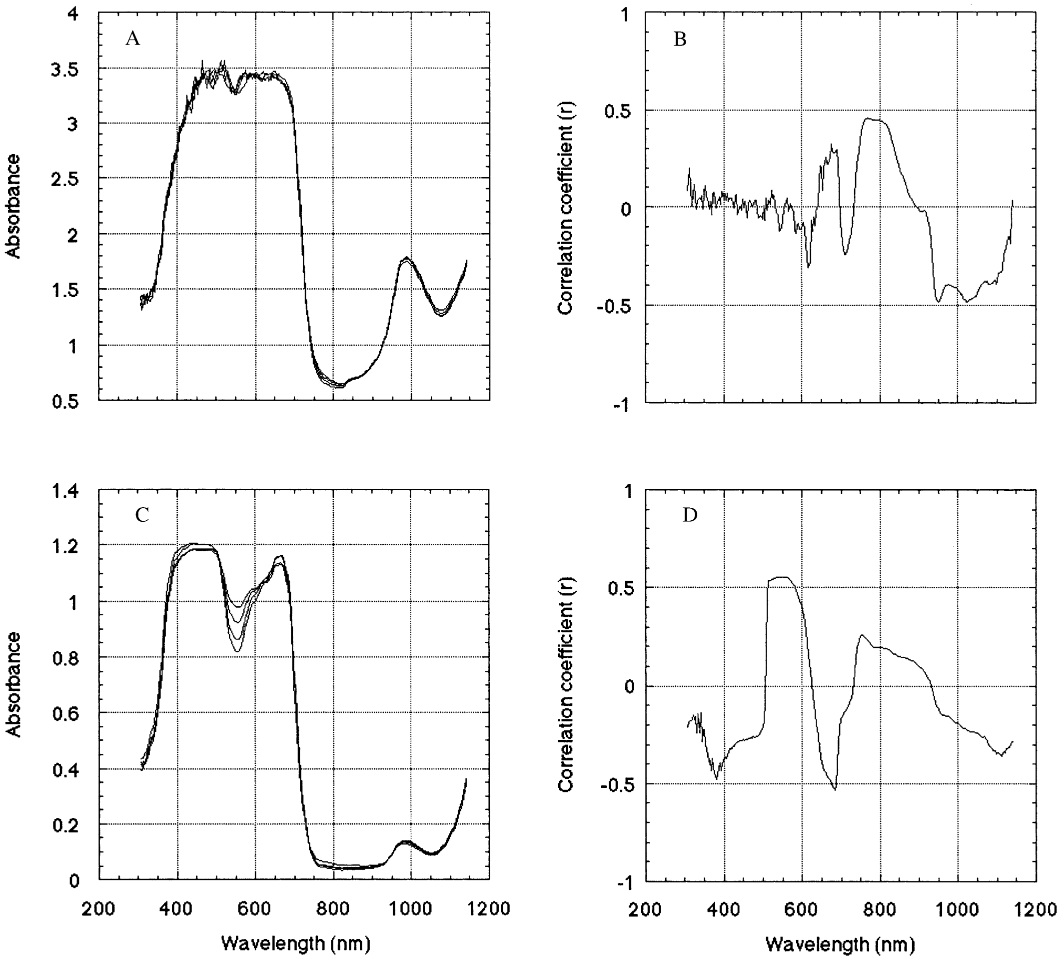

Fig. 2. Mean absorbance spectra (A, C) and correlation coefficients (B, D) for DM for interactance (A, B) and reflectance (C, D) measurements. The four curves in A and C are mean spectra taken from normalised and smoothed spectra (multiplicative scatter correction, 9-point Savitsky–Golay smoothing) at each harvest.

many latent variables. The largest regression coefficient occurred at 914 nm. This is consistent with the presence of the third overtone of the carbohydrate CH absorbance band that nominally occurs in the 900–920 nm range (Williams and Norris, 1981). This absorbance band is often cited as the most important band for DM and/or sugar determination as it is removed from troublesome interferences from the water absorbance bands that typically dominate spectra of fruit. Band assignments can be suggested for other strong coefficients in the vicinity of ~890 nm (carbohydrate), 930 nm (oil), 760, 840 and 960 nm (water). However, these assignments can only be tentative because of other peaks and troughs present hereabouts in the β-coefficient display.

To gain further insight, with less ambiguity, a global search was conducted to find the best 1-, 2-, 3- and 4-wavelength combination to use in a MLR model. The spectral pre-processing options used were the same as the corresponding PLS model. The results revealed that at least a four-wavelength model was necessary to give comparable validation statistics to the best PLS model (Table 3).

