NIR technologies for high speed fruit grading

A thesis submitted to the
University of Waikato
for the degree
of
Doctor of Philosophy in Physics
by

Daniel Grant Fraser

The
University
of Waikato
Te Whare Wananga
o Waikato

September 2001

Abstract

Today's consumers seek information about food products so they can purchase with confidence, even if it translates to higher priced items. It is the aim of this thesis to improve the instruments available for the prediction of food quality, in particular, to advance the technology used for the rapid, low cost, non-destructive analysis of the internal quality of fruit. This thesis contains both experimental and theoretical work covering two main areas of research: the first deals with issues related to the practical grading of fruit, the second with the NIR light penetration and distribution within a fruit.

A 'rotary grader' has been built to enable the rapid, non-destructive assessment of fruit. Up to twenty fruit can be placed on a rotating disk, which can be spun at a controlled speed or moved to fixed positions for stationary studies. A variety of optical arrangements can be mounted to enable the acquisition of spectra from individual fruit in different 'modes' of operation. A comparison between the different optical arrangements is presented. It was found that transmission mode (where the light passes right through the fruit) performs better than reflectance mode (where the detected light comes primarily from near the surface of the fruit). This result has been previously predicted and has now been demonstrated with the equipment used in this work. Different angles of transmission are compared as well as different fruit orientations for their influence on fruit quality predictions. It was found that 90° transmission (where the light is detected at right angles to the illumination of the fruit) might work best for randomly orientated kiwifruit; if they are hand placed then 180° or 150° may be more effective. A study of mandarins suggested that the acid content cannot be assessed directly, but it can be inferred from surface properties such as colour and chlorophyll content. This is contrary to several published reports and questions the ability of NIR spectroscopy to measure low concentrations of acid in fruit. A third experiment showed that soluble solids content (SSC) and dry matter (DM) can be assessed from moving apples with a prediction error of 0.54 ± 0.07 °Brix (R_P^2 = 0.65) and 0.66 ± 0.08 % (R_P^2 = 0.80) respectively. A unique perspective is presented showing where the useful information is attained from the continuously moving

fruit. Transmission and reflectance spectra from apples moving at speeds of up to 3 fruit/s have been recorded, and the optimal spectra from the sequence were combined for improved predictions.

To investigate the NIR light distribution and penetration in fruit a fibre optic probe was built to take direct measurements of the light levels inside fruit. This enabled, for the first time, the degree of light penetration to be assessed without altering the optical boundaries of the fruit (for example cutting away sections of the fruit). Corresponding simulations using Monte Carlo photon tracing, match the experimental measurements. It was found that fruit such as mandarin exhibit strong internal reflection from the skin, a result that has not previously been reported. The influence of the core and other features in the fruit on the light penetration and distribution is also studied showing perturbations in the internal light levels mapping out these features.

Acknowledgements

Firstly, I would like to thank the Technology Development Group (TDG) of HortResearch, where this project was undertaken. TDG provided the equipment, facilities, staff expertise, and financially supported this project.

Enormous thanks must go to my supervisors: Rainer Künnemeyer and Bob Jordan for their constant availability, feedback, support, encouragement and training. I am also indebted to Andrew McGlone and Peter Schaare for their considerable input and knowledge, which they so freely provided throughout my time at TDG. I would like to thank the workshop, in particular Lynn Cate, for the tolerance, training, and help building the equipment used in this thesis. Thanks also to the many other staff at TDG who have helped me during my time here and offered me their friendship.

I would like to thank and acknowledge my scholarship benefactors without whom this work would not have been possible: AGMARDT (Agricultural and Marketing Research and Development Trust) and the University Vice-Chancellors Committee (William Georgetti Scholarship).

Thanks to my family and those friends who have supported me over the (almost ten!) years I have been studying at university. It has been a long road and without your encouragement, praise, and constant support I may not have made it. And finally, a special thank you to my wife Katrina, for the sacrifices, support and love.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABS	Absorbance
A/D	Analogue to Digital
CCD	Charge Coupled Device
DM	Dry Matter
F	Factors (number of)
IPLS	Iterative Partial Least Squares
MA	Massachusetts (state of America)
MLR	Multiple Linear Regression
MSC	Multiple Scatter Correction
N	Number of samples
NA	Numerical Aperture
NH	New Hampshire (state of America)
NIR	Near Infra-Red
NN	Neural Network
NZ	New Zealand
PC	Principle Component or Personal Computer
PCA	Principle Component Analysis
PCR	Principle Component Regression
PDA	Photo Diode Array
PLS	Partial Least Squares
PTFE	Polytetrafluroethylene (Teflon)
RMSEC	Root Mean Standard Error of Calibration
RMSECV	Root Mean Standard Error of Cross Validation
RMSEP	Root Mean Standard Error of Prediction
PRESS	Predicted Residual Error Sum of Squares
SA	Scale by Area
SECV	Standard Error of Cross Validation
SEP	Standard Error of Prediction
SG	Savitzky-Golay
SNV	Standard Normal Variate
SSC	Soluble Solids Content
TDG	Technology Development Group, HortReseach, Hamilton, New Zealand.
USA	United States of America
UVS	Unit Variance Samples

1

Introduction

The work presented in this thesis has been supervised by the University of Waikato but was undertaken at the Technology Development Group (TDG), HortResearch, Hamilton, New Zealand, as part of an established and on going NIR technologies program. This work builds on my own recent involvement (Schaare and Fraser, 1999), the work of Osborne (1997), and others at TDG who have been investigating methods of non-destructive testing of fruit using NIR since 1993. This chapter outlines the rationale for the research and gives a brief outline of each chapter.

1.1 Rationale

Because of ever increasing competition in the fresh horticultural products market it is important for producers to differentiate their produce. By marketing superior fruit, premium prices and entry to more lucrative markets may be attained. Near infrared spectroscopy has the potential to allow fruit to be sorted based on internal quality attributes such as sugar content, acid level and firmness. This means that not only can the fruit be graded by size and external appearance but the purchaser can also be assured of internal quality as well. Already there are marketing initiatives to differentiate fruit sorted by sugar content, both in Japan (McGlone, et al. 1997), and elsewhere (e.g., the TasteMark™ brand developed by Compac Sorting Equipment, New Zealand).

The international fresh fruit market is huge with 410 million tonnes being consumed around the world in the 1999-2000 year (O'Rourke, 2000). Apples make up 15%, citrus 24%, and kiwifruit less than 1%. However, kiwifruit is a particularly important crop for New Zealand, one of the three biggest exporters of kiwifruit (along with Italy and Chile). Apple and kiwifruit exports together made

up 52% of New Zealand horticultural exports, which in turn were 6.5% of New Zealand's total produce exports for the 1999-2000 year (Kerr and Aitken, 2001).

The online grading of fruit using near infra-red (NIR) spectroscopy has been used in Japan for many years (Kawano, 1998) but has still not been widely utilised elsewhere. In 1997 only 100 grader lines were in use in Japan (McGlone et al., 1997). Other countries, including New Zealand (Compac Sorting Equipment Ltd., Onehunga, Auckland, NZ), have begun commercial trials of the technology. In spite of this there is still considerable potential for improving the process and for understanding of what is actually measured.

Most of the early systems utilised reflectance mode for fruit assessment where the detected light penetrates the tissue in a limited manner. With improving technologies, such as more sensitive detectors and improved designs, use of transmission mode where the detected light has passed right through the fruit is becoming practical even for thick-skinned fruit. Japan at least has moved towards transmission mode. It is expected that other countries will follow the trend. This is a very important development because defect detection (e.g., brown heart in apples) is potentially possible using transmission mode. Such a system, if successful, could catapult the acceptance of NIR technology into common use by saving whole lines of fruit in a severely afflicted year.

Very little is known about the distribution of visible or NIR light *inside* fruit. Knowing the light distribution may enable the selection of more effective modes of measurement for a given fruit type or shape. Indeed, the ability to detect internal defects or to estimate the blocking effect of a fruit stone could be assessed directly if the light distribution profiles were known. The development of reliable Monte Carlo simulation tools could allow computation of optimal illumination and detection geometries without the need for extensive experimental measurements.

There are considerable difficulties still to be overcome regarding the grading of moving fruit with NIR using transmission mode at high speed. The illumination levels need to be much higher for transmission so that the detected signal can be maximised to improve the signal to noise ratio and the ultimate speed of the system. Randomly oriented non-spherical fruit and fruit with varying fruit size lead to a large variation in detected signal strength which needs to be handled. For

these reasons optical layout and components are critical. The fact that the fruit are moving means that in transmission mode the object being examined is constantly changing and a large amount of data is collected. How this data is collected, synchronised and processed for the best possible prediction has yet to be fully studied. The paths that the light travels inside intact fruit has not been examined in detail previously. Simulation of the light levels will lead to greater understanding of the transmission process. This thesis works towards solving some of these problems.

1.2 Original Contributions

NIR Spectroscopy of fruit is an established field. The areas where this thesis covers novel or new work include:

- The comparison of grading performance using different angles of transmission for kiwifruit and mandarins,
- Investigations into the effect of fruit orientation on calibration performance,
- Analysis of best predictive positions/timings for spectra acquisition on moving fruit using transmission and reflectance,
- Fundamental studies of light propagation and distribution in fruit tissue by direct measurement, and
- Monte Carlo simulation of fruit tissue scattering properties as applied to intact fruit.

Publications to date that directly relate to aspects of this research include:

- Fraser, D.G, Jordan, R.B., Künnemeyer, R., McGlone, V.A., Light distribution inside mandarin fruit during internal quality assessment by NIR spectroscopy. Postharvest Biology and Technology (in press).
- Fraser D.G., Jordan R. B., Künnemeyer R., McGlone V. A., 2002. A New Method for mapping the visible-near infrared light levels in fruit. Journal Near Infrared Spectroscopy, Proceedings of the 10th international conference, Korea 10-15 June 2001 (in press).
- Fraser, D.G., McGlone, V.A., Künnemeyer, R., Jordan, R.B., 2001. NIR (Near Infra-Red) light penetration into an apple. Postharvest Biology and Technology. 22 (3), 191-194.

1.3 Description of Thesis

This thesis is divided into five chapters. In Chapter 2, background information for this work is described. An introduction to non-destructive testing is given, along with those aspects of fruit physiology that are relevant to this study. The statistical processes and tools for extracting information from fruit spectra are outlined, and previous relevant work on fruit grading is described. The last section of Chapter 2 describes diffuse scattering and absorption in fruit tissue. Monte Carlo photon tracing used in this study is explained, and previous work on light distribution and penetration in fruit is described.

Chapter 3 describes three fruit grading experiments and their results. The first trial investigates the degree by which fruit orientation influences fruit quality predictions, and attempts to establish an optimal orientation for assessment of kiwifruit properties. The second experiment measures acid and sugar levels in mandarins using NIR spectroscopy in reflectance, interactance and transmission modes with different wavelength ranges. The third experiment investigates using NIR spectroscopy on moving apples for both reflectance and transmission modes. The predictive quality of spectra measured at different locations as the fruit passes the detector are studied and ways to improve them by combining the best sequential spectra explored.

In Chapter 4, a custom-made fibre optic probe to sample light levels inside the fruit is described and experimental measurements presented. Monte Carlo simulation of simple geometric models and the predicted light levels are then compared with the experimental results. The depth of visible-NIR light penetration into 'intact' apples was measured, and the rate of light extinction in apple, nashi, mandarin, and kiwifruit tissue are compared. An investigation into the internal reflective properties of mandarin skin was made, and the experimental results are compared to Monte Carlo simulations.

In Chapter 5 a summary of the results is presented.

The Appendix contains a description of the Monte Carlo simulation code which is contained on the disk included with this thesis.

2

Background

Sorting and grading fruit for its quality immediately after harvest is very important to maximise returns for growers and the associated industries. Quality standards must be maintained to meet the increasingly demanding market and achieve the premium prices. Not only must the poor quality fruit be removed, but also the best of the crop must be selected for marketing to the highest paying customers. Grading for external appearance is important but so is the *internal* quality. Traditional methods for assessing internal fruit quality require the fruit to be destroyed or damaged in the assessment process. This means that in order to estimate the qualities of fruit on a grading line, random samples are taken, and the information is then inferred statistically from the samples. This requires a separate sample to be taken for each new population (e.g., a different grower line) which is a time consuming and expensive process and assumes that such lines have consistent properties. In addition there is often a lengthy delay until the lab results are processed. A much better solution is to assess each fruit non-destructively so that all fruit can be of guaranteed quality, and so no fruit are lost to the testing process. NIR spectroscopy can non-destructively test *every* fruit, at fruit grader speeds (up to 10 fruit/s).

The principle underpinning an NIR grader is to assess the chemical constituents within the fruit by measuring white light at different wavelengths, which have been reflected from or transmitted through each fruit. Using this technique it has been shown that fruit can be assessed for sugar content, dry matter, acidity, maturity level, and even internal defects. However, the technique is still under development for moving fruit with a number of difficulties still to be overcome.

Except in Japan, where the first commercial system was installed in 1989 (Kawano, 1994) the technique is struggling for widespread acceptance. Currently graders are being trialed in many other countries including Italy, New Zealand (NZ), Australia, South Korea and USA. Because of the huge volume of fruit produced worldwide there is enormous market potential for this technology.

In this chapter an outline of the basic fruit physiology of the main fruit studied in this thesis is outlined, an introduction to NIR spectroscopy is presented, and the theory of light transmission in highly scattering tissue is covered.

2.1 Relevant aspects of kiwifruit, mandarin, and apple physiology

2.1.1 Kiwifruit

Kiwifruit (Actinidia deliciosa (A. Chev) C.F. Liang et A.R. Ferguson 'Hayward') were commercialised in NZ with exports to the USA beginning in 1953 by Turners and Growers Ltd. using the name 'kiwifruit' (Earp, R.W., 1990). Kiwifruit are grown on vines and prefer sheltered areas with well-drained soils, which are free from spring and early autumn frosts (Ferguson, 1990). They are unique for their relatively high vitamin C content (Thomas and Corden, 1970) and bright green internal flesh colour and appearance. Until 1999 'Hayward' kiwifruit was the most common commercially produced cultivar. Hayward has an excellent long-term storage life (Beever and Hopkirk, 1990). Recently, a new kiwifruit called Zespri gold has been released which has a yellow flesh colour and a larger soluble solids content (SSC) when fully ripened. Only the green variety is used for the fruit experiments in this thesis.

Kiwifruit are ovoid in shape and weigh typically 80-160 g when mature. There are three major sections within the fruit: the core, the inner pericarp, and the outer pericarp (Fig. 2-1).

The outer pericarp is made up of elliptical parenchyma cells of two sizes. Individual or small groups of large cells (diameter 0.5-0.8 mm) are dispersed in a matrix of smaller cells (0.1-0.2 mm). The cell walls range in thickness from 1-5 μm. The inner pericarp is made up of locules and locule walls. The locule area contains both seeds and radially elongate thin-walled cells (0.2-0.4 mm in diameter to greater than 1 mm in length). The locule wall consists of smaller

thicker-walled radially elongate cells (0.05-0.1 mm × 0.2-0.4 mm). Core tissue is relatively uniform, where most cells are spherical/ellipsoidal, of cross sectional dimensions 0.1-0.2 mm (Hallet et al., 1992). Within the inner and outer pericarp are chlorophyll containing plastids which give the fruit its characteristic bright green colour (since chlorophyll absorbs in the red wavelength region as shown in later chapters).

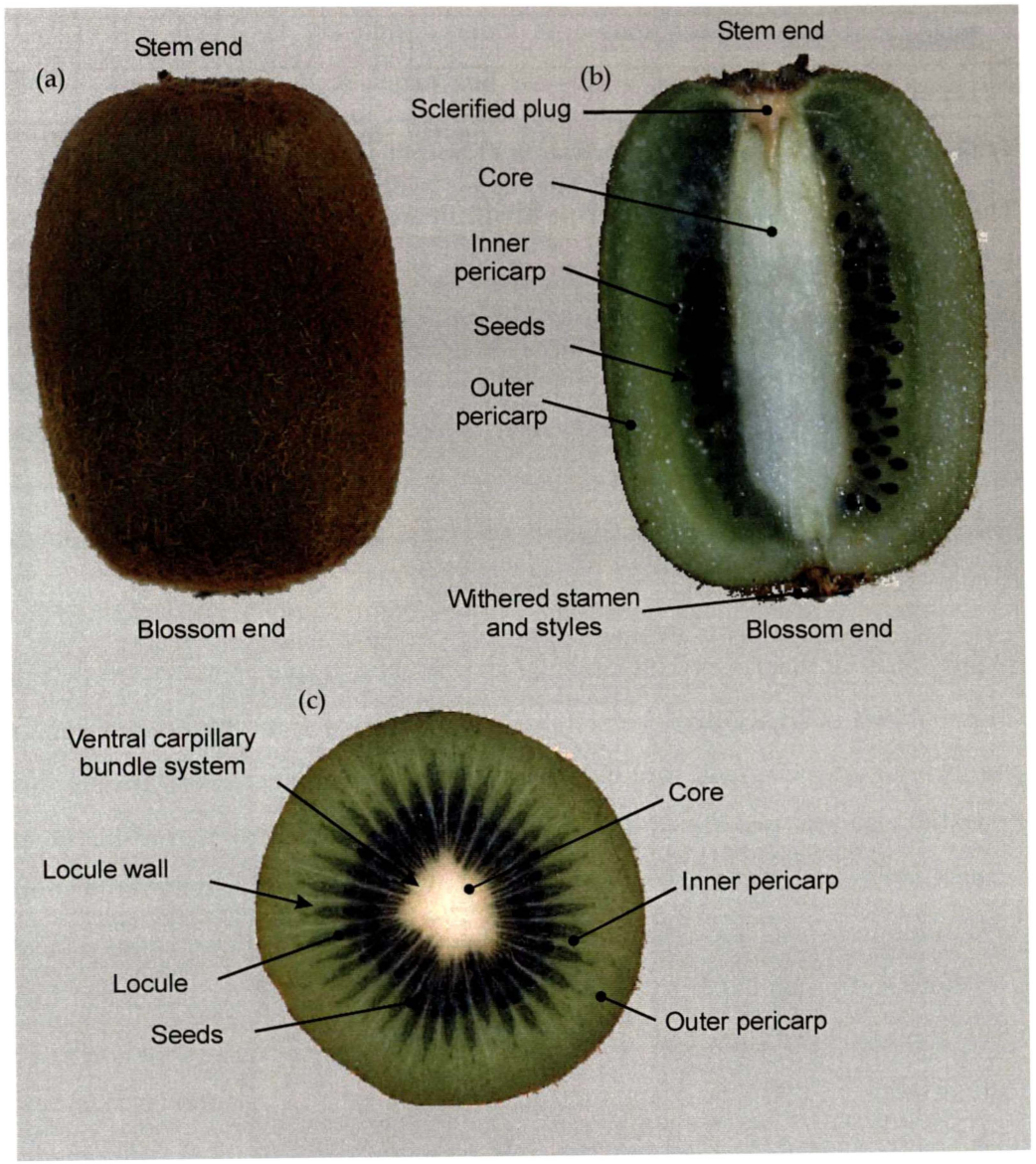

Fig. 2-1: Mature Hayward kiwifruit. (a) External. (b) Section along axis. (c) Section through equator. (Martinsen, 1999)

Kiwifruit, like most fruit, predominantly consist of water; Table 2-1 shows some of the most significant constituents and their proportions.

Component	Proportion (% fresh weight)
Water	80-88
Carbohydrate	10-15
Acids (% as citric)	1.0-1.6
Fibre	1.1-3.3
Minerals (ash)	0.45-0.74
Protein	0.11-1.2
Lipid	0.1-0.9

Table 2-1: Typical constituent proportions in a kiwifruit. (Beever and Hopkirk, 1990)

The main carbohydrates found in ripe kiwifruit are sugars (glucose 2-6%, fructose 1.5-8% and sucrose ~2%). At harvest, the soluble solids (sugars and acids) typically make up 6-8% (6-8 °Brix measured by a refractometer) of the fruit, and the remaining carbohydrate is in the form of granulised starch. At this time the fruit are very hard, having a penetrometer firmness of 60-90 N. After harvest and during transport and storage, the starch is gradually converted to sugars via hydrolysis, softening starts (reaching 6 N) as the cell walls thicken (Beever and Hopkirk, 1990).

Within the fruit there is an SSC imbalance between the two ends of the kiwifruit (blossom end is higher than the stem), which increases from the time of harvest and then diminishes as the fruit becomes fully ripe (Harris et al., 1972). The final ripe SSC typically reaches 12-16 °Brix and is a strong determinant of consumer acceptability (Jordan et al., 1996). At harvest time the fruit have only reached 50-75% of the final SSC depending on the progress of the starch to sugar conversion. Therefore sorting on harvest time SSC will not give an accurate estimate of final SSC. However, the final SSC is roughly equal to final DM - 3 where the 3% (Jordan et al., 2000) accounts for the acid and other structures in the fruit (effectively constant). Because the DM of the kiwifruit does not change significantly from harvest time to eating time, the DM at harvest time (which equals the unconverted starch + SSC + 3%) will be a good indicator of final SSC. Hence the fruit can be sorted for taste using harvest time DM content.

An alternative indicator of kiwifruit quality at harvest time is density, which is typically 1020-1060 kgm^{-3}. Density is positively correlated to both fruit DM at

harvest (r^2 = 83%) and the ripe SSC (r^2 = 85%) of kiwifruit (Jordan et al., 2000) and so therefore a good indicator for consumer acceptability.

2.1.2 Mandarins

The Satsuma mandarin (Citrus reticulata) is believed to have originated in Japan 350 years ago (Morton, 1987). Mandarins make up more than half of the citrus fruit grown in NZ but are only a minor crop compared to other fruit. Most NZ mandarins are exported to Japan and were less than 0.6% of total fresh fruit exports in 2000 (Kerr and Aitken, 2001). One of the major cultivars is 'Miyagawa', which is a seedless, easy peel variety and is the variety used in chapter 3 for non-destructive assessment tests for acid level.

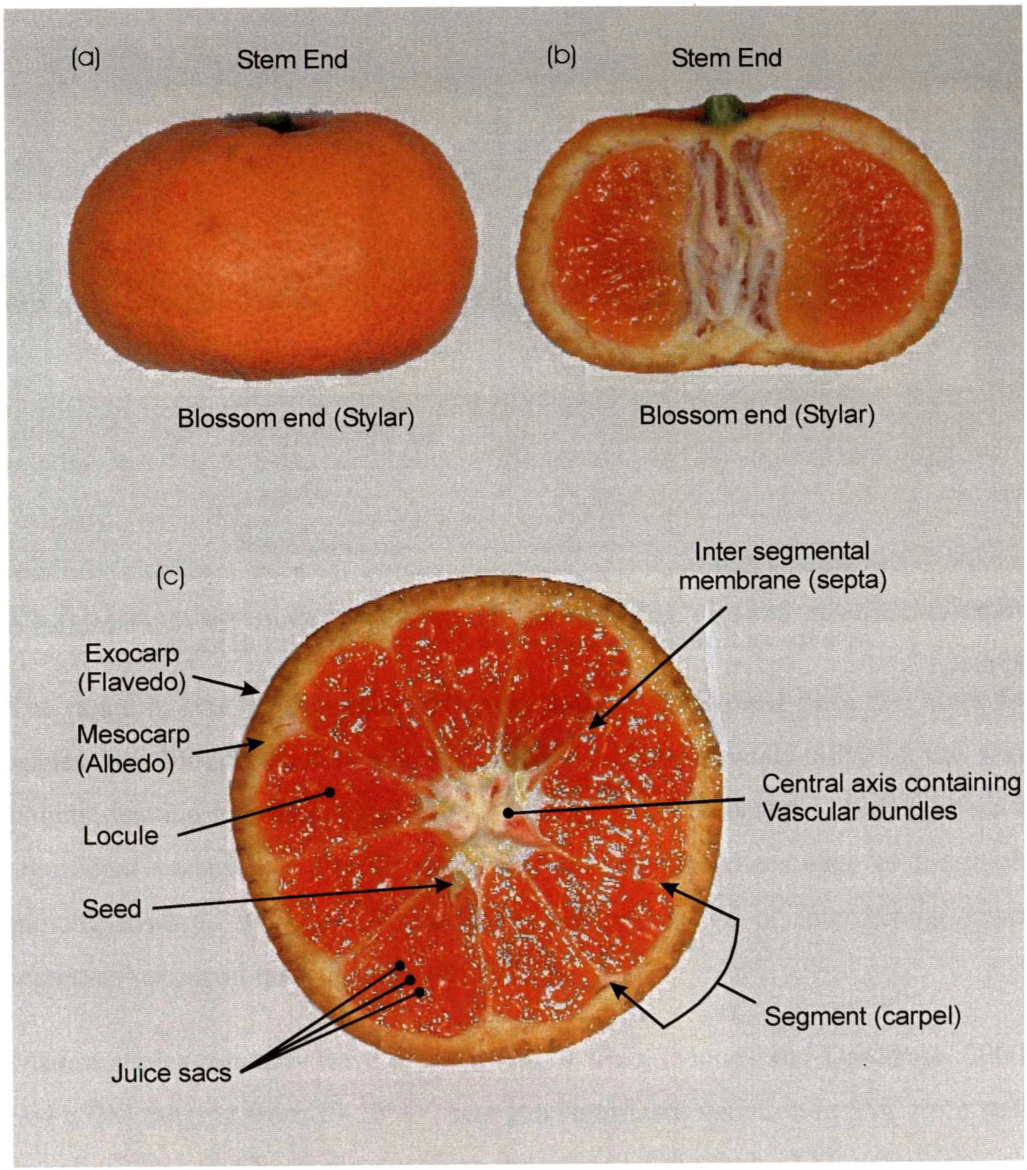

Fig. 2-2: Mature Satsuma mandarins. (a) External. (b) Section along axis. (c) Section through equator.

Typical fruit sizes used in this thesis were small, ranging from 30-50 mm in diameter and 50-110 g for mass. Just before harvest, mandarin acid levels are high (2%), and SSC is low (~8 °Brix). As the fruit mature the acid level reduces, and the SSC increases. In New Zealand when the mean ratio of SSC (in °Brix) to acid (in %) of a sample of fruit reaches 7 (the recommended minimum) the fruit may be harvested (Harty, 1995).

Within the mature fruit's thick (3-5 mm) peel there are segments containing juice sacs and the central axis or core, which takes up a significant proportion of the fruit volume (Fig. 2-2). The juice sacs store sucrose and hexoses and are the location for citric acid formation in the fruit (Goldschmidt and Koch, 1996). The amount and type of sugars stored in juice sac can differ markedly from sac to sac (Ting, 1969; Bartholomew and Sinclair, 1941). This contributes to the variation observed between segments and to differences between stem and stylar ends of a segment (Koch and Avigne, 1990).

2.1.3 Apples

Apples (Malus domestica) have been grown for thousands of years and 7500 varieties are grown throughout the world. The worldwide apple production was 46.4 million tonnes in 1999 (compared to less than 1 million tonnes of kiwifruit) with approximately 5 million tonnes (11%) traded internationally. New varieties, Braeburn, Royal Gala, and Pacific Rose, represent 75% of New Zealand's 310,000 tonnes of apple exports. NZ exports more than 60% of its apple crop (Kerr and Aitken, 2001). Fig. 2-3 shows the internal structure for two varieties of apple.

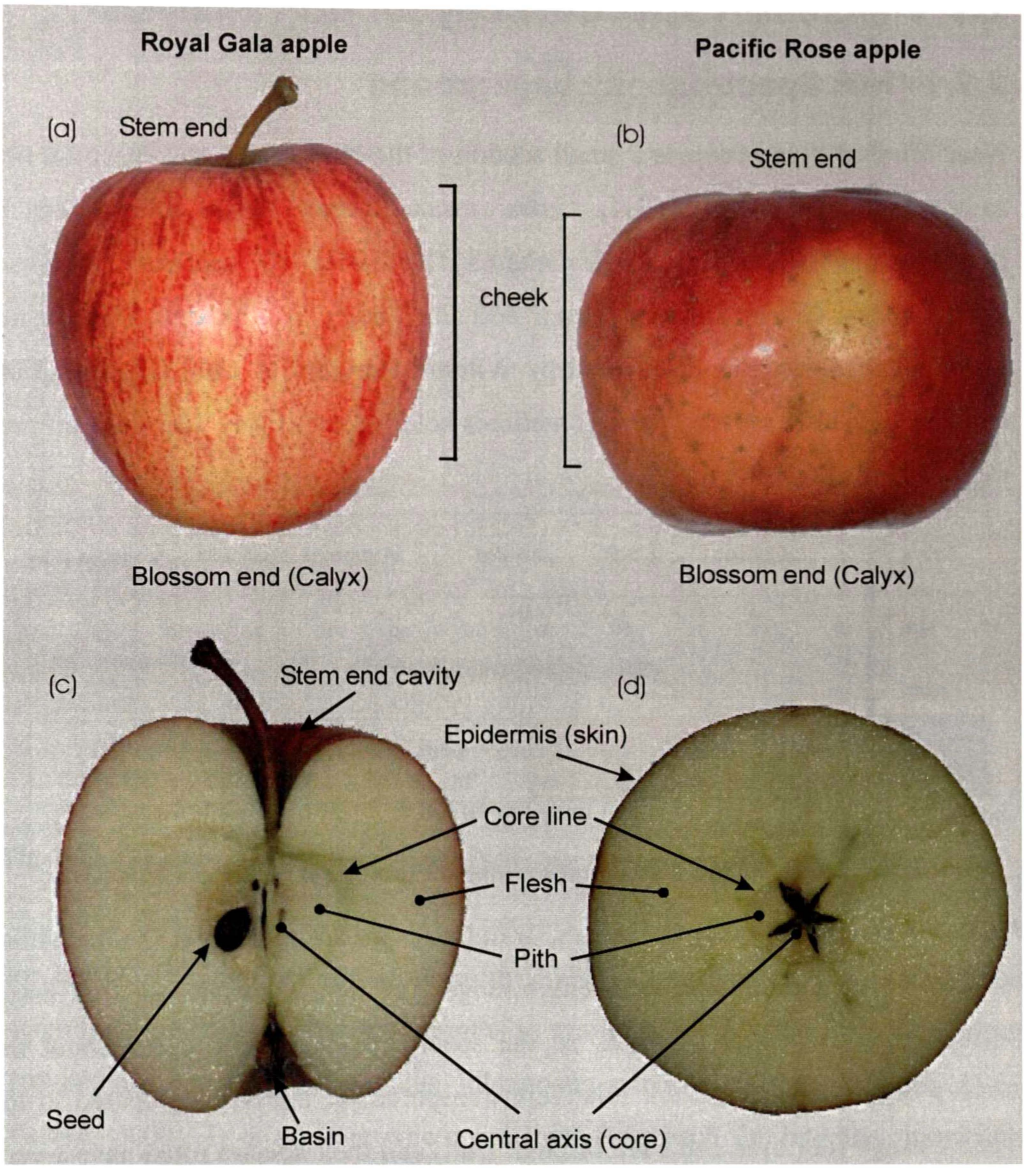

Fig. 2-3:Two varieties of apple. (a) Royal Gala external. (b) Pacific Rose external. (c) Section along Royal Gala axis. (d) Section through Pacific Rose equator.

The range in size used in this thesis was 140-170 g. Typical flesh cell sizes for apples are 200 μm in diameter and make up approximately 80% of the fruit volume leaving a large amount of intercellular space (Mohsenin, 1986). There is significant variation in the air space density from the outer edge of the apple moving towards the core. Furthermore, the percentage of intercellular space increases with fruit mass (Mohsenin, 1986).

Mature apples typically have a SSC of 9-19 °Brix (Moons and Dardenne, 2000) and a DM ranging from 10-20%. There is a significant variation in SSC from stem to calyx (3%) and from the centre to the edges of the fruit (2.8%) (Peiris et al., 1999). Over 80% of the acid in apples is malic acid (Sohn and Cho, 2000).

2.2 Visible-NIR spectroscopy for fruit analysis

2.2.1 NIR Spectroscopy background

Near infrared light occupies a small section of the electromagnetic spectrum next to the visible range (Fig. 2-4). 'Infra' means 'below' hence infrared can be interpreted as 'below the red' wavelengths. The visible-NIR spectrum range used in this thesis covers the range from 400 nm though to 1700 nm. The infrared wavelength region was discovered by William Herschel in 1800 (Davies, 2000) and the 780-1100 nm region is sometimes referred to as the 'Herschel infrared region' in his honour.

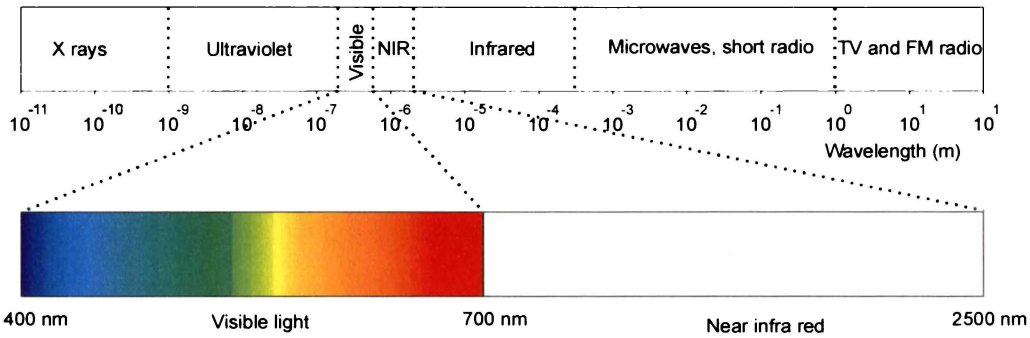

Fig. 2-4: Electromagnetic spectrum showing the location of visible and NIR light.

Spectroscopy is the study of spectra, in our case we are looking at the intensities of light in the Visible-NIR wavelength range. The light absorption of fruit tissue influences the spectral shape as so the spectra contain information about the tissues constituents. Fruit are relatively transparent in the NIR compared to the visible range (our eyes can't see through fruit) and their spectra often have a very low intensity in the visible region. In the NIR region, water is the primary absorber, and its absorption profile has been measured by many people (Prahl, 2000). The measurements presented by Segelstein (1981) are shown in Fig. 2-5.

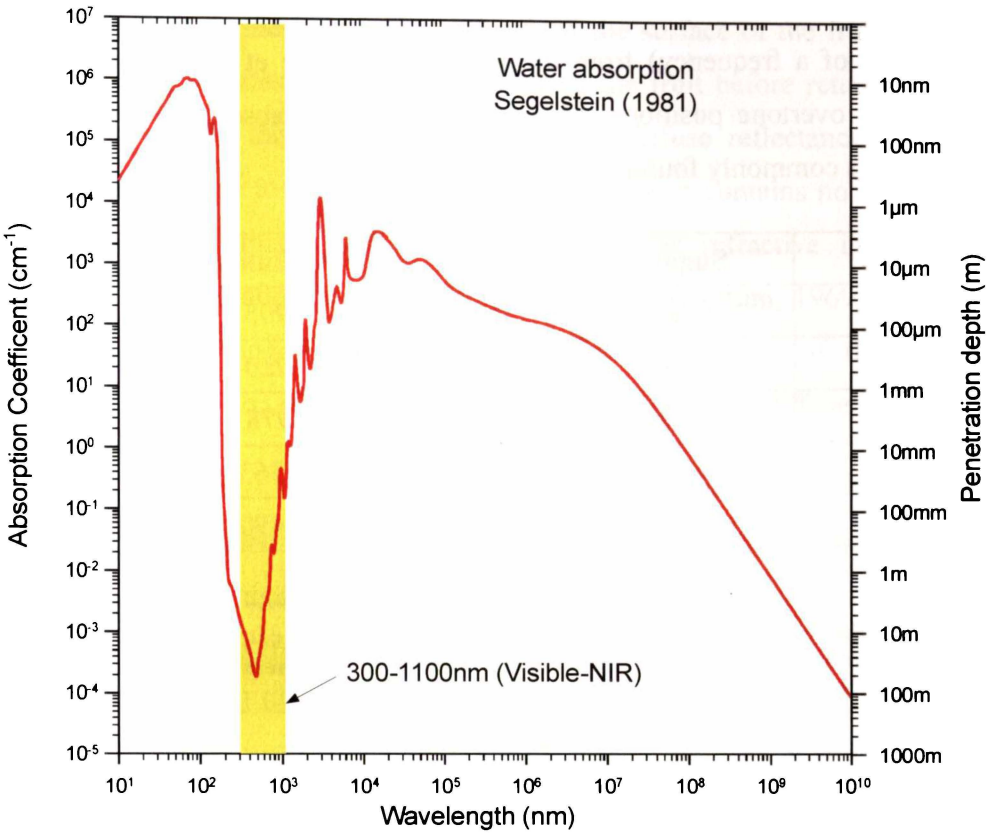

Fig. 2-5: The absorption profile of water shows low absorption in the visible-NIR region.

The large dip in the water absorption permits the NIR spectroscopy measurement for tissue. The penetration depth from 200-800 nm is more than 4 orders of magnitude greater than for wavelengths in the infrared (2.5-500 µm). The region 700-800 nm has been termed the 'diagnostic window' (Osborne et al., 1993; Tuchin, 2000). It is an area where the light can reach far into the tissue (see Chapter 4), contains useful information, and is conveniently within the sensitive region of low cost silicon detectors.

Spectra such as that observed for water (Fig. 2-5) arise because all molecules interact with electromagnetic radiation. When a photon of a characteristic wavelength is absorbed a molecule is raised from one of its quantum mechanical states to another. These transitions correspond to electronic, vibrational or rotational state changes in the molecule depending on the wavelength (energy) absorbed. In the ultraviolet and visible wavelength ranges the photon energy is often sufficient to excite an electron to a more energetic energy level, while in the far infrared region the energy is only sufficient to cause rotational transitions. Absorption lines due to vibrational transition are usually observed in the infrared between 2500 and 50,000 nm. All these absorption lines are quite distinct, but in the Vis-NIR range investigated in this thesis the spectra are highly convoluted,

due to the combination of a large number of broad vibrational overtones (multiples of a frequency) from the infrared (Osborne et al., 1993). Table 2-2 shows the overtone positions for some of the strong absorption lines of some compounds commonly found in fruit.

Water	Sugar	Starch	Cellulose	Protein
760	978	878	905	909
834	1380	918	920	1018
958	1440	979	978	1143
1153	1490	990	1363	1187
1409		1440	1490	1510
1460		1450		

Table 2-2: Principal absorption lines (in nm) identified in some constituents in fruit. Water and sugar were measured by transmission, the others by diffuse reflectance. (Williams and Norris, 1987; Osborne et al., 1993)

These absorptions are broad and closely spaced compared to the narrow bands found at longer wavelengths in the infrared region. As a consequence, there is a very large overlap of the absorption bands in the NIR spectra of fruit. One of the major computational challenges is the extraction of the constituent absorption bands from the convoluted spectra. Because of this computational need the NIR spectroscopy field has thrived since the introduction of computers and correlation transform techniques for the data extraction (Osborne et al., 1993).

2.2.2 Optical configurations

We define three basic modes of NIR spectrum measurement, reflectance, interactance and transmission, which are illustrated in Fig. 2-6.

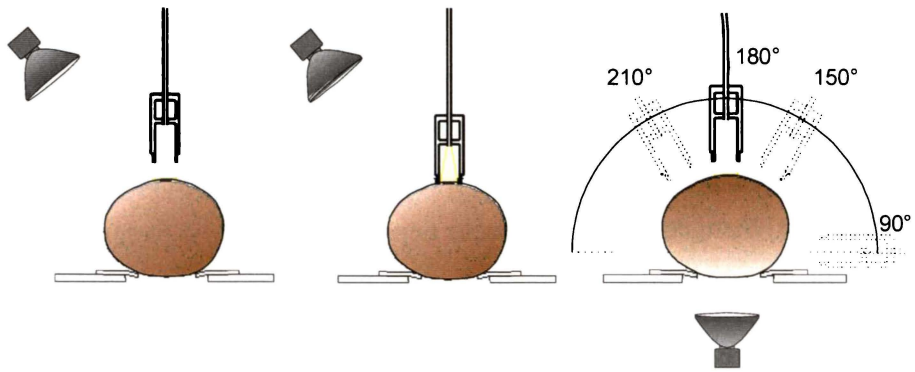

Fig. 2-6: Three different modes of measurement for NIR spectroscopy on fruit: Reflectance (left), interactance (middle), and transmission (right).

In reflectance mode the light either reflects off the surface of the fruit (specular reflection) or penetrates a small distance inside the fruit before returning to the surface and exiting the fruit on the same side (diffuse reflectance). Specular reflection adds noise to the measured spectra because it contains no information about the fruit chemistry and depends only on the refractive index at the air/surface interface and the roughness of the surface (Kortüm, 1969). Generally the detector is arranged to minimise or exclude detection of the specular component so as to maximise the measurement of the diffuse reflectance component.

In the interactance mode, the light is restricted so that the detector sees only light that has passed through some of the fruit tissue ensuring a diffuse reflectance measurement. The restriction is commonly achieved using a light seal (Fig. 2-6, middle) or bifurcated fibre probes which press up against the sample, forcing the light through at least some of the tissue before entering a return fibre.

With transmission mode the detected light passes through a significant amount of tissue. We define transmission mode in this thesis as 90° or more angle between the illuminated and sensed positions for the fruit. The measurement can either be non-contact (Fig. 2-6, right) or a light seal is used around the detector to ensure no stray light is measured. Stray light can arise from an external light source or a geometry that allows reflections from other objects to reach the detector indirectly.

The relative positioning of the light source, fruit, and detector is a critical factor for the successful acquisition of good quality spectra in NIR spectroscopy. The primary concern is to maximise the signal to the detector without saturating it yet ensuring maximum information is obtained from the fruit being analysed. Reflectance measurements are arranged to minimise specular reflection, by setting the angles of illumination and detection or shadowing the specular region in some way. For interactance and transmission it is very important to have good light sealing at the illumination point or other suitable geometrical arrangements to avoid detection of stray light that hasn't passed through the fruit.

Once the optical arrangement is set it must not be changed while recording spectra both for calibration and for using a model developed to predict fruit properties later. Slight geometric adjustments alter the region viewed by the detector, which

will lead to systematic spectral changes and overall signal strength changes. If the light source or associated optical components (such as a condenser lens or reflector) are altered then there will again be a corresponding change in the detected signal. A calibration will be difficult to develop with such a large unaccounted variation in the spectra. If the calibration has already been developed and the system is altered then a large bias is likely to be introduced to the predictions. Hence attention must be paid to ensure light source stability, repeatable sample positioning, adequate light sealing for all the samples, and firm mountings for the source, sample and detector.

After collecting a raw spectrum it is commonly corrected for systematic instrumentation and experimental effects by subtracting the dark current and perhaps dividing by a reference signal to isolate the measurements from environmental changes (e.g., the gradual reduction in source intensity due to bulb degradation). If the raw spectra are divided by a reference spectrum they are called 'relative spectra'.

It is desirable to maximise the detected signal to optimise both the signal to noise ratio and the speed of spectrum acquisition. In this thesis the detectors are non-contact fibre based, and so the challenge is to get as much of the available light into the fibre as possible. Because fibres have a limited acceptance angle and fixed core diameter and the light being detected is diffuse, it does not help placing a lens in front of the fibre, which just collects a different set of rays (but may alter the part of the fruit viewed). The maximum amount of light that can be collected is defined by the system étendue which is the product of the image area times the solid acceptance angle and is limited by the fibre parameters (Welford and Winston, 1989).

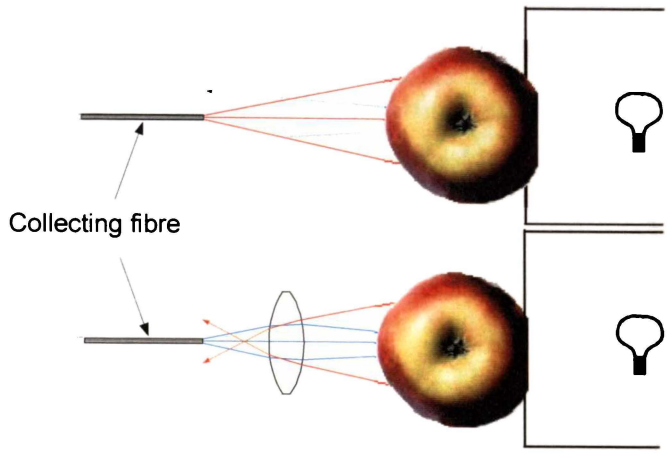

Fig. 2-7: A lens does not help a fibre bundle to collect more light from a fruit.

It may however be possible to increase the light collected by a fibre system from an illuminated fruit using non-imaging optical components, such as a compound parabolic concentrator (Welford and Winston, 1989) though this has not been tested in this thesis.

2.2.3 Chemometrics

The extraction of information from the NIR region of the spectrum is challenging for several reasons. The absorption bands are broad vibrational overtones, which overlap. This leads to co-linearity across the spectrum as the energy at nearby wavelengths will be strongly related. Many of the absorption bands are weak, especially compared to the dominant effect of water, making them difficult to resolve. In addition the relationship between light intensity and concentration is non-linear. Beer's law states there is a linear relationship between the concentration and the absorbance, which is defined as

$$A = -\log\left(\frac{I}{I_0}\right), \qquad 2\text{-}1$$

where I is the light intensity after it has passed through a sample with a constant concentration, and I_0 is the initial light intensity. This means that the intensities from the spectra will be non-linear with concentration. A source of non-linearity arises from the scattering variations from different particle sizes.

To solve these complications, techniques for extracting the information have been developed. Chemometrics has become an established scientific discipline; it involves the use of mathematics and statistics on chemical data to extract useful information. The convoluted spectra are analysed using a multivariate calibration

algorithm such as partial least squares (PLS). This process computes the relationship between constituent information and spectrum shape, enabling the use of spectra alone to predict fruit properties. The analysis can be applied to the entire spectrum or to a selected subset of the wavelength intensities. As an additional step to remove noise on the spectra and to help correct non-linearity effects, pre-treatment algorithms can be applied prior to the multivariate algorithm.

In practice, a set of fruit is used to perform the calibration. The fruit are measured spectrally and are then destructively assessed for the response values of interest (e.g., DM, SSC). Pre-treating the spectra and performing a multivariate analysis with the response values produces a calibration model. Spectra from other fruit can then be used with the model to predict the response values of those fruit without the need to destroy them. Destructively testing a set of fruit and comparing the actual response to the predicted values tests the predictive performance of the calibration model in a process called validation. Fig. 2-8 illustrates the calibration / validation process with an example; 50 apples are used to build the calibration and 50 apples are reserved to validate and assess the model.

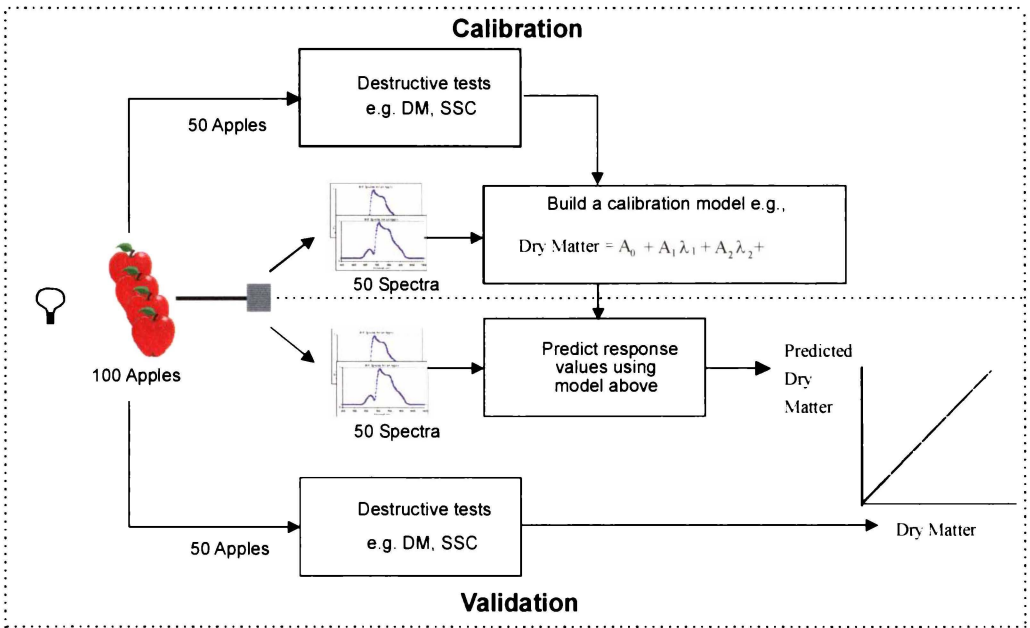

Fig. 2-8:Calibration and validation process.

The spectra (or the relative spectra) once collected are usually improved by undergoing one or more pre-treatment steps prior to applying the multivariate analysis. Table 2-3 lists and describes the pre-treatment used in this thesis.

Pre-treatment	Description
ABS	Converts the spectrum (I) to absorbance form = -log($1/I$).
SG(x,y,z)	Savitzky-Golay smoothing (Savitzky and Golay, 1964), where x is the width of the smoothing, y is the order of the smoothing polynomial, and z is the order of the derivative to be applied to each spectrum.
SA	Scale Area, each spectrum is scaled to have unit area. (Osborne et al., 1993)
MSC	Multiple scatter correction (Geladi et al., 1985) calculates the average spectrum and then adjusts the slope and offset of each individual spectrum to bring it as close as possible to the average.
SNV	Standard normal variate (Barnes et al., 1989) adjusts the spectrum of each sample so that the intensities are mean centred and have unit variance.
UVS	Unit variance samples, adjusts the intensity at each wavelength, across the set of samples, so that there is a zero mean and unit variance. (McGlone, 2001)

Table 2-3: Pre-treatment processes used in this thesis.

By converting the spectra to absorbance form using ABS, the data is adjusted so that they will respond linearly as asserted by Beer's law (Equation 2-1). By taking derivatives of the spectra, minor absorption peaks are given greater emphasis, and the baseline and slope of the spectra are removed. Derivatives were computed using the SG function in combination with smoothing. The spectra often require smoothing if derivatives are applied which may blur the profiles of small overlapped peaks (Holler, et al., 1989). Scaling the spectra to uniform area attempts to correct for variations in signal intensity for different sized samples. MSC and SNV both try to remove the multiplicative interferences of scatter and particle size. The unit variance samples (UVS) pre-treatment performs the transpose of SNV and was investigated to see if it offered any benefit for the calibration.

There are many different calibration methods in addition to PLS (e.g., neural networks (NN), multiple linear regression (MLR), principle component regression (PCR)). Many will give similar performance but some will give quite different results. Partial least squares (PLS) is a particularly versatile technique giving satisfactory performance in a wide variety of situations (Martens and Næs, 1996) and is used throughout this thesis for calibration modelling. PLS was first

proposed by Wold (1966) and several variations of the basic algorithm can be found in the literature (Young, 1994). The aim is to find a mathematical relationship between a set of independent variables, the X matrix ($N_{objects} \times K_{wavelengths}$), and a set of dependant variables, the Y matrix ($N_{objects} \times M_{measurements}$). The resulting model has the form:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \qquad 2\text{-}2$$

where **B** is the matrix of regression coefficients obtained from PLS analysis, and **E** is the matrix of residuals (Osborne et al., 1997).

Reducing the number of variables (wavelength positions) before applying a PLS analysis can often improve a calibration. The challenge for this wavelength selection approach is to know which wavelengths to select and how should it be done. For a 256-wavelength spectrometer it is not possible to build and test calibrations with every possible combination of these wavelengths due to the shear number of possible permutations. A popular manual or semi-manual approach involves selecting wavelengths based on chemical absorption knowledge and experience, followed by a multiple linear regression (MLR) search (performed by a computer) to find complementary wavelengths. A correllelogram can be used to aid manual selection (Osborne et al., 1993). In a correllelogram the x-axis is wavelength and the y-axis the correlation between optical data at a given wavelength and the response value for which the calibration is being sought. Alternatively a computer can be used to search through all of the wavelengths in a patterned way, which is the technique predominantly using in this work. Another method of wavelength selection is iterative variable selection (IPLS), developed by Osborne et al. (1997). This technique has been used in some sections of this thesis. Recently, Roger and Bellon-Maurel (2000) used genetic algorithms for wavelength selection and achieved a three-fold improvement in prediction error in sugar content for cherries.

Having generated a model from the calibration process, response values can then be predicted from a given spectrum. If the same spectra that built the model are individually tested using the model we can estimate the standard error of calibration,

$$SEC = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{N-(F+1)}}, \qquad i = [1, 2, ..., N] \qquad 2\text{-}3$$

where y_i is the predicted response value, $\hat{y}_i$ is the actual value, N is the number of samples, and for PLS F is the number of factors used so that F+1 is the number of degrees of freedom (Martens and Næs, 1989; Bellon et al., 1993; Kramer, 1998). By computing the error between the computed response and the measured response we get a better estimate of performance than a simple correlation coefficient. The correlation coefficient for a test of calibration data is denoted R_C in this thesis.

An approximation to SEC is the closely related, root mean standard error of calibration,

$$RMSEC = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{N}}, \qquad i = [1, 2, ..., N] \qquad 2\text{-}4$$

where, as with SEC, the tested spectra are from the calibration set. When N is large, the difference between SEC and RMSEC is negligible.

It is important to realise that SEC and RMSEC only give an indication of how well the regression was able to fit the calibration data set. They do not indicate how well a calibration will perform when predicting samples independent from those in the calibration set.

If we take a new independent set of fruit and make predictions using a previously formed calibration model and compare the predictions to the actual response values we get a much improved assessment of the model performance. To compute this measure of predictive ability the standard error of prediction (Martens and Næs, 1989; Kramer, 1998),

$$SEP = \sqrt{\frac{\sum(\hat{y}_i - y_i - bias)^2}{N-1}}, \qquad i = [1, 2, ..., N] \qquad 2\text{-}5$$

is often used. The bias is defined as,

$$bias = \frac{\sum(\hat{y}_i - y_i)}{N}, \qquad i = [1, 2, ..., N] \qquad 2\text{-}6$$

which is a systematic error in the model being applied to the data. Other authors (Bellon et al., 1993; McClure, 1994) have defined SEP slightly differently. In this thesis we call their definition the root mean standard error of prediction,

$$RMSEP = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{N}} \qquad i = [1, 2, ..., N] \qquad 2\text{-}7$$

which is commonly used in reporting prediction errors in this field. Again this measure is only for testing a calibration with independent sets. RMSEP measures the mean standard error without compensating for any possible bias. If the bias term in Equation 2-6 is small the difference between the two definitions is minimal. If it is not Equation 2-5 will give a noticeably smaller standard error. In this thesis the ambiguity is avoided by always using RMSEP. The correlation coefficient for measured versus predicted response values from an independent set is termed, R_P.

Another definition is necessary for the cross validation (Stone, 1974; Snee, 1976) procedure. This is where a set of data is evaluated for its predictive ability using only data internal to the set. The advantage of cross validation is that it can be used when there is only a small set of data. A separate independent set is not required for testing. In full cross validation (leave one out) the calibration of a set of N samples is repeated N times, each time leaving a different sample out of the calibration and using that sample for prediction. To reduce computation time the data can be split into groups (cross validation split), and each group is used in turn for prediction from the calibration produced from the remaining groups. In this case care must be taken to randomise the data so that no group contains all of the samples with a particular property in the spectral data (Martens and Næs, 1989). The root mean standard error of cross validation, RMSECV is calculated from the predicted and actual pairs as each is left out in turn using a formula identical to 2-7.

A key issue with calibration is the complexity (number of factors for PLS) of the model and the potential in modelling for over fitting. The complexity is the number of independent factors that have to be estimated. As the number of factors increases so does the chance that the estimation process draws noise and other spurious phenomena from the calibration data into the resulting calibration model (Martens and Næs, 1989). When this happens the fit will look good but, when

tested on an independent set, will perform poorly. This phenomenon is called over fitting. By monitoring the RMSECV and RMSEC as the complexity is increased, an appropriate level can be chosen. Typically the RMSEC decreases as more factors are added. RMSECV decreases to a minimum and then begins to increase again as the data becomes over fitted. This is the main reason why SEC or RMSEC is a poor indicator of prediction performance. The minimum point for RMSECV may still not indicate the best number of factors to use. A simpler model, with less factors, offering only slightly larger RMSECV may be a better choice. Once the number of factors has been decided, the calibration should be rebuilt at that level of complexity and tested with an independent set to give RMSEP.

2.2.4 Progress of NIR spectroscopy on fruit

Karl Norris and his co-workers started to use near-infrared light to detect the difference in moisture content in grain in the 1960's (Norris and Hart, 1965; Massie and Norris, 1965). Since then there have been many studies on agricultural products for which NIR spectroscopy has been able to provide useful information for quality factors. Williams and Norris (1987) compiled a comprehensive book on the basic fundamentals of NIR in the agricultural and food industries.

NIR spectroscopy on fruit is difficult for a number of reasons:

- Fruit have high water content resulting in strong absorption above 950 nm.
- There is substantial variation in produce size and shape causing spectral intensity variations and in the case of reflectance, specular noise.
- There is often considerable spatial variation of the target attributes within the fruit.
- There is a substantial variation in the light scattering within fruit.

In spite of these difficulties many authors (Table 2-4) have used NIR spectroscopy to assess a huge and still increasing range of fruit. The papers listed are not a complete set, but key papers that have been quoted in this thesis have been listed. Kays et al. (2000) have surveyed some additional papers that use NIR spectroscopy for fruit analysis.

Fruit	Study
Apple	Bellon and Sévila 1993; Lammertyn et al. 1998 [I] (A, SS, F); Moons and Dardenne 2000 [R] (A, SS, F, DM).
Cantaloupe	Dull 1989.
Date	Dull et al. 1991 [T] (M).
Kiwifruit	Slaughter et al. 1996 (SS,S); McGlone and Kawano 1998 (DM, F, SS); Osborne et al. 1999 [I] (SS); Mowat and Poole 1997a [R]; Schaare and Fraser 2000 [R,I,T] (DM, SS).
Honeydew Melon	Dull et al. 1992 (SS).
Mandarin	Kawano et al. 1993 [T} (SS); Fujiwara et al. 1995 (A, SS); Miyamoto et al. 1998 [T] (SS, A).
Nectarine	Slaughter 1995 (SS).
Papaya	Birth et al. 1984; Slaughter et al. 1999 (SS).
Pawpaw	Greensill and Newman 1999 [T].
Peach	Bellon 1993; Kawano et al. 1992 [I] (SS); Peiris et al. 1998a [T] (SS).
Persimmon	Mowat and Poole 1997b [R].
Pineapple	Guthrie and Walsh 1998 [I, T] (SS).
Rock melon	Greensill and Walsh 2000 [I] (F, SS).
Tomato	Chen and Nattery 1980; Peiris et al. 1998b, 1999b (SS).

Table 2-4: Selected publications reporting NIR spectroscopy results on fruit. Where relevant, the mode and parameter investigated is given in brackets: [R=Reflectance, I=Interactance, T=Transmission], (A=Acid, F=Firmness, DM= Dry Matter, SS=Soluble Solids)

There is an increasing trend towards use of transmission mode NIR because it is thought to be necessary for thick-skinned fruit such as mandarins (Kawano et al., 1993) and is expected to be much better at detecting internal defects (McGlone et al., 1997). Birth and Olsen (1964) used NIR transmission for watercore assessment in apples, and Peiris et al. (1998) successfully used NIR transmission to detect section drying in tangerines. Upchurch et al. (1997) were able to measure internal breakdown in apples using 90° NIR transmission measurements. Recently, Clark et al. (2001) were able to predict the brownheart internal disorder found in Braeburn apples using transmission mode.

Another possibility is that NIR transmission may be more sensitive because the detected light traverses a greater amount of tissue. Previous studies of kiwifruit have reported SSC RMSEP = 0.39 °Brix (R_p^2 = 0.90) and DM RMSEP = 0.42%

($R_P^2 = 0.90$) using interactance (McGlone and Kawano, 1998). An earlier study by Slaughter and Crisosto (1996) used an interactance fibre optic probe and found SSC SEC = 0.72 °Brix ($R_{C} = 0.99$). Schaare and Fraser (2000) compared reflectance, interactance and transmission modes for the assessment of gold kiwifruit properties. In this first attempt, we showed interactance to be better (SSC SEP = 0.80 °Brix with R_P^2=0.93) than reflectance but were unable to show transmission to be better than interactance. It was suggested in the paper that a more powerful light source used for the transmission mode would have improved the transmission mode result.

If transmission mode is more sensitive it may enable the measurement of small constituents (< 2%) such as fruit acids. Miyamoto et al. (1998) used transmission mode on mandarins and reported acid predictions with SEP = 0.147% and R_{P} = 0.83. Sohn and Cho (2000) reported the assessment of malic acid in Fuji apples, which ranges between 0.3-0.5%, with an SEP = 0.037% ($R_P = 0.82$). Moons et al. (1997) reported malic acid prediction for apples using reflectance, SECV = 0.087% ($R_{CV}^2 = 0.94$). Note SECV is approximately equal to RMSECV. Steinmetz and Biavati (1998) reported that it is not possible to predict acid levels in oranges with low-cost non-destructive sensors. Other studies e.g., Fujiwara and Honjo (1995) have measured acid levels in fruit juice, but the wavelength range was extended to 2500 nm, a range which is difficult to use for intact fruit due the absorption of water. Acidity, SSC, and chlorophyll content can all be strongly correlated to each other and to fruit maturity. Since both SSC and chlorophyll (Bittner and Norris, 1976) are known to be measurable spectrally it is possible that the acid level is being inferred from one of these other properties without directly measuring it.

In 1989, Mitsui Mining and Smelting Co. of Japan developed and installed the first automatic sweetness-sorting machine for peaches (Kawano, 1994). It operated at 3 fruit/s and used NIR reflectance. Over the last ten years, improved sensitivity, speed and the reduced price of detectors such as charge coupled devices (CCD) and photodiode arrays (PDA) have lead to more on-line systems for fruit grading. However, there have been very few studies published on moving fruit. Bellon et al. (1993) reported a fibre based system that assessed peaches on-line with SEP = 1.04 °Brix (R_{P} = 0.81) at up to 3 fruit/s. Moons and Dardenne

(2000) developed an on-line system for assessing apples using reflectance; they reported SSC SECV= 0.49 °Brix (R^2_{CV} = 0.88), DM SECV = 0.78% (R^2_{CV} = 0.84), and acid SECV = 0.093% (assuming malic acid) with R^2_{CV} = 0.75. Hwang and Noh (2000) developed a reflectance PDA based system while attempting to minimise the variability due to fruit size. They reported SSC SEP = 0.473 (their SSC range was 9.4-18.2 °Brix). All these systems require alignment or hand placement of the fruit which will increase the cost of processing and thereby limit the technique to high valued fruit only. Automatic alignment of apples is difficult but is becoming more reliable; Throop et al. (2001) reported a system capable of 94% alignment for red delicious, 82% alignment for large Fuji apples, and between 60% and 100% for other varieties. The performance of NIR spectroscopy on randomly oriented fruit is an area than needs further investigation.

Commercial NIR graders are set up all over the world, though still in very small numbers outside of Japan. Fruit-grading equipment in Japan has continued to improve and transmission grading is now being used on at least mandarins and apples (McGlone et al., 1997). Peaches and pears are also being measured commercially. In these systems, the fruit are hand placed onto the conveyers, which move at 5 fruit/s (Kawano, 1998). Fantec, another Japanese company has installed their 'F5' transmission grader in Italy under the company name Sacmi. There are currently four large scale commercial graders in Korea: Two are from Korean companies, Hungsun Co. Ltd. and Spectra Science Co., Ltd., one is imported from Japan, and one is being trialed from Compac Sorting Equipment Ltd., New Zealand (Lee, Kang-Jun, personal communication, 2001). In the United States, Agricultural Innovations Ltd. is selling a transmission grader for fruit (Kays et al., 2000), and at least one other company (Compac) is also selling fruit sorting equipment there. In spite of the number of companies trialing the technology, fruit graders are still plagued by many problems including the need for constant recalibrations and unexplained losses in performance. Hence there is a continuing need for further research to improve future systems.

2.3 Light propagation in fruit tissue

2.3.1 Theory of diffuse scattering and absorption

The study of NIR light interaction in tissue has been under investigation for many years, mainly in the medical and biomedical fields. The theory presented here is well established and much of it is comprehensively covered by Tuchin (2000). The characteristic absorption and scattering of light in the tissue determines the spatial distribution of light absorbance (or the detectable transmitted light). In the medical field this is important for therapeutic uses such as laser surgery and photodynamic therapy (Wilson and Patterson, 1986). For fruit, knowing where the light penetrates may allow better design of spectra acquisition modes and provide knowledge about how internal defects may be detected at different wavelengths.

NIR light in fruit is predominantly scattered at the cell walls (Birth, 1978; McGlone and Kawano, 1998) due to the discontinuities in refractive index between the aqueous regions and the predominantly cellulose walls (Wilson and Jacques, 1990). Cell size therefore has a strong influence on the scattering as is illustrated by the classic ground glass experiment by Birth and Hecht (1987), where green glass becomes more scattering and appears whiter as it is ground more finely.

