CSIRO PUBLISHING
www.publish.csiro.au/journals/ajar *Australian Journal of Agricultural Research*, 2004, **55**, 471–476

Assessment of quality defects in macadamia kernels using NIR spectroscopy

John Guthrie[A,D], *Colin Greensill*[C], *Ray Bowden*[A], *and Kerry Walsh*[B]

[A]Agency for Food and Fibre Sciences, Queensland Department of Primary Industries, Rockhampton, Qld 4702, Australia,
[B]Plant Sciences Group, Central Queensland University, Rockhampton, Qld 4702, Australia
[C]Faculty of Engineering and Physical Systems, Central Queensland University, Rockhampton, Qld 4702, Australia.
[D]Corresponding author; email: john.guthrie@dpi.qld.gov.au

Abstract. Spectral data were collected of intact and ground kernels using 3 instruments (using Si-PbS, Si, and InGaAs detectors), operating over different areas of the spectrum (between 400 and 2500 nm) and employing transmittance, interactance, and reflectance sample presentation strategies. Kernels were assessed on the basis of oil and water content, and with respect to the defect categories of insect damage, rancidity, discoloration, mould growth, germination, and decomposition. Predictive model performance statistics for oil content models were acceptable on all instruments (R^2 >0.98; RMSECV <2.5%, which is similar to reference analysis error), although that for the instrument employing reflectance optics was inferior to models developed for the instruments employing transmission optics. The spectral positions for calibration coefficients were consistent with absorbance due to the third overtones of CH_2 stretching. Calibration models for moisture content in ground samples were acceptable on all instruments (R^2 >0.97; RMSECV <0.2%), whereas calibration models for intact kernels were relatively poor. Calibration coefficients were more highly weighted around 1360, 740, and 840 nm, consistent with absorbance due to overtones of O-H stretching and combination. Intact kernels with brown centres or rancidity could be discriminated from each other and from sound kernels using principal component analysis. Part kernels affected by insect damage, discoloration, mould growth, germination, and decomposition could be discriminated from sound kernels. However, discrimination among these defect categories was not distinct and could not be validated on an independent set.

It is concluded that there is good potential for a low cost Si photodiode array instrument to be employed to identify some quality defects of intact macadamia kernels and to quantify oil and moisture content of kernels in the process laboratory and for oil content in-line. Further work is required to examine the robustness of predictive models across different populations, including growing districts, cultivars, and times of harvest.

Additional keywords: macadamia, moisture, NIR, quality, oil.

Introduction

Payment to growers for macadamia nut-in-shell is based on the weight of nut-in-shell at 10% moisture content and on the percentage of sound (defect-free) kernels, determined by the processor on representative samples from each consignment. Penalties are imposed if the percentage of unsound kernels exceeds 3.5%. Sorting of kernels into sound and the various unsound categories for grower payment is currently done by subjective visual evaluation. The industry would benefit from objective tests that could be used in the processing plant quality assurance laboratory.

A major defect of macadamia nuts is kernel immaturity. Immaturity is currently qualitatively identified using visual criteria for the purpose of grower payments. This defect can be quantified by measuring oil content (Ripperton *et al*. 1938; Himstedt 2002). Other quality defects include insect damage, rancidity, mould growth, decomposition, germination, and discoloration. These defects are also assessed visually. Discoloration in this context is that which is due to causes other than insect damage, mould growth, decomposition, or germination. Moisture content is also an important quality parameter, having a major effect on shelf life (Cavaletto *et al*. 1966).

Processing of macadamia nuts requires factory drying from approximately 10% nut-in-shell (4% kernel) moisture to 4% nut-in-shell (1.5% kernel) moisture. Kernel moisture

 10.1071/AR03179 0004-9409/04/040471

content is currently assessed during the drying process by using an oven drying method. Determination of kernel moisture by NIRS would be much more rapid and would therefore be beneficial to the industry.