Table 2
Calibration and validation statistics for DM predictions using the interactance and reflectance NIR modes

Harvest	N_c	LV	RMSECV (% DM)	N_p	RMSEP (% DM)	R_p^2
Interactance						
H1	180	12	1.3	60	1.3	0.88
H2	179	13	1.7	60	1.7	0.83
H3	180	14	1.7	60	1.8	0.84
H4	179	13	2.3	59	2.3	0.63
Hall	718	16	1.8	239	1.8	0.88
Reflectance						
H1	180	18	1.8	60	1.8	0.78
H2	179	17	2.3	60	2.4	0.63
H3	180	16	2.8	60	2.8	0.55
H4	179	13	2.8	59	2.7	0.47
Hall	718	20	2.6	239	2.6	0.75

Data displayed are based on application of a PLS model. Abbreviations are number of observations in the calibration (N_c) and validations sets (N_p), number of latent variables (LV) used, errors of cross validation (RMSECV) and prediction (RMSEP), and goodness of fit (R_p^2).

Three of the four wavelengths selected were clustered in the vicinity of 900–920 nm. The fourth wavelength, with the weakest coefficient, was found at 765 nm, which is very close to the third overtone region for the water/carbohydrate OH absorbance band that nominally occurs at 760 nm (Williams and Norris, 1981). Hence these results confirm that the relevant spectral information is obtained primarily from a carbohydrate/lipid CH absorbance band in the 900–920 nm range, with only a minor role for water-related absorbances.

4. Discussion

The linear models produced above are suited to incorporation into on-line NIR grading systems for the determination of maturity of individual

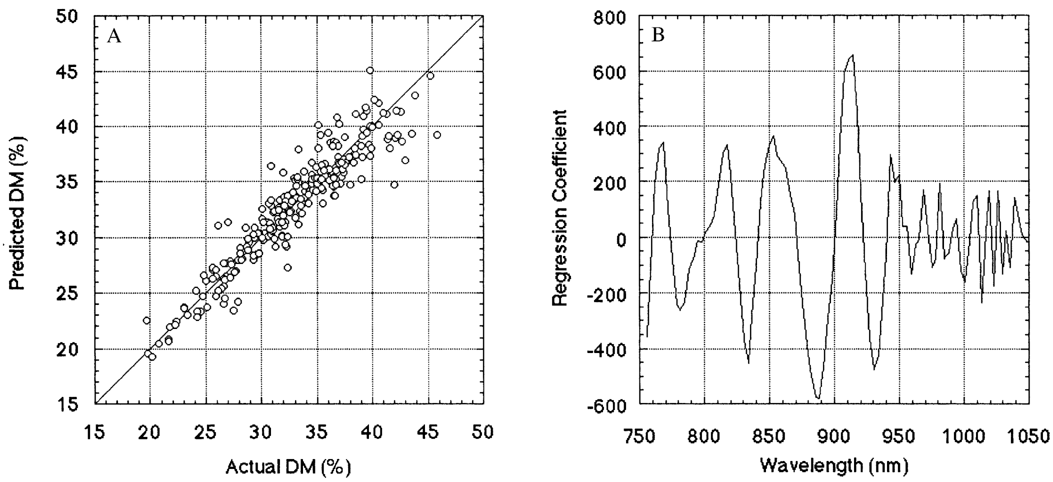

Fig. 3. Scatter-plot of interactance mode NIR predictions against actual DM measurements (A), and β coefficients for the DM model (B), for the combined data set (Hall). Data displayed are based on application of a PLS model.

Table 3
Interactance DM prediction results for 1-, 2-, 3- and 4-wavelength MLR models on the Hall data set

Model	λ (nm)	r ($\times$ 100)	RMSECV (% DM)	RMSEP (% DM)	R_p^2
1	846	−5	4.1	4.0	0.41
2	768, 875	−5, −12	3.2	3.3	0.59
3	800, 914, 1053	13, 8, 16	2.4	2.4	0.79
4	765, 901, 914, 924	−3, −40, 56, −24	2.0	2.0	0.85

The regression coefficients (r) are approximate, to indicate relative magnitude, and are listed in the wavelength order. Abbreviations are defined in Table 2.

fruit in lines of harvested avocado. This would lead to improved oil content and thus taste acceptability, especially for lines early in the season where fruit variability leads to problems with marketing. This recommendation is consistent with the preliminary findings of Schmilovitch et al. (2001).