Radiative transport theory (RTT) models the light as the propagation of single photons, which may be absorbed or elastically scattered (Wilson and Jacques, 1990). RTT is valid for an ensemble of scatterers located far from one another and has been successfully used to work out some practical aspects of tissue optics (Tuchin, 2000), and therefore the justification is largely empirical. The fundamentally more rigorous electromagnetic wave theory has not yet lead to practical solutions (Wilson and Jacques, 1990). The rates of radiant energy loss are defined by the absorption (μ_a) and scattering (μ_s) coefficients, which are in units of mm^{-1}. Typically in the 650-1000 nm wavelength range, μ_s for fruit is of the order 1.0-2.5 mm^{-1}, implying a mean distance between scattering events of 0.4-1.0 mm, and shows a gradual linear reduction with increasing wavelength (10-35% over this wavelength range). The absorption coefficient, μ_a, varies from 0.002-0.04 mm^{-1}, implying a mean distance between absorption of 25-50 mm, and is very strongly influenced by water and chlorophyll, two dominant absorbers in

fruit (Cubeddu et al., 2001; Budiastra et al., 1998). The ratio of scattering to absorption is often called the single scattering albedo,

$$\Lambda = \frac{\mu_s}{\mu_t} = \frac{\mu_s}{\mu_a + \mu_s} \qquad 2\text{-}8$$

where μ_t is the total attenuation coefficient. The mean free path,

$$mfp = \frac{1}{\mu_t} \qquad 2\text{-}9$$

is the average distance travelled by a photon before a scattering or absorption event and in fruit is typically 0.4-1.0 mm as suggested above by the dominating influence of scattering. NIR light still penetrates deeply into the fruit tissue because most of the interactions are scattering events (due to the relative sizes of absorption and scattering coefficients), and the scattering is highly forward directed (as will be shown in Chapter 4). The degree of forward scattering is defined by the scattering anisotropy factor, g, which specifies the scattering phase function which is typically the Henyey-Greenstein function (see section 2.3.2). In mammalian tissue, g is typically in the range 0.8-0.95 (highly forward scattering) and is essentially wavelength independent (Wilson and Jacques, 1990). A study of g variation in fruit tissue has not been reported to date.

The similarity principle (Graaff et al., 1993) basically states that the value of g is not important, defining the transport mean free path as

$$mfp' = \frac{1}{\mu_a + \mu_s(1-g)} \qquad 2\text{-}10$$

to be used instead of *mfp*. This is the (larger) distance between isotropic scattering events as opposed to forward biased scattering events. The difference between *mfp* and *mfp'* is illustrated in Fig. 2-9, after a finite number of forward biased scatter events the new direction is essentially isotropic so using *mfp'* and $g = 0$ creates the same effect.

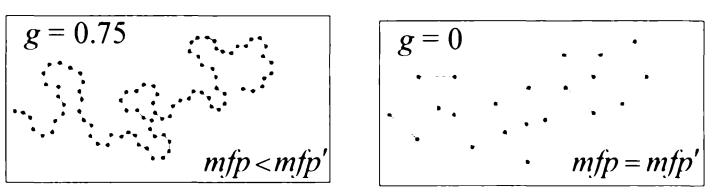

Fig. 2-9: Left: forward bias scattering at each step. Right: Isotropic scattering at each step.

The similarity principle holds as long as there is not a rapidly changing light level, which might be expected near any model boundaries and the light source input location. Equation 2-10 dictates that when g is very close to 1, indicating extreme forward scattering, mfp' can become very large, up to the path length expected if there was only absorption (25-50 mm).

Using diffusion theory the penetration into the tissue can be characterised. The effective attenuation coefficient, μ_{eff}, and the $\frac{1}{e}$ penetration depth, δ, are calculated with

$$\delta = \frac{1}{\mu_{\text{eff}}} = \frac{1}{\sqrt{3\mu_a(\mu_a + \mu_s(1-g))}} \qquad 2\text{-}11$$

The $\frac{1}{e}$ penetration depth, δ, is the distance where the light intensity is reduced by a factor of e which is not dissimilar to a time constant (read distance constant) in electrical engineering. Equation 2-11 only holds far from boundary effects and when the light is diffuse, i.e., far from the input surface where the light is initially anisotropic before gradually becoming isotropic as scattering proceeds. When this criteria is met the light level reduces exponentially with depth x below the surface

$$I(x) = I_i \exp(-\mu_{\text{eff}} x), \qquad 2\text{-}12$$

where I_i is the intensity just inside the illuminated tissue. The intensity at this point will be equal to the incident intensity (less the reflected specular component) plus the intensity due to the backscattered light (Jacques, 1989).

Boundary effects can be only ignored if the object radius exceeds 12δ (Wilson and Jacques, 1990). Fruit are relatively small objects compared to the typical penetration depth implied by the range of δ found in fruit (2-13 mm, see Chapter 4). Most fruit modelling therefore requires the definition of boundary conditions. When the photons within the model intercept the internal boundary at an oblique angle internal reflection occurs, reducing the externally observable diffuse reflectance signal. This reflection is governed by Fresnel's equations (Anderson et al., 1989) and therefore by the relative refractive indices of the materials as well as the geometry involved. There are varying (40-73%) theoretical and experimental (mammalian tissue) estimates for the proportion of total internal reflectance that occurrs at the boundary (Anderson et al., 1989; Freund, 1992; Zhu et al., 1991).

2.3.2 Monte Carlo simulations

Monte Carlo simulations involve tracing photons through a random path based on the scattering and absorption properties of the tissue. They have been successfully used to model light levels inside tissue, the main drawback being that it is extremely computationally expensive.

The Matlab (The Mathworks Inc., Natick, MA, USA) code written for this thesis was derived from the code created by Wang et al. (1995), which is available on the internet (Jacques, 1998a). The original C code has been completely rewritten and expanded on. The original equations have been converted to standard transformations for clarity. Boundary conditions have been added to allow for spherical, cylindrical and prism shaped partially reflective surfaces. The interaction occurring at the boundary can be specular or diffuse reflection; the photons can be randomly bounced back or can be removed from the simulation (escaped). In addition sections of the boundary can be removed completely. The ability to insert photons as a beam rather than at a single spot was added. Monitoring of the light distribution was enhanced so that the light collection by an embedded fibre bundle could be simulated. Additional routines were written to distribute and recombine the computations across multiple computers. Further details of the code are provided in the appendix and attached CD.

Similar but less complex simulations have been completed for mammalian tissue, e.g., Wilson and Adam (1983), Jacques (1998b), Flock et al. (1989). The Monte Carlo method is the more appropriate modelling technique for fruit because it can handle the complex illumination geometry, complicated boundary conditions, and can simulate the experimental measurements using an embedded fibre reported in Chapter 3.

In the Monte Carlo simulations used in this thesis a homogenous tissue region is defined either with no boundaries, a single planar boundary (semi-infinite), or by an enclosing geometric shape (sphere, cylinder, or rectangular). Photons are then launched at a specified position and direction and are traced until they are fully absorbed or leave a user defined enclosing boundary.

It is computationally more efficient to trace groups of photons through the tissue rather than individual photons, which terminate at the first absorption event

(Profio, 1979). A weight, representing the total amount of light available for absorption or scattering, is assigned to each group. By removing a fraction of the group's weight at each interaction point the statistical precision is improved for a given number of input photons. The probability of an interaction (absorption or scattering) occurring over a distance s is $\exp(-\mu_t s)$. So the probability of a photon being scattered in a length less than s is $P(s) = 1 - \exp(-\mu_t s)$, where $P(s)$ must be between 0 and 1. Replacing $P(s)$ with a random number, R_{rand}, in the interval (0,1) results in

$$R_{rand} = 1 - \exp(-\mu_t s), \tag{2-13}$$

and inverting 2-13 gives,

$$s = \frac{-\ln(1 - R_{rand})}{\mu_t}. \tag{2-14}$$

So the photon group travels a distance s before depositing a portion of its weight proportional to the ratio of absorption to the total attenuation rate,

$$\text{fraction reduced} = \frac{\mu_a}{\mu_t} = \frac{\mu_a}{\mu_a + \mu_s}. \tag{2-15}$$

A new direction is then chosen using the Henyey-Greenstein scattering function (Henyey and Greenstein, 1941),

$$\cos(\theta) = \frac{1}{2g}\left(1 + g^2 - \left(\frac{1 - g^2}{1 - g + 2gR_{rand}}\right)\right) \tag{2-16}$$

to generate a new scattering angle, θ, away from the current direction of motion. Another random number is used to set the azimuth angle, φ, to an angle in the interval $(0,2\pi)$ randomly rotating the new vector about the old trajectory path (Fig. 2-11).

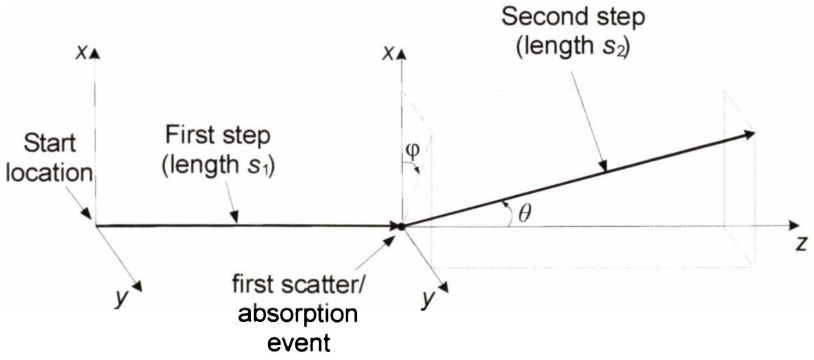

Fig. 2-10: Partial absorption and scattering occurs after each step of length *s*.

The weight that is removed as absorption at each absorption/scattering location is then deposited into a three-dimensional array, which creates a map of the region the photons travel in. The grid size of this array is critical, The cells need to be small enough in order to give adequate resolution of the light distribution pattern but not so small to create excessive statistical noise in each cell (since there would be fewer events in each).

At each scattering event the direction changes, and it is necessary to rotate the coordinate system as the photon is traced through the tissue. Fig. 2-11 shows how the coordinate system is rotated for each change in trajectory.

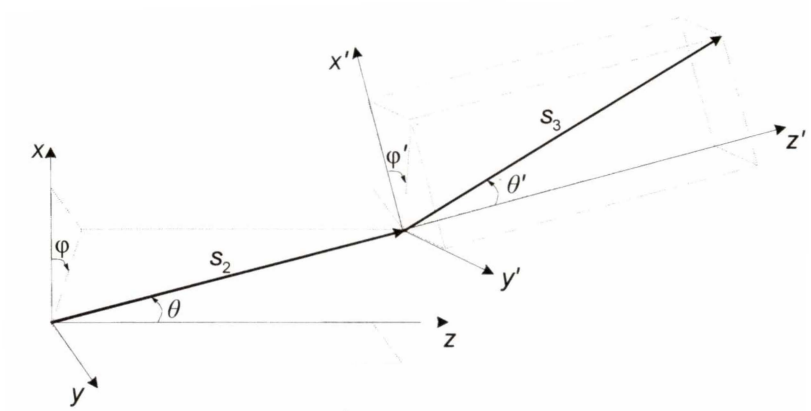

Fig. 2-11: Coordinate rotation as the photon trajectory is altered.

Before the scatter event at the point between s_2 and s_3 the current direction is defined by the cosines of the x, y, and z components,

$$\mathbf{d} = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = \begin{pmatrix} \sin\theta.\cos\varphi \\ \sin\theta.\sin\varphi \\ \cos\theta \end{pmatrix}. \qquad 2\text{-}17$$

To compute the next direction, $\mathbf{d}'$, on the x,y,z axes (as opposed to the x',y',z') we need to combine the current direction with the latest scattering parameters

(θ',φ'). This is achieved by transforming the new scattering direction onto the same axis as the current direction,

$$\mathbf{d}'_{x,y,z} = \left(\mathbf{R}_z(-\varphi).\mathbf{R}_y(\theta)\right)\mathbf{d}'_{x',y',z'}, \qquad 2\text{-}18$$

where $\mathbf{R}_z(-\varphi)$ is a rotation about the z-axis by $-\varphi$, and $\mathbf{R}_y(\theta)$ is a rotation about the y-axis by θ.

$$\mathbf{R}_z(-\varphi).\mathbf{R}_y(\theta) = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta.\cos\varphi & -\sin\varphi & \sin\theta.\cos\varphi \\ \sin\varphi.\cos\theta & \cos\varphi & \sin\theta.\sin\varphi \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \qquad 2\text{-}19$$

This is just one of many possible transformations which achieve the same result, the calculation of the new direction in terms of the original x,y,z axes.

The groups of photons continue on their trajectories until they escape from the boundary, are reflected by it or are completely absorbed. Clearly a partially reflective boundary can also be implemented. In simulations in this work, partially reflective boundaries of different shapes were implemented; when the photons intercept the boundary there is a percentage chance of reflection. Fig. 2-12 illustrates just ten photon group trajectories into a mandarin model where there is considerable internal reflection (see Chapter 4). For the model shown in Fig. 2-12 the computation was greatly reduced by setting $g = 0$ so that the trajectories undergo completely isotropic scattering events after each step of length mfp'. If the photons are not reflected they escape and are counted and removed from the model (red circles in Fig. 2-12). If the photons are reflected (green circles in Fig. 2-12) then depending on the definition of the boundary they are either specularly reflected, diffusely reflected or reflected randomly at the point of interception.

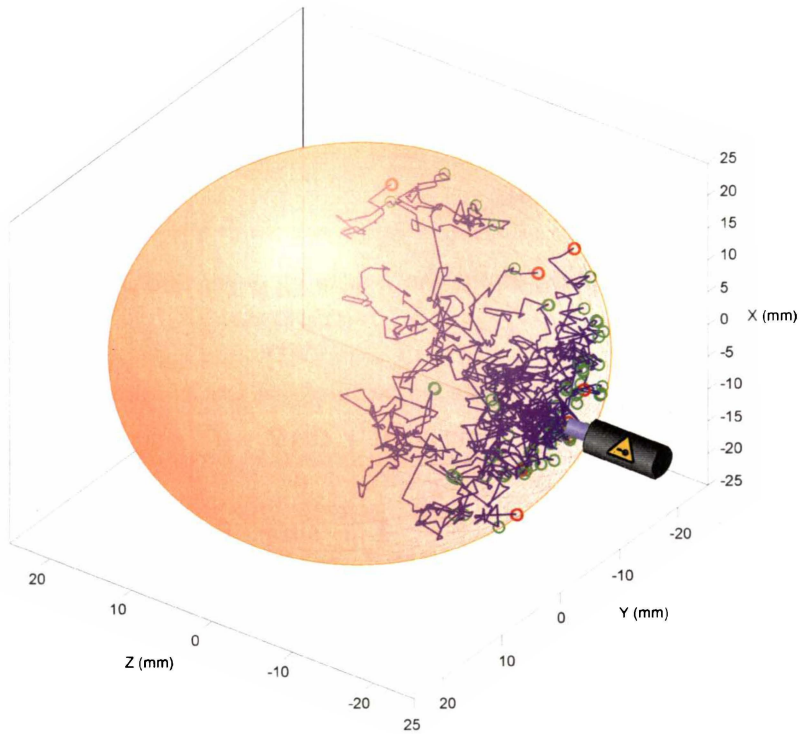

Fig. 2-12: Monte Carlo simulation of photons entering diffuse tissue with a partially reflective boundary. The green circles indicate an internal reflection; the red circles indicate an exit location.

A technique to minimise computation time is to stop tracking groups that have nearly been totally absorbed: nine out of ten low weight groups of photons are removed, and one out of ten increased tenfold in weight to compensate. This is implemented by randomly removing from the simulation (with a 90% chance) the groups whose weights have dropped to 1% of the original value. The remaining low weight groups (10% chance) have the weight increased by a factor of 10 (i.e., to 10% of the original starting weight). Hence the total weight injected into the model by the photons will equal the weight deposited within the model plus the weight of the photons that escape through the boundary, thus conserving the total input and output energies.

If the size of the model is expressed in terms of *mfp* (or *mfp'*) then the calculated model can be scaled for different values of μ_t, provided the Albedo (Equation 2-8) and *g* are constant (Wilson and Adam, 1983; Graaff et al., 1993).

Monte Carlo simulations are very suitable for parallelism; multiple computers can work on the same problem. Because each photon group's trajectory is independent of the other trajectories, simulations can be performed on different computers, and the resulting absorption distributions subsequently combined. For this work the simulations were performed on multiple Microsoft Windows PCs and Unix

servers (run out of hours when under-utilised) in order to speed up the simulation time for a given level of statistical performance.

2.3.3 Previous modelling and experimental work on fruit

Very little is known about the actual path that the light takes inside an intact fruit and the attenuation it experiences in different regions of that fruit. For optimisation of optical arrangements for internal defect detection research such knowledge would be very valuable.

Only a few investigations have been made on the optical properties of 'intact' fruit. Studies on the optical properties of potato tissue have been published by Birth (1978). Upadhyaya and O'Brien (1989) presented a model for the light diffusion in potato flesh. Chen and Nattuvetty (1980) assessed the light transmission properties for apples, oranges, and green tomatoes using a black knife to block light through progressive regions of the fruit. Hother et al. (1995) and Lammertyn et al. (2000) used similar diffuse reflectance techniques to estimate light penetration depth for apples; they measured changes at the intact surface while slices from the opposite side of the fruit were removed. Greensill and Walsh (2000) made transmission measurements to estimate the light penetration for rock melon. They measured light distribution across the cut surface at different depths into the fruit as 10 mm slices were sequentially cut away, progressing towards the illuminated side. The methods of Hother et al. (1995), Lammertyn et al. (2000) and Greensill and Walsh (2000), significantly alter the optical geometry of the investigated fruit.

Previous authors (Birth, 1978; Chen and Nattuvetty, 1980; Lammertyn et al., 2000; Greensill and Walsh, 2000) have shown that the intensity of the detected light in a fruit decreases exponentially with distance from the source. The optical properties of apples and pears have been measured by Budiastra et al. (1998) and compared to maturity level and sugar content. Recently time-resolved reflectance spectroscopy has been applied to fruit by Cubeddu et al. (2001), who provide estimates for the reduced scattering and absorption coefficients of 'intact' fruit. These estimates can now be used in Monte Carlo simulations to map out how the light levels could be changing inside fruit.

In the next chapter (Chapter 3), three experiments will be described and discussed. The investigations assess: The effect of orientation of fruit on quality predictions, show the necessity to move towards transmission measurements, consider the inability of NIR to assess acid content in mandarins, and look at the difficulties in grading moving fruit. In Chapter 4 light penetration and internal light distribution experiments, along with Monte Carlo simulations, will be described and discussed. The studies show that internal light distributions can be better assessed with a probe rather than conventional slicing methods and that the depth of penetration into fruit is highly wavelength dependant.

3

Fruit Quality Prediction Experiments

This chapter begins with an introduction to the chemical and physical methods used to assess the properties of fruit used in this study for quality assessment (dry matter (DM), soluble solids content (SSC), and acid). In this thesis, fruit parameters obtained chemically (e.g., dry matter) are referred to as response values. Using NIR light spectra and the response values, chemometric techniques can be developed to enable fruit grading. The equipment used to acquire spectra is described, and the results of three major trials are discussed. The first trial investigates the ability of NIR spectroscopy to predict kiwifruit properties. The results from a variety of orientations and optical configurations are compared. The second trial assesses soluble solids content and acid levels in mandarins using NIR spectroscopy, and the third investigates SSC and dry matter levels in apples using spectra acquired from moving fruit.

3.1 Chemical methods

Because there is often significant variation in SSC around a fruit it is often measured from two locations. In the case of kiwifruit the standard method is to remove 10 mm end caps (blossom and stem) and express the juice onto a refractometer from each in turn (Fig. 3-1). The refractometer gives a reading in °Brix which is a close approximation for the SSC of the fruit. For apples, two 10-16 mm (depending on fruit size) sections are removed from opposite shoulders (Fig. 3-2) of the fruit, and the juice is expressed from the cut surfaces. For the mandarins, the fruit is completely juiced, and a sample of the total extracted juice is used.

Fig. 3-1:Measuring the SSC for kiwifruit with a refractometer.

To assess the acid level of mandarins, 5 mL samples of juice were titrated against NaOH using phenolphthalein as an indicator.

The dry matter of a sample is calculated by dividing the sample's dry weight by the fresh (undried) weight. For kiwifruit and apples approximately 4 mm thick slices from the equator and shoulder respectively (Fig. 3-2) are placed on pre-weighed plastic petri dishes to determine the fresh weight prior to drying.

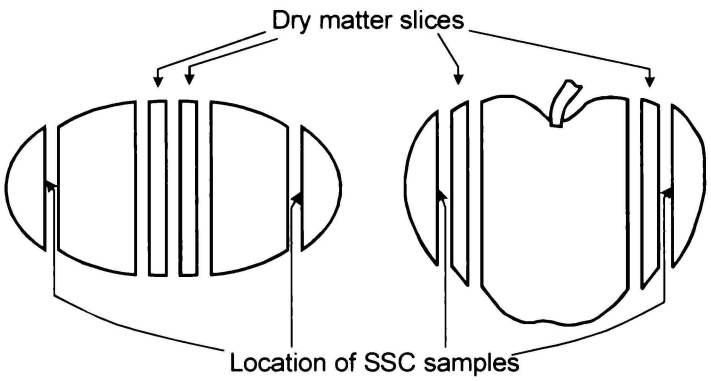

Fig. 3-2:Dry mater and SSC sample locations for kiwifruit and apples.

A significant mass is desirable in order to minimise the error from the balance reading. However, large pieces of tissue will not dry completely therefore representative slices of tissue are used. By using two slices a significant mass of tissue is sampled with minimal labour required and, in the apple case, gives better spatial sampling. The dry weights of the slices are measured with a balance after 36 hours of drying at 65°C. If the slices are thinner, 24 hours can be sufficient, but the drying time should be consistent throughout a trial since the final weight is slightly asymptotic with drying time.

To assess the whole fruit dry matter and the variation across the longitudinal axis, kiwifruit were sliced completely with a 911 series Onion Witch (Prince Castle

Inc., Illinois, USA) with 4.8 mm blade spacing (Fig. 3-3, left). For soft kiwifruit an ordinary knife can be used, however a custom made multiple blade knife (Fig. 3-3, right) saves time and improves slice thickness repeatability.

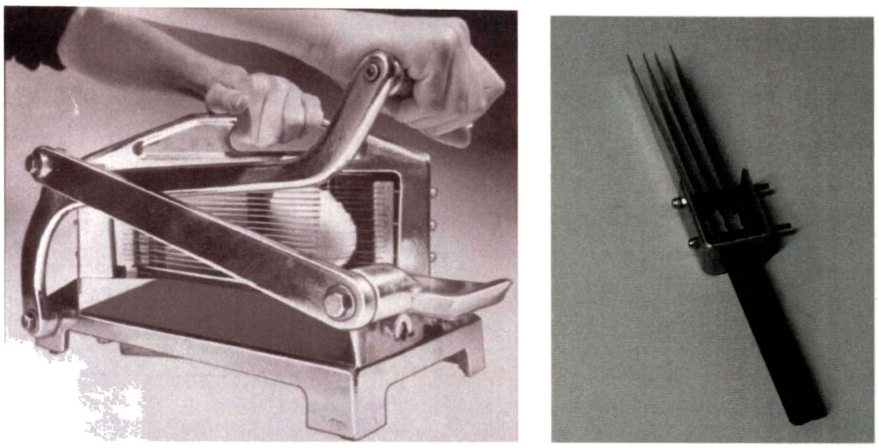

Fig. 3-3: Left: 911 series Onion Witch. Right: Custom made, multiple blade knife for cutting soft kiwifruit.

Density of kiwifruit is measured using Archimedes principle, recording the fruit weight and the weight of the fruit when it is immersed in water at a known temperature (Fig. 3-4). The system used to measure density in this work was designed by Seelye et al. (1997). Care is taken to pre-wet the fruit to reduce air bubbles on the surface of the hairy kiwifruit.

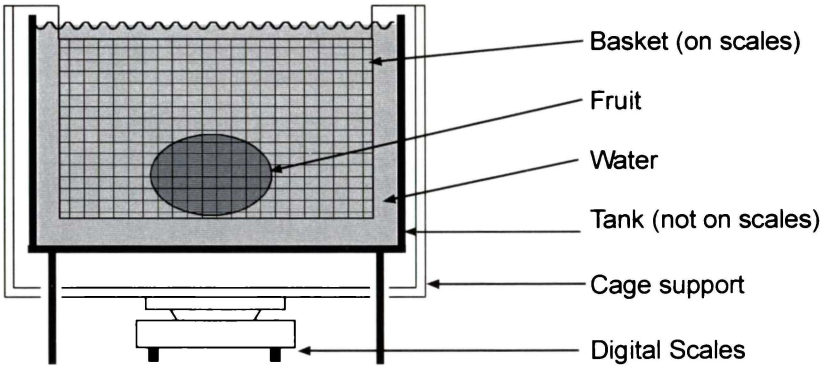

Fig. 3-4: Density measuring equipment.

3.2 Collecting NIR spectra

3.2.1 Rotary grader system

An experimental grading rig was built throughout the course of this work and is illustrated in Fig. 3-5. It could be adapted to allow different optical configurations (reflectance and transmission) at different detection angles.

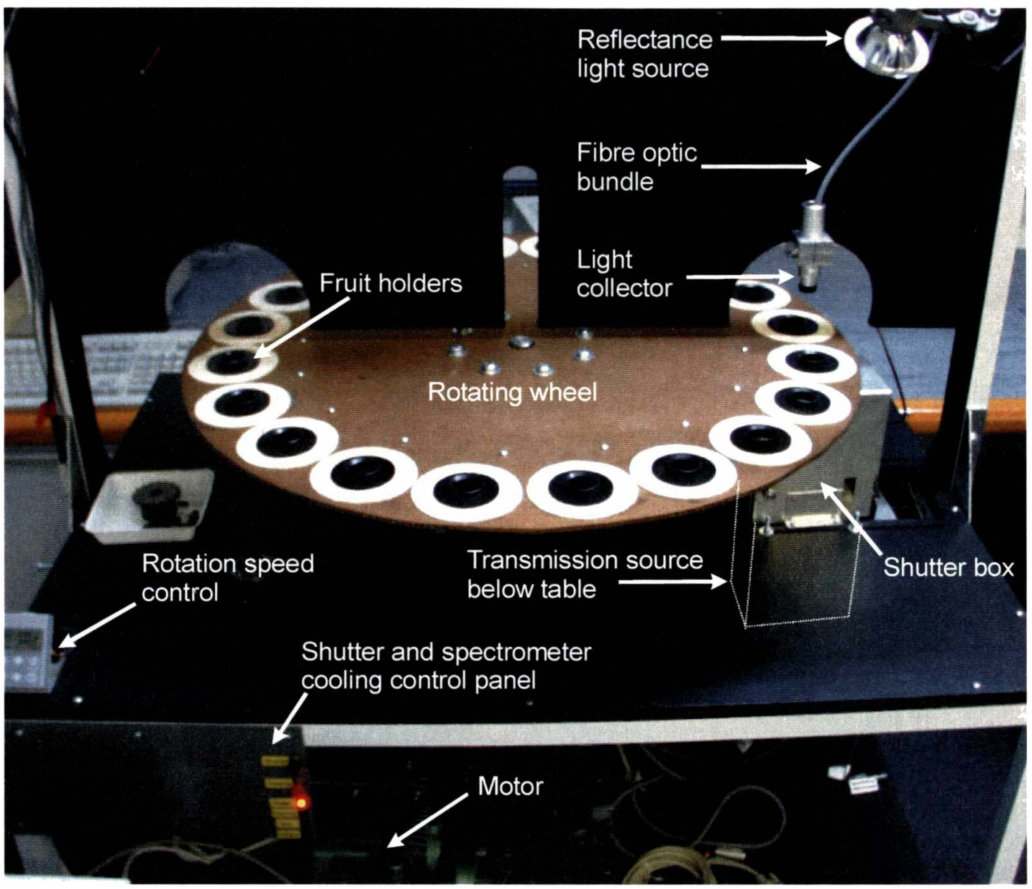

Fig. 3-5: Rotary grader system.

Light sources could be mounted both below the fruit for transmission and above for reflectance. The light was detected with a fibre-based spectrometer; the collecting end of that fibre was mounted in an aluminium holder and is shown on the right hand side of Fig. 3-5. The fruit could be moved into the measurement location in a step-wise manner using a 40-toothed cog and spring loaded indexing arm mounted directly below the fruit holding wheel (Fig. 3-6). The indexing system allowed accurate alignment for each of the 20 fruit at the correct position under the detector.

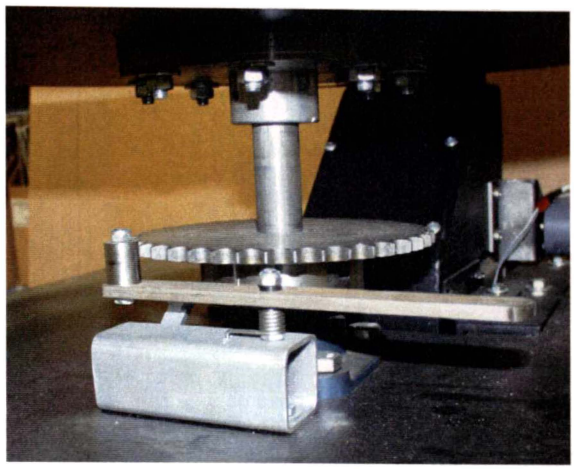

Fig. 3-6:Notching cog to allow the accurate positioning of fruit for the rotary grader.

The fruit wheel could also be spun when required at controlled speeds ranging from 11 fruit per second down to approximately 11.5 seconds per fruit using a programmable logic controller (ACS 100, ABB Industry Oy, Helsinki, Finland), and a 90 W motor (LS56, Leroy Somer, Angouleme Cedex, France).

The light source used for transmission measurements was a 250 W projector (WideoScope, AB Wiktors Mekaniska, Jarfälla, Sweden), which had been modified to use a stable DC power supply. To minimise fruit heating a computer-controlled shutter blocked the light when measurements were not being taken. The fruit were positioned on circular light seals, which could be varied in size depending on the fruit being investigated. Light was measured by a silicon-based solid-state spectrometer (MMS-1, Carl Zeiss Jena Gmbh, Jena, Germany). The spectrometer provided 15-bit resolution intensity measurements via an A/D card from Tec5 (Sensorik und Systemtechnik GmbH, Steinbach, Germany) at 256 wavelengths from 300-1150 nm with 3.3 nm resolution and ~10 nm bandwidth. A peltier cooling system and housing was designed for the spectrometer (Fig. 3-7). and kept the spectral sensor at a constant temperature of 2 ± 0.1 °C.

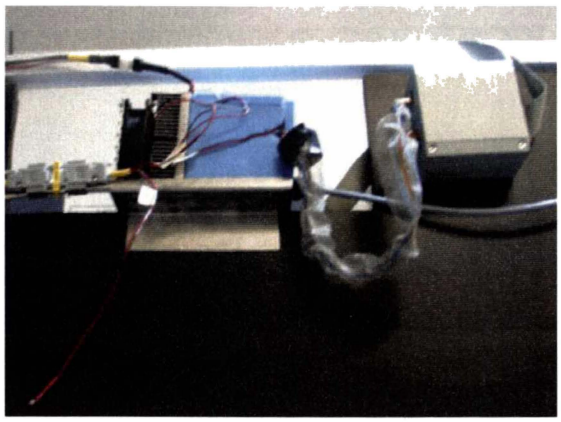

Fig. 3-7: Cooling for the Zeiss spectrometer (left, within blue foam). The wires are connected to temperature sensors. The A/D unit is shown (right) linked to the spectrometer by a shielded cable.

The spectrometer receives the light via a 1 m fibre optical cable consisting of 30 individual fibres. Each fibre had a diameter of 70 μm, and the resulting bundle collectes light over a circular region 500 μm in diameter with a NA (Numerical Aperture) of 0.2. The holder for the light input end of the fibre bundle could be positioned over an arc of 180° around the fruit and at an adjustable distance from the fruit (Fig. 3-8).

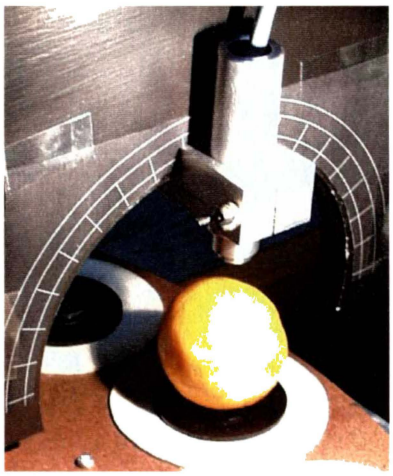

Fig. 3-8: Fibre holder to collect light from a fruit at different angles and distances from the fruit.

A 50 W tungsten filament lamp (RJT 5012 FL, Radium halogen, Germany) was used to illuminate kiwifruit and apples for reflectance measurements. For mandarins, a 70 W lamp (P14-5s, Halogen Globe, Wagner Lighting, VA, USA) was used. Both lamps were powered by a regulated DC power supply. The lamps could be mounted at different locations and angles relative to the fruit. The illumination source used for measuring the reflectance spectra of moving apples is shown in Fig. 3-9.

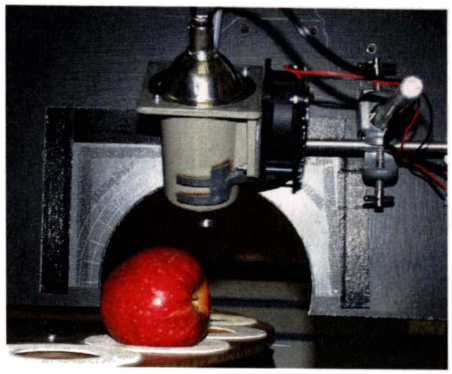

Fig. 3-9: Reflectance light source used for the moving-apple experiment.

Synchronisation of spectrum acquisition and fruit location was achieved by an optical interrupter (EX-10 series, SunX Trading Co, Tokyo, Japan), which was linked to a Tec5 PD-ISA16V3 data acquisition card, which also controlled the spectrometer. A Microsoft Visual C++ program was written to store the rapidly collected spectra and save them into a Microsoft Access database for subsequent analysis. This program used some procedures that were written by Paul Martinsen (Technology Development Group (TDG), Hamilton, New Zealand).

3.2.2 Interactance system

Previous work at TDG led to the development of a single fruit bench top NIR interactance measurement system (Fig. 3-10) (McGlone et al., 2001). This system was used to obtain comparative measurements for the other spectrum acquisition systems.

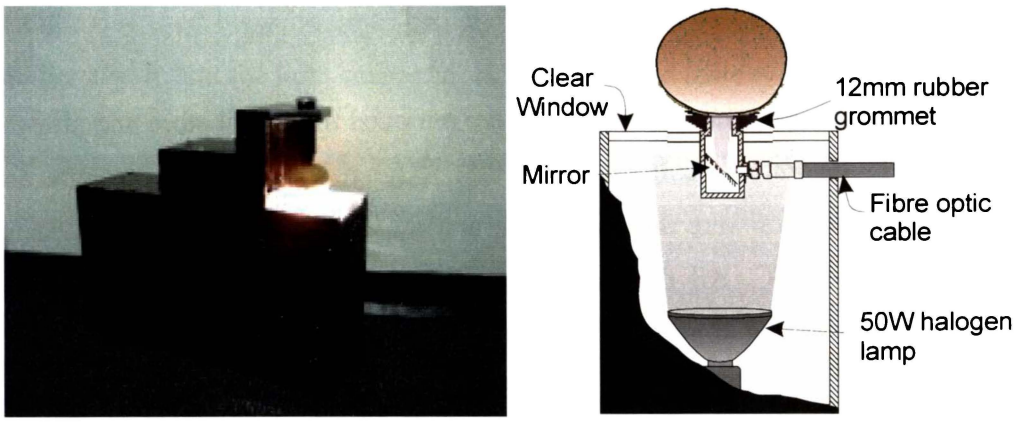

Fig. 3-10: Left: Bench top interactance system. Right: Cut away diagram (McGlone et al., 2001).

The interactance system passed the light to a single, uncooled, MMS-1 spectrometer. The kiwifruit trial described in section 3.3 used this arrangement. Later improvements made by others at TDG allowed for the light to be measured by two solid-state spectrometers connected via a Y branching fibre bundle. One of

the spectrometers was the silicon detector based MMS1 as above, the other a MMS NIR tc II spectrometer (also from Zeiss) with InGaAs detector. The InGaAs spectrometer provided 15-bit intensity resolution at 128 wavelengths over the wavelength range 940-1700 nm. Both spectrometers were cooled to constant temperature. This arrangement was used for both the mandarin and the motion trials reported in sections 3.4 and 3.5.

3.3 Kiwifruit trial

3.3.1 Experimental procedure

Kiwifruit[1] were obtained from 2 different growers via a packhouse (Baypack, Te Puke, New Zealand) on 13th and 26th May 1999. Fruit were taken from 3 different size categories: large (~150 g), medium (~112 g), and small (~84 g). Half of the medium sized fruit came from the first grower (and collection date), all the other fruit came from the second grower (and second collection date). The small and large kiwifruit were placed immediately into a 0 ± 0.5 °C cool store. The medium sized fruit were first sorted based on their density, dried and then placed into the cool store. The pre-sorting by density allowed the use of a 400 fruit sub sample of the 1000 medium sized fruit while still maintaining a desirable variation in density (and the corresponding dry matter, cf. Jordan et al., 2000). The medium sized fruit along with 190 randomly selected fruit from each of the small and large size sets were assessed spectrally using both the rotary grader (in stepwise mode) and the bench top interactance system over a period of two weeks after collection.

Prior to each assessment day, the fruit were removed from cool store and allowed to warm to laboratory temperature, ~19 °C. The reflectance, interactance and transmission spectra were then acquired, and the fruit immediately assessed for dry matter and SSC. All the fruit were measured using the industry standard method for dry matter (2 equatorial slices). Additionally, 219 of the 784 fruit were assessed using whole fruit dry matter where the entire fruit was sliced longitudinally using the Onion Witch into 4.8 mm slices. A subset consisting of 38 fruit from each of the large and small sets were also analysed for SSC using a refractometer (Palette PR-100, Digital Refractometer, Atago co. Ltd., Tokyo, Japan). Two random and separate sub samples of the medium sized fruit (one for

[1] The description of the sample sets used in the kiwifruit trail is very complicated, and it may be useful to refer to the map of sample sets given in the appendix.

each grower), were used to estimate the SSC of the spectrally analysed fruit at analysis time.

The three optical geometries used for transmission measurements on the fruit, horizontal-1, vertical and horizontal-2, are shown in Fig. 3-11. Transmission spectra for different angles (90°, 120°, 150°, 180°, 210°, 240°, and 270°) were collected for the horizontal-1 and vertical orientations. The horizontal-2 oriented fruit were measured at three positions only (150°, 180°, and 210°).

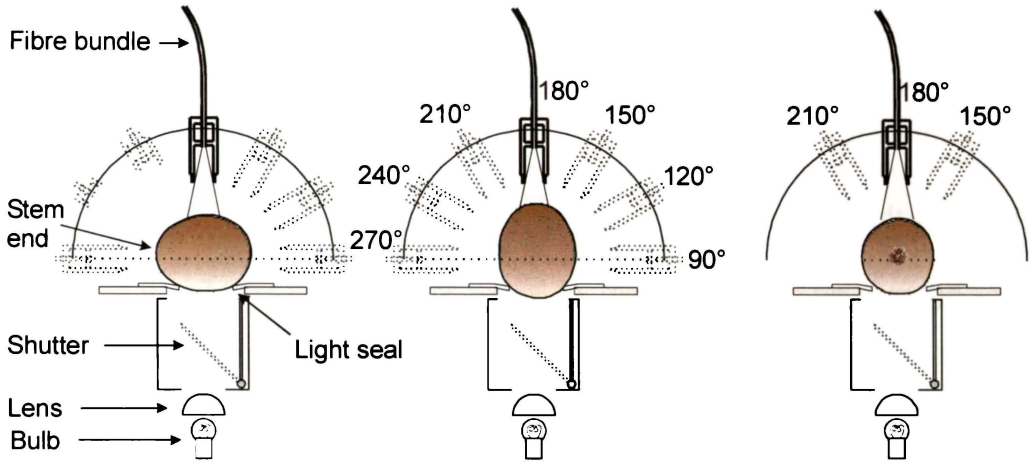

Fig. 3-11: Transmission modes of detector and fruit orientations. From left to right: Horizontal-1 (stem at left), vertical (stem at top), and horizontal-2 (stem at front).

The fruit rested on a light seal, which had an aperture 48 mm in diameter. Care was taken not to over expose the fruit to the transmission light source. No damage was observed when the exposure time was kept to less than 10 seconds. The shutter was used between times of light acquisition to protect the fruit. Having a shutter also meant the light source could be left on and thereby remain stable.

For reflectance mode (Fig. 3-12) two different orientations of the kiwifruit (horizontal-1 and vertical) were investigated. A fixed angle of 40° was used between the source and detector.

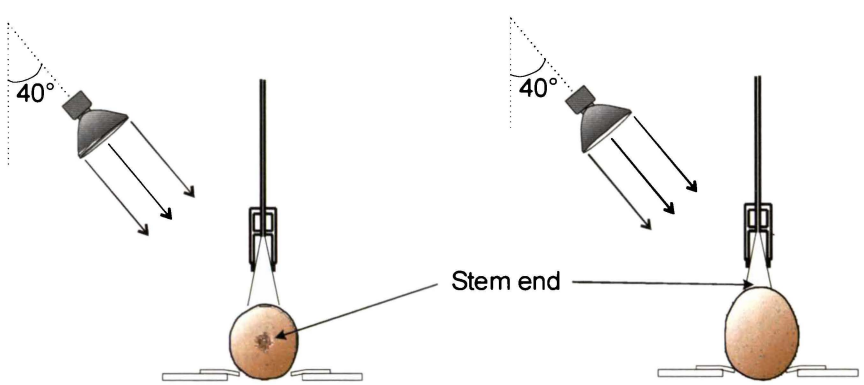

Fig. 3-12: Reflection mode fruit orientations. The source-detector angle was 40°.

The distance between the tip of the fibre and the kiwifruit varied due to the differences in kiwifruit size. The field of view when the distance was greatest (i.e., small fruit in a horizontal orientation) led to a maximum collection region of approximately 30 mm in diameter.

For a single orientation and mode, a reference and a dark spectrum were recorded prior to the measurement of each group of 19 fruit. The reference for the small and medium sized fruit was a white 70 mm polytetrafluroethylene (PTFE) sphere (supplied by Kawano, S., National Food Research Institute, Ministry of Agriculture, Kannondai, Tsukuba, Japan) which was located in the place of the 20^{th} fruit. For the large fruit an 80 mm PTFE sphere (Ceelon Plastics Ltd., Penrose, Auckland, New Zealand) was used. Dark spectra were obtained by blocking the aperture to the optical fibre.

For the interactance system (Fig. 3-10) the fruit were placed on top of a 12 mm rubber seal and illuminated from below. Light passing through the fruit and emerging within the light seal then reflects onto the entrance aperture of the fibre via a 45° mirror. The reference used for the interactance system was a flat PTFE surface (sealed from all light except the lamp) placed above the detector window.

The integration times used for each of the three modes of operation are given in Table 3-1. For the shorter integration times up to 10 spectra were continuously acquired and then averaged to improve the signal to noise ratio. For the transmission measurements only two 1000 ms spectra were measured to ensure the fruit was not over exposed to the intense light source.

	Small Fruit	Medium Fruit	Large Fruit
Reflectance	10 x 75 ms	10 x 50 ms	10 x 60 ms
Interactance	10 x 150 ms	10 x 150 ms	10 x 150 ms
Transmission	2 x 500 ms	2 x 1000 ms	2 x 1000 ms

Table 3-1: Number of spectra averaged and integration times used for kiwifruit measurements.

The integration times were chosen to maximise the signal while avoiding the case where an exceptionally transparent fruit may saturate the detector. On the rare occasion when saturation did occur the integration time was shortened. For these exceptional cases the dark current corrected spectra were scaled so they were

comparable with others within the same mode. Relative spectra were also computed by dividing the fruit spectra by the reference measurement.

It is a valid criticism that it is not fair to compare predictions from the different modes when they have not had equivalent input power. However, by choosing these integration times, the best the particular mode (given the equipment limitations) can do in a realistic time is selected. Averaging many more reflectance spectra to represent the same input power as for transmission mode improves the signal to noise ratio and therefore may improve the performance of the reflectance mode but many not be realistic. It may be more efficient to grade fruit faster instead. In these studies it was decided to set the integration time so that the detector's dynamic range is fully utilised. Perhaps future equipment improvements may allow more light to be transmitted through the fruit without damaging it and hence reducing the necessary integration time for near full scale signal on the detector used in transmission mode.

The chemometric analysis was performed using Matlab (The Math Works Inc., Natick, MA, USA) in conjunction with the PLS_Toolbox ver. 2.0 (Eigenvector research, USA). In order to obtain an optimal calibration a range of spectral transformations, pre-processing sequences, and wavelength selections were compared. Initially a 'brute force' exhaustive search was implemented to assess the effectiveness of the pre-processing on the predictive ability. Starting with either the raw spectra with dark current correction, or the relative spectra (raw divided by reference spectra) chemometric algorithms (described in section 2.2.2) were applied in different combinations. The wavelength range was restricted to 600-1000 nm. An optional transformation to absorbance units could be made followed by an optional smoothing and differentiation step using the Savitzky-Golay (SG) algorithm.

A range of parameters for the SG algorithm were tested; for a 2^{nd} order polynomial the width was varied from 3-11 points (~10-36 nm) and differentiation order varied from 0-2. Up to two pre-treatment algorithms (from a set of five) could then be applied in any combination. The algorithms tried were: scale by area (SA), multiple scatter correction (MSC), standard normal variate (SNV), second derivative, and unit variance samples (UVS).

After pre-processing, Partial Least Squares (PLS) with a four way cross validation was used to assess the predictive ability of the transformed spectra. Previous work (Schaare and Fraser, 2000) showed PLS to be the most effective algorithm for generating a calibration with similar kiwifruit data. The four way cross validation procedure splits the data into four groups and predicts on each group from a model built with the other three groups using PLS. A PRESS plot was calculated based on the mean of the four RMSEP and RMSEC curves at each number of factors. From this PRESS plot the appropriate number of factors can be determined. As the point where the prediction error is minimised without over-fitting the data.

With the degree of model complexity determined (number of factors to use) the cross validated predictions vs. actual response values then gives an estimate of the predictive ability of the model. By ranking the models by RMSECV and R^2_{CV} (cross validation prediction vs. response values) the best performing transformation and pre-treatment sequences were compared using the method of Fearn (1996). Having established the most promising transformation sequences, RMSEP was estimated by reserving one third of the samples as an independent set. Additional wavelength selection, using both manually set ranges and iterative partial least squares (IPLS, Osborne et al., 1997) were also used. Randomising the sample positions and repeating the prediction process 20 times gave an estimate for RMSEP variation.

3.3.2 Results and discussion

The fruit[2] used in this trial had been sorted into three size categories by the grader in the packhouse. Due to this categorisation the mass distribution is tri-modal and is shown in Fig. 3-13 (left). The large sized fruit exhibited a wider mass distribution compared to the smaller sizes, due to the packhouse grader bin size settings. The medium sized fruit were ranked on their density values, and 60% of the total fruit were then removed from the middle densities to reduce the number of fruit to be analysed while maintaining a large variation in dry matter. The truncated density distribution of the remaining medium sized fruit is shown in Fig. 3-13 (right), the density mean and standard deviation was 1035 ± 7 kgm^{-3}.

[2] The distinct subsets of the kiwifruit trial used to produce each result graph in this section have been listed in the appendix in a diagram to aid in interpretation of the figures in this section.

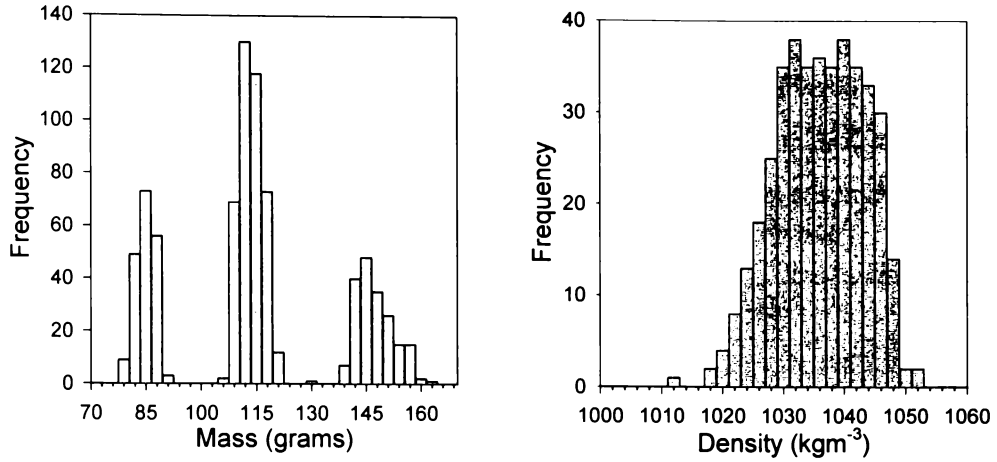

Fig. 3-13: Left: Mass distribution for three sizes of fruit. Right: Density distribution for selected medium sized fruit.

The industry standard dry matter (DM) for the kiwifruit was measured as described in section 3.1. The resulting distribution is shown in Fig. 3-14 (left) for all (~780) kiwifruit tested, the mean and standard deviation was 16.5 ± 1.7 %. The mean SSC is calculated by averaging the SSC recorded at each of the stem and blossom end caps. For Grower 1 (May 13th) a random sub samples of 20 medium sized fruit gave SSC = 8.1 ± 0.9 °Brix showing only slight progression of the starch to sugar conversion at the time of analysis. For grower 2 (May 26th) SSC = 10.3 ± 1.5 °Brix was found for 20 medium sized fruit. The fruit from grower 1 was harvested 2 weeks earlier than the fruit from the second grower, which will have contributed to the lower SSC level measured at assessment time. The distribution for the mean SSC level at assessment time for 38 fruit from each of the small and large fruit sets (both from grower 2) are shown in Fig. 3-14 (right).

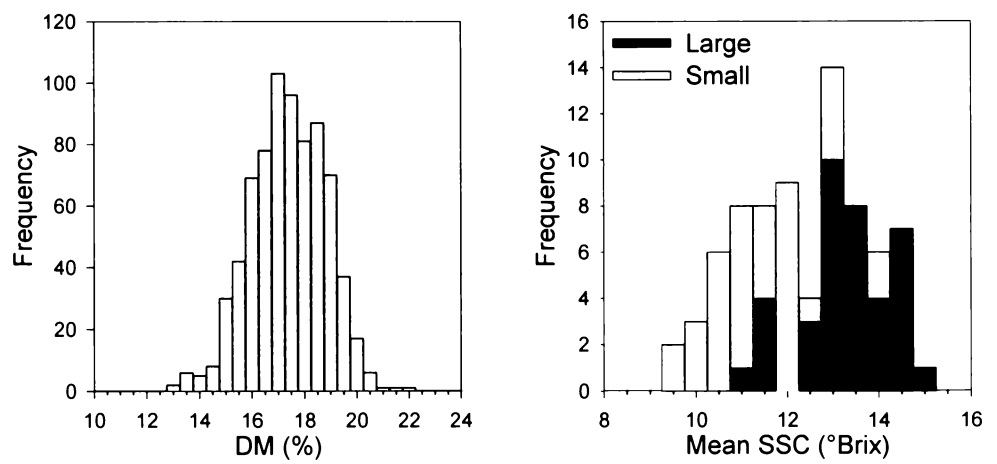

Fig. 3-14: Left: Representative dry matter distribution for all kiwifruit tested. Right: Mean SSC distribution for a sub sample of the small and large kiwifruit sets.

The large fruit had a larger mean SSC (13.1 ± 1.0 °Brix) level than the small fruit (11.3 ± 1.1 °Brix) possibly indicating a more advanced maturity. For the subset tested, the correlation between SSC readings on the two end caps was $R^2 = 0.83$. The mean SSC level for the blossom end cap was 1.2 ± 0.2 °Brix higher than the mean SSC at the stem (Fig. 3-15).

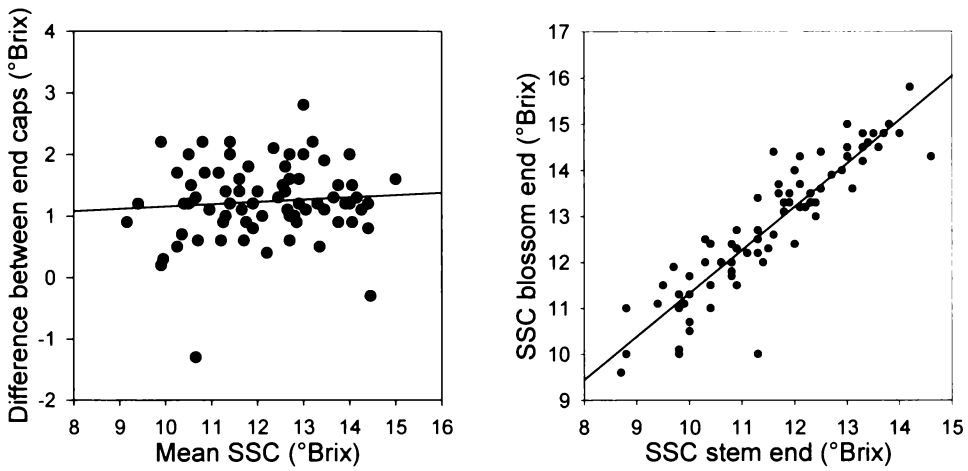

Fig. 3-15: End cap SSC analysis for a subset of the small and large sized kiwifruit.

A selection of the kiwifruit were completely sliced to enable the assessment of the whole fruit dry matter and the longitudinal variation. Fig. 3-16 shows the variation in the flesh regions for four sets of kiwifruit. The first set consisted of medium sized fruit with densities in the middle to high range and showed on average practically no variation in dry matter along the longitudinal axis. The other three sets of kiwifruit came from a different grower (thus harvested from a different location) than the first set and showed on average a slight increase in dry matter from the blossom end (slice 1) to the stem end (slice 14). This is the opposite from the trend observed for the SSC in Fig. 3-15, where the higher SSC was measured at the blossom end. Slices 1 and two were discarded from this analysis because they had a much higher dry matter due to the increased ratio of skin and the presence of withered stamens and styles found at the blossom end. Similarly slices at the stem end were also discarded, the hard (sclerified) plug strongly contributes to the dry matter at this end.

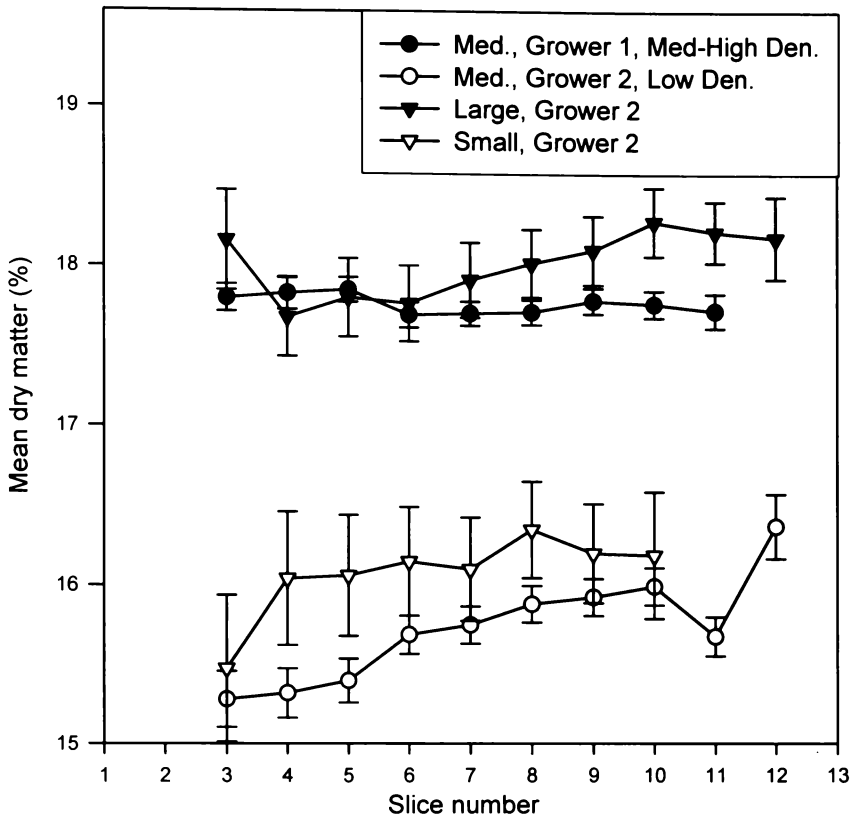

Fig. 3-16: Variation in dry matter along the longitudinal axis for four sets of kiwifruit. Blossom end is slice number 1, Stem end is slice number 14.

Table 3-2 shows the results from fitting a linear model to the longitudinal dry matter (LDM) variation for each fruit in each of the four sets and recording the slope of the fit. The standard deviation (SD) of the slopes for each set is also shown and indicates a large amount of variation found for individual fruit.

	Num. Samples	Mean LDM (%)	Mean slope (%LDM/Slice)	SD for slopes (%LDM/Slice)
Med, G1, M-H	113	17.8	0.058	1.27
Med, G2, L	69	15.7	0.593	1.06
Large, G2	18	17.9	0.475	0.38
Small, G2	19	16.2	0.503	1.76

Table 3-2: Longitudinal variation in dry matter across four sets of kiwifruit; Med, Large, and Small refer to the fruit size, G1 = Grower 1, G2 = Grower 2, M-H = Medium to high-density fruit, L = Low density fruit.

For the first set used in the analysis an investigation of the dry matter sampling method was made. Fig. 3-17 compares the dry matter computed from 2 slices or 4 slices to the dry matter from the whole fruit.

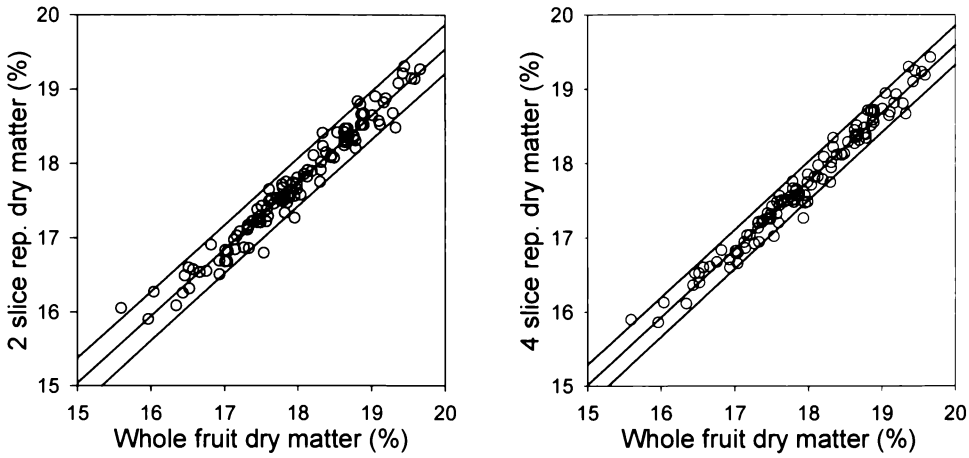

Fig. 3-17: Correlation between whole fruit dry matter and two slice (left) and four slice (right) dry matter estimates of dry matter for medium sized fruit with medium to high densities.

The two slice dry matter (the industry standard) is strongly related, $R^2 = 0.960$, to the whole fruit dry matter. For the fruit in this comparison the industry standard method under-estimates the whole fruit dry matter by 0.25 ± 0.05 %. Using four slices a slightly stronger relationship, $R^2 = 0.974$, is found however, it is more labour intensive.

The density values for the medium sized fruit correlated to the representative dry matter values (Fig. 3-18) with an $R^2 = 0.76$. This justifies our earlier preselection using density to obtain a range of dry matters and is consistent with other studies (Jordan et al., 2000).

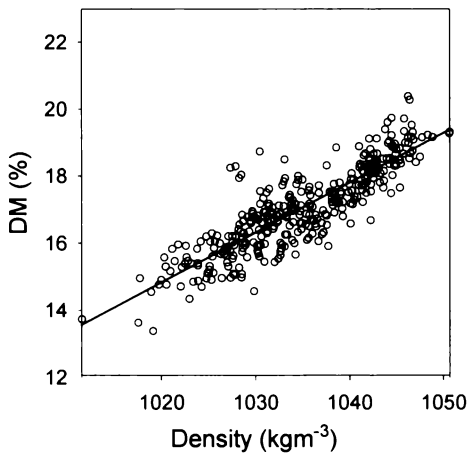

Fig. 3-18: Correlation between the representative dry matter and the density for medium sized kiwifruit.

Fig. 3-19 shows typical reflectance, interactance, 120° transmission and 180° transmission relative spectra shapes for a kiwifruit in the horizontal-1 orientation.

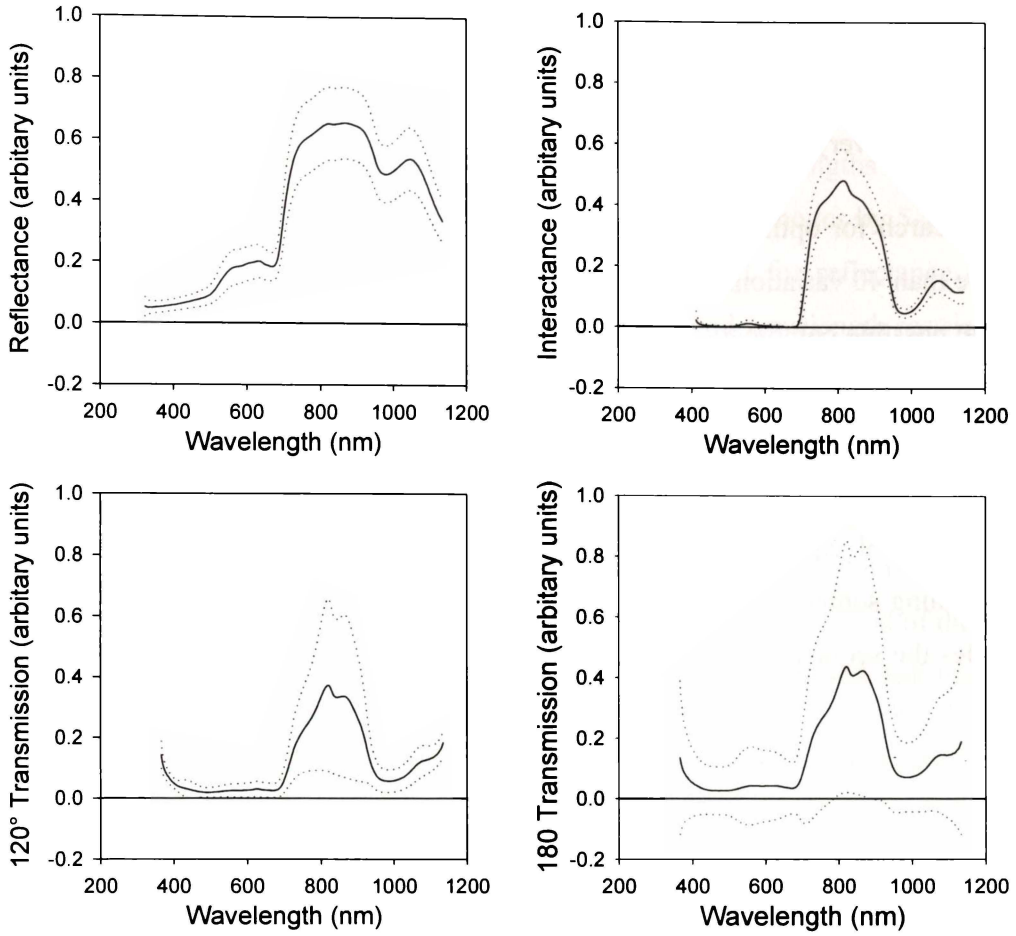

Fig. 3-19: Typical spectra measured for kiwifruit in different modes. The solid lines are the mean spectra recorded; the dotted line indicates 2 standard deviations from the mean spectrum.

The solid line is the mean spectrum recorded for all of the fruit sampled in each mode of operation. The dotted lines indicate a distance two standard deviations each side of the mean spectrum; most of the recorded spectra had shapes that fit between the two dotted lines.

The transmission modes showed the greatest amount of spectral variation, which is expected since both the fruit size and the maturity will have a large influence on the detected signal (Schaare and Fraser, 2000). The interactance spectra shape can be viewed as intermediate between those of reflectance and transmission modes. Both the interactance and transmission modes exhibit the greatest light transmission in the wavelength range 700-950 nm which is consistent with the presence of a 'diagnostic window' in fruit (Fraser et al., 2000; Osborne et al., 1993; Tuchin, 2000). A large trough attributed to water absorption is observed in all the spectra approaching 958 nm. An additional smaller trough near 834 nm is also attributed to water absorption. Chlorophyll absorbs strongly at 645 nm and

663 nm, and an absorption trough is observed in this region. Because of the reduced light levels reaching the detector and the decreased sensitivity of the silicon detector the regions below 450 nm and above 1100 nm have a rapidly decreasing, very poor signal to noise ratio.

The search for optimal pre-treatment sequences for each optical mode produced more than 40 variations for the best performing pre-treatments from the full set of 1200 possible combinations. Within this best set no sequence seemed to be better than any other with certainty ($P > 0.05$) (Fearn 1996). For the reflectance and interactance modes, where the predictions were poorer, there were even more (~80) pre-treatments that performed well. Although many were only slight variations of each other. For example, the width of the Savitzky-Golay (SG) smoothing sometimes had only a slight effect on the RMSECV. For transmission modes the sequences predominantly started with the relative spectra (as opposed to raw spectra or absorbance spectra) and included a second derivative step. The interactance mode sequences favoured the raw spectra and did not use derivatives. Of all the variations tested the best usually included area scaling and either MSC or SNV. Table 3-3 shows for each mode a selection of the sequences, which gave good performance. These pre-treatments all produced RMSECV that could not be distinguished from each other as per the criterion described by Fearn (1996). The pre-treatment highlighted in bold was selected for all subsequent calibrations for that type of spectra.

	Pre-processing sequence
Reflectance	**Relative spectra, SG(9,2,0), area scaling, MSC**
	Relative spectra, SG(5-9,2,0-1) or SG(11,2,2), area scaling
	Absorption, SG(7-11,2,0), area scaling
Interactance	**Raw spectra, SG(3,2,0), area scaling**
	Raw spectra, no SG smoothing, area scaling
	Raw spectra, SG(3-11,2,0), area scaling, MSC
Transmission	**Relative spectra, SG(9,2,2), area scaling, MSC**
	Relative spectra, SG(9-11,2,2), area scaling, MSC or SNV
	Relative spectra, SG(9-11,2,2), MSC or SNV

Table 3-3: Pre-treatment procedures producing the best RMSECVs. The numbers following SG indicate the width, order and derivative order respectively; if a range is given then the all parameters within that range are also suitable.

After choosing the pre-treatment, limiting the wavelength range led to further improvement in prediction performance. Models that used only the range 800-1000 nm produced simpler models (in terms of the number of factors needed) and in most cases a lower RMSEP. Reducing the number of wavelengths further with IPLS reduced the RMSEP for transmission and interactance modes by 5-10%, but a large bias was often introduced into the validation set for reflectance mode predictions. With the chosen pre-treatment sequences and wavelength restrictions the PRESS plots were examined manually to determine the appropriate number of factors, F, for the calibration.

Table 3-4 compares four modes of operation, IPLS was not used on these simulations, but the wavelength range was restricted to 800-1000 nm. The fruit were orientated in the horizontal-1 position. For the DM predictions all of the fruit sizes are included. For density only, the medium size fruit were used, and for SSC only 76 of the small and large fruit had response values. F is the number of factors used in the calibration, N is the number of validation samples, R_P^2 is the correlation between the predicted independent validation spectra against the fruit's corresponding response values (DM, density or SSC).

		DM			Density			SSC		
Mode	F	N	R_P^2	RMSEP (%)	N	R_P^2	RMSEP (kgm^{-3})	N	R_P^2	RMSEP (°Brix)
Reflectance	7	238	0.69	0.78	115	0.62	4.7	26	0.64	0.86
Interactance	5	244	0.79	0.64	115	0.82	3.1	26	0.87	0.51
Transmission120°	5	244	0.87	0.52	121	0.80	3.4	26	0.86	0.56
Transmission180°	5	244	0.90	0.45	121	0.83	3.1	26	0.91	0.47

Table 3-4: Prediction performance for various modes of spectra acquisition.

The results from Table 3-4 are plotted in Fig. 3-20. The error bars indicate the standard deviation for 20 repeated prediction tests with randomised sample splitting; they give an indication of the variability that can be expected for RMSEP.

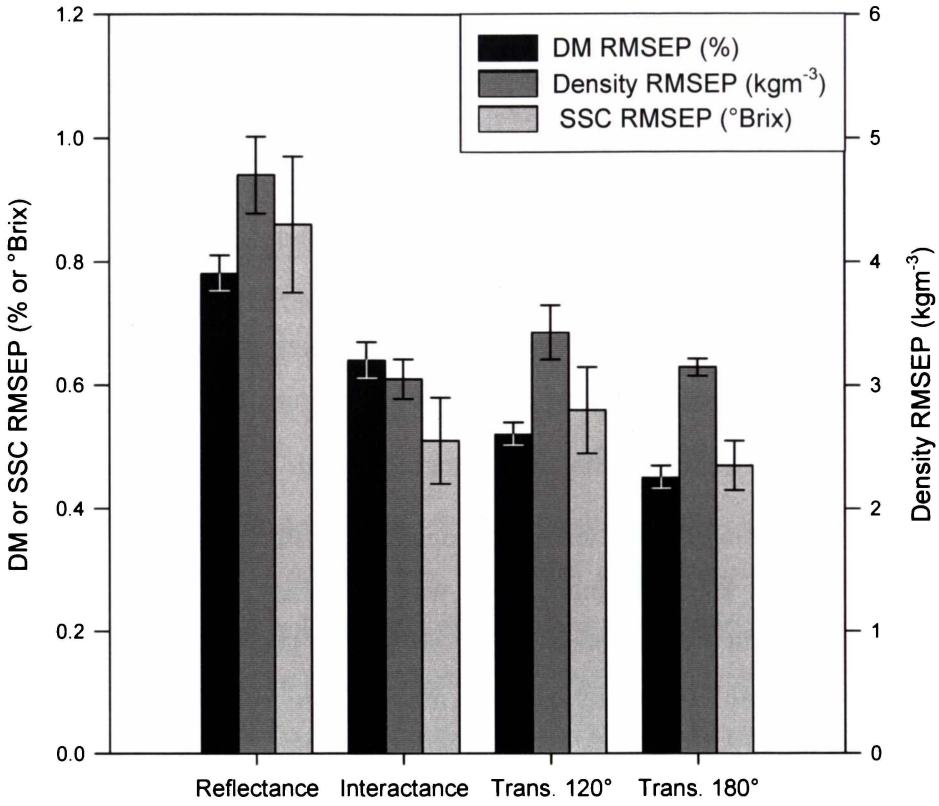

Fig. 3-20: Prediction RMSEPs of DM, SSC, and density for four different optical modes of measurement. The error bars indicate the standard deviation for 20 repeated RMSEP calculations.

Fig. 3-20 shows that for dry matter predictions, the transmission 180° mode was better than the interactance mode, which was better than the reflectance mode. Schaare and Fraser (2000) found transmission mode to perform worse than interactance for SSC and density calibrations. The improved performance for transmission mode compared to the other modes shown in Fig. 3-20 is attributed to the use of a stable light source (due to continuously running the lamp from a DC supply), better light sealing around the fruit, and increased total exposure time. These effects combine to improve signal to noise ratio, and thereby improve the predictive ability of the spectra. For SSC and density the reflectance mode performed poorer than the others. The RMSEP achieved using reflectance mode for the SSC is not as low as reported by others (e.g., McGlone and Kawano (1998), measured 0.39-0.47 °Brix). Similarly, other authors have reported even lower RMSEP using interactance mode (e.g., Osborne et al. (1999), measured 0.27 °Brix for SSC). This may be explained in some part by the small number of fruit and the narrower range of SSC's used in the calibration. Because the total integration time for the light acquisition in each mode varied the different modes should be compared with caution. The integration times selected for each mode

were chosen such that the signal was near full scale (for the extreme case of fruit and orientation selection). For reflectance 10 spectra were averaged, for transmission only 2 spectra were averaged. This means that a larger signal to noise ratio is expected for the mean reflectance signal. However, the value of the spectra itself may actually be influenced by the mode under investigation. Transmission may be getting better information from inside the fruit even though the signal to noise ratio is lower.

An attempt was made to determine the fruit mass from the spectra. Using spectra from all three size categories measured in the 180° transmission mode a prediction R_P^2 of 0.75 is obtained for mass. However, a similar degree of correlation is achieved for predicting the size category (small, medium or large) from the same spectra. This means an effect other than mass is adding to this classification such as varying instrument performance or fruit handling effects since the spectra for each size was measured on different days. Attempting a size prediction on just one of the size category (using the best pre-treatment) gave R_P^2 of 0.2 implying a very poor prediction correlation. To improve the fruit mass prediction, it would be preferable to obtain a range of fruit sizes rather that the three tight size categories used in this investigation. An attempt was also made to find a strong correlation between mass and the individual wavelengths in the second derivative spectra as Kawano et al. (1993) discovered for mandarin. Several wavelength regions of the interactance and transmission 120° spectra gave an R^2 of around 0.2, too poor to enable their utilisation to remove the size variation in the samples. Limiting the set to just the small and large fruit showed some better regions of correlation with mass (R^2 = 0.65) but no wavelengths with significant correlations to dry matter or SSC were available in these ranges.

A selection of 160 medium sized fruit from the second grower, which had a larger range of dry matters (mean 16.82%, standard deviation 1.51), was measured spectrally at different transmission angles and fruit orientations (using the same integration time at each angle). Table 3-5 shows the resulting RMSEP for dry matter predictions. To give a variability estimate for RMSEP, 20 predictions were made, each time randomly selecting one third of the samples to use as a validation set.

Angle (°)	Horizontal-1		Vertical	
	R^2	RMSEP	R^2	RMSEP
90	0.87	0.55	0.90	0.50
120	0.84	0.59	0.91	0.45
150	0.81	0.70	0.91	0.44
180	0.78	0.73	0.92	0.42
210	0.78	0.74	0.91	0.45
240	0.84	0.62	0.91	0.47
270	0.88	0.53	0.90	0.49

Table 3-5: DM prediction ability of models generated at different transmission angles for 160 medium sized fruit (~53 used for validation).

Fig. 3-21 compares the RMSEP produced at varying angles for the horizontal-1 and vertical orientations. The error bars show the 95% confidence interval for the mean RMSEP.

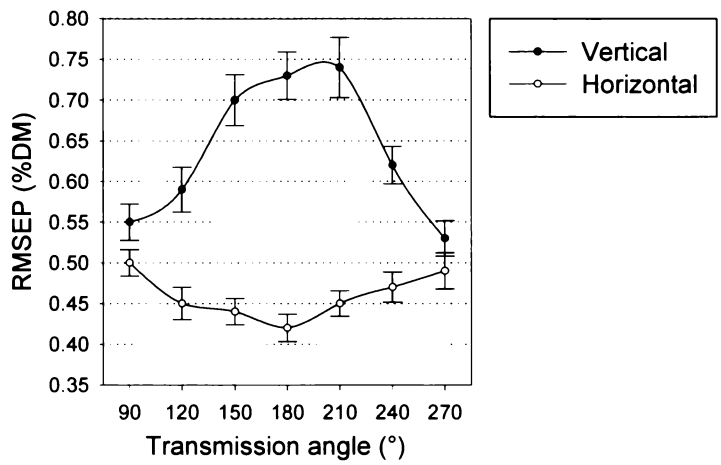

Fig. 3-21: Comparison of the DM prediction ability at different transmission angles for two orientations of medium sized kiwifruit. The error bars are the 95% confidence interval for the mean.

The variation in RMSEP estimates from the random sample selection is large in terms of the differences in performance between the modes tested (especially between the different horizontal angles). This emphasises the need to make multiple predictions or include statistical comparisons. For the horizontally oriented fruit we see in Fig. 3-21 that 180° performed better than the angles further around the fruit where there is a larger detected signal. This may be due to the extraction of more information about each fruit's dry matter since the light has traversed the entire width of the fruit. For the vertically orientated fruit better predictions are made from the sides, where the signal was larger. The lower levels

when light transmitted over the length of the fruit would have contributed to lower signal to noise ratios. From Fig. 3-21 it can be seen that the better angles for the horizontal orientation are the worst angles for the vertical orientation and vice versa. Measurements made at 90° (or 270°) produced good models with either orientation suggesting that this may be an appropriate optical geometry for randomly orientated fruit. A small subset of 19 fruit was also measured in the horizontal-2 position at three angles. The results have not been included in this comparison because (due to the low number of samples) the calibrations were poor and the variation in RMSEP at each angle was too great (e.g., 210° gave RMSEP of 0.85 ± 0.25 % for dry matter).

The orientation problem was further investigated using the same set of 160 medium sized fruit from grower 2. A calibration was performed at each angle for the horizontal-1 orientated fruit, and the resulting model was then used to predict the dry matter from spectra recorded from vertically oriented fruit (Fig. 3-22, left). The process was then reversed, predicting with the horizontal spectra using a model formed from the vertical spectra (Fig. 3-22, right). This time the error bars are showing the standard deviation for the 20 replicates and are an indication of the variation involved in estimating RMSEP.

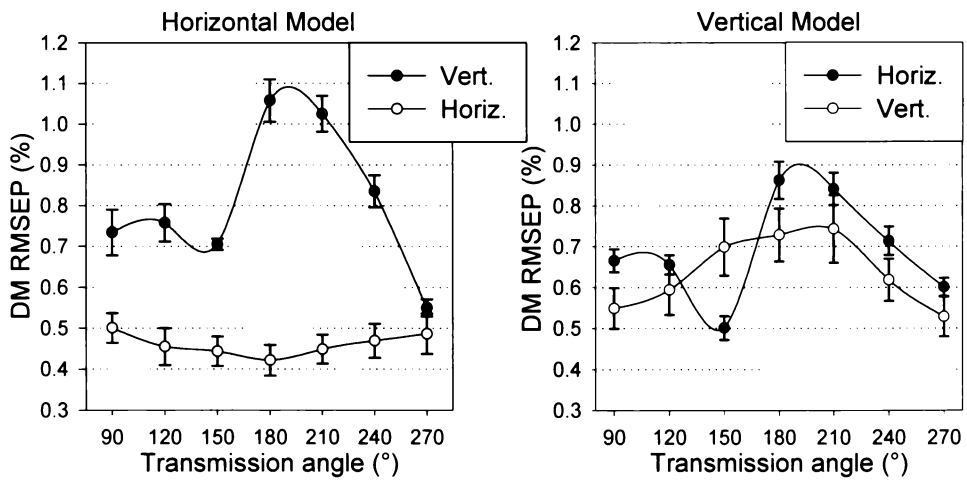

Fig. 3-22: Using a model built with one orientation to predict DM from spectra measured at a different orientation (predicting from the same orientation is also shown for comparison). The error bars show the standard deviations..

Fig. 3-22 shows that even when the models are made from spectra measured with the worst orientation mismatch, acceptable predictions can still be obtained. The vertical model generally performed better with the horizontal spectra (indicated by the lower RMSEP values) than the horizontal model did with the vertical spectra. The exceptionally good performance at 150° was surprising since a symmetrical

graph is expected. One possibility to explain the asymmetry is that the vertical model in Fig. 3-21 shows 150° to be slightly better (RMSEP = 0.7) than 210° (RMSEP = 0.74) this may account for some of the extra performance and hence the dip at 150°.

Using both horizontal and vertical spectra in the calibration, a combined model was made. This model can then be used to assess the expected prediction performance if the orientation is not known. Fig. 3-23 shows the results from using this model to predict vertically or horizontally oriented fruit as well as a mixture of these orientations.

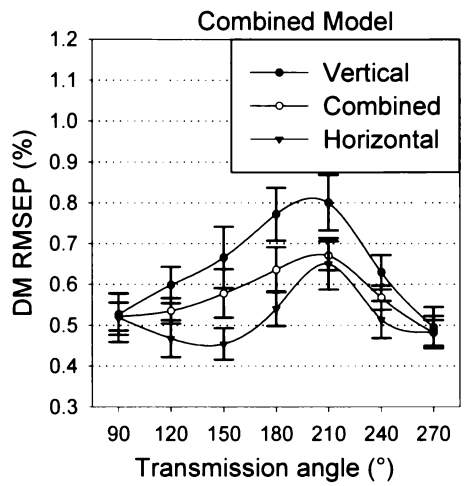

Fig. 3-23: Investigating prediction performance from a combined model on vertically and horizontally oriented fruit under transmission mode.

The combined model (Fig. 3-23) generally predicted better, from either type of spectra, than the models build solely from one type of spectra (Fig. 3-22). This is to be expected because it incorporates knowledge about both kinds of spectra, enabling better predictions. At 210° the combined model performed poorly, this was also noticed for the cross orientation application of the spectra in Fig. 3-22 at this angle. It may be that some of the spectra recorded in the vertical orientation at 210° were defective or there was physically something different about that illumination geometry resulting in a poor model and predictions. The choice of the best transmission angle to use will depend on the orientations presented by a practical grader. If the grader predominantly orients the fruit in a horizontal position then (from Fig. 3-21) 150° may work best since it has the added advantage of not looking directly at the source when a fruit is not present (as would 180°). If the grader presents a mixture of vertical and horizontal oriented fruit then (from Fig. 3-23) best results may be obtained at the smaller transmission angle 90° (or 270°) which gives good performance for both orientations.

Because the transmission spectra at different angles are generally similar in shape it may be possible to apply spectra recorded at one transmission angle to models generated at another angle. Using the medium sized fruit in the horizontal-1 orientation, models were generated at each angle from 90° through to 270°. Each calibration was then tested using spectra from each of the angles (in the special case when a model is to be tested with spectra from the same angle an independent set was split off). The testing was repeated 20 times using random splits of the samples to give a mean RMSEP and standard deviation estimate (Fig. 3-24, left). When a model is used on spectra that are different than those used in the calibration, there is often a bias (a constant offset for the prediction). The bias results from the prediction tests with spectra from different angles are shown in Fig. 3-24, right.

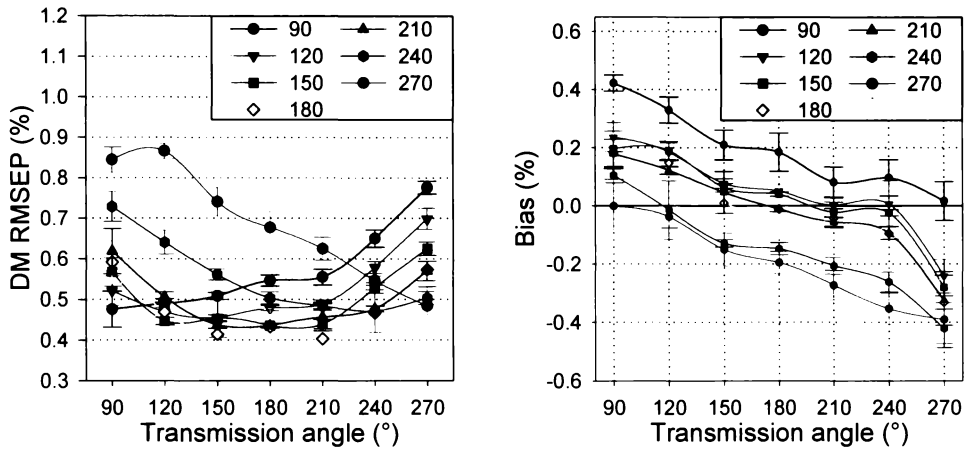

Fig. 3-24: Predictive ability of models formed at one angle (labelled in the legend) on spectra measured at different angles (horizontal axis). Left: RMSEP for the prediction. Right: The bias (constant offset) measured for the predictions.

The best RMSEP is found at the angle where calibration was carried out and increases with increasing difference in angle. There is a similar effect for the bias. Essentially zero bias is observed at the angle of calibration. The absolute bias again increases with increasing difference. It is surprising that the model built at 90° is poor at predicting spectra measured at 270°, which is a very similar optical geometry, and similarly 120° to 240°. This suggests that the spectra measured must be slightly different at the two different sides of the fruit. This is of concern because the angles 90°-180° are symmetric to the angles 270°-180° and should therefore be equivalent. One possibility is that the illumination/selection arrangement was somehow biased to one side. However, there must be significant

similarities in the spectra because the models were still able to generate usable predictions for spectra measured from different transmission angles.

3.4 Mandarin trial

3.4.1 Experimental procedure

In order to get a large variation in acid and SSC, Satsuma mandarins were harvested specifically for this trial over a long period and from different locations. Approximately 90 mandarins each week were picked over a period of 8 weeks commencing 13th April 2000. Each set of 90 mandarins was taken from three different orchards (Davies, Kapiro, and Kwan) with similar management style in the Kerikeri region of New Zealand. At each orchard the same 5 adjacent trees were uniformly sampled throughout the experiment. Immediately after each picking, the fruit were transported overnight to TDG, unboxed and were held in an air-conditioned room for 2 hours. The spectroscopic measurements of the fruit were made in the morning and the destructive chemical measurement in the afternoon. Each fruit was first weighed and then completely juiced, and the juice was assessed for SSC and acid level.

Fig. 3-25 shows the optical arrangements used to measure the spectra of mandarins. The interactance system is that used for the kiwifruit trial in section 3.3. The fruit were placed on top of a 12 mm rubber seal and illuminated from below. Light passing through the fruit and emerging within the light seal then reflects onto the entrance aperture of a Y-branch fibre bundle via a 45° mirror. The Y-branch bundle was used to convey the light to a silicon spectrometer (as used in section 3.3) as well as an InGaAs spectrometer which has greater sensitivity at the longer wavelengths.

For transmission and reflectance modes the mandarins rested on a light seal, which had an aperture 25 mm in diameter. For the transmission (which needs a powerful light source) care was taken not to overexpose the fruit to the intense light (always < 5 s).

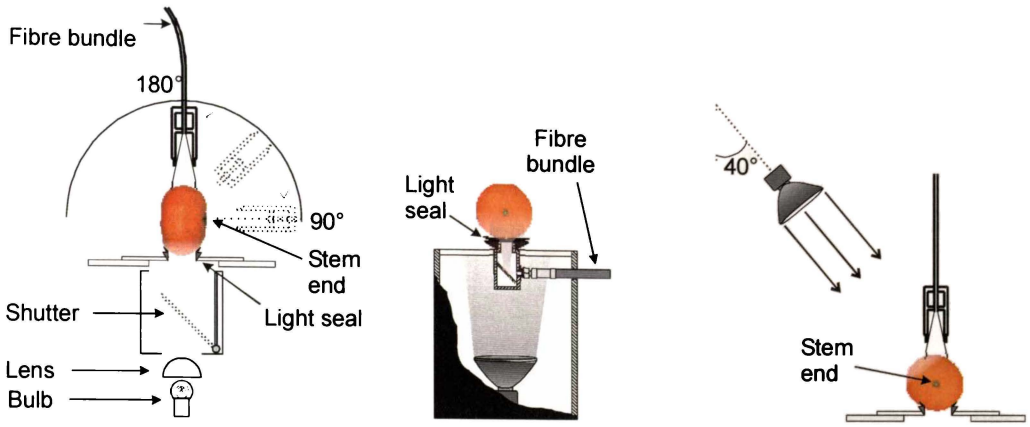

Fig. 3-25: Optical arrangements used for measuring spectra from mandarins. Left: Transmission mode (up to 10 angles from 90° to 180° were used). Centre: interactance mode. Right: reflectance mode.

An investigation of heat tolerance for mandarins showed that they are more susceptible to heat damage than kiwifruit. Furthermore, the damage is not immediately visible; fruit illuminated for 10 s showed no visible damage until several hours after exposure, at which point slight skin decolourisation was apparent, becoming worse after 2 days. Dissection of the burnt area showed (even in the extreme case of 60 s exposure) that the visible damage extended less than 1 mm into the fruit skin.

The distance between the tip of the fibre and the mandarin varied due to the variation in fruit size. The field of view when the distance was greatest (i.e., for the smallest fruit ~28 mm in diameter) led to a maximum collection region approximately 20 mm in diameter. This reduced to approximately 15 mm for large fruit (~55 mm diameter)

A reference and a dark current spectrum were recorded prior to measurement of each group of 15 fruit, (30 for interactance mode) in each optical configuration. The reference used for the transmission and reflectance modes was a 60 mm polytetrafluroethylene (PTFE) sphere located in the place of the fruit (as with the kiwifruit). For the interactance system a flat white PTFE surface (sealed from all light except the lamp) was placed above the collector window. For each configuration the dark current measurements were recorded by blocking the aperture to the optical fibre.

For both transmission (all angles) and reflectance modes an integration time of 200 ms was used and an average of five spectra (to improve the signal to noise ratio) was stored for subsequent analysis. For the interactance mode longer

integration times were necessary due to the signal lost in the Y-branch fibre optic cable. For the silicon spectrometer 2000 ms was necessary and for the InGaAs spectrometer 5000 ms. Towards the end of the harvest period it was necessary to reduce the integration time for the silicon spectrometer in the interactance system to 1600 ms to avoid detector saturation.

As with the previous section (3.3) a search of different pre-treatments and wavelength ranges was carried out for each of the modes. Having established a suitable pre-treatment, the predictive ability of different wavelength ranges was compared.

3.4.2 Results and discussion

Fig. 3-26 shows that as mandarins mature on the tree the SSC increases and the acid level reduces. Fruit from the Kwan orchard had significantly lower SSC and higher acid content. Sampling began before the orchards harvested their crops and continued after the rest of each orchard had been picked (which occurred approximately half way through the trial).

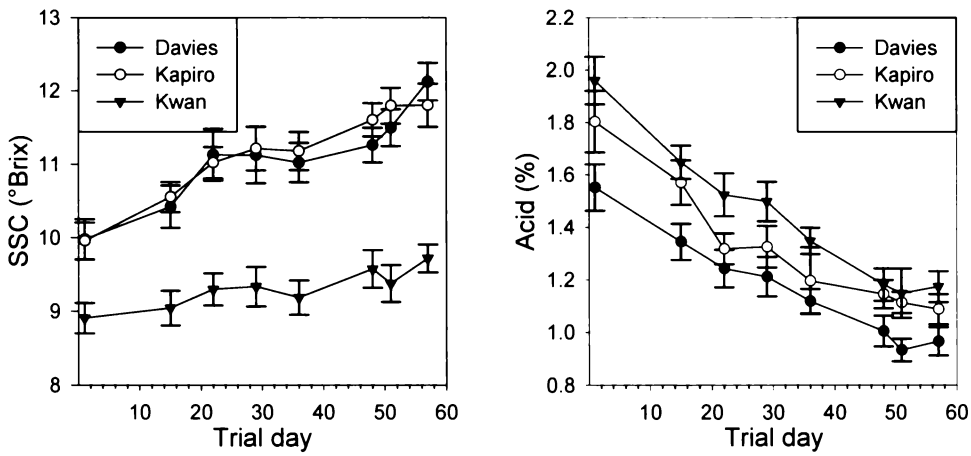

Fig. 3-26: Changes in the mean SSC (left) and acid (right) content in mandarins sampled from three different orchards over an eight week period. The error bars indicate the 95% confidence interval for the mean.

Fig. 3-26 (right) shows how the mean acid levels in the fruit plateau towards the end of the sampling period; extending the trial further would not have given a significant increase in the number of low acid fruit. Fig. 3-27 shows the distribution of SSC (left) and acid (right) for the entire set of fruit (all orchards and harvest dates). The acid distribution shows a long tail on the high acid side indicating some extremes in acid level in some fruit.

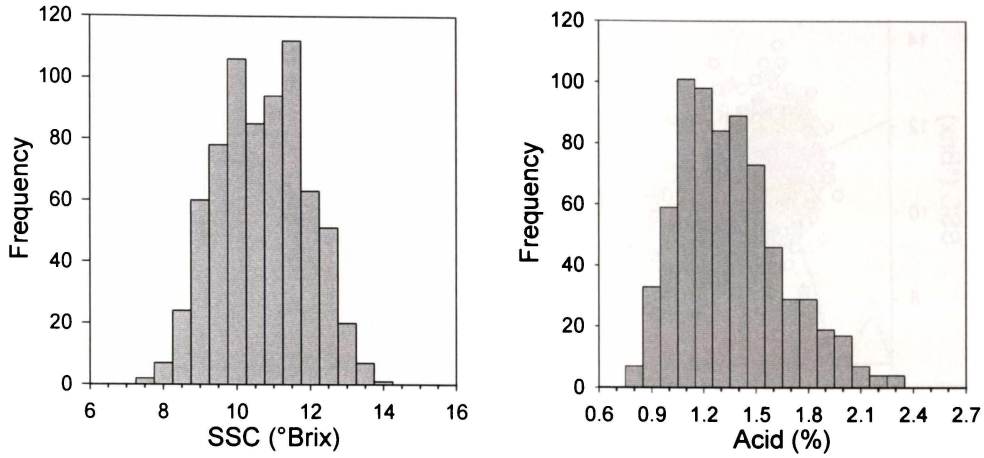

Fig. 3-27: The SSC (left) and acid (right) distribution for all three orchards over all the sample dates.

The mean ± standard deviation for the SSC was 10.4 ± 1.2 °Brix and for the acid measurements 1.3 ± 0.3 %. This large range in SSC and acid is a result of the long sampling period (8 weeks) and sampling from different orchards.

The mass distribution for the entire set (Fig. 3-28) shows a large spread in fruit size (69 ± 16 g) However, there was only a slight increase in mean fruit size over the sampling period (Fig. 3-29).

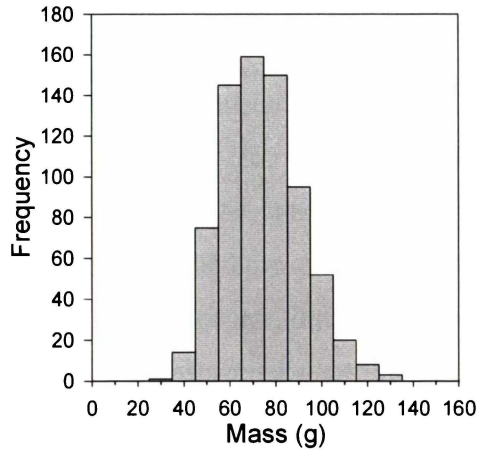

Fig. 3-28: Mass distribution for the entire set of mandarins.

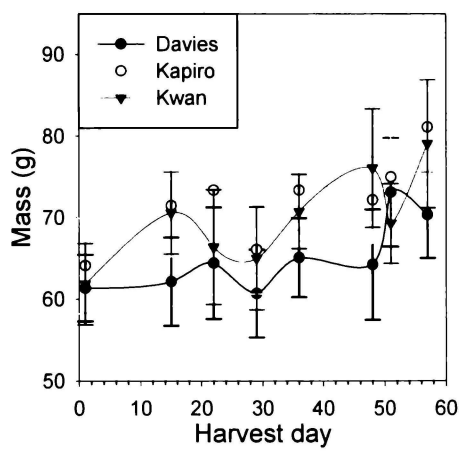

Fig. 3-29: Mean mass for each orchard over the sampling period, including measurement standard deviations (error bars).

The dip in the mean fruit mass on harvest day 29 is a result of using a different person to pick the fruit on that day.

For this fruit set there is a very weak ($R^2 = 0.22$) negative correlation between SSC and acid in mandarins. This is shown in Fig. 3-30. This relation is expected since the SSC increased, and the acid level decreased as the trial progressed.

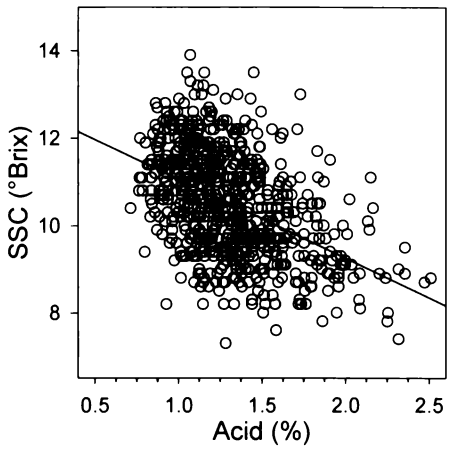

Fig. 3-30: SSC vs. acid for the entire set of mandarins.

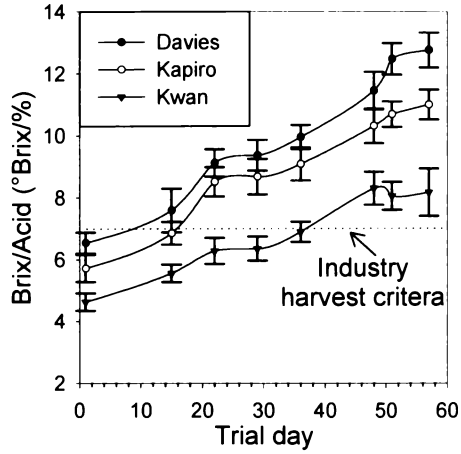

Fig. 3-31: Changes in SSC/Acid ratio as mandarins mature. In NZ when the mean ratio reaches 7 the crop may be harvested

Fig. 3-31 displays the change in SSC/Acid ratio in the mandarins throughout the sampling period. The SSC/Acid ratio is one measure for maturity. In NZ when the mean ratio reaches 7 the fruit may be harvested. The fruit from Kwan were generally less mature by this criteria than the fruit from the other two orchards.

Fig. 3-32 shows typical reflectance, interactance, 90° transmission and 180° transmission spectra shapes for mandarin, relative to their respective reference measurements. The solid line indicates the mean spectrum recorded for all of the fruit samples in each mode of operation, the dotted plot one standard deviation each side of the mean spectrum.

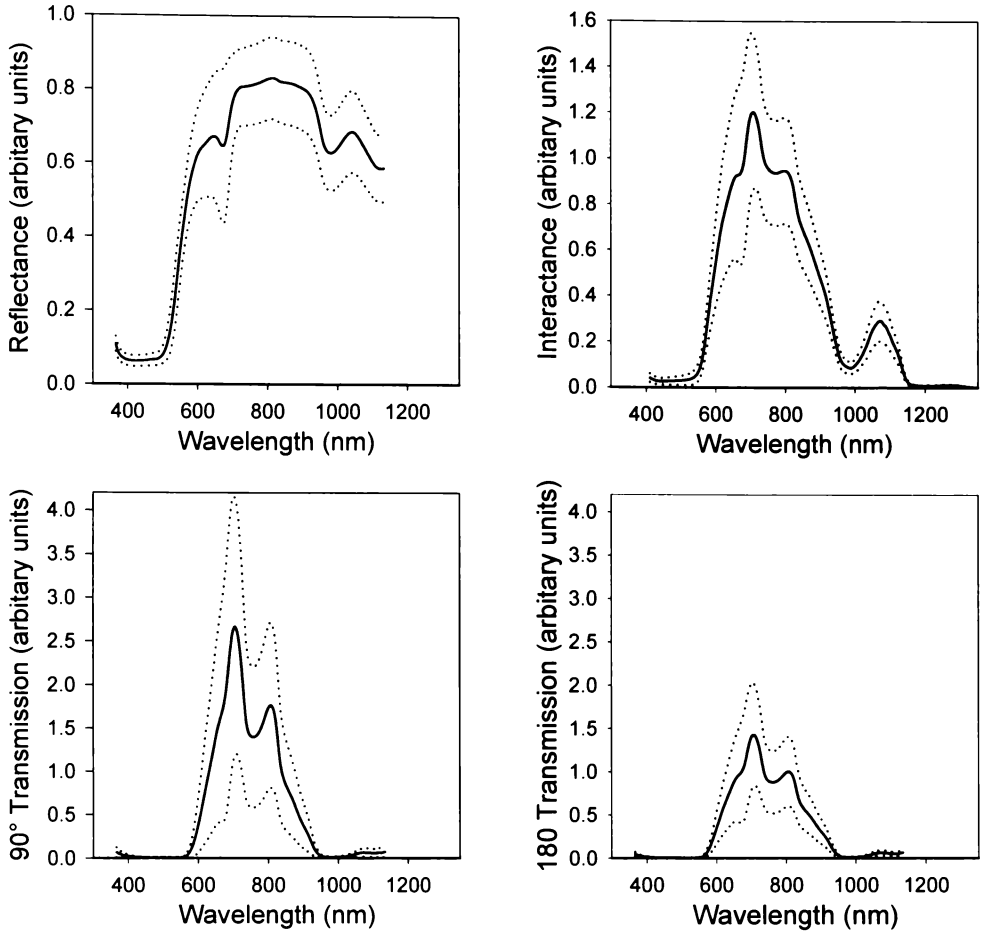

Fig. 3-32: Typical relative spectra (mean ± one standard deviation) measured for mandarin in different modes.

The spectra are generally similar to the kiwifruit spectra shown in Fig. 3-19. The reflectance curve for mandarin shows a larger signal around 600-650 nm corresponding to the orange colour of the peel. The dip in the relative reflectance spectra around 660 nm arises from chlorophyll absorption. Water absorption is seen in all the spectra, dips at the expected 760 nm and 834 nm positions and a much larger absorption near 958 nm. The relative spectra for 180° transmission are small compared to the 90° transmission signal. However, the raw spectra were of similar magnitude for all angles, the reduction of the relative spectra at the larger transmission angles is because the reference spectra systematically reduced as the angle increased from 90° to 180°. The relative interactance spectra showed a very low signal in the InGaAs extended region above 1050 nm. This suggests that there is no advantage in using the second spectrometer for interactance measurements on mandarins.

In this work particular interest is taken in the ability of NIR spectroscopy to assess the mandarin's acid level. Detection of constituents of low concentrations (~2% or

less by mass) is difficult because the absorption bands of lesser components are often masked by those of the major components, particularly water. Acids, while only a minor constituent, play a major role in consumer perception of fruit taste. It is therefore highly desirable to be able to assess them with NIR spectroscopy. Intuitively, transmission has the best chance for assessing acid content in mandarins. Since acid is a very small constituent of the fruit, the NIR light interaction should be maximised. Because acid is in the internal flesh tissue, using transmission mode maximises the acid interaction. Fig. 3-33 shows the correlation of the second derivative intensity at individual wavelengths for both transmission and reflectance spectra against acid.

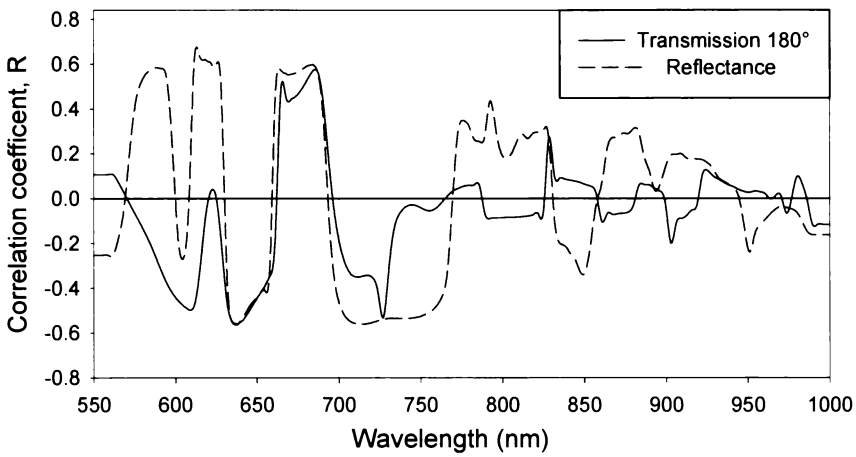

Fig. 3-33: Correlation of acid level with the intensity at each wavelength in the second derivative of both reflectance and 180° transmission spectra.

For the transmission spectra the strongest correlations at single wavelengths are below 730 nm with only weak correlations at the higher wavelengths. This suggests there is information relating to the acid content of mandarins at these lower wavelengths. The reflectance signal also shows wavelength regions with significant correlations. The reflectance measurements may not have penetrated the fruit flesh at all of the wavelengths that are strongly correlated with acid, (see chapter 4) so the spectra could be measuring other surface properties, which are correlated to acid such as colour or other maturity information.

Multivariate calibrations and predictions were performed on spectra from each mode of spectral measurement. Approximately 150 pre-treatment combinations were tested for their effect on predictive performance. The sequences tested included those that were successful with the kiwifruit trial described in section 3.3. The best pre-treatment sequences were then tested with different wavelength

ranges for each type of spectra. The parameters that led to the lowest RMSEP values are shown in Table 3-6.

Predicted property	Mode	# Valid. Samples	Pre-treatment sequence	Wavelength range (nm)	# Factors
SSC	Reflect.	235	SG(9,2,2), AS	600-1000	8
	Interact.	205	SG(9,2,2), AS	800-1000	7
	Trans. 90°	235	SG(9,2,2), AS	800-1000	3
	Trans. 180°	235	SG(9,2,0), AS, SNV	800-1000	8
Acid	Reflect.	233	SG(9,2,2), AS	630-695	5
	Interact.	204	SG(9,2,2), AS	600-1000	8
	Trans. 90°	233	SG(9,2,0), AS, SNV	630-695	6
	Trans. 180°	233	SG(9,2,0), AS, SNV	630-695	5

Table 3-6: Pre-treatments, wavelength range and model complexity used different modes. SG=Savitzky-Golay (width, order, derivative), AS = Area Scaling, Valid. =Validation.

The reflectance (Reflect.) and interactance (Interact.) modes performed well with second derivative Savitzky-Golay smoothing followed by area scaling (AS). The transmission modes (Trans. 90° and Trans. 180°) performed well with and without the second derivative, and preferred area scaling and SNV applied.

The acid prediction models were simpler (less factors) and performed strongly using only the 630-695 nm wavelength range, an area in which both fruit colour and chlorophyll would be expected to aid the process. To show the wavelength selection process in more detail Fig. 3-34 compares the acid prediction using the different wavelength regions of the transmission 180° spectra.

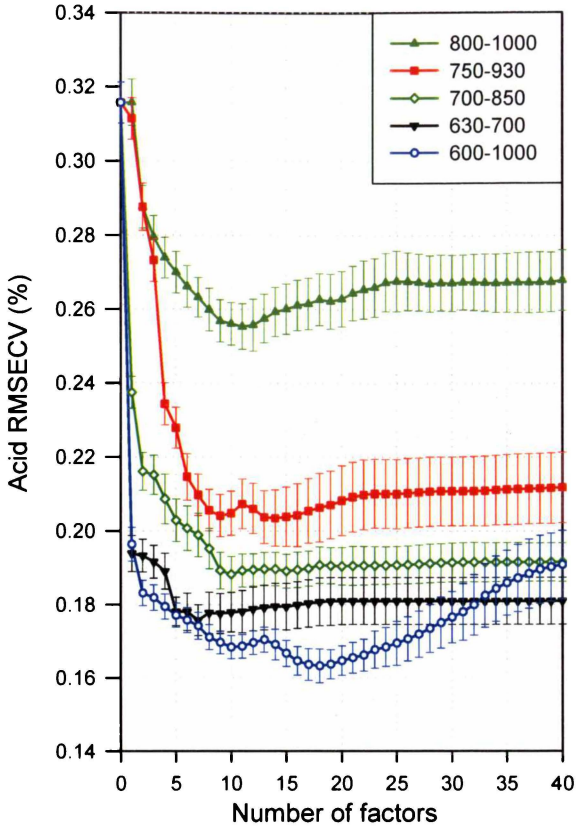

Fig. 3-34: Cross validation prediction errors (RMSECV) for acid level in mandarins against number of factors, for calibrations that use different wavelength ranges.

The RMSECV at 0 factors is defined as the standard deviation of the response values in the validation set. The RMSECV reduces with increasing PLS factors, until the data begins to be over fitted, and sometime after this point RMSECV increases again. The error bars indicate the standard deviation for the RMSECV from 20 repeated calculations with randomised splitting of the samples between the validation and calibration sets. These standard deviations can be seen to increase as the RMSEP increases due to the over fitting, this is particularly clear on the 600-1000 nm range.

From Fig. 3-34 the wavelength range 600-1000 nm performs very well for acid prediction, with just one factor the RMSEP is reduced to 0.195 % which corresponds to a R_P^2 of approximately 0.62. With two factors R_P^2 improves to 0.67, beyond that there is a steady improvement until 10 factors are used. The standard deviation, which indicates the variation from the replicates at each RMSECV value, increases as the data begins to be over fitted. The much smaller wavelength range 630-700 nm performs almost as well as 600-1000 nm. Therefore the wider range is not offering a significant amount of additional information. The primary information may well be coming from the 630-700 nm

wavelengths. In order to minimise the complexity of the model, 5 factors using the 630-700 nm range may well be the best choice since at this point it performs as well as the larger wavelength range.

The wavelength range 800-1000 nm (which performed well for the SSC measurements) was unable to produce good acid predictions. Miyamoto et al. (1998) found that the range 730-930 nm was the most suitable for acid predictions. However, the range 750-930 nm in Fig. 3-34 did not perform particularly well. One explanation for this discrepancy may be that the fruit used by Miyamoto et al. (1998) may have had a high correlation between SSC and acid. So their model may be indirectly measuring acid through its association with SSC. Because the correlation between SSC and acid for the fruit in this experiment is very poor ($R^2 = 0.22$) we get a poor prediction for acid using this wavelength range. The model reported in this thesis was unable to utilise the acid bands described by Miyamoto et al. (1998). It seems that assessment of the mandarin acid level can be done as well with the chlorophyll region of the spectra (and inferring the acid content indirectly) as can be done by including the longer wavelengths. Grading mandarins by chlorophyll content is not a new idea, Chuma et al., 1974 and 1976 describe the automatic sorting of mandarins for maturity and colour using wavelengths from 550-700 nm.

Interactance mode was the exception for acid prediction, performing well using the broad wavelength range (including the chlorophyll range) with only 8 factors in the model. Fig. 3-35 shows the prediction errors from the parameter settings in Table 3-6.

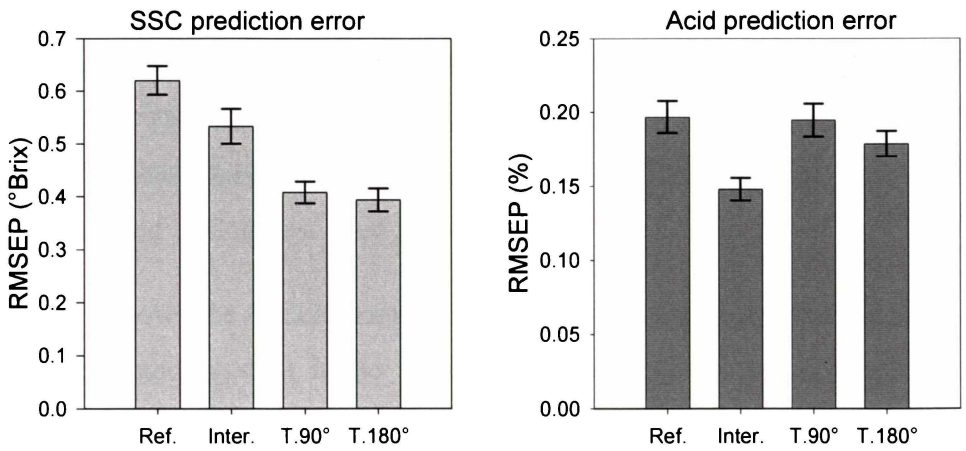

Fig. 3-35: Comparing the predictive performance of different modes for acid and soluble solids contents in mandarins.

The error bars indicate the standard deviation from 20 estimates of RMSEP using random data splits between calibration and independent validation sets.

From Fig. 3-35 (left) the transmission modes (T. 90° and T. 180°) were able to predict the SSC of mandarins better than the interactance mode (Inter.), which in turn was better than reflectance mode (Ref.). This is consistent with our result for kiwifruit (Fig. 3-32). The lowest prediction error for SSC was 0.39 ± 0.02 °Brix ($R_P^2 = 0.89$) using transmission at 180°. This is higher than the value 0.32 °Brix reported for a smaller sample by Kawano et al. (1993), but is less than 0.5 °Brix reported by Miyamoto and Yoshinobu (1995) for intact mandarins.

For the acid predictions, (Fig. 3-35, right) with the exception of the interactance mode, there was no improvement for transmission mode compared to reflectance mode. This suggests that the predictions for acid may only be based on near surface properties (measured by reflectance) rather than internal properties (which are better assessed by transmission). This is further evidence that acid itself is not directly contributing to the acid prediction. The lowest acid prediction error was 0.148 ± 0.007 % ($R_P^2 = 0.78$) using the interactance mode, which is comparable to 0.147 % measured by Miyamoto (1998) using transmission spectra. From the transmission 180° spectra, the lowest acid RMSEP was 0.177 ± 0.01 % ($R_P^2 = 0.69$) using only the 630-700 nm wavelength range. Considering that the standard deviation for the acids is 0.3 %, most of the modes give predictions that are only good enough for removal of extremes of acids from a line of fruit on a grader rather than categorising into levels.

Additional transmission spectra were recorded every 10° from 90° to 180° on a subset of 150 fruit from the Davies orchard. Using the same pre-treatments as used for the transmission spectra in Table 3-6, calibration/prediction tests were performed to compare the predictive ability at the different transmission angles (Fig. 3-36).

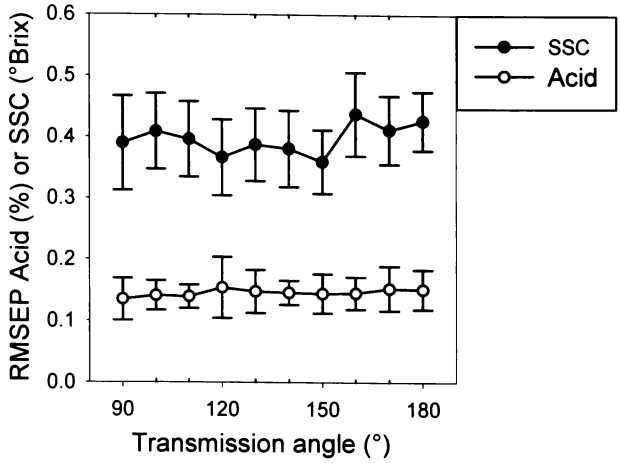

Fig. 3-36: Prediction performance of acid and SSC levels from mandarin spectra measured at different transmission angles. The error bars indicate the standard deviation for RMSEP from 20 random calibration/prediction set splits.

Fig. 3-36 shows that there is no preferred orientation for mandarin assessment. These fruit were hand placed, and the spectra recorded at angles varying from 180° (directly looking at the stem end of the fruit), around to 90° (looking at the equator of the fruit). If these extremes in orientation can all produce good predictions then it may be possible to grade randomly oriented fruit with little loss in performance.

3.5 Motion trial

3.5.1 Experimental procedure

Apples were obtained from a local packhouse (Sunfruit Orchards Ltd., New Zealand) on the 28th March 2001. Three different grower lines were screened for bruised or scarred fruit leaving 120 fruit for analysis. The fruit were sorted for size at the packhouse and were within the range 144-164 g. Upon arrival at TDG the fruit were stored overnight in an air-conditioned room to stabilise their temperature. The fruit were assessed spectrally in the morning, and the destructive chemical measurements were made in the afternoon. Each fruit was assessed for dry matter and SSC as described in section 3.1.

Fig. 3-37 shows the optical arrangements used to measure the spectra of apples. The interactance system is the same as the one used for the kiwifruit trial in section 3.3.

For the reflectance and transmission arrangements, the apples rested on light seals, which consisted of 2 mm thick rubber 'washers' with a 48 mm internal diameter.

Twenty light seals were evenly located around the perimeter of a large disk (see Fig. 3-5 at beginning of chapter), the centres of the fruit being located on a circle 600 mm in diameter. The disk was rotated at a steady speed, allowing the fruit to be moved under the detector for the rapid acquisition of spectra.

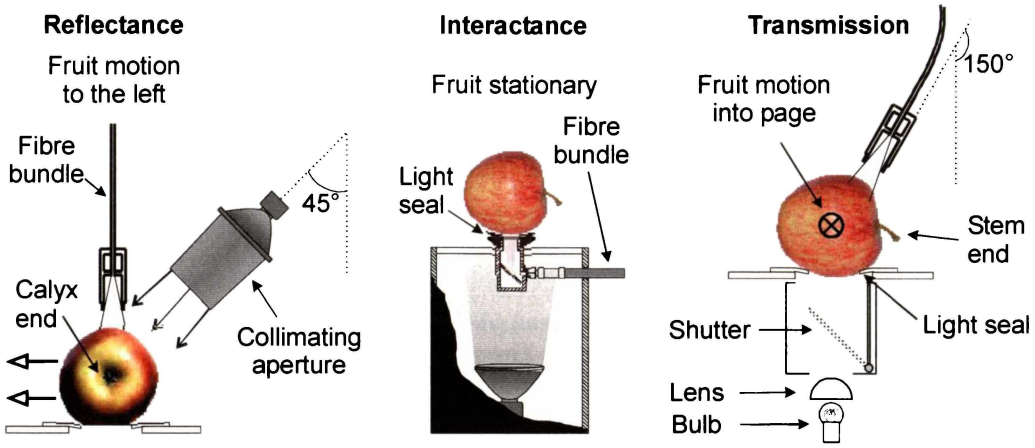

Fig. 3-37: Modes of spectral acquisition for apples. Left: Reflectance, apple moving left. Centre: Interactance mode, stationary apple. Right: Transmission, apple moving into the page.

For reflectance mode, 10 fruit were placed in alternate positions around the disk so that there was a large gap between fruit to avoid gathering light that may have reflected from another fruit. Reference measurements were taken every 20 fruit using a flat 99% reflectance tile (SRS-99-020, Labsphere, North Sutton, NH, USA) as illustrated in Fig. 3-38, left.

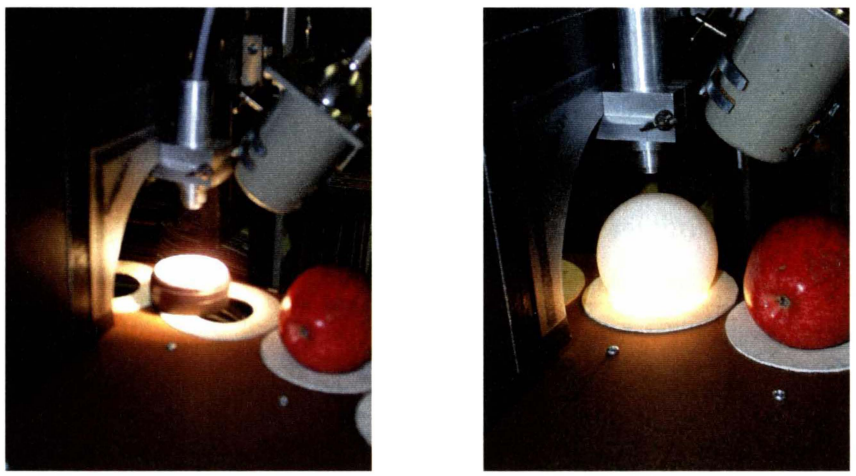

Fig. 3-38: Reference measurements. Left: Reflectance measurement (illumination from top right). Right: Transmission measurement (illumination from below).

For reflectance mode, the integration time was set as high as possible while ensuring the detector was not saturated by any of the fruit in the set. The number of useful spectra acquired from each fruit is inversely proportional to the disk's speed. For reflectance mode the integration time was fixed at 10 ms and two constant speeds were used: approximately 3 fruit/s and 1 fruit/s. The absolute

speeds are not important because they are scalable with a brighter light source or more sensitive detector. Less than ⅓ of the spectra are actually useful since the detector is often looking at the space between apples and some of the spectra must be discarded during processing to ensure that the detected light belongs entirely to the apple being investigated. The 3 fruit/s and 1 fruit/s gave 15 and 61 useful spectra per fruit respectively so these sets of spectra are named (R.15) and (R. 61).

For transmission mode, 19 fruit and the 80 mm diameter PTFE reference sphere (Fig. 3-38, right) were loaded onto the disk. An integration time of 1000 ms ensured there was no saturation of the detector, and the disk was spun slowly (~10 seconds between fruit) giving 7 spectra/fruit (only 5 useful spectra). The fruit were then measured again with a 500 ms integration time, reducing the signal size but giving 11 useful spectra/fruit. These sets of spectra are called (T. 7) and (T. 11) respectively.

For interactance mode, 5 sequential spectra, each with an integration time of 150 ms were averaged to form a single spectrum with an improved signal to noise ratio. The reference used for the interactance system was a flat PTFE surface (sealed from all light except the lamp) placed above the detector window.

For each of the modes a dark current was also recorded each time a reference measurement was made. Subtracting the dark current and dividing each 'raw' spectrum by a reference spectrum produced relative spectra for the chemometric analysis.

As with the previous two sections (3.3 and 3.4) a search of different pre-treatments and wavelength ranges was carried out for each of the modes and speeds. Having established a suitable pre-treatment, a comparison was made between the spectra taken at different positions over the fruit. A further investigation was then made to consider the predictive performance gained by combining the spectra taken at the best performing positions.

3.5.2 Results and discussion

Fig. 3-44 shows the distribution of SSC (left) and DM (right) for the 120, similar sized, Royal Gala apples used in this trial.

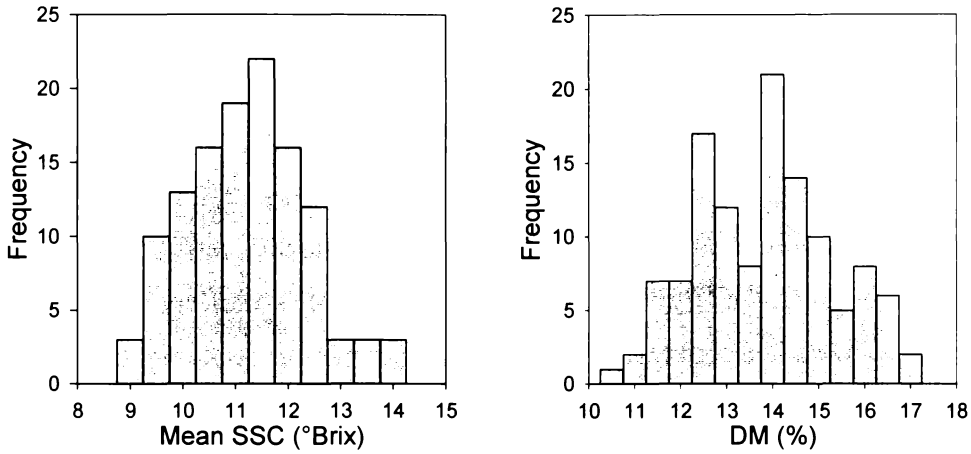

Fig. 3-39: The SSC and DM distributions for 120 Royal Gala apples.

The mean and standard deviation of the SSC was 11.0 ± 1.1 °Brix and for DM 13.6 ± 1.5 %. The SSC measurement used is the average of two samples on opposite shoulders of the apple (Fig. 3-2). These two measurements are correlated to each other with $R^2 = 0.50$ (see Fig. 3-40).

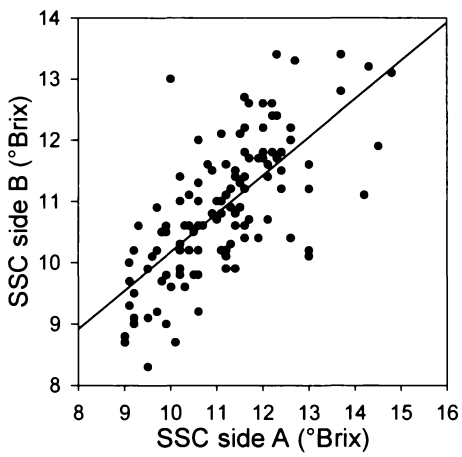

Fig. 3-40: Comparing the SSC on opposite sides of an apple.

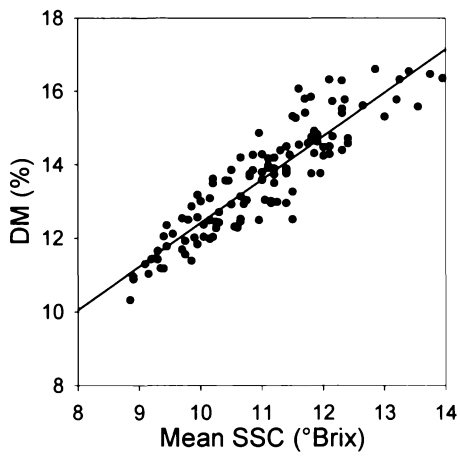

Fig. 3-41: Correlation between DM and mean SSC for apples.

Fig. 3-41 illustrates the strong relationship between DM and mean SSC, with $R^2 = 0.81$. Note that the SSC measurements from the kiwifruit end caps from our earlier trial (Fig. 3-15, right) show a much stronger correlation than the apples here.

Fig. 3-42 (a, c, and d) show typical reflectance, interactance, and transmission spectra shapes relative to their respective reference measurements for the apples in this trial. Fig. 3-42 (b) shows a sequence of reflectance spectra as an apple moves past the detector.

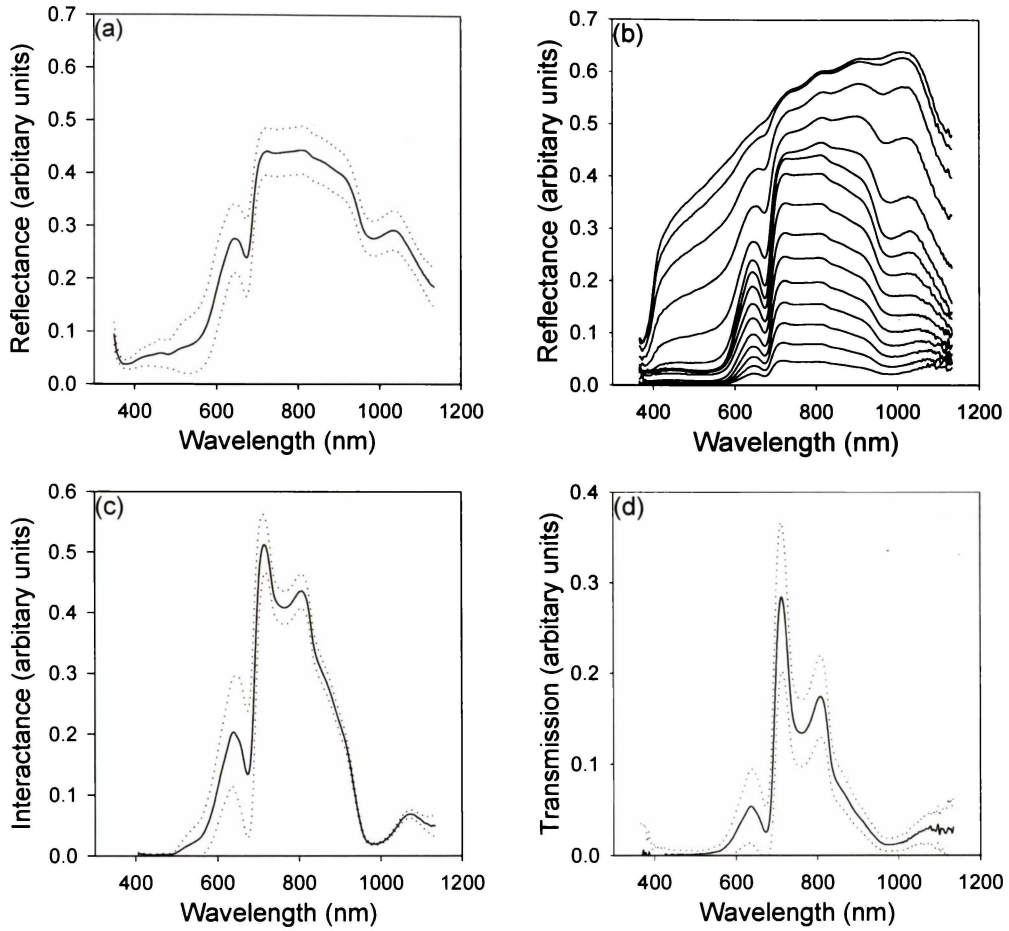

Fig. 3-42: Typical relative spectra measured for apples in different modes. For (a), (c), and (d) the solid line is the mean spectrum recorded, and the dotted line indicates 1 standard deviation from the mean. (b) shows reflectance spectra corresponding to different measurement positions as a fruit moves under the detector.

The spectra are generally similar to those of kiwifruit and mandarin (Fig. 3-19 and Fig. 3-32 respectively). The dip around 660 nm is attributed to chlorophyll absorption. Water absorption in seen in all the spectra with dips near 760 nm and 834 nm and a much larger absorption close to 958 nm.

Fig. 3-42 (b) shows the relative spectra from 15 sequential reflectance spectra as a single fruit moves under the detector. Initially the relative reflectance signal is very small, gradually increasing in magnitude; the maximal relative reflectance signal is very flat with strong similarity to the reference signal. At this maximal point the collected, reflected light has a very large specular reflection content (as opposed to diffuse reflectance alone) and is not expected to produce useful

calibrations for SSC or DM. The relative reflectance and transmission spectra shown in Fig. 3-42 (a,c,d) correspond to a spectrum recorded at a position central in the range of useful spectra recorded for each fruit.

Multivariate calibrations and predictions were performed on the relative spectra from each mode of operation. Approximately 150 pre-treatment combinations were tested for their effect on prediction performance, including those that were successful with the earlier kiwifruit and mandarin trials. Table 3-7 shows the pre-treatments, which produced low RMSEP for both SSC and DM response values for each of the modes.

Mode	Pre-treatment sequence	Wavelength range (nm)	Number of factors
Reflectance	SG(9,2,0), SA, SNV	800-1000	8
Interactance	SG(9,2,0), SA, MSC	600-1000	12
Transmission	SG(9,2,0), SA, SNV	600-1000	12

Table 3-7: Model parameters used for the apple calibrations for each of the modes investigated.

For reflectance and transmission, the raw (dark current corrected) spectra performed just as well as the relative spectra, for the interactance mode the relative spectra were better. When the wavelength range was narrowed, simpler models (less factors) could be formed, but they generally had a larger RMSEP. All modes did slightly better without the second derivative pre-treatment which was surprising considering it generally gave improved RMSEPs for both kiwifruit and mandarins.

Using the pre-treatments and wavelength range shown in Table 3-7, the interactance spectra gave an RMSEP of 0.47 ± 0.06 % for DM and 0.39 ± 0.04 °Brix for SSC. This compares well with stationary fruit measurements by others. Lammertyn et al. (1998) used diffuse reflectance to find an SEP of 0.61 °Brix and Moons et al. (1997) reported an SECV of 0.33 °Brix.

Calibration and prediction tests were performed using the spectra recorded at each of the positions on the moving fruit to allow the fruit positions yielding the most information to be identified. Fig. 3-43 shows the peak intensity of the transmission spectra as a typical fruit passes under the detector, as well as the corresponding prediction results from the full set of spectral data at each position

(enlarged in horizontal scale). The situation can be visualised as though the fruit are stationary, and the source and detector move past the fruit measuring the intensity trace shown in the figure. The position index (0-100) is defined to map the useful spectrum region. For transmission, this is the sequence of spectra commencing just beyond the peak intensity recorded between fruit, and just before the next similar peak.

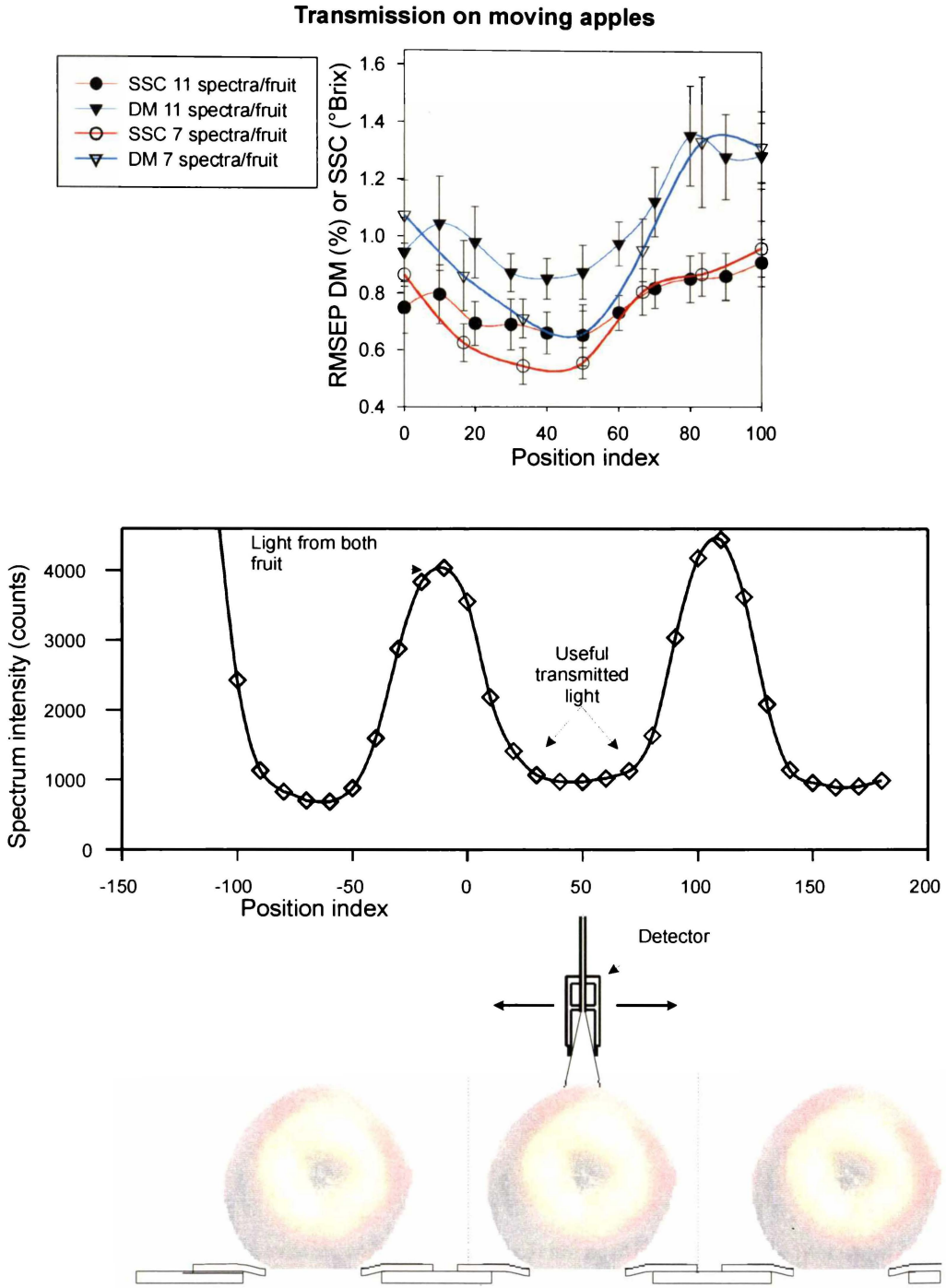

Fig. 3-43: Visualising the fruit as stationary, the detector and illumination source trace out a transmission intensity corresponding to fruit position. The upper plot shows the predictive ability at the best fruit positions using spectra recorded for the complete set of moving apples.

As the fruit, illuminated from below, pass under the detector the signal significantly reduces. In spite of the resulting reduction of signal to noise ratio from the lower spectral amplitudes the predictions actually improve (as shown by the reduced RMSEP in the upper graph of Fig. 3-43). This is because the spectra, though reduced in magnitude, contain more information relevant to the predictions for SSC and DM. Spectra recorded between fruit (peaks in Fig. 3-43) predict the fruit's parameters poorly because the detected light (reflecting off the disk surface) contains information relating to both fruit.

Fig. 3-44 shows the peak intensity of the reflectance spectra as a typical fruit passes under the detector, and the corresponding prediction results from the full set of spectral data at each of the locations indicated (position 0-100).

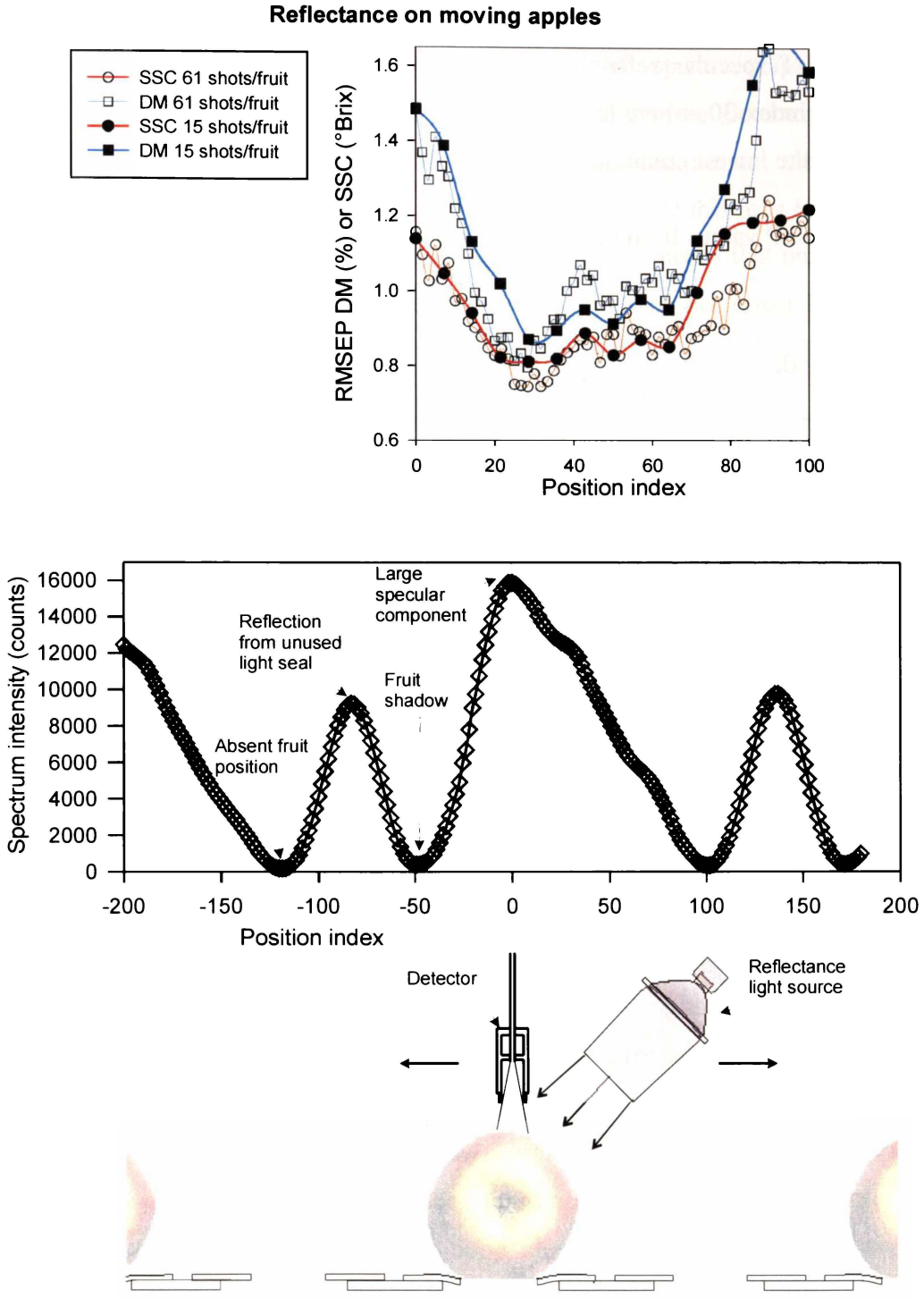

Fig. 3-44: Reflectance intensities corresponding to fruit position. The upper plot shows the predictive ability at the best fruit positions using spectra recorded for the complete set of moving apples.

Fig. 3-44 shows that as the light from the fruit (initially via reflections from the disk surface) gradually comes into the detector's view at position 100, the signal

increases, and so does the predictive ability of the spectra. The best prediction position is around position index 30, at lower position indexes, the predictive ability reduces even though the intensities of the spectra are increasing; this is because the spectra are becoming distorted with specular reflection. From plot (b) in Fig. 3-42 the largest spectrum corresponds to position index 0, with the largest amount of specular reflection. The 5th largest spectrum is approximately at position index 30, where high signal and minimal specular reflection combine to produce the lowest prediction error.

The RMSEP values from the best position in each of the modes are given in Table 3-8.

Mode	Speed (Spectra/fruit)	RMSEP SSC (°Brix)	RMSEP DM (%)
Transmission	7	0.54 ± 0.07	0.66 ± 0.08
	11	0.65 ± 0.10	0.85 ± 0.07
Reflectance	15	0.81 ± 0.09	0.87 ± 0.09
	61	0.74 ± 0.08	0.79 ± 0.09

Table 3-8: RMSEP results from the best positions for each of the modes and speeds of moving apples.

Transmission mode was better at predicting the SSC and DM for the apples than the reflectance mode. However, the transmission measurements were made at a much slower speed than the reflectance mode - a difficulty for a practical fruit grading system. For reflectance mode, taking 61 spectra/fruit instead of 15 was achieved by slowing down the grader speed and keeping the integration time constant. The slower speed led to improved predictions from the measured spectra even though the signal magnitudes at both speeds were comparable. The improvement could have been a result of the reduced blurring of the image as the fruit moves slowly under the detector or the reduction of the influences of the mechanical grader (e.g., vibrations were smaller at slower speeds) on the measured spectra. For the transmission mode, taking 11 spectra/fruit instead of 7 meant that the signal was lower (due to the shorter integration times and consequently the predictive ability decreased as indicated by the larger RMSEP values.

The prediction errors found for moving fruit compare well with the values reported by Bellon-Maurel et al. (1997), using a Zeiss MMS1 detector on Golden

Delicious apples. The authors found an SEP of 1.25°Brix with Rp = 0.94. Moons and Dardenne (2000) reported RMSECV for SSC of 0.49 °Brix and DM 0.78% for moving apples.

In Japan it was thought (McGlone, et al., 1997) that using transmission mode may increase the grading speed. The reasoning was that spectra acquired using transmission are useful right across the passage under the detector whereas significant numbers of spectra from the reflectance sequence are not useful due to the influence of specular reflection and fruit shading. While this may be true, the large variation in transmission intensity as a fruit passes means that only a small amount of the detector's dynamic range can be used. The net effect is that the transmission mode arrangement is not likely to be able to operate as fast as the reflectance mode in the near future.

By averaging the spectra at the best locations some of the predictions were improved. Table 3-9 shows which positions were averaged and the RMSEP achieved (compared to those reported for the single positions in Table 3-8).

Mode	Speed (spectra/fruit)	Position range	# Spectra averaged	RMSEP for SSC (°Brix)	RMSEP for DM (%)
Transmission	7	33-50	2	0.53 ± 0.06	0.73 ± 0.09
	11	30-50	3	0.61 ± 0.06	0.78 ± 0.08
Reflectance	15	64-79	3	0.69 ± 0.08	0.80 ± 0.08
	61	66-75	6	0.73 ± 0.08	0.79 ± 0.1

Table 3-9: RMSEP results from combining the best positions in each of the modes and speeds.

Fig. 3-45 compares the single spectrum prediction errors to the combined spectra errors. The error bars are the standard error of the mean generated from 20 repeated estimates of RMSEP using different random calibration/prediction set splits.

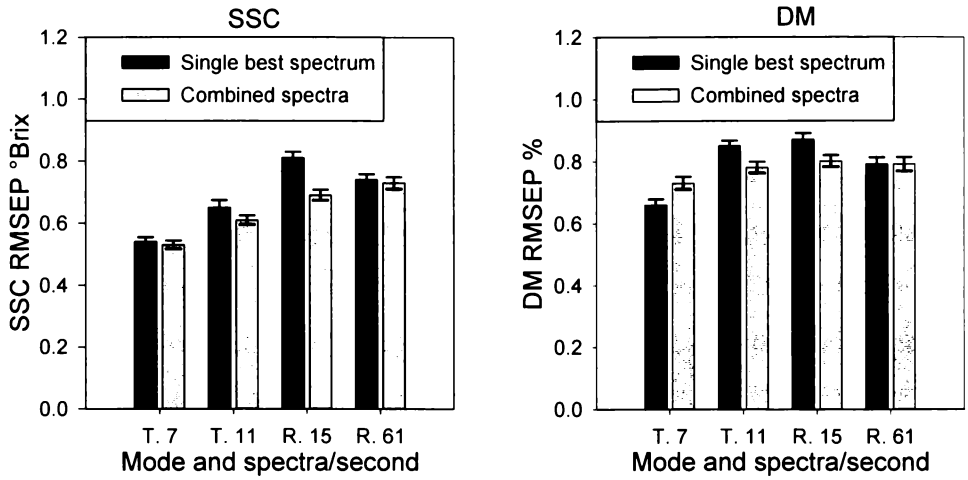

Fig. 3-45: Comparing the prediction error from the single best position and combined spectra for each mode (R=reflectance, T=transmission) and speed (7, 11, 15, or 61 spectra/fruit).

It can be seen from Fig. 3-45 that for the fast reflectance mode (R. 15) the averaging improved both the SSC and DM predictions. For the slow reflectance mode (R. 61) no improvement is found. For the 11 spectra/fruit transmission (T. 11) mode the averaging improved the prediction for DM with only slight improvement on SSC. The 7 spectra/fruit transmission mode (T. 7) showed no change on averaging for SSC prediction and reduced the predictive ability for DM. It may be that the 7 spectra/fruit transmission mode does not sample frequently enough to facilitate predictive improvement by averaging two sequential spectra. It is surprising that the average spectrum produced by 6 very similar spectra, from the best positions of the slow reflectance mode (R.61), did not improve the prediction, the influence of the larger RMSEPs from the 1st and 6th spectra must have been as great as the gain in signal to noise from averaging the spectra.

4

Light levels inside fruit

This chapter describes a custom-built fibre optic probe designed and developed for this work. The probe is used to make direct measurements of the light levels inside fruit. The degree of light penetration in apple is assessed for a range of wavelengths, and the light extinction rates for different fruit are compared at 808 nm. The effect of the mandarin skin on internal light levels is examined and is compared with results from Monte Carlo simulations.

4.1 Measuring the light distribution in fruit

A fibre optic probe was designed and built to measure the light intensity at any point within a fruit illuminated with either a white light source or a diode laser. Fig. 4-1 shows the typical experimental arrangement to investigate a mandarin. Essentially, a reinforced fibre probe is inserted into a fruit using translation stages.

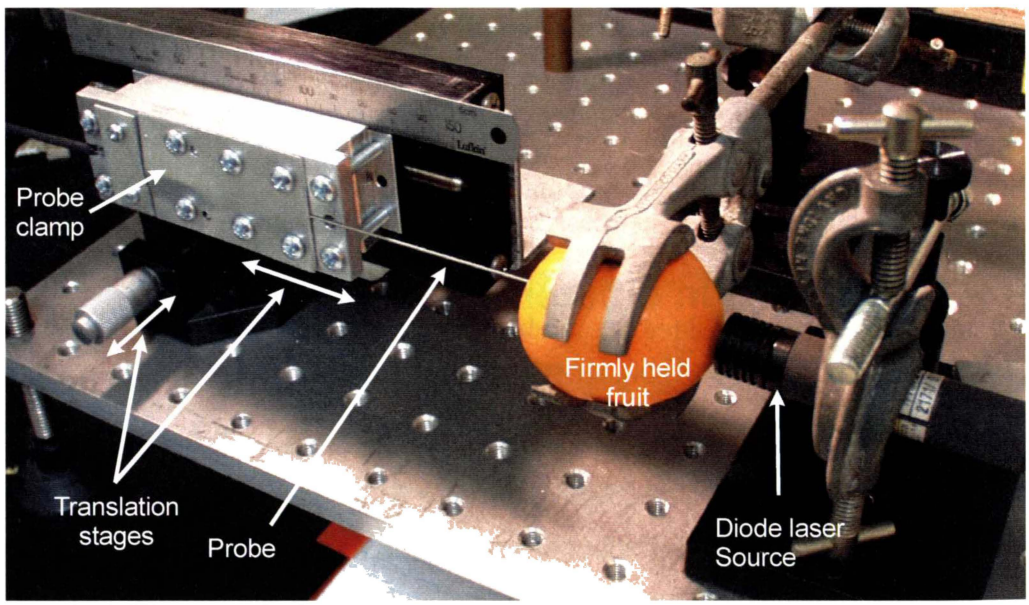

Fig. 4-1: Light penetration measurement arrangement for a mandarin illuminated with a diode laser.

The probe consists of a fibre optic cable with a 400 μm diameter core (425 μm including cladding) and terminated with an SMA connector. The numerical aperture (0.39) of the fibre (Thorlabs Inc., Newton, NJ, USA) gives a total acceptance angle of 46° in air reducing to about 34° in mandarin juice. The fibre has particularly low water absorption properties, providing greater sensitivity over the water absorbing wavelength ranges. At the measurement end of the probe, the polished bare fibre was glued inside a 150 mm length of stainless steel sheath (500 μm i.d., 780 μm o.d.). The probe assembly was clamped to a two-axis translation stage to enable insertion into the fruit in small controlled steps at different horizontal locations.

The light collected by the probe was conveyed to a MMS-1 solid-state spectrometer which has a wavelength range 300-1150 nm. In some experiments an additional InGaAs spectrometer (coupled with a Y-branch cable) with a range 940-1700 nm (MMS NIR tc II) was also used. The spectrometers were cooled to constant temperature. Because there is a large variation in light intensity across the fruit, the integration times for light acquisition were varied from 4-6500 ms. The spectra were corrected for dark current and were scaled by the integration time to normalise them.

In concept, the probe could be positioned and orientated at any position within the fruit. Sideways and backscatter could have been measured as described by Wilson et al. (1984) for mammalian tissue. However, in the studies presented in this thesis all measurements were made with the probe pointing towards the illuminated side of the fruit, parallel to the illumination direction of the light source. This minimises the influence of the probe on the measurement since the probe's shaft is not between the source and the sampling tip.

4.2 Light levels in different fruit

4.2.1 Light penetration depth in apples

Several varieties of apples ('Braeburn', 'Granny Smith', 'Royal Gala', and 'Pacific Rose') were purchased from a local fruit store and held in an air-conditioned room for 1 day to thermally equilibrate. Fig. 4-2 illustrates the experimental arrangement for measuring the light intensity at different depths and wavelengths through the centre of an apple.

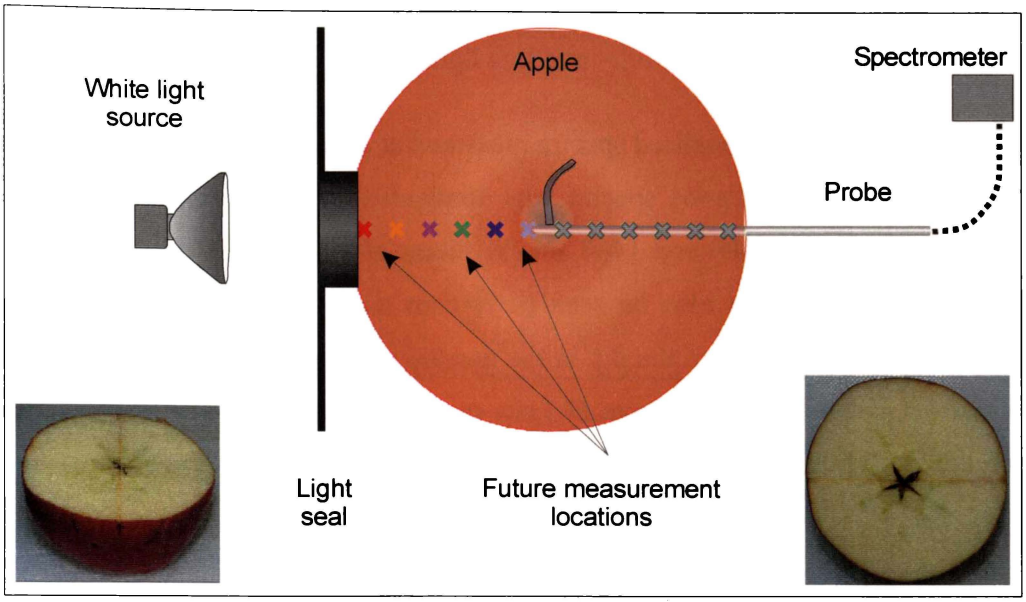

Fig. 4-2: Experimental arrangement for measuring the depth of light penetration into an apple.

The light source was a 250 W slide projector (WideoScope®, AB Wiktors Mekaniska, Jarfälla, Sweden), which had been modified to use a stable DC power supply. A shutter blocked the light when measurements were not being taken. The circular, 22.5 mm diameter light seal enabled illumination of the lower shoulder of the apple. The light was collected by the probe as it was gradually inserted into the 'Pacific Rose' apple moving towards the illuminated side. Measurements were taken every 3.175 mm from the far surface to within 1 mm of the illuminated surface for a total of 22 measurements inside the fruit. The procedure was then repeated on the other three varieties of apple (sampling more densely in the source side of the fruit) in order to verify the trends measured for the 'Pacific Rose'.

The light collected by the probe was passed to the two spectrometers via a Y-branching fibre cable. The integration time was 6500 ms for the majority of the distance but, to avoid saturation of the detectors, was reduced to 1000 ms for penetration depths less than 10 mm. A reference value was recorded by sampling the light level at the fruit entry point without the apple present at an integration time of 100 ms. Dark currents were recorded at each integration time. Spectra were normalised by subtracting the dark current, scaling for integration time and dividing by the reference. The spectra from the two spectrometers were combined using the 500-952 nm range from the silicon spectrometer and the 957-1690 nm from the InGaAs spectrometer. Wavelengths below 500 nm and above 1690 nm

were discarded due to the very low signal to noise ratio arising from the combination of falling sensitivity and low light levels.

Fig. 4-3 shows the normalised spectra measured at each depth inside the fruit. The shape of the transmittance spectra was dominated by water absorption lines at 958 nm, 1153 nm, 1409 and 1460 nm (Williams and Norris, 1987). Minor water absorption effects can also be seen at 760 nm and 838 nm. The absorption of chlorophyll-a gives rise to the dip in the transmission spectra around 663 nm, and chlorophyll-b with absorption at 645 nm can be seen to create a slight deviation on the side of this chlorophyll dip.

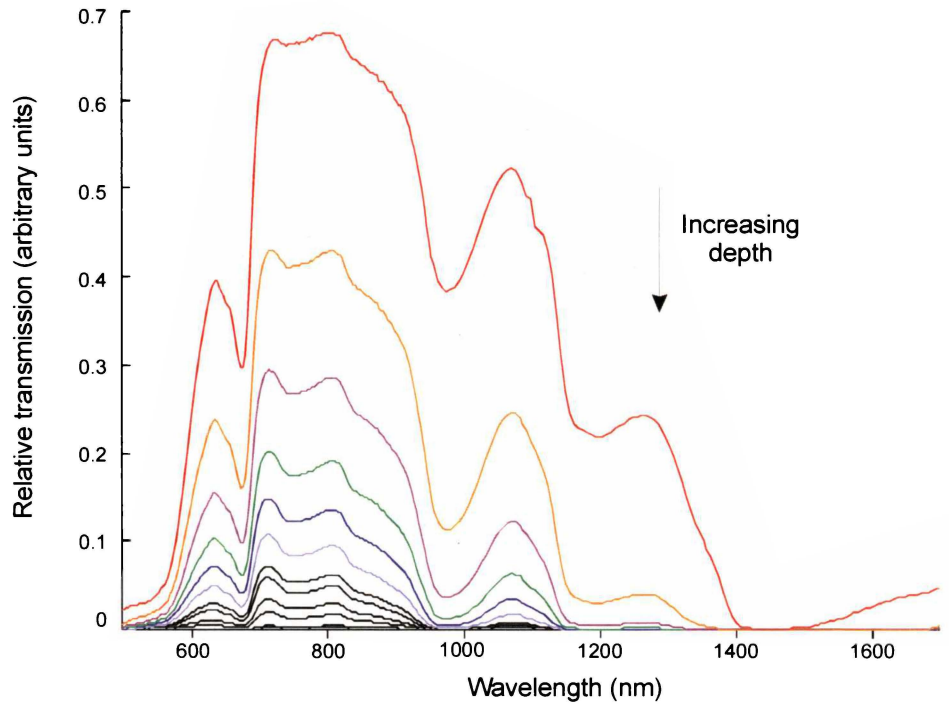

Fig. 4-3: Transmission spectra recorded inside an apple which is illuminated by a white light source. The largest spectrum was taken at a position 1 mm from the illuminated surface, each subsequent measurement is 3.175 mm deeper into the apple.

The rate of light extinction is approximately log-linear down to a relative transmission intensity of ~1% (Fig. 4-4). Beyond that, tissue discontinuities and/or the sensor noise floor lead to departure from this form. An abnormally low intensity value was recorded at approximately 33 mm into the fruit for all of the wavelengths that were measurable at that depth. Subsequent dissection of the apple showed this to be an area of the apple's core where there was some very hard tissue was surrounding a seed.

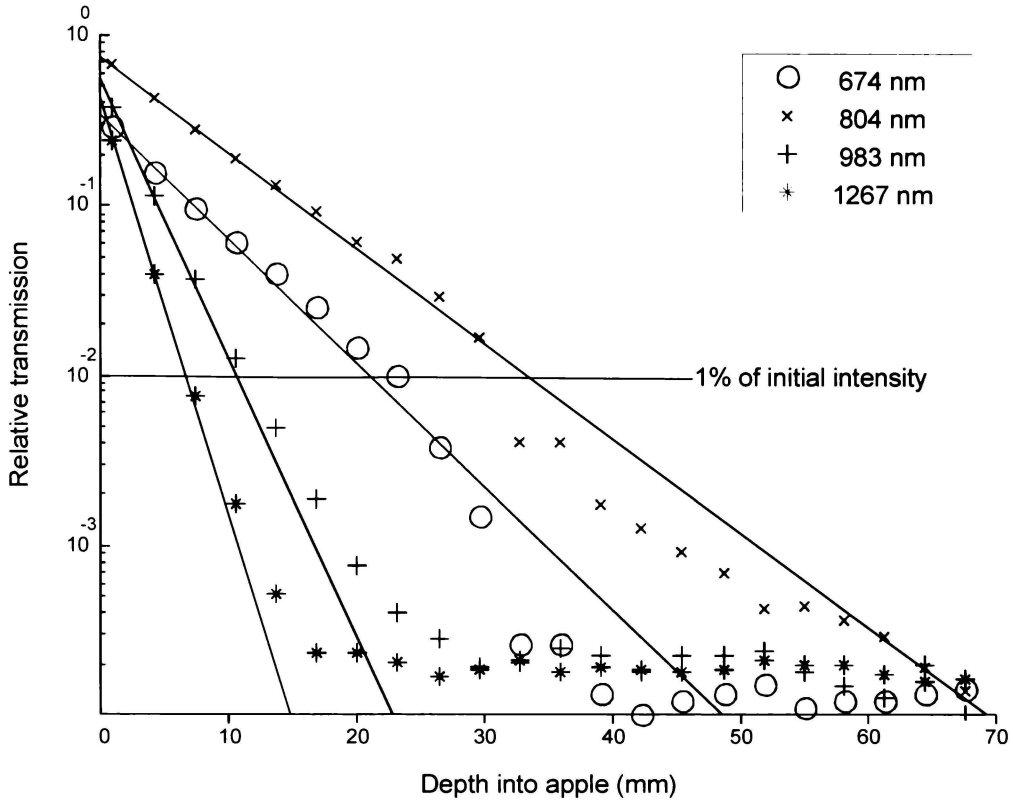

Fig. 4-4: Rates of light extinction for several selected wavelengths. There is a discontinuity at d = 33 mm, where the probe was against a seed cavity in the apple. The straight lines illustrate log-linear form of the first few data points before the noise level and/or discontinuity come into affect.

For each wavelength, an exponential function of the form,

$$I = I_0 \exp(-\mu_e d) \qquad 4\text{-}1$$

was fitted to the first nine data points (or less if the threshold of 0.002 relative transmission is reached). I is the intensity of the light at a given depth into the fruit, I_0 is the intensity of the light just inside the surface of the fruit, μ_e is the light extinction coefficient, and d is the depth into the fruit from the measurement position which is closest to the illuminated edge. The skin of the fruit may influence the experimentally measured μ_e, whereas calculations of the effective attenuation coefficient, μ_{eff} (Equation 2-12) do not take into account the influence of boundaries.

The 1% light penetration depth, defined as the depth at which the incident light intensity is reduced by 99%, was calculated from the fitted extinction coefficients, μ_e for each wavelength. The penetration depth (Fig. 4-5) follows the same general pattern with wavelength as the transmitted spectra and was dominated by water and chlorophyll absorption.

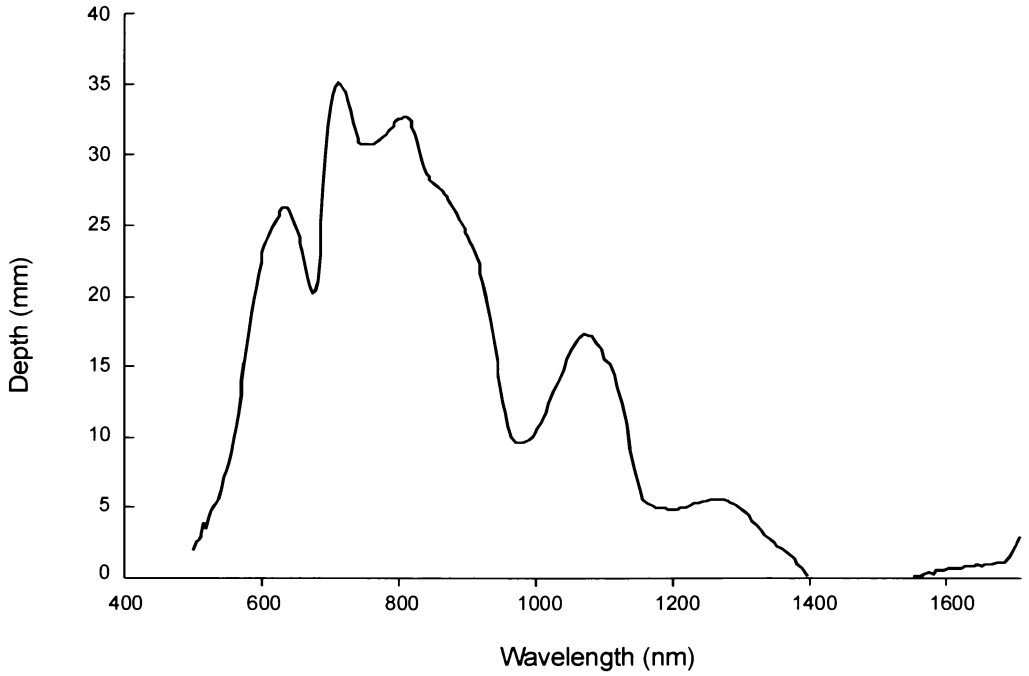

Fig. 4-5: The depth inside an intact apple where the NIR light has been reduced to 1% of the incident intensity.

Measurements on fruit of three other varieties, 'Braeburn', 'Granny Smith' and 'Royal Gala' produced similar results, suggesting that the degree of light penetration against wavelength shown here is general for apples at this stage of maturity. The results confirm the presence of a diagnostic window in the range 700-900 nm for which the 1% penetration depth was more than 25 mm, reaching a peak of 35 mm at 713 nm. This window is a direct result of reduced water absorption (Segelstein, 1981) and is the only such window giving more than a 10 mm "1% penetration depth" over all electromagnetic wavelengths from well below 10 nm to well above 10 mm. At wavelengths higher than 1000 nm the penetration is greatly reduced, due to the water absorption, and reaches less than 1 mm in the range 1400-1580 nm. The light in this wavelength range is therefore not capable of sampling flesh greater than 1 mm into the fruit at this level of sensitivity. The fact that this diagnostic window coincides with the sensitive region of relatively cheap silicon detectors (300-1050 nm) and the good depth of penetration make this region attractive for NIR assessment of fruit quality.

4.2.2 Light distributions in fruit

4.2.2.1 Extinction curves

By using a laser instead of the intense tungsten lamp the fruit can withstand far more extensive measurements without being damaged by excessive heat. Using a similar arrangement to Fig. 4-2, the fruit was illuminated over an area of 2.5 mm by 4.5 mm (long side parallel to the stem-calyx axis of the fruit) with an 808 nm diode laser (βeta-TX, Vector Technology Ltd, Abertillery, Gwent, UK) operating at 11 mW. The laser wavelength was chosen because it gives good transmission in fruit (Fig. 4-5; Osborne et al., 1993) and coincides with the most sensitive region of our detector.

By sampling the light level at probe locations parallel and in line with the source laser beam through the centre of a mandarin, a light extinction curve is formed as a function of penetration depth. Fig. 4-6 shows the average light extinction curve for four separate mandarins, the error bars indicate ± 2 standard errors of the mean at each depth.

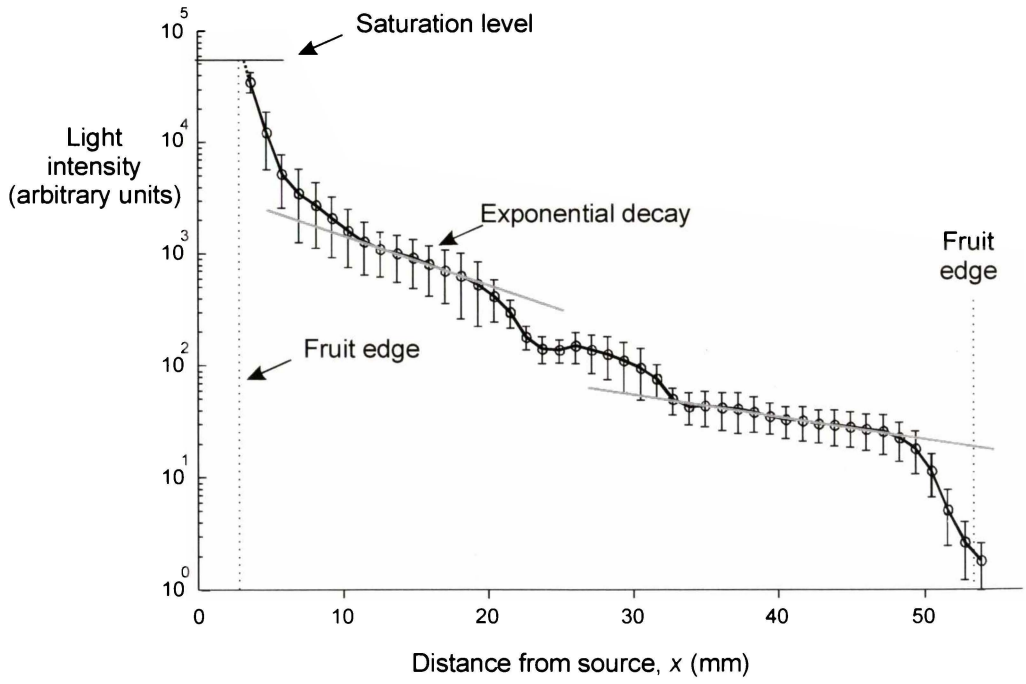

Fig. 4-6: Mean light extinction rate for 4 mandarins. The error bars indicate the 95% confidence interval for the mean.

Apparent in the curve is the rapid reduction in light level across the 4 mm thick illuminated skin, a less rapid but still high reduction as the light continues to diffuse into the flesh to about 10 mm into the fruit (x = 13 mm), two generally

exponential (constant slope on the log-linear scale) decay regions through the fruit flesh (separated by the perturbation from the central core region containing pith and air at $x = 22\text{-}32$ mm), and finally a rapid drop across the distal skin ($x = 50$ mm). The reason mandarins have two different rates of decay is due to the reflective skin of the mandarin and is considered in detail in section 4.2.3.

Monte Carlo simulations show the initial rapid reduction in light level is attributable to the diffusion process and will be influenced by the optical properties of the tissue, the diameter and incidence angle of the irradiation beam, as well as the boundary (skin) properties (Wilson and Jacques, 1990; Wilson and Adam, 1983). At 808 nm the attenuation due to the skin will probably be dominated by the scattering (and consequent dispersion) in that tissue rather than absorption per se. The rate of exponential decay will depend on the scattering and absorption properties of the tissue and will be influenced by the skin and other features in the fruit such as the core and segment boundaries. On each side of the central core there are two distinct regions with exponential decay (Fig. 4-6, straight lines). The region closest to the illuminated side has an extinction coefficient of 0.041 mm^{-1} ($\frac{1}{e}$ penetration depth, $\delta = 24$ mm) and the more distant region 0.014 mm^{-1} ($\delta = 71$ mm). Both these rates are lower than the rate experimentally measured for an apple (0.173 mm^{-1}, $\delta = 5.8$ mm), a nashi (0.159 mm^{-1}, $\delta = 6.3$ mm), and the inner pericarp of a kiwifruit (0.241 mm^{-1}, $\delta = 4.1$ mm), which are compared to an individual mandarin extinction curve in Fig. 4-7. Greensill and Walsh (2000) reported a rate of 0.0788 mm^{-1} ($\delta = 12.7$ mm), for melon tissue at 812 nm. The steep decline in light levels when the probe passes through the skin at both ends of the mandarin shows that the mandarin skin region is highly scattering/absorbing compared to the mandarin flesh. This has been asserted without data, by other authors (e.g., Kawano et al., 1993).

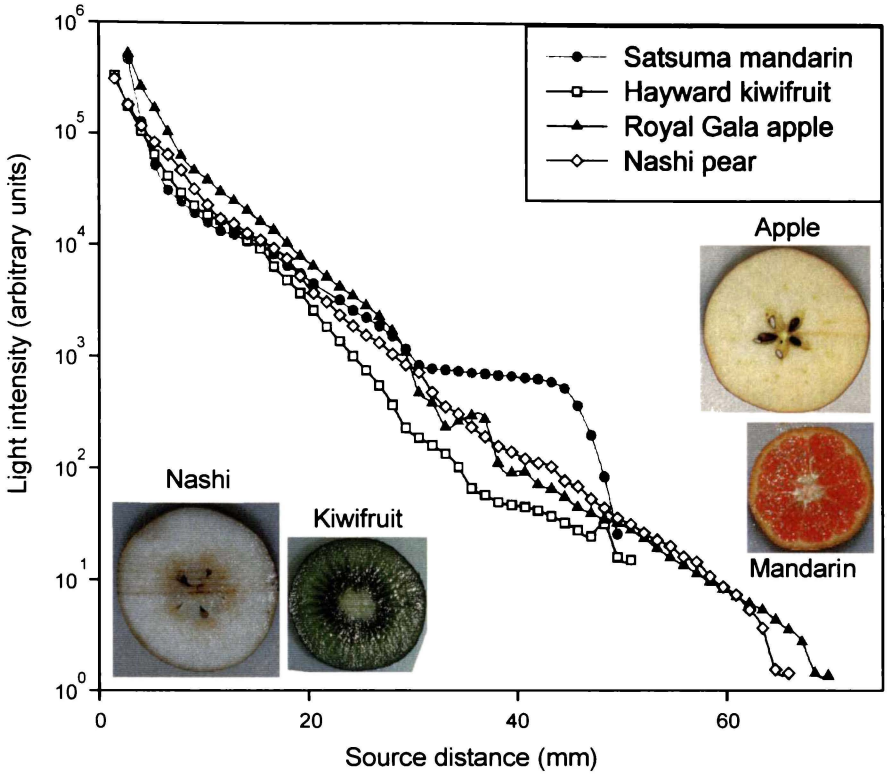

Fig. 4-7: Rates of light extinction for different fruit.

Fig. 4-7 illustrates that each fruit follows an initial power law rate of light attenuation, which reduces to an exponential rate after a characteristic diffusion distance. In this diffusion region the initially forward oriented light becomes isotropic, and the distance over which this occurs depends on the tissue's absorption and scattering properties as well as the degree of scattering anisotropy. It can be seen that different regions in each fruit have different exponential rates of light extinction. The rate of reduction depends on the tissue scattering and absorption properties as well as on the influence of the fruit skin. The mandarin in particular stands out as having an elevated light level in the region distant from the illuminated side. The kiwifruit also showed a marked change between inner and outer pericarp (at 16 mm and 36 mm).

Fig. 4-8 shows light level curves measured by the probe in a single mandarin. The three curves correspond to the levels recorded when the probe first enters the mandarin (initial insertion), as it is being removed (retraction), and when it re-enters the previously formed 'channel' (repeat insertion). Normal mapping measurements would only involve the initial insertion curve as that curve is least affected by tissue damage created by the probe pr earlier insertions (e.g., development of a light channelling path). Multiple insertions were made in order

to assess the effect of damage caused by the probe on the light levels. In the measurements illustrated in Fig. 4-8 a seed has interfered with the insertion path (confirmed by a subsequent dissection) and consequently the deviation between the initial and repeated insertion curves was the worst observed in any experiment. However, the overall shape of the curves still shows the same trend as Fig. 4-6.

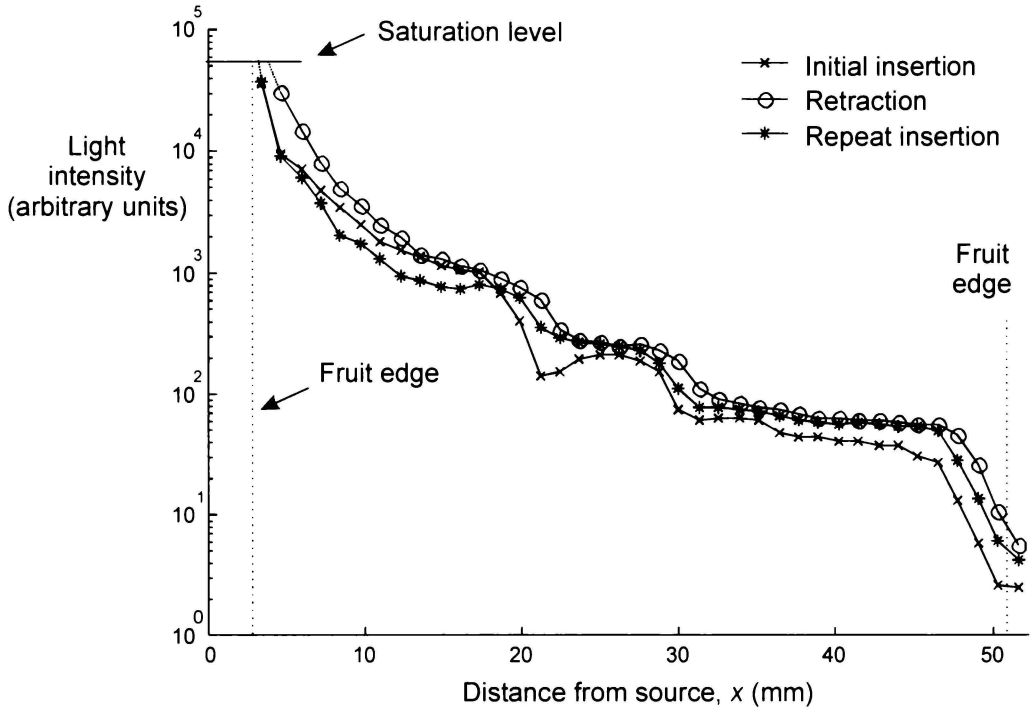

Fig. 4-8: Testing the influence of the probe on light level measurements. If the probe is reinserted down the same path increased light levels are recorded. There is a horizontal shift when comparing measurements made in the forward and backward directions.

Beyond 18 mm, the repeated measurements showed higher light levels (approximately twice) than those of the initial insertion. This may be due to a channelling effect with light moving more freely along the path that was previously opened up, or to the absence of some of the scattering/absorbing tissue for the repeated insertion. The first insertion is more likely to encounter obstacles such as seeds and core material, which may be dragged a short way before being pushed aside and may even deflect the probe. Such interference has probably occurred at position 18 mm, from this point to a point just before the source side, the second insertion shows an abnormally reduced light level compared to the initial insertion. In this region the second insertion may be breaking new ground instead of following the first insertion's channel (because the first insertion was deflected off centre by the seed for a short distance) before intercepting the channel again closer to the source.

The direction of the probe movement during measurements is important. When the probe is first inserted it presses against the mandarin skin but doesn't actually puncture it until it is moved to the next step of 1.27 mm. Similarly when the probe reaches the other side of the mandarin the skin deflects a little before the probe breaks through. Both effectively shift the fruit to the left thus shifting the light curve in the negative x direction depending on the elasticity of the tissue. Similar effects probably also occurs to some extent at all measurement points in the fruit. This is illustrated in Fig. 4-8 by the difference between the repeated and retraction measurement; they are practically the same except for the approximately 1.5 mm shift in the x direction. This helps define some limits on the accuracy of the position coordinates because compression or stretch in the tissue is unlikely to be more than this when travelling in one direction.

4.2.2.2 Flux maps

To map the light levels, the probe was inserted at six different off-axis positions located along the equatorial plane (Fig. 4-9) from 0 to 15 mm in steps of 3 mm from the centre line of the fruit. Light measurements were recorded every 1.27 mm along each insertion path to map the light flux.

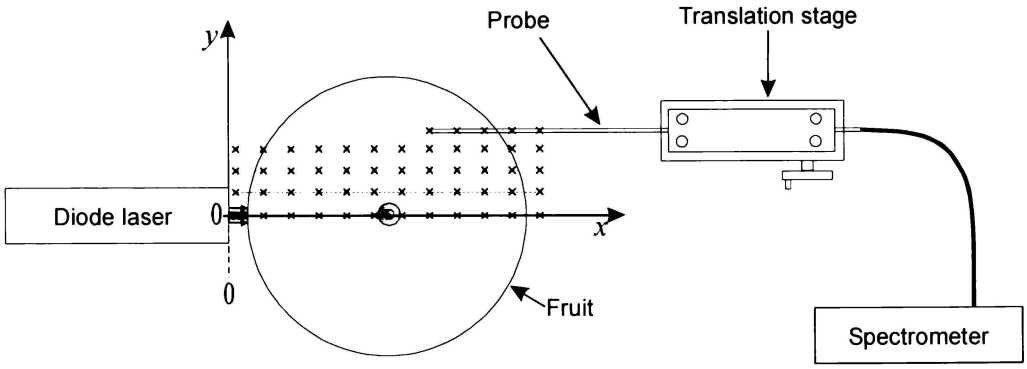

Fig. 4-9: Experimental arrangement for measuring the light distribution in fruit. The crosses illustrate some of the measurement locations within the equatorial plane of the fruit.

The log of the light intensity at each measurement location in an 'intact' apple was used to compute the contour lines shown in Fig. 4-10. The apple was subsequently dissected and photos were taken, The photo could then be used as a reference to the contour lines in Fig. 4-10. The initial diffusion region, where the contours are not emanating spherically, and the perturbations around the seeds can be clearly seen.

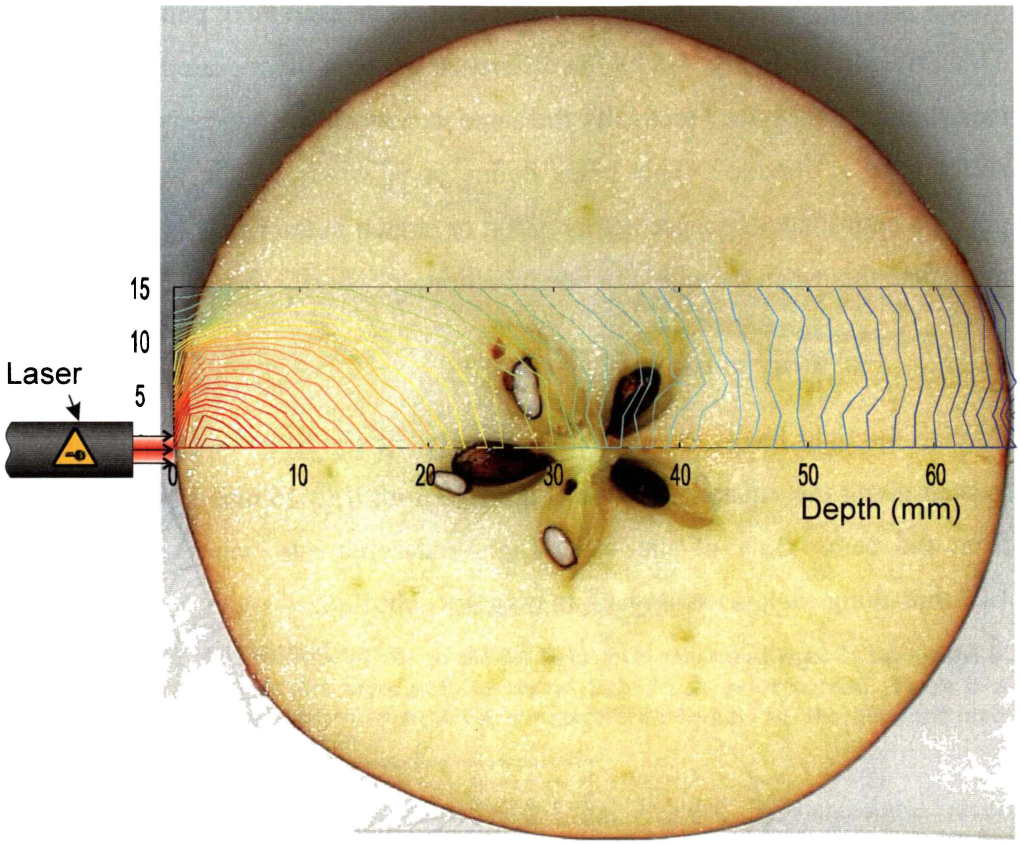

Fig. 4-10: Mapping 808 nm laser light through an intact apple (the cross section of the apple is overlaid for illustration purposes).

A mandarin was examined in greater detail by plotting the initial insertion curves along three different mapping paths. The influence of the different regions of the mandarin can be seen in Fig. 4-11.

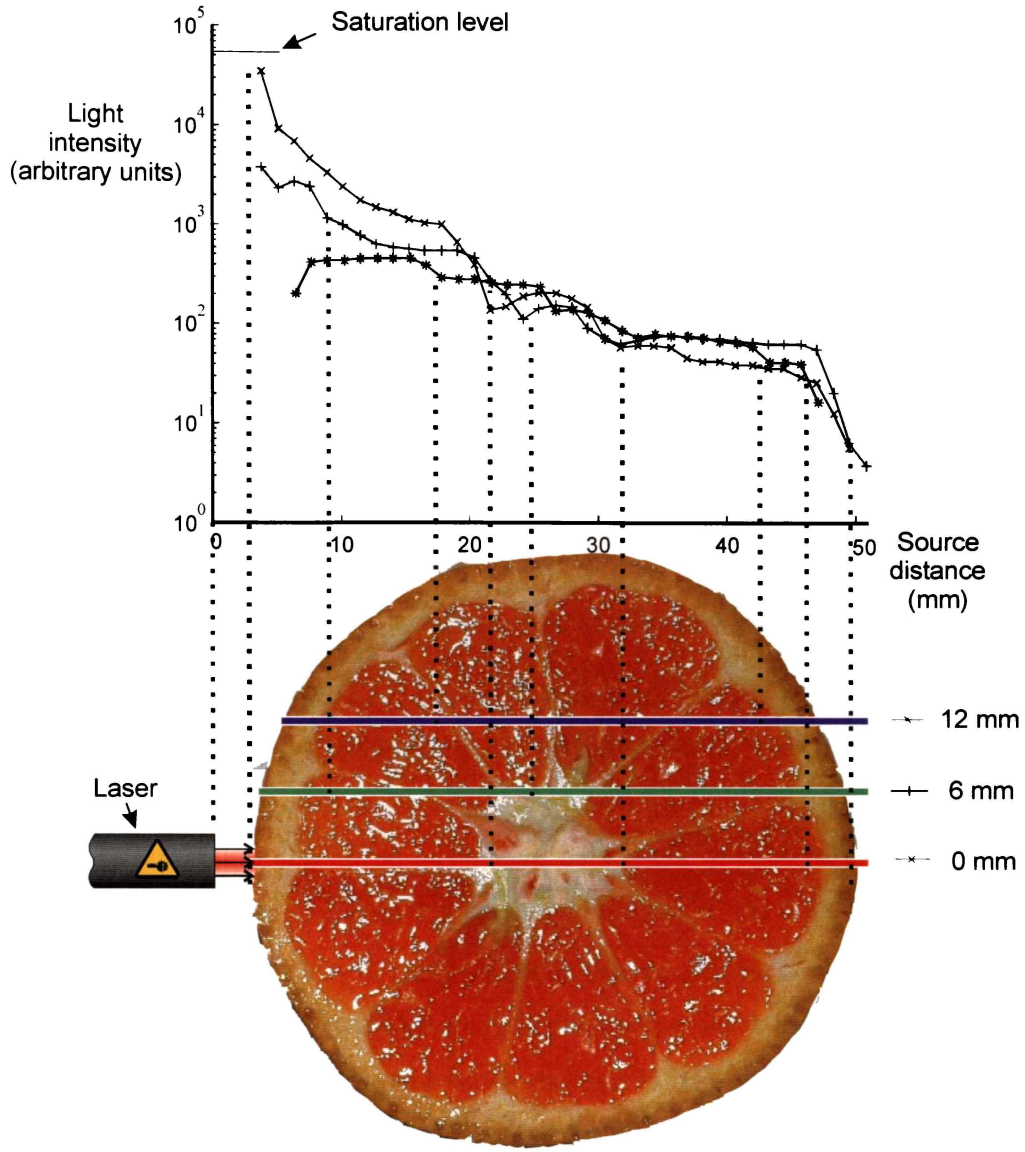

Fig. 4-11: Mapping the light levels in a mandarin for three penetration paths: $y = 0$, 6 and 12 mm. Perturbations in the light level curves can be attributed to the different material the probe is passing through, e.g., skin, section walls, and the core.

When the paths intersect segment boundaries some perturbation is observed. Wilson et al. (1984) reported similar perturbations in their exponential curve for the muscle-muscle boundaries in rabbit tissue. For paths that intercept the core there seems to be a dip in the insertion curves just before and after entry, while within the core a slight elevation of the light level is recorded. Some increase in light collection is expected within the core (and possibly skin also) due to the change in the fibre's acceptance angle as the probe moves from a juice filled region of the flesh to regions containing air. The change of refractive index from juice to air would mean an increase in acceptance angle from 34° to 46° thereby increasing the light collection by 82% possibly explaining the increased light level

observed inside the core region. Other perturbations may be attributed to tissue compression or tissue dragged in front of the probe tip.

4.2.3 Mandarin's internally reflective skin

There is a much slower decay and the light level almost plateaus for the second half of the mandarin that is far from the light source. This was surprising as other fruits such as apples and nashi (Fig. 4-7) exhibit a near continuous exponential decay, which is steeper and extends over a greater portion of the fruit. One possible explanation for the elevated light level for mandarins is that the thick white mesocarp section of the skin may be back-reflecting some of the light that would have escaped had the skin been less reflective.

4.2.3.1 Experimental measurements

The possibility of significant skin reflectance was investigated on a separate mandarin. Segments of skin from the side opposite the source were carefully removed incrementally using a scalpel. Fig. 4-12 defines d as the distance from the original fruit edge to the peeled plane. The probe was repeatedly inserted at the central location ($y = 0$) for four values of d: 0 mm, 5 mm, 15 mm and 25 mm.

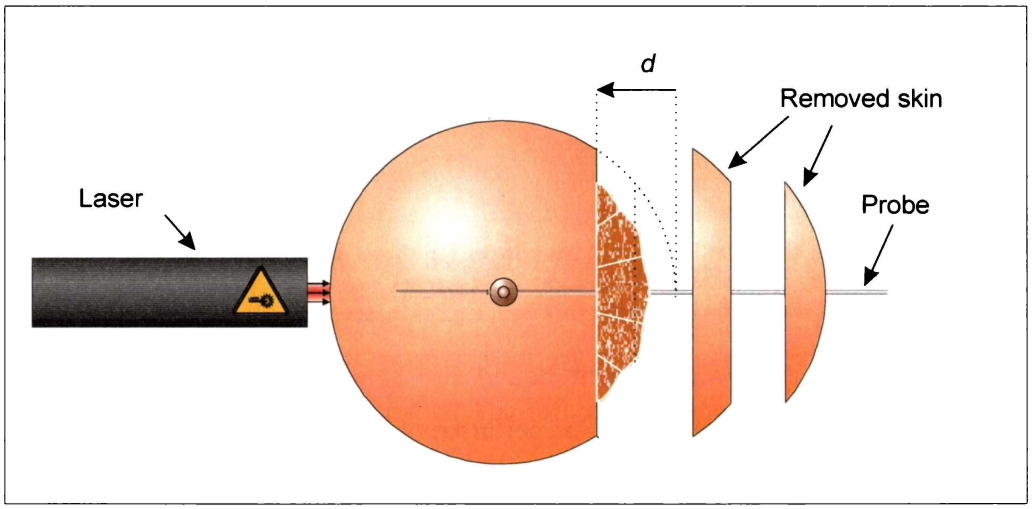

Fig. 4-12: Measuring the effect of mandarin skin by removing amounts of the skin from the flesh of the mandarin, where *d* is the distance from the original fruit edge.

For the first case the probe was inserted twice with the skin intact ($d = 0$ mm), and the first measurement was ignored so that all subsequent insertions would occur along a previously created channel. It can be seen (Fig. 4-13) that the measurement is repeatable in the unpeeled region for the four reinsertions of the probe. As more skin is removed, the rate of decay on the peeled side of the fruit from the core to the edge of the mandarin increases, and the plateau disappears.

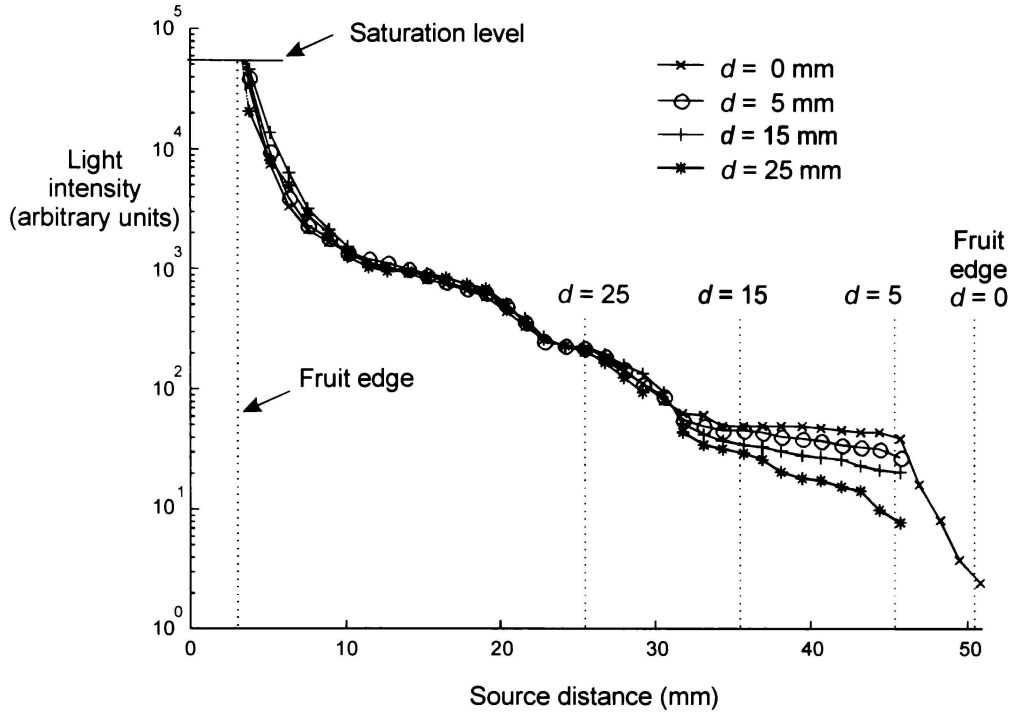

Fig. 4-13: Light intensity resulting from partial skin removal for a 48 mm diameter mandarin.

The light levels measured in the core are relatively unaffected by the peeling, even when half of the skin is removed ($d = 25$ mm).

A theory to explain the loss of the 'plateau' is that when a small amount of skin is removed a hole has been made where light can escape. As the hole size increases the light level inside the fruit reduces suggesting the light is escaping out through the hole in the skin layer and conversely that the skin must be reflecting light back into the interior rather than absorbing it. This explains the light level reductions observed in Fig. 4-13 as incremental amounts of peel are removed. This theory may also explain the observation that exponential decay rates observed in the fruit on the source side of the core are steeper than the exponential decay rate observed on the far side of the core. In the region of the fruit near to the light source where the sides of the fruit are diverging from the sampled path we might expect the light level to drop off quicker than in the region far from the source where the sides are converging towards the probe path.

4.2.3.2 Monte Carlo investigation

To investigate the skin reflection theory above a Monte Carlo simulation of the light levels inside a mandarin was undertaken. Computation was based on code created by Wang et al. (1995), which is available via the Internet (Jacques,

1998a). Similar simulations have been completed for mammalian tissue, e.g., Wilson and Adam (1983) and Jacques (1998b). The Monte Carlo method is the most appropriate modelling technique for this problem because it can handle complex illumination geometry and boundary conditions, and because the light collection for our fibre probe can be easily simulated.

The mandarin is modelled as a 50 mm diameter sphere of tissue surrounded by a boundary which specularly reflects 90% of the incident radiation internally. Based on the modelling of Flock et al. (1989), Wilson and Jacques (1990) suggested that 50% of light inside mammalian tissue is internally reflected. We propose a much larger value because the white reflective mesocarp that lines the inside of the mandarin skin should have significantly higher reflectance. In addition, a large refractive index mismatch can be expected between the juice-filled mandarin tissue and the air-containing mesocarp region, as well as between the exocarp (outer skin) and the external environment.

In our simulation large numbers of photons enter at one edge of the sphere and are tracked throughout the volume. The simulated tissue has absorption and scattering coefficients μ_a and μ_s, respectively. The degree of forward scattering for a scattering event is defined by the anisotropy factor, g, which is not known for mandarin tissue. However, except close to the model boundaries and the source where there may be a rapidly varying spatial distribution of the light level, the degree of scattering anisotropy is unimportant, and μ_s and g may be combined via the similarity relation (see section 2.3.1). This reduces computation time by more than an order of magnitude, but causes less accurate light level assessment close to the illuminated side where the light is not fully diffuse (i.e., where the light is still predominantly anisotropic). Further details of the Monte Carlo simulations are described in section 2.3.2.

Measurements by Cubeddu et al. (2001) show the reduced scattering and absorption coefficients for 808 nm light in an apple to be: $\mu_a = 0.003$ mm^{-1} and $\mu_s' = 2.3$ mm^{-1} and for a peach: $\mu_a = 0.003$ mm^{-1} and $\mu_s' = 0.7$ mm^{-1}. These coefficients imply absorption path lengths of the order 300 mm and scatter path lengths of less than 1 mm. Birth, (1978) in discussing the light scattering in food, reported that the light scattering elements in plant leaves were the cell walls. It is probably the same with fruit, and the cell wall interfaces will be the dominant

light scatterers (McGlone et al., 1997c). The juicy flesh region of the mandarin consists of large, fluid filled sacs with relatively large cell wall spacing (see section 2.1.2), implying a lower scattering coefficient, compared to some other fruit, is appropriate for the inner tissue. The absorption coefficient in tissue is dominated by the absorbing constituents such as water, sugar, and chlorophyll and as such is highly wavelength dependent. At 808 nm we have chosen the whole fruit parameters for a mandarin to be $\mu_a = 0.001$ mm^{-1} and $\mu_s' = 1$ mm^{-1}. These values do not correspond to any individual tissue type within the mandarin (for which there may be variation) but are the optical properties portrayed as if the fruit were a homogenous object. Because the absorption coefficient is considerably lower than the scattering coefficient, use of the similarity relation is appropriate since there will be many interactions before absorption occurs (Jacques, 1998b). The boundary conditions are critical for this model because the penetration depth will be significant compared to the fruit size. Based on diffusion theory the penetration depth required to reduce the light level by a factor of 1/e is given by,

$$\delta = \frac{1}{\sqrt{3\mu_a(\mu_a + \mu_s')}} = \frac{1}{\sqrt{0.003 \times 1.001}} \approx 18\,\text{mm}. \qquad (4\text{-}2)$$

We cannot ignore boundary conditions, which would allow use of the much simpler semi-infinite model (an air/tissue interface plane that has an infinite tissue depth on one side), because our simulated fruit radius does not exceed 12δ (Wilson and Jacques, 1990). For the parameters used in our simulation, that simplification would require fruit with radii greater than 200 mm. Modelