

---

**C S I R O      P U B L I S H I N G**

---

# **Australian Journal of Experimental Agriculture**

Volume 37, 1997

© CSIRO Australia 1997



*... a journal publishing papers (in the soil, plant and animal sciences)  
at the cutting edge of applied agricultural research*

**[www.publish.csiro.au/journals/ajea](http://www.publish.csiro.au/journals/ajea)**

All enquiries and manuscripts should be directed to  
*Australian Journal of Experimental Agriculture*  
**CSIRO PUBLISHING**  
PO Box 1139 (150 Oxford St)  
Collingwood  
Vic. 3066  
Australia

---

Telephone: 61 3 9662 7614  
Facsimile: 61 3 9662 7611  
Email: chris.anderson@publish.csiro.au  
lalina.muir@publish.csiro.au

Published by  
**CSIRO PUBLISHING**  
in co-operation with the  
**Standing Committee on Agriculture  
and Resource Management (SCARM)**



## Non-invasive assessment of pineapple and mango fruit quality using near infra-red spectroscopy

J. Guthrie<sup>A</sup> and K. Walsh<sup>B</sup>

<sup>A</sup> Department of Primary Industries, Rockhampton, Qld 4702, Australia.

<sup>B</sup> Department of Biology, Central Queensland University, Rockhampton, Qld 4702, Australia.

**Summary.** The potential of near infra-red (NIR) spectroscopy for non-invasive measurement of fruit quality of pineapple (*Ananas comosus* var. Smooth Cayenne) and mango (*Magnifera indica* var. Kensington) fruit was assessed. A remote reflectance fibre optic probe, placed in contact with the fruit skin surface in a light-proof box, was used to deliver monochromatic light to the fruit, and to collect NIR reflectance spectra (760–2500 nm). The probe illuminated and collected reflected radiation from an area of about 16 cm<sup>2</sup>. The NIR spectral attributes were correlated with pineapple juice Brix and with mango flesh dry matter (DM) measured from fruit flesh directly underlying the scanned area. The highest correlations for both fruit were found using the second derivative of the spectra ( $d^2 \log 1/R$ ) and an additive calibration equation. Multiple linear regression (MLR)

on pineapple fruit spectra ( $n = 85$ ) gave a calibration equation using  $d^2 \log 1/R$  at wavelengths of 866, 760, 1232 and 832 nm with a multiple coefficient of determination ( $R^2$ ) of 0.75, and a standard error of calibration (SEC) of 1.21 °Brix. Modified partial least squares (MPLS) regression analysis yielded a calibration equation with  $R^2 = 0.91$ , SEC = 0.69, and a standard error of cross validation (SECV) of 1.09 °Brix. For mango, MLR gave a calibration equation using  $d^2 \log 1/R$  at 904, 872, 1660 and 1516 nm with  $R^2 = 0.90$ , and SEC = 0.85% DM and a bias of 0.39. Using MPLS analysis, a calibration equation with  $R^2 = 0.98$ , SEC = 0.54 and SECV = 1.19 was obtained. We conclude that NIR technology offers the potential to assess fruit sweetness in intact whole pineapple and DM in mango fruit, respectively, to within 1° Brix and 1% DM, and could be used for the grading of fruit in fruit packing sheds.

**Additional keywords:** Brix, non-destructive, specific gravity.

### Introduction

Fruit of the pineapple [*Ananas comosus* (L.) Merrill] is anatomically complex, being a multiple fruit (i.e. comprised of many fruitlets, each resulting from an individual flower) which also contains accessory tissue, involving expansion of bracts associated with each fruitlet. The pineapple fruit is non-climacteric (Wills *et al.* 1989). Starch storage occurs in the leaf and stalk. The fruit accumulates soluble sugars and acids, and little change in internal chemical composition occurs after harvest. Chemical composition (e.g. sugar content) varies from base to apex, and from sun to shade sides of the fruit. Fruit eating quality is related to total soluble solids (Brix), acidity, pH, Brix:acid ratio, and ester concentrations (Bowden 1967; Smith 1988b). Smith (1988b) reported that the variable best correlated with eating quality was Brix (linear relationship, coefficient of determination,  $R^2 = 0.70$ ). Fruit of less than 14 °Brix was considered unacceptable for the fresh market. Wills *et al.* (1989) report that the primary sugars in pineapple fruit are sucrose, glucose and fructose (70, 20 and 10% of total sugars, respectively).

Crop maturity is currently assessed before harvest by sacrificing fruit to measure juice Brix. External skin colour is used to assess fruit quality in the packing shed, and flesh translucency is used in processing operations. Both measures are accepted as subjective and approximate, and vary between seasons. Measurement of Brix of the juice of a cut fruit is destructive, and cannot be used to grade all fruit. A quadratic relationship between eating quality and fruit specific gravity was described by Smith (1984). This relationship may reflect the contribution of sugar content to fruit density. For example, an increase in the Brix content of a solution from 7 to 15° will increase specific gravity by 8%. Smith (1988a) recommended specific gravity as the best single index for grading intact fruit for eating quality. However, the relationship was dependent on season and geographic location, with an average  $R^2$  of only 0.28. We interpret this result as the contribution of air spaces within the fruit (infertile carpels and floral areas). Maturation of the fruit is likely to involve expansion of bract and carpel tissues, and thus a decrease in air content and an increase in specific gravity of the fruit.

In contrast, fruit of the mango (*Magnifera indica*) is a simple, single-seeded berry. As such, chemical composition varies little within the mango fruit, relative to pineapple. The fruit is climacteric (Wills *et al.* 1989), with starch stored in the fruit converted to sugars during ripening. Eating quality of the fruit is related to dry matter (DM) content, an index of starch content (Bradley and Scudamore-Smith 1987). The current Queensland industry standard for fruit harvest is a minimum of 14% DM, which is estimated by a subjective visual assessment of the colour of the flesh of a cut fruit, or determined by oven drying in a laboratory. Ripeness and suitability for processing is determined by firmness of the fruit.

Near infra-red (NIR) spectroscopy, nuclear magnetic resonance (NMR), and acoustic techniques all offer potential for the non-invasive assessment of the internal composition of intact fruit. NIR spectroscopy is the most advanced with regard to instrumentation, applications, accessories and chemometric software packages. NIR spectroscopy was first used to detect HCl in solution (Gordy and Martin 1939). Kaye (1954) developed the analytical application of NIR spectroscopy with work on spectral identification of organic compounds. Ben-Gera and Norris (1968) recognised the potential of NIR reflectance spectroscopy, in which light reflected from the sample rather than transmitted through the sample was detected, for non-invasive analysis of cereal grains and oilseeds. Protein, oil and moisture in both cereal grain and soybean seed was correlated with absorbances at 1680, 1940, 2100, 2180, 2230 and 2310 nm. Norris *et al.* (1976) correlated forage quality with absorbances at 1672, 1700, 1940, 2100, 2180 and 2336 nm. These authors suggested that (inexpensive) discrete filter instruments, as opposed to instruments incorporating a scanning monochromator, would be suitable for such analyses. Adoption of NIR technology by the food and fodder industries followed in the late 1970s to early 1980s, with application to quality assessment of flour, baked products, dairy products and forage (Shenk and Westerhaus 1993).

A useful review of NIR spectroscopy and its application in the food industry is presented by Osborne *et al.* (1993), and Shenk and Westerhaus (1993). Briefly, in discrete filter instruments, measurement of reflectance is highly influenced by light scattering by the sample, and uniform sample particle size is required. Use of the first and second derivative data treatments alleviates the problem of light scattering by an irregular surface, enabling consideration of intact fruit samples. However, such a procedure requires a scanning monochromator and increases the level of calculations required. During the mid 1980s, commercial companies began marketing NIR spectroscopy instruments incorporating a scanning monochromator or scanning filters ('tilting filters') for the analysis of agricultural products. More recently,

NIR spectroscopy has benefited from increasing sophistication of electronics, fibre optics, software and calibration techniques. In particular, better chemometric software packages have been developed which are necessary for calibrations involving the first and second derivatives of spectral absorbance.

There are more than 500 NIR reflectance and transmittance instruments in use in Australia for the off-line analysis of grain and other produce (Ronalds pers. comm.). Applications have been reported for the off-line analysis of organic acids, alcohol, amino acids, cellulose, chlorophyll, fibre, oil, fungal spores, moisture, protein, starch and sucrose in such products as biscuits, beer, cheese, flour, meat, milk, pasta and wine (Osborne *et al.* 1993), and for the in-line analysis of sugar mill streams (N. Berding pers. comm.).

Near infra-red spectroscopy has been used to non-invasively assess DM in onions (Birth *et al.* 1985), soluble solids in intact cantaloupe (Dull *et al.* 1989), peach (Kawano *et al.* 1992), mandarin (Kawano *et al.* 1993) and pineapple fruit (Shiina *et al.* 1993), and sugar content, acidity and hardness of plum fruit (Onda *et al.* 1994). Commercial (Mitsui Mining Corp., Omiya and Maki Manufacturing Co., Hamamatsu) in-line NIR spectroscopy units are used in Japanese packing sheds to assess the sweetness, ripeness and acidity of thin-skinned simple temperate fruit (citrus, apples, pears and peaches) at 3 pieces per second per lane (Kawano 1994). These units operate in either reflectance or transmittance modes, using incident white light with post-dispersive spectral analysis, rather than incident monochromatic light. NIR spectroscopy technology has not been used for grading intact fruit in Australia. In this study, we consider the application of NIR spectroscopy for the non-invasive assessment of pineapple and mango fruit eating quality. Results for pineapple are contrasted with those obtained by Shiina *et al.* (1993). These fruit present the challenges of large size, thick skin, multiple fruit and the correlation of NIR absorbance data with fruit DM (starch) content.

## Materials and methods

The experimental approach of Kawano and co-workers was adopted in this study. Kawano *et al.* (1992, 1993) reported the use of NIR spectroscopy for non-invasive assessment of Brix of peach and mandarin fruit, in reflectance and transmittance modes, respectively. The steps followed in the development of a calibration equation (choice of simple absorbance data, or first or second derivative of data, selection of 'gap' and 'segment' size, and selection of the form of the calibration equation) are standard. However, the conclusions are not universally applicable to all fresh fruit because of the effect of fruit surface characteristics and internal chemistry on the NIR reflectance spectra. We therefore anticipated recommendation of a different set of

data manipulations, as well as specific wavelengths, to that reported by Kawano and co-workers (1992, 1993).

Pineapples (var. 'Smooth Cayenne') were grown commercially at Yeppoon, Central Queensland. Fruits used for specific gravity work were derived from a field trial involving different levels of potassium fertiliser. Fruits were harvested in November 1994 (specific gravity measurements) and February 1995 (NIR spectroscopy measurements), and transported to the laboratory on the day of harvest. Specific gravity of developed fruit was determined from fruit weight and volumetric displacement of distilled water. Scanning commenced after the fruit had equilibrated to 23–25°C, and within 3 days of harvest. Mangoes (var. Kensington) grown commercially at Alton Downs, Central Queensland, were harvested 'hard green', but unripe in late December 1995. Scanning of the fruit was undertaken within 3 days of harvest at 23°C.

A scanning spectrophotometer (Model 6500, NIRSystems, Silver Springs, MD, USA) using NSAS software (Version 3.25, NIRSystems), was connected via a 1.5 m fibre optic cable to a remote reflectance probe. In this configuration, incident monochromatic light travels through the fibre optic cable to the sample, and visible and NIR detectors in the probe monitor the intensity of reflected light. With each scan, reflected light intensity was assessed over 2-nm intervals across the spectral range 400–2500 nm. Each sample's spectrum was the mean of 50 scans (<1 min). Intact fruit were held in a light-proof box with a 60 mm diameter opening. A laboratory jack held the fruit against the opening so that the fruit skin was in direct contact with the quartz glass window of the probe. For each pineapple fruit, spectra were measured at each of 4 equidistant locations around the middle circumference of the fruit. With mango, a spectrum was taken from each of the 2 faces (cheeks) of the fruit. A reference spectrum (mean of 50 scans) was obtained using teflon as the reference before collection of each sample spectrum. Teflon is reported to have low absorption and similar light-scattering characteristics to plant material (Birth *et al.* 1985; Dull *et al.* 1989; Kawano *et al.* 1992, using onion, cantaloupe and peach, respectively) and is therefore used

to assess 'incident' light intensity relative to that reflected from the sample surface.

After scanning, a 60 mm diameter stainless steel corer was used to excise both skin and underlying flesh to 2 cm depth. For pineapple, the skin was removed and about 20 mm of the flesh homogenised in a laboratory stomacher for 2 min. The resulting slurry was centrifuged at 14 000 rpm for 4 min, and Brix of the supernatant determined in duplicate using an Erma (Tokyo, Japan) digital refractometer (accuracy  $\pm$  0.2 °Brix). Mango fruit tissue samples were obtained in the same manner as pineapple, but finely diced (<3 mm) and DM determined by drying in a convection oven at 70°C for 48 h.

Calibration equations were derived from spectral data using multiple linear regression (MLR) and modified partial least squares (MPLS) analysis (ISI Version 3.0, Infrasoft International, PA, USA). NSAS software was found to offer superior instrument operation and diagnostics, whereas ISI software algorithms for MPLS were considered superior to the NSAS PLS algorithms. The 94 pineapple samples were ranked by analyte level, and a prediction set (10% of population) was randomly chosen from each of 9 equal Brix ranges. This procedure produced a prediction set equally weighted for analyte level across the available range. The remaining population ( $n = 85$ ; ranging from 6.9 to 16.5 °Brix, mean  $\pm$  s.d.  $12.0 \pm 2.34$ ) was used as a calibration set. The prediction set was characterised by a range of 7.0–16.6 °Brix, mean  $12.0 \pm 3.32$ . The fitness of the MLR regression was characterised within the calibration and prediction populations by standard error of calibration (SEC) and prediction (SEP), respectively.

A cross validation procedure (Shenk and Westerhaus 1993) was used in the MPLS regression analysis. Cross validation calculates validation errors by partitioning the population into several groups (6 in this case). Each group is sequentially removed from the population and a calibration established on the remaining population (i.e. 6 calibrations undertaken). Each calibration is applied to all members of the removed group until every sample is predicted once. Validation errors are combined into a standard error of cross validation (SECV). The SECV

**Table 1. Spatial variation in total soluble solids (Brix) within pineapple fruit**

Fruit were divided into northern and southern sides (sun and shade sides), crown (top), middle and base, and flesh adjacent to skin (outer), flesh (middle) and stalk tissue (inner) segments  
Data are presented as means  $\pm$  s.e. ( $n = 5$ )

Side	Northern side			Southern side		
	Skin	Flesh	Stalk	Skin	Flesh	Stalk
Crown	8.3 $\pm$ 1.2	9.5 $\pm$ 1.1	9.1 $\pm$ 0.8	7.6 $\pm$ 0.9	8.6 $\pm$ 1.2	7.6 $\pm$ 1.5
Middle	10.7 $\pm$ 1.2	12.4 $\pm$ 1.2	9.8 $\pm$ 1.8	9.9 $\pm$ 1.2	11.8 $\pm$ 1.6	9.1 $\pm$ 2.3
Base	11.7 $\pm$ 0.5	14.2 $\pm$ 0.9	12.0 $\pm$ 2.0	11.5 $\pm$ 0.7	14.0 $\pm$ 1.9	10.6 $\pm$ 1.7

value ideally should be close to the SEC value. MPLS procedure is efficient in that all samples are used for both calibration and validation.

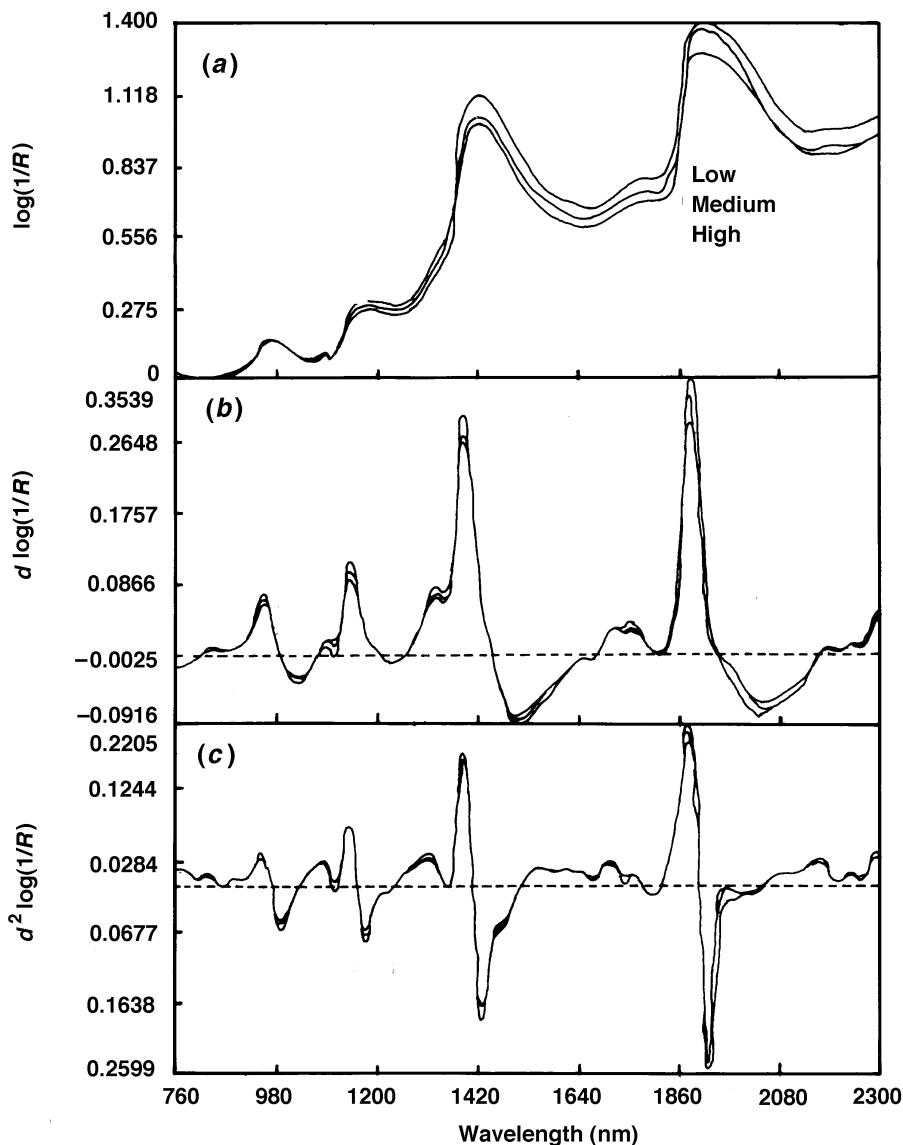
Near infra-red spectra of mango fruit were assessed as for pineapple, before assessment of DM. The 2 sides of the fruit were not different in terms of DM content (data not shown). Consequently data from only one side of the fruit was used in predictive calculations. A calibration set of 62 mango samples with a range of 11.0–24.5% DM, mean  $\pm$  s.d.  $18.0 \pm 3.65$ , was used in MLR analysis. The prediction set consisted of 7 samples, with a range of

12.8–22.2%, mean  $\pm$  s.d.  $17.9 \pm 3.34$ . The prediction samples were selected as indicated for pineapple.

## Results and discussion

### Pineapple

*Specific gravity as a measure of Brix in pineapple.* Fruit specific gravity was found to be poorly correlated with Brix (quadratic relationship:  ${}^{\circ}\text{Brix} = 58 - 95 \text{ SG} + 48 \text{ SG}^2$ ,  $R^2 = 0.2$ ; linear relationship:  ${}^{\circ}\text{Brix} = 9.5 + 1.5 \text{ SG}$ ,  $R^2 = 0.03$ ; where SG is specific gravity). We conclude that specific gravity is an unreliable indicator of Brix



**Figure 1.** Near infra-red spectral data of three intact pineapple, representative of high, medium and low Brix fruit. (a) Original [ $\log(1/R)$ ], (b) first derivative [ $d \log(1/R)$ ] and (c) second derivative [ $d^2 \log(1/R)$ ], spectra plotted against wavelength. The dashed line (b and c) represents the zero baseline.

content with fruit grown in Central Queensland. This result is consistent with that of Smith (1984, 1988b), who reported a poor overall relationship between eating quality and specific gravity, with variation between seasons and localities. For fruit at a given locality, Smith (1988a) described a quadratic relationship, but noted the relationship also depended on the 'uniformity of fruit in each farm block'. In our case, fruit were derived from plants grown under various potassium fertilisation regimes and a more tropical climate, which may have upset the relationship between Brix and specific gravity expected by Smith. As noted earlier, the air content of the fruit may vary in response to a range of factors unrelated to Brix content, overriding the contribution of Brix to specific gravity. Our attention was then focussed on the use of NIR spectroscopy for the non-invasive assessment of Brix in fruit.

*Brix variation within pineapple fruit.* According to Smith (1984), the Brix content of the bottom of the pineapple fruit was always about 3 °Brix higher than the top. We found a similar variation between top and bottom of the fruit (Table 1). Also, the northern side of the fruit ('sun side') was generally 1 °Brix higher than the 'shade side' (Table 1). The Brix content of the inner vascular tissue (stalk) of the fruit was generally several units less than the central flesh, and slightly lower than that of the flesh adjacent to the skin. The Brix of tissue adjacent to the skin was highly correlated with that of

**Table 2. Norris regression (NSAS) of first and second derivative pineapple spectral data (760–1300 nm) ( $n = 85$ )**

The correlation coefficient ( $R$ ) for each regression is presented with an associated wavelength (nm) in parentheses as calculated for a range of gap and segment sizes

Gap size is the wavelength change between successive means; segment size is the wavelength range for which the absorbance data are averaged for each data point in the correlation

Gap size (nm)	Segment size (nm)			
	8	12	16	20
<i>First derivative [d log(I/R)]</i>				
0	0.730 (784)	0.730 (784)	0.731 (786)	0.731 (788)
4	0.730 (784)	0.730 (786)	0.731 (786)	0.732 (788)
8	0.730 (784)	0.731 (786)	0.732 (788)	0.732 (790)
12	0.731 (786)	0.731 (788)	0.732 (788)	0.732 (790)
16	0.731 (786)	0.732 (788)	0.732 (790)	0.731 (790)
<i>Second derivative [d<sup>2</sup> log(I/R)]</i>				
0	-0.711 (800)	-0.729 (802)	-0.743 (802)	-0.749 (802)
4	-0.725 (800)	-0.740 (802)	-0.748 (802)	-0.753 (804)
8	-0.737 (802)	-0.747 (802)	-0.752 (804)	-0.755 (806)
12	-0.746 (802)	-0.751 (804)	-0.754 (806)	-0.755 (808)
16	-0.750 (804)	-0.754 (804)	-0.755 (808)	-0.774 (1220)

adjacent flesh (data not shown), being generally 1–2 units lower than adjacent flesh (Table 1).

*Near infra-red spectra of intact pineapple fruit.* The rough skin of pineapple presents difficulties, relative to

**Table 3. Multiple linear regression (NSAS) of first and second derivative of pineapple spectra for calibration ( $n = 85$ ) and prediction ( $n = 9$ ) populations**

The calculated multiple regression coefficient (MR) and associated standard error of calibration (SEC), and a standard error of prediction (SEP) and associated bias are presented for each additional wavelength term used in each of two models

The form of the regression equations is explained in the text (equations 1 and 2)

$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	MR	SEC (°Brix)	SEP (°Brix)	Bias (°Brix)
<i>d log(I/R) for equation 1</i>							
788				0.68	1.68	1.28	-0.63
788	816			0.72	1.60	0.98	-0.25
788	816	850		0.75	1.53	0.91	0.02
788	816	850	1082	0.76	1.51	0.91	0.04
<i>d log(I/R) for equation 2</i>							
788	970			0.72	1.65	1.26	-0.62
788	970	816	1216	0.76	1.57	2.97	-5.01
<i>d<sup>2</sup> log(I/R) for equation 1</i>							
866				0.73	1.61	2.03	-0.09
866	760			0.82	1.37	0.77	-0.46
866	760	1232		0.86	1.23	0.79	-0.06
866	760	1232	832	0.86	1.21	0.75	0.10
<i>d<sup>2</sup> log(I/R) for equation 2</i>							
866	778			0.68	1.73	1.64	-0.57
866	788	864	1234	0.81	1.41	1.37	0.42

the smooth skin of mango, for reflectance spectroscopy. The original reflectance spectra of the skin of pineapple fruit between 760 and 1300 nm was subject to baseline shift in response to varying levels of radiation reflected from the sample's irregular surface and other factors, such as skin thickness. Consequently, the actual absorbance readings ( $\log 1/R$ , where  $R$  is reflectance) are difficult to relate to Brix content (Fig. 1a). The slope (Fig. 1b), and rate of change in slope (Fig. 1c) (first and second derivatives, respectively, of absorbance), were found to minimise such background effects as discovered by earlier workers. A first order derivative results in a curve containing peaks and valleys that correspond to the point of inflection on either side of the absorbance peak. This presentation is difficult to interpret. The second order derivative results in the display of peaks pointing down instead of up, with peak location similar to the  $\log (1/R)$  plot. Note that features associated with these peaks may be positively or negatively correlated with the analyte of interest. For example, an increase in specific water content of a fruit could be associated with a decrease in specific sugar content.

*Calibration between near infra-red spectroscopy absorbance and Brix of pineapple fruit.* The NSAS Norris regression program was applied to varying gap (wavelength change between successive mean values) and segment (absorbance data averaged over a range of wavelengths) sizes for spectral data between 760 and 1300 nm to find the best derivative condition for the multiple linear regression analysis. Spectral data (2-nm resolution) were averaged over 4 'segment' sizes (i.e. for a 20 nm segment, 10 values of the original spectral data set were averaged), and for each of 5 'gap' sizes for correlation analysis of both first and second derivative spectral data against Brix (Table 2). Second derivative data in the majority of cases yielded better correlations than first derivative data. In general, varying segment and gap size had little effect on the correlation coefficient ( $R$ ) or the wavelength in the correlation. The small change in wavelength (784–790 nm for first derivative data, and 800–806 nm for second derivative data) was not significant considering the resolution of the original data (2 nm). The highest correlation coefficient was obtained with a gap size of 16 nm and a segment size of 20 nm, using the second derivative of absorbance at 1220 nm. However, optimal data smoothing was achieved by the combination of a gap size of 0 and a segment size of 20 nm. Using these parameter settings, the highest correlation coefficient was obtained using the second derivative of absorbance at 802 nm. These settings were therefore used in subsequent analyses (J. Guthrie and K. Walsh unpublished data). In contrast, Kawano *et al.* (1992) recommended the use of a segment size of 16 nm and a gap size of 4 nm in

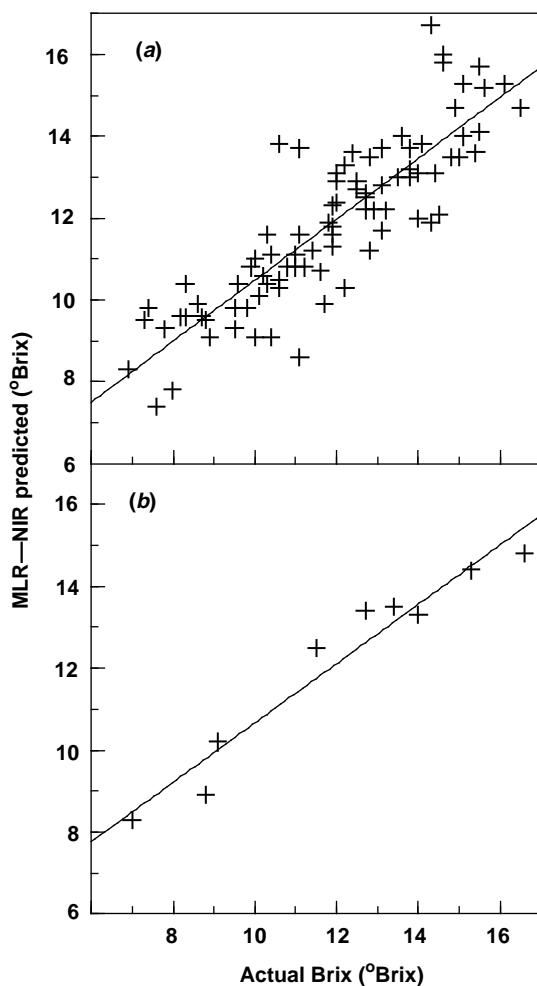
their study of intact peaches. The optimal parameter settings for spectral data treatment are expected to vary between fruits, as noted in the Introduction.

The NSAS multiple linear regression procedure was applied to first and second derivative data, using up to 4 wavelength terms within 2 equation forms for the relationship between analyte and wavelength data:

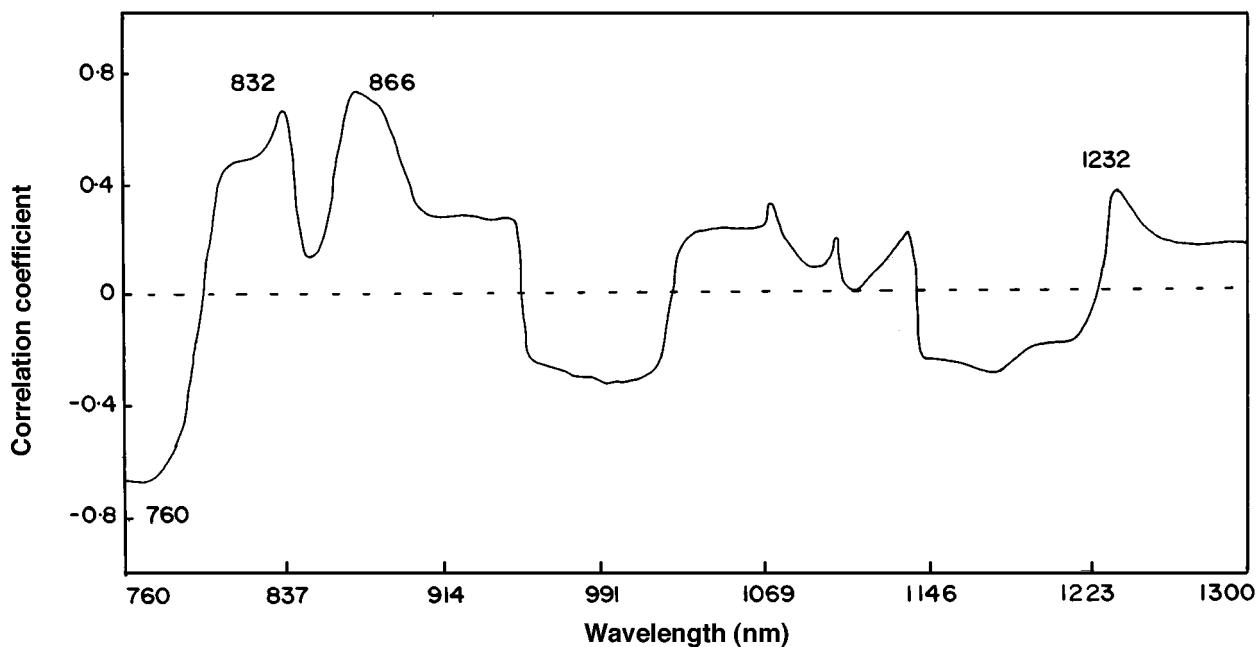
$$\text{Predicted Brix} = k_0 + k_1 d^2 \log(1/R\lambda_1) + k_2 d^2 \log(1/R\lambda_2) \\ + k_3 d^2 \log(1/R\lambda_3) + k_4 d^2 \log(1/R\lambda_4) \quad (1)$$

$$\text{Predicted Brix} = k_0 + k_1 d^2 \log(1/R\lambda_1)/k_2 d^2 \log(1/R\lambda_2) \\ + k_3 d^2 \log(1/R\lambda_3)/k_4 d^2 \log(1/R\lambda_4) \quad (2)$$

where  $k_0$  represents the intercept of the regression



**Figure 2.** Relationship between Brix predicted from a MLR-derived calibration and measured Brix in pineapple fruit for (a) calibration and (b) prediction populations. The MLR procedure involved the second derivative of absorbance data, following equation 3 (in text). Data of population used in Figures 3 and 5. In the calibration regression  $R^2 = 0.76$  and SEC = 1.21 °Brix; in the prediction regression  $R^2 = 0.95$  and SEP = 0.75 °Brix.



**Figure 3.** Correlation coefficients between the second derivative of  $\log(1/R)$  (where  $R$  is reflectance) and Brix values, plotted against wavelength (NSAS program) for pineapple. Fruit was harvested in February 1995. Data of population used in Figures 2 and 4.

equation,  $k_1$ ,  $k_2$ ,  $k_3$  and  $k_4$  represent constants, and  $d^2 \log(1/R)\lambda_n$  represents the second derivative of absorbance ( $\log$  of inverse of reflectance) data for wavelengths  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$ . In empirical terms, the second equation form (equation 2) can allow for a corrective term (e.g. to accommodate thickness of the object in transmission studies). Dull *et al.* (1989) developed a calibration for Brix in cantaloupe which used equation 2, whereas Kawano *et al.* (1992) used equation 1 in a calibration for Brix in peach. Again the optimal data treatment is expected to vary between fruit types and must be decided for the fruit of interest.

A higher multiple correlation coefficient (MR) and a lower SEC were achieved using second rather than first derivative data, and using calibration equation 1 rather than equation 2 (Table 3). Most explanation of data variance in the additive calibration (equation 1) of the second derivative data was achieved with the use of 2 wavelengths (866 and 760 nm). Incorporation of extra wavelengths (1232 and 832 nm) into the correlation equation (equation 3, also Fig. 2a) improved the regression marginally, yielding a multiple regression coefficient of 0.87 (coefficient of determination  $R^2 = 0.76$ ) with a SEC of 1.21. The equation is:

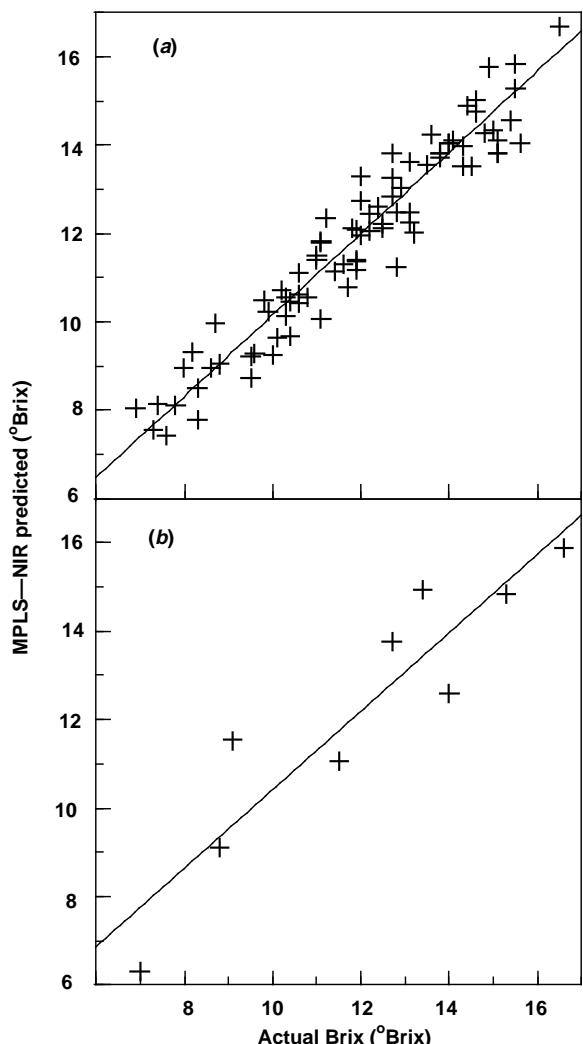
$$\begin{aligned} \text{Predicted Brix} = & 8.45 + 1596 d^2 \log(1/R\lambda_{866}) \\ & - 42 d^2 \log(1/R\lambda_{760}) - 2182 d^2 \log(1/R\lambda_{1232}) \\ & + 464 d^2 \log(1/R\lambda_{832}) \end{aligned} \quad (3)$$

Applied to a prediction set ( $n = 9$ , Fig. 2b), this equation yielded a SEP of 0.75 with a bias of 0.10. This

prediction set was randomly chosen, but was of a small sample number. The low value of the SEP, relative to SEC, reflects the size of the validation set in relation to the calibration set. Nonetheless, the SEC and SEP of around 1 °Brix is promising for the application of this technology to the grading of fruit quality. The decrease in SEP with each added wavelength term (for second but not first derivative data) indicates the data has not been overfitted.

The plot of correlation coefficients between second derivative data and Brix against wavelength (Fig. 3) illustrated the strong negative correlation at 760 nm, and positive correlations around 800–832, 866 and 1232 nm. As noted above, in the second derivative a positive correlation between absorbance and Brix is indicated by a negative correlation between the second derivative data and Brix. Thus absorbance around 760 nm is positively correlated with Brix, while absorbance at the other wavelengths is negatively correlated with Brix. Kawano *et al.* (1992) suggested that the negative correlation is likely to involve water, as water content decreases as sugar content increases.

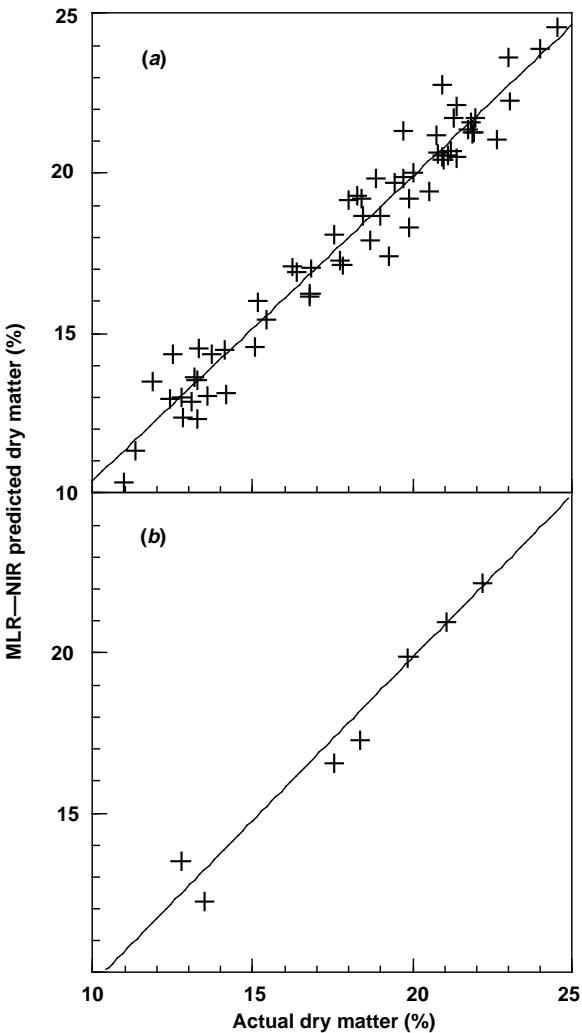
An absorption spectrum was obtained for sucrose in solution that matched that reported by Kawano *et al.* (1992). Three characteristic absorption bands were noted at 918, 1002 and 1175 nm and a weak band was noted around 760 nm. The latter 2 bands overlapped those of water (peaks at 1002 and 1175 nm) and therefore represent the presence of water in the sugar solution (Kawano *et al.* 1992). The positive correlation observed



**Figure 4.** Relationship between Brix predicted from a MPLS-derived calibration and measured Brix in pineapple fruit for (a) calibration and (b) prediction populations. The modified partial least squares (MPLS) regression procedure involved the second derivative of absorbance data. Data of population used in Figures 2 and 4. In the calibration regression  $R^2 = 0.91$ , SEC = 0.69 and SECV = 1.100  $^{\circ}$ Brix; in the prediction regression  $R^2 = 0.85$  and SEP = 1.2  $^{\circ}$ Brix.

between Brix and NIR spectroscopy absorbance at 760 nm and the negative correlation at 1232 nm is expected from the absorbance spectra of sugar and water, respectively. The presence of negative correlations between absorbances at 800–832 and 866 nm implies that factors other than water content may be inversely related to Brix level.

A MPLS regression analysis was performed with second derivative data of the same population as used for calibration in the MLR procedure. The resulting



**Figure 5.** Relationship between dry matter predicted from a MLR-derived calibration and measured dry matter content of mango fruit for (a) calibration and (b) prediction populations. The MLR procedure involved the second derivative of absorbance data, following equation 3 (in text). In the calibration regression  $R^2 = 0.90$  and SEC = 0.8% DM; in the prediction regression  $R^2 = 0.96$  and SEP = 0.79% DM.

calibration was applied to the same prediction set as used in the MLR procedure and yielded 8 terms in the equation. The calibration and prediction populations were as used for the MLR procedure. The resulting calibration ( $R^2 = 0.91$ , SEC = 0.69 and SECV = 1.1) and prediction regressions ( $R^2 = 0.85$ , SEP = 1.2) (Fig. 4a and b) better described the data than that derived by MLR (Fig. 2).

The MPLS regression assigns a weighting to each wavelength term in the spectrum, for each term used. Thus this type of regression requires a full NIR spectrum to be acquired on each sample assessed. Use of the

**Table 4. Assessment of the depth of pineapple fruit tissue contributing to the near infra-red spectroscopy spectra**

Predicted Brix is calculated (equation 3 in text) of spectra collected from two fruit samples as flesh was sequentially trimmed from the side away from the skin, which was against the spectroscopy remote reflectance unit

Thickness of flesh under skin (mm)	Fruit sample 1 (Predicted °Brix)	Fruit sample 2 (Predicted °Brix)
90	12.314	12.206
80	11.906	11.920
70	12.208	12.173
60	12.295	12.661
50	11.971	12.069
40	12.130	12.336
30	12.432	12.240
20	12.385	12.193
15	12.108	12.281
10	11.809	12.129
5	11.567	11.949

MLR-derived calibration (i.e. second derivative data of absorbance at 4 wavelengths) would allow fewer detectors to be used in an in-line instrument. Application of NIR spectroscopy for the detection of fruit quality in an in-line packing or processing setting will require a rapid detection system. For example, the video-based colour and size sorting technology developed by Colour Vision Systems can assess 10 items of fruit per second per lane (C. Esson, Colour Vision Systems, pers. comm.). Therefore the MLR-derived regression is more suitable for adoption into an in-line setting, using an array of photodiodes to simultaneously record the intensity over desired wavelength ranges.

Shiina *et al.* (1993) reported on the use of NIR reflectance spectroscopy to assess pineapple Brix and acidity, using MLR (NSAS) and the pineapple cultivar (Smooth Cayenne) used in the current study. NIR spectra of the skin side of pineapple slices were correlated with Brix of associated flesh, and a multiple correlation coefficient of 0.89 ( $R^2 = 0.79$ ) with a bias-corrected SEC of 1.06 °Brix was obtained using second derivative data of reflectance at 8 wavelengths (788, 1604, 1140, 932, 2004, 1884, 1684 and 1652 nm). These wavelengths were considerably different to those determined in this study. These differences will reflect the use of different instrumentation (e.g. different light intensity-depth of penetration of fruit), data treatment conditions (e.g. wavelength range in the current study was restricted between 760 and 1300 nm), and pineapple growing conditions. For example, in a comparison of pre- and post-dispersive optical configurations (J. Guthrie and K. Walsh unpublished data), we attempted a calibration of a population (208 samples) of pineapples harvested in December 1995. The optimum segment and gap size as

determined in the current study were used in the calibration of the December group. This calibration yielded a MR of 0.72 with a SEC of 1.30 using second derivative data of 766, 878, 1018 and 760 nm wavelengths (data not shown). These wavelengths were similar to, but not identical with, those selected with respect to the February population (Table 3). These variations in wavelengths highlight the need to establish a robust calibration, or series of calibrations, to accommodate seasonal and locality variation.

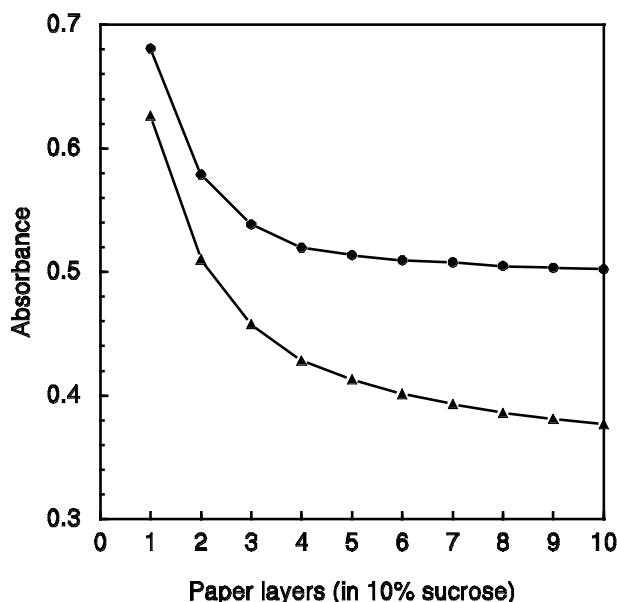
As noted above, different instrumentation will involve different levels of incident radiation, and thus effectively different volume and depth of the fruit will be assessed. We found calibrations produced for the fruit surface to differ from that derived of the internal flesh, and also of the flesh side of <10-mm thick strips of the fruit skin (data not shown). We conclude that the calibration is specific to the instrumentation used, in terms of the depth and area of the fruit which is assessed.

#### Mango

Data treatment conditions were optimised following the procedure outlined above for pineapple fruit. Multiple linear regression of second derivative absorbance data against DM content yielded a calibration equation using wavelengths at 904, 872, 1660 and 1516 nm with  $R = 0.95$  ( $R^2 = 0.90$ ) and SEC = 0.849 (Fig. 5a). Applied to the prediction population, the equation yielded SEP = 0.79% DM with a bias of 0.39 (Fig. 5b). MPLS analysis gave a calibration equation with  $R^2 = 0.98$ , SEC = 0.538, and SECV = 1.186 (data not shown). Applied to the prediction population, this relationship yielded a  $R^2 = 0.96$  and a SEP = 0.79% DM. As for pineapple, the MPLS regression procedure yielded a calibration that better described the data than MLR.

*Depth of wavelength penetration into sample.* Two attempts were made to assess the depth of tissue that contributed to the NIR spectrum of an intact fruit. Near infra-red spectra were obtained from a pineapple fruit as the fruit was sequentially sliced from beneath the skin. Predicted Brix (applying equation 3) was expected to change rapidly once the sample was cut to a size less than the effective depth of penetration of the relevant wavelengths. There was no significant change in predicted Brix as the sample depth was decreased from 90 to 5 mm (Table 4). Production of slices thinner than 5 mm was not possible.

In another approach, NIR spectra were obtained from accumulated layers of cellulose filter paper soaked in 10% sucrose solution. Absorbances at 866 and 965 nm were used. Absorbance at 866 nm was the primary term in the pineapple calibration (equation 3). Good separation of second derivative spectral data of the filter papers soaked in sucrose solution was evidenced at 965 nm (data not shown), and is probably absorbance by water. Absorbance was initially high due to absorbance by the



**Figure 6.** Relationship between absorbance ( $\log 1/R$ , where  $R$  is reflectance) and the number of layers of filter paper soaked in a 10% sucrose solution. Absorbance at 965 (●) and 866 nm (▲) is shown.

background. As layers were added, absorbance decreased (Fig. 6). A steady reading was obtained with addition of greater than 4 layers at 965 nm, and 10 layers at 866 nm, equivalent to 2.2 and 5.4 mm in depth, respectively. Apparently depth of penetration was dependent on wavelength. We conclude that all spectral information obtained from fruit was derived from within 5 mm of the fruit surface, with an exponential decrease in information with increasing depth within the 5 mm zone.

### Conclusion

The irregular skin surface, thick skin and spatially variable composition of the pineapple fruit and the large seed and complex chemistry of the mango fruit were anticipated to present problems in the analysis by NIR spectroscopy. The problems of variable surface reflection were reduced by use of the second derivative procedure. The intensity of light used allowed collection of spectral data from only the surface millimetres of the fruit, avoiding interference from the mango seed. However, the Brix/DM of this surface layer was correlated with that within the fruit (i.e. the measured value).

Near infra-red reflectance spectroscopy allowed a prediction of Brix in pineapple to within 1 °Brix unit, and of DM in mango to within 1%. Indeed, if samples were classified into 3 groups using a calibration of  $R^2 = 0.76$  (pineapple) or 0.90 (mango), 77.5 and 81.4%, respectively, of the samples would be correctly classified (interpolation of table presented in Shenk and Westerhaus 1993). The Brix of pineapple fruit in the fresh market

varies from 8 to 21°. Grading of fruit into a nominal 3 grades (low 8–10 °Brix; medium 10–13 °Brix; high 13–18 °Brix) would be useful for marketing. Grading of mango fruit into 2 groups, relative to the industry minimum standard of 14% DM is suggested as a useful exercise. To cope with variation in quality within a fruit (e.g. Brix within the pineapple fruit), NIR spectroscopy assessment could be made at either a set location on the fruit or a scan of the whole fruit.

We conclude that NIR spectroscopy is appropriate to the non-invasive assessment of pineapple and mango fruit quality. Further work is required to establish a robust calibration to accommodate seasonal and locality variations, followed by development of an in-line assessment procedure capable of assessment at a rate of items per second.

### Acknowledgments

This work was made possible through the technical support and the loan of instrumentation from Linbrook and the Sugar Research Institute. We also thank Nils Berding (BSES), Sumio Kawano (National Food Research Institute, Japan), Randolph Pax (SRI-USQ) and Hilton Deeth (UQ) for advice on NIR spectroscopy, Allan Lisle and Kerry Bell for statistical advice, and Bing Liu and Brett Wedding for technical assistance. Funding support from Central Queensland University, and the supply of fruit by Yeppoon Pines (Mick and Vince Cranny) and Mango Magic (Graham and Meredith Fossett) is acknowledged. Total soluble solids variation within the pineapple fruit was assessed by Sandra Coleman as part of an undergraduate project.

### References

- Ben-Gera, I., and Norris, K. H. (1968). Direct spectrophotometric determination of fat and moisture in meat products. *Journal of Food Science* **33**, 64–7.
- Birth, G. S., Dull, G. G., Renfroe, W. T., and Kays, S. J. (1985). Nondestructive spectrophotometric determination of dry matter in onions. *Journal of the American Society of Horticultural Science* **110**, 297–303.
- Bowden, R. P. (1967). Translucency as an index of ripeness in pineapples. *Food Technology in Australia* **19**, 424–7.
- Bradley, B. F., and Scudamore-Smith, P. D. (1987). Dry matter, a maturity standard for freezing mangoes. *Food Technology in Australia* **39**, 298–300.
- Dull, G. G., Birth, G. S., Smittle, D. A., and Leffler, R. G. (1989). Near infra-red analysis of soluble solids in intact cantaloupe. *Journal of Food Science* **54**, 393–5.
- Gordy, W., and Martin, P. C. (1939). The infra-red absorption of HCl in solution. *Journal of Chemistry and Physics* **7**, 99–102.
- Kawano, S. (1994). Present condition of nondestructive quality evaluation of fruits and vegetables in Japan. *Japan Agricultural Research Quarterly* **28**, 212–16.
- Kawano, S., Fujiwara, T., and Iwamoto, M. (1993). Nondestructive determination of sugar content in satsuma mandarin using near infra-red (NIRS) transmittance. *Journal of the Japanese Society of Horticultural Science* **62**, 465–70.

- Kawano, S., Watanabe, H., and Iwamoto, M. (1992). Determination of sugar content in intact peaches by near infra-red spectroscopy with fibre optics in interactance mode. *Journal of the Japanese Society of Horticultural Science* **61**, 445–51.
- Kaye, W. (1954). Near infra-red spectroscopy. 1. Spectral identification and analytical application. *Spectrochimica Acta* **6**, 257–87.
- Norris, K. H., Barnes, R. F., Moore, J. E., and Shenk, J. S. (1976). Predicting forage quality by infra-red reflectance spectroscopy. *Journal of Animal Science* **43**, 889–97.
- Onda, T., Tsuji, M., and Komiyama, Y. (1994). Possibility of nondestructive determination of sugar content, acidity and hardness of plum fruit by near infra-red spectroscopy. *Journal of the Japanese Society for Food Science and Technology* **41**, 908–12.
- Osborne, B. G., Fearn, T., and Hindle, P. H. (1993). 'Practical NIR Spectroscopy with Applications in Food and Beverage Analysis.' 2nd Edn. (Longman Scientific and Technical: United Kingdom.)
- Shenk, J. S., and Westerhaus, M. O. (1993). 'Analysis of Agriculture and Food Products by Near Infra-red Reflectance Spectroscopy.' NIRS Handbook. (Infrasoft International, marketed by NIRSystems, Inc.: Silver Springs, MD, USA.)
- Shenk, J. S., Workman, J. J., and Westerhaus, M. O. (1992). Application of NIRS spectroscopy to agriculture products. In 'Handbook of Near Infra-red Analysis'. (Eds D. A. Burns and E. W. Ciurczak.) pp. 383–481. (Marcel Dekker Inc.: New York.)
- Shiina, T., Ijiri, T., Matsuda, I., Sato, T., Kawano, S., and Ohoshiro, N. (1993). Determination of Brix value and acidity in pineapple fruits by near infra-red spectroscopy. *Acta Horticulturae* **334**, 261–72.
- Smith, L. G. (1984). Pineapple specific gravity as an index of eating quality. *Tropical Agriculture (Trinidad)* **61**, 196–9.
- Smith, L. G. (1988a). Indices of physiological maturity and eating quality in Smooth Cayenne pineapples. 1. Indices of physiological maturity. *Queensland Journal of Agricultural and Animal Sciences* **45**, 213–18.
- Smith, L. G. (1988b). Indices of physiological maturity and eating quality in Smooth Cayenne pineapples. 2. Indices of eating quality. *Queensland Journal of Agricultural and Animal Sciences* **45**, 219–28.
- Wills, R. B. H., McGlasson, W. B., Graham, D., Lee, T. H., and Hall, E. G. (1989). 'Postharvest: an Introduction to the Physiology and Handling of Fruit and Vegetables.' (New South Wales University Press: Sydney.)

Received 6 March 1996, accepted 27 November 1996