Of the two operating modes investigated, interactance was better than reflectance for the determination of DM. Some commercial instruments, such as those from Compac (Hey, 2001), currently operate in the interactance mode.

Measurement precision would be a major consideration when transferring these findings to a commercial setting. Our error of prediction amounted to 1.8% DM over a DM range of 20–45%. The only comparative data are those of Schmilovitch et al. (2001) in which two ‘greenskin’ varieties of avocado were investigated, but not ‘Hass’, the industry standard. In their study evaluating only reflectance NIRS, errors of prediction for ‘Ettinger’ and ‘Fuerte’ were 0.9 and 1.3%, respectively, for fruits having DM contents in the range 14–24%. These prediction errors are considerably better than those we achieved. However, ‘Ettinger’ and ‘Fuerte’ are both smooth- to medium-textured, thin-skinned varieties of avocado, whereas ‘Hass’ has a thicker, peebly-textured surface, less conducive to light transmission. Prediction errors would necessarily be greater in this latter situation.

Fruit inspection times for on-line grading need to be of the order of 100 ms. Long integration times, such as those required here (1.5 s), may be avoided by using more powerful light sources and/or lens systems to efficiently focus the exit light onto the detector (Martinsen, 2002). Two- or three-wavelength detection systems further reduce the necessity for long integration times. While it might appear that dynamic measurement of avocado DM content is already feasible using NMR (Kim et al., 1999), thereby making an NIR alternative redundant, investigations to date have employed laboratory-based instruments. The obstacles to be overcome before a commercial NMR packhouse instrument was ever produced remain formidable (Hills and Clark, 2003). Further investigation of our results with existing on-line NIR technologies is, therefore, clearly warranted.

Acknowledgements

The authors would especially like to thank Hugh Moore of NZ Kiwifruit Ltd for co-ordinating fruit harvests. This work was funded by the New Zealand Foundation for Science, Research and Technology, Contract C06X0003.

References

Abbott, J.A., Lu, R., Upchurch, B., Stroshine, R., 1997. Technologies for nondestructive quality evaluation. Hort. Rev. 20, 1–120.

Armstrong, R., 2000. ‘Sweetness guaranteed’ fruit arrives in Europe. Eurofruit Magazine August, 44, 46.

Chen, P., McCarthy, M.J., Kauten, R., Sarig, Y., Han, S., 1993. Maturity evaluation of avocados by NMR methods. J. Agric. Eng. Res. 55, 177–187.

Hey, J., 2001. Compac makes strides with sweetness sorter. Eurofruit Magazine, August, 70–71.

Hills, B.P., Clark, C.J., 2003. Quality assessment of horticultural products by NMR. Ann. Rev. NMR Spectrosc.,Volume 50, in press.

Kim, S.M., Chen, P., McCarthy, M.J., Zion, B., 1999. Fruit internal quality evaluation using on-line nuclear magnetic resonance sensors. J. Agric. Eng. Res. 74, 293–301.

Lee, S.K., Young, R.E., Schiffman, P.M., Coggins, C.W., Jr, 1983. Maturity studies of avocado fruit based on picking dates and dry weight. J. Am. Soc. Hort. Sci. 108, 390–394.

Lewis, C.E., 1978. The maturity of avocado—a general review. J. Sci. Food Agric. 29, 857–866.

Malagoli, S., 2002. Sacmi gauges performance of NIR technology. Eurofruit Magazine, February 92.

Martinsen, P., 2002. A time-delayed integration spectrometer. Meas. Sci. Technol. 13, 1280–1283.

Peiris, K.H.S., Dull, G.G., Leffler, R.G., Kays, S.J., 1999. Spatial varibility of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: implications for the development and use of NIR spectrometric techniques. HortScience 34, 114–118.

Schmilovitch, Z., Hoffman, A., Egozi, H., El-Batzri, R., Degani, C., 2001. Determination of avocado maturity by near-infrared spectrometry. Acta Hortic. 562, 175–179.

Schroeder, C.A., 1985. In: Physiological Gradient in Avocado Fruit. California Avocado Society Yearbook, pp. 137–143.

Weisberg, S., 1985. Applied Linear Regression, second ed. Wiley, New York.

Williams, P.C., Norris, K.H., 1981. Qualitative applications of near-infrared reflectance spectroscopy. In: Williams, P.C., Norris, K.H. (Eds.), Near-Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemists, St Paul, MN, pp. 241–246.