Near infrared spectroscopy (NIRS) is widely employed for oil and moisture determination in the oil seed and grain industries. For example, Tillmann *et al*. (2000) reported an R^2 of 0.95 and a standard error of cross validation (SECV) of 0.83% for oil content of whole canola seed, and Williams and Sobering (1993) $R^2 = 0.95$ and standard error of prediction (RMSEP) of 0.35% for moisture content of whole barley seed. Further, Ha *et al*. (1998) reported an $R^2 > 0.9$ on a range of specific fatty acids in sesame seed oil. NIRS has also been used to assess various attributes of oil seed quality (e.g. acid value, peroxide value as indicators of rancidity) (Cho *et al*. 1998; Ha *et al*. 1998). Other workers have used NIRS to assess malting barley grain for fungal contamination resulting in grain discoloration (calibrating against grain brightness—L*) (G. Fox, Qld Dept of Primary Industries, pers. comm.). The technique can also be used for qualitative purposes (e.g. the detection of insect damage in whole wheat kernels by measuring the amount of reflected *v.* absorbed light) (Chambers *et al*. 1994).

For oil, strong electromagnetic absorption is reported around 2200–2400 nm (CH_2 stretch bend and combinations), with weaker absorption around 1750, 1200, and 900 nm (first, second, and third overtones of CH_2 stretching) (Osborne *et al*. 1993). However, shorter wavelengths allow better penetration of biological samples (Kawano *et al*. 1993), and as such, shorter wavelengths should be useful in assessment of whole macadamia kernels.

The aim of this project was to assess the feasibility of using NIRS as an objective analytical method to replace the existing subjective methods for detection of kernel defects. Typically, the feasibility of using NIRS for a given application is assessed using laboratory grade instrumentation that is unsuitable for industry use with regard to cost and complexity. In this project we have drawn on our prior expertise with modular, low cost instrumentation (Greensill and Walsh 2000*b*; Walsh *et al*. 2000), comparing their performance with laboratory grade instrumentation.

Materials and methods

Sampling

Raw kernels were collected over several weeks of the 2002 season by a commercial processor. Three populations of kernels were assembled for calibration development with respect to oil, moisture, and defects, respectively. Samples were bulk packed into evacuated, heat sealed foil laminate bags and stored at 4°C between collection and analysis.

A population of 200 kernels (100 mature and 100 immature and free from other defects) was utilised for oil content models. The use of both mature and immature kernels greatly increased the range of oil contents studied (Table 1), desirable for reliable predictive modelling. A second population of sound, mature kernels was used for the moisture content study. Kernels with a range of moisture contents were obtained by taking nut-in-shell samples at various stages during the factory (approximately 5–8 days) drying process (Table 2). Of this population, 105 kernels were used intact, while 35 kernels were ground to pass a 2-mm screen in a Zyliss CH3250 grinder. Intact kernels were removed from the bulk packs and packaged individually in heat-sealed low-density polyethylene (LDPE) bags (from a single manufacturing batch), and were stored refrigerated in individual sample jars, to maintain kernel moisture levels. Spectra were acquired of kernels in their individual LDPE bags. Spectra were also acquired of ground samples, packed to an approximate depth of 5 mm, in individual LDPE bags. The contribution of the LDPE bags to the acquired spectrum of each sample was assumed to be constant.

Table 1. Oil content, kernel weight and kernel height of 100 mature and 100 immature kernels, and of the combined (n = 100) population

Attribute	Mature	Immature	Total
Oil (%)			
Mean ± s.d.	75.7 ± 2.1	46.2 ± 10.7	60.7 ± 16.84
Range	67.8–81.0	18.9–70.2	18.9–81.0
Weight (g)	2.4	1.3	1.85
Height (mm)	14.1	11.3	12.7

Table 2. Moisture content (%) of 105 intact and 35 ground samples

Attribute	Intact	Ground
Mean ± s.d.	1.92 ± 0.33	1.98 ± 0.34
Range	1.4–4.2	1.5–2.9

A third population consisted of 20 mature kernels of each of the following categories: sound, affected by mould, brown centres, insect damage, decomposition, germination, discoloration, and rancidity (n = 160).

Instrumentation

Three instruments, operating over different areas of the electromagnetic spectrum (Table 3), were used to collect spectra of intact and ground kernels. Instrumentation was powered on 2 h before spectral acquisition to ensure operational stability of both light source and detector.

The Si-PbS system was used in 2 differing modes of reflectance. One incorporated the use of the Foss NIRSystems remote reflectance probe (operating at 400–1900 nm) with intact kernels, and the other, a spinning cup module (operating at 400–2500 nm) with ground samples.

The InGaAs and Si systems used a full transmission configuration, on intact kernels with respect to oil, moisture, brown centres, and rancidity. A single 50 W Philips 402004 tungsten halogen lamp was mounted at 180° with respect to the sample and the detector fibre optic. Ground samples sealed in plastic bags were placed on an aluminium plate with a 7-mm-diameter hole, between detector and light source.

The InGaAs and Si systems also used an interactance configuration (Ocean Optics bifurcated optical fibre held 1 mm from kernel surface), on intact kernels affected by mould, brown centres (part kernels), insect damage, decomposition, germination, and discoloration. In this sample probe, a ring of illumination fibres is approximately 1 mm from a central detection fibre. As such, this technique is a highly localised one and useful in exploratory studies such as this where defect area is small in comparison to non-affected areas.

The Si-PbS instrument was operated using NSAS software, while the Si and InGaAs instruments were operated with in-house developed Labview based software. The NSAS default setting (average of 32 scans)

Table 3. Description of instrumentation

Detector type	Instrument	Wavelength (nm)	Optical configuration
Silicon-Lead Sulfide (Si-PbS)	Foss NIR Systems 6500	400–2500	Diffuse reflectance (remote reflectance and spinning cup accessories)
Silicon (Si)	Zeiss MMS1-NIR Enhanced	300–1100	Transmittance, interactance
Indium Gallium Arsenide (InGaAs)	Zeiss MMSNIR	800–1700	Transmittance, interactance

was adopted in collecting each spectrum for the Si-PbS system, whereas 20 scans were averaged for the InGaAs and Si systems (for a discussion for signal averaging, signal to noise ratio and its effect on calibration model performance, see Guthrie and Walsh 1999, Greensill and Walsh 2000*a*). Integration times were adjusted to achieve count levels above 50% of detector saturation (20–80 ms) for the InGaAs and Si systems.

Reference analysis (oil content, moisture, and defect)

Reference analysis was undertaken using a micro-soxhlet apparatus (s.e. of ±3% from the mean) for oil content and a TGA-601 Thermogravimetric Analyser (LECO Corporation) for moisture content (s.e. of ±2% from the mean). Defect samples were supplied by the commercial processor. Samples were assessed for defect category by staff accredited for subjective assessment of macadamia kernel defects by the Australian Macadamia Society.

Chemometric analysis

WinISI (ver. 1.5) software package was used for chemometric analyses. Calibration models were developed using absorbance, first and second derivatives of absorbance, with or without scatter correction (single normal variance and/or detrend) (12 combinations of treatments, as per Guthrie *et al*. 1998). Calibration development was based on the ISI modified partial least squares (MPLS) regression technique. Calibration performance is reported in terms of the R^2 of the calibration (R_c^2), the R^2 of a 6-group cross validation procedure (R_v^2), the root mean square of the standard error of the cross validation (RMSECV), and the ratio of population standard deviation to RMSECV (SDR).

In the assessment of different instruments, the same population of kernels was presented to all instruments. However, sample mis-presentation or signal saturation resulted in unequal sample numbers (Table 4). Further, different numbers of outlier spectra-reference values were identified and removed from each calibration set. Therefore, comparison of R_c^2, R_v^2, and RMSECV between calibrations must be tempered by consideration of s.d. The SDR statistic, as a ratio of s.d. to RMSECV, is useful in this connection. To achieve a statistically valid comparison between datasets acquired using 2 instrument platforms, all samples giving rejected and outlier values were eliminated from both population sets, and residuals used to compare RMSECV values at a 95% confidence interval using Fearns criteria (Fearn 1996). A Microsoft Excel spreadsheet was developed to implement this procedure and is available from the first author (john.guthrie@dpi.qld.gov.au).

The ISI discriminant analysis routine was used in an attempt to differentiate spectra of whole kernels scored for soundness, rancidity and brown centres, and spectra of part kernels affected by insect damage, discoloration, mould growth, germination, and decomposition.

Results and discussion

Kernel characteristics

Immature kernels were smaller, with an average weight of only 54% of mature kernels, and contained only 62% of the specific oil content (% w/w) of mature kernels (Table 1). Thus, immature kernels contained only 34% of the oil content of mature kernels (on a per kernel basis). The mean diameter of mature kernels used in this study was 16.3 mm and the mean diameter of immature kernels was 9.6 mm.

The current industry standard for kernel maturity is 72% or higher oil content (Ripperton *et al*. 1938; Himstedt 2002). The supplying processor selected obviously mature and immature kernels for this study, and a low error rate was achieved (only 3 classified as mature had <72% oil content, at 68, 69, and 70%), and none of the kernels classified as immature had >72% oil content. This selection process is reflected in the clear separation of means for the 2 populations (76 and 46% oil content, respectively) (Table 1).

The assessed moisture contents (population 2, kernels taken through the drying process) varied between 1.4 and 4.2% (Table 2). These moisture contents are as expected for Australian macadamia production and processing (Himstedt 2002).

On visual assessment, membership of the mould, decomposition, insect damaged, and discoloration groups appeared overlapped. Germinating kernels were distinguished by a brown, orange, or green coloration on the micropyle. Insect damaged kernels exhibited a small discolored and sometimes mouldy area centred on a wound site. Rancid kernels were not visually distinguishable from sound kernels.

Oil MPLS

Calibration statistics for oil content models for intact kernels were acceptable on all instruments (Table 3) based on SDR (SDR >3.0, RMSECV similar to reference method error). However, the calibrations developed on the InGaAs instrument (SDR 9.3) were significantly (based on Fearns criterion) superior to those from the Si instrument (SDR 6.5) and to those from the Si-PbS reflectance system (SDR 3.1) (Table 4, Fig. 1).

The better performance of the InGaAs unit, relative to the Si unit, for the assessment of oil content was attributed to detection of wavelengths relevant to lower order overtones of the oil CH_2 bond. This advantage presumably outweighs the disadvantage of the more limited penetration of wavelengths above 1000 nm and may indicate that the kernel is relatively homogenous in relation to oil content.

The Si-PbS instrument collected data over the full 700–2500 nm range; however, reflectance optics were used.

Table 4. Oil and moisture PLS calibration results for three instruments

Calibration models were developed for each attribute/instrument using a factorial combination of derivative condition (0, 1, 2, i.e. raw absorbance data, first or second derivative of absorbance data), and scatter correction routines (standard normal variance, detrend, or both (SC)). Calibration result is reported in terms of number of spectra, number of outliers removed, standard deviation of the assessed population, R^2 of the calibration and validation sets (R_c^2, R_v^2, respectively), RMSECV, and SDR (s.d./RMSECV)

Instrument	Math	# PC	N (# outliers)	s.d.	R_c^2	R_v^2	RMSECV	SDR
			Oil %—intact					
SiPbS	1 SC	7	199 (7)	16.3	0.94	0.91	5.3	3.1
Si	0 SC	11	199 (19)	15.4	0.98	0.98	2.4	6.5
InGaAs	0 SC	8	199 (16)	15.8	0.99	0.99	1.7	9.3
			Moisture %—intact					
SiPbS	1	6	105 (5)	0.25	0.89	0.79	0.11	2.2
Si	2 SC	9	98 (11)	0.25	0.78	0.59	0.16	1.6
InGaAs	2 SC	4	102 (4)	0.34	0.37	0.29	0.29	1.2
			Moisture %—ground					
SiPbS	1 SC	3	35 (5)	0.28	0.97	0.95	0.06	4.4
Si	2 SC	1	34 (1)	0.30	0.54	0.45	0.22	1.3
InGaAs	0 SC	2	32 (1)	0.33	0.58	0.31	0.28	1.2

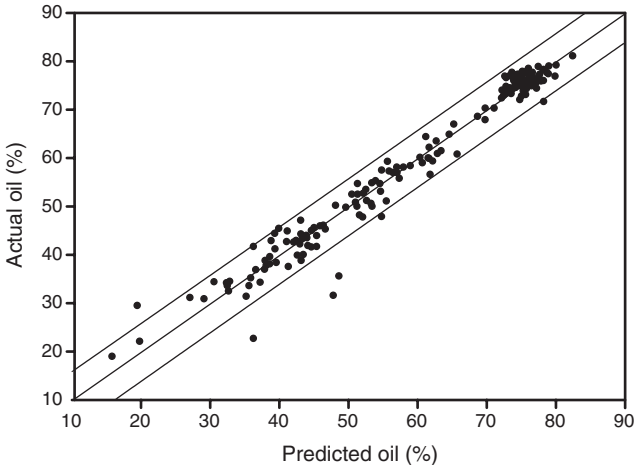

Fig. 1. MPLS regression (calibration) for actual and predicted oil content in intact macadamia kernels, based on spectral data from the Si (Zeiss MMS1) operated in transmission geometry and over the wavelength range 700–1100 nm. Regression performed on absorbance data, pre-treated with SNV and detrend. Dotted lines represent 95% confidence interval. Calibration statistics (from Table 4): n = 199, s.d. = 15.4, R_c^2 = 0.98, RMSECV = 2.4%.

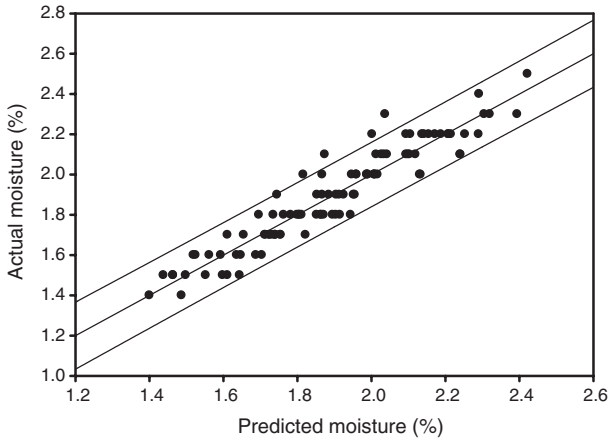

Fig. 2. MPLS regression (calibration) for actual and predicted moisture content in intact macadamia kernels, based on spectral data from the Si-PbS (Foss NIRSystems 6500) operated in reflectance geometry and over the wavelength range 700–1900 nm. Regression performed on first derivative absorbance data, without scatter correction. Dotted lines represent 95% confidence interval. Calibration statistics (from Table 4): n = 105, s.d. = 0.25, R_c^2 = 0.89, RMSECV = 0.11%.

The transmission optical path employed with the photodiode instruments was presumably advantageous for this application.

For spectral data collected with the Si photodiode instrument (used in transmission mode), the best model performance was achieved using absorbance spectra, treated with detrend (data not shown). Typical statistics for calibration models on oil were: R_v^2 = 0.98; RMSECV = 2.4% (Fig. 1). McGlone and Kawano (1998) suggested that an SDR of >3 is adequate to support sorting into 3 classes. The SDR of >6 reported here is indicative that the calibration would support a useful sorting function.

Calibration loadings were consistent with absorbance due to the third overtones of CH_2 stretching (e.g. heavy weightings around 930 nm) (data not shown).

Moisture MPLS

Calibration models for moisture content of intact kernels were not as reliable as those for oil content. Best results were obtained with the Si-PbS instrument (R_v^2 = 0.79, RMSECV 0.11%) (Table 4, Fig. 2).

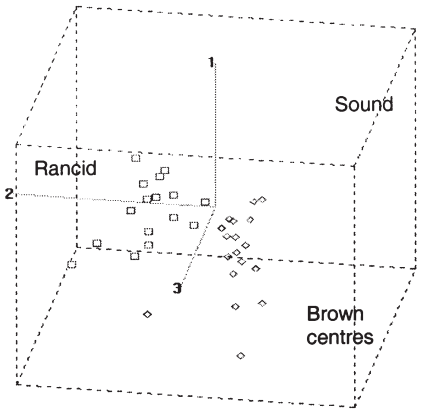

Fig. 3. Discriminant analysis (PCA) of whole macadamia kernels using the InGaAs (Zeiss MMSNIR) in transmission mode. Plot uses the first 3 principal components.

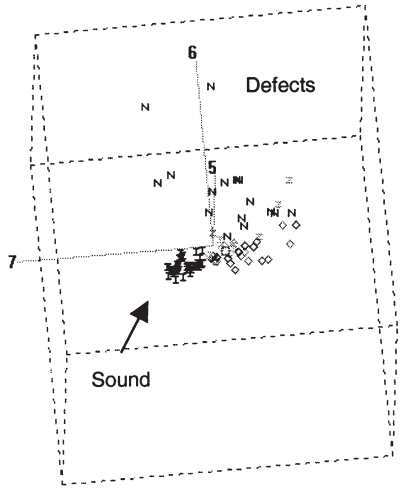

Fig. 4. Discriminant analysis (PCA) of macadamia kernels using the Si (Zeiss MMS1) in interactance mode. Plot of the last 3 (fifth, sixth, seventh) principal components.

Calibration model performance based on spectra acquired using the Si instrument (R_v^2 = 0.59, RMSECV 0.16) (Table 4), were superior to that obtained with the InGaAs instrument. The better performance of the Si, relative to InGaAs, photodiode array based unit is consistent with a relatively non-homogenous distribution of moisture content in the kernel, with the better relative penetration of SW-NIR (700–1100 nm) radiation, supporting a stronger calibration model. Conversely, best results were obtained with the Si-PbS instrument, which was operated in a reflectance mode. Reflectance optics would optically 'weigh' the surface layers of the kernel. These results are not consistent, and we anticipate further work to resolve this issue.

It was necessary to limit the number of principal components (PC) to 3 to avoid over-fitting of data with the ground samples. ISI (1998) suggest 1 PC per 10 samples to avoid over-fitting. Within this constraint, the Si-PbS unit (R_v^2 = 0.97, RMSECV 0.06) (Table 4), operated with the spinning cup reflectance module, supported better model statistics than that obtained from intact kernels with the same unit, operated with a remote reflectance probe (Table 4). This result is consistent with a level of inhomogenity of moisture within intact kernels. Grinding reduces this variation, and the spinning cup module allows scanning of a large proportion of the sample. In contrast the predictive models developed for the Si and InGaAs units were relatively poor (e.g. R_v^2 <0.5). The transmission optics employed for the Si and InGaAs units with ground samples allowed spectral assessment of only a small proportion of the sample. A change to this arrangement is recommended in future work.

Calibration coefficients were weighted around 1360, 740, and 840 nm (data not shown), consistent with absorbance due to overtones of O-H stretching and combination.

Discriminant analysis

Sound kernels and those with brown centres and rancidity could be discriminated from each other using principal component analysis of the spectral data obtained in transmission mode (Fig. 3).

Part kernels affected by insect damage, discoloration, mould growth, germination, and decomposition could be discriminated from sound kernels (Fig. 4). However, discrimination among these defect categories was not distinct and could not be validated on an independent set (data not shown). This was probably due to the overlapping of defect categories (e.g. decomposed also exhibited discoloration and possibly mould growth). This study was also based on relatively low sample numbers, and further replication is recommended.

Conclusion

The low cost Si photodiode array instrument achieved an RMSECV of 2.4% on oil content of intact kernels, and 0.2% on moisture content of ground kernels, and also allowed discrimination between kernels affected with a number of defects. This study has shown encouraging results indicating that this technology should prove useful to the processing industry as an assessment tool of some quality attributes both in the processor laboratory and in-line. Further work should consider optical configurations to optimize sampling of the product, and also the robustness of the calibration models across different populations, including growing districts cultivars and times.

Acknowledgments

We thank CQU/DPI for financial and equipment support, the technical support of Peter Martin, David Pytko, and Clayton Lynch, and the supply of samples from SunCoast Gold Macadamias, Gympie. We acknowledge the financial support of Horticulture Australia Ltd and the Australian Macadamia Society.

References

Cavaletto CA, Cruz Dela E, Ross E, Yamamoto HY (1966) Factors affecting macadamia nut stability. 1. raw kernels. *Food Technology* **20**, 1084–1087.

Chambers J, Ridgway C, van Wyk CB, Baker CW, Barnes RJ (1994) 'The potential of near-infrared spectroscopy for the rapid detection of pests in stored grain.' (Central Science Laboratory (MAFF): Slough, Berkshire)

Cho S, Kim J, Rhee C (1998) Determination of rancidity of soybean oil by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* **6**, A349–A354.

Fearn T (1996) Comparing standard deviations. *NIR News* **7**, 5–6.

Greensill CV, Walsh KB (2000*a*) Optimisation of instrumentation precision and wavelength resolution for the performance of NIR calibrations of sucrose in a water-cellulose matrix. *Applied Spectroscopy* **54**, 426–430.

Greensill CV, Walsh KB (2000*b*) A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit. *Measurement Science and Technology* **11**, 1674–1684. doi:10.1088/0957-0233/11/12/304

Guthrie J, Walsh K (1999) Influence of environmental and instrumental variables on the non-invasive prediction of brix in pineapple using near infra-red spectroscopy. *Australian Journal of Experimental Agriculture* **39**, 73–80. doi:10.1071/EA96127

Guthrie JA, Wedding B, Walsh K (1998) Robustness of NIR calibrations for soluble solids in intact melon and pineapple. *Journal of Near Infrared Spectroscopy* **6**, 259–265.

Ha J, Koo M, Ok H (1998) Determination of the constituents of sesame oil by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* **6**, 371–373.

Himstedt SR (2002) Oil content and other components as indicators of quality and shelf-life of Macadamia integrifolia. PhD thesis, The University of Queensland, Qld.

Kawano S, Abe H, Iwamoto M (1993) Novel applications of non-destructive techniques for quality evaluation of fruits and vegetables in Japan. In 'Proceedings of the International Workshop on Non-destructive Technologies for Quality Evaluation of Fruits and Vegetables'. Spokane, Washington DC.

McGlone VA, Kawano S (1998) Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. *Postharvest Biology and Technology* 131–141.

Osborne BG, Fearn T, Hindle PH (1993) Theory of near infrared spectrophotometry. In 'Practical NIR spectroscopy with applications in food and beverage analysis'. (Ed. D Browning) pp. 13–35. (Longman Singapore Publishers Ltd: Singapore)

Ripperton JC, Moltzau RH, Edwards DW (1938) Methods for evaluating the macadamia nut for commercial use and the variation occuring among seedlings plantings in Hawaii. Hawaii Agricultural Experiment Station Bulletin No. 79.

Tillmann P, Reinhardt TC, Paul C (2000) Networking of near infrared spectroscopy instruments for rapeseed analysis: a comparison of different procedures. *Journal of Near Infrared Spectroscopy* **8**, 103–107.

Walsh KB, Guthrie JA, Burney J (2000) Application of commercially available, low-cost, miniaturised NIR spectrometers to the assessment of the sugar content of intact fruit. *Australian Journal of Plant Physiology* **27**, 1175–1186. doi:10.1071/PP99111

Williams PC, Sobering DC (1993) Comparison of commercial near infrared transmittance and feflectance instruments for analysis of whole grains and seedsa. *Journal of Near Infrared Spectroscopy* **1**, 25–32.

Manuscript received 1 September 2003, accepted 14 January 2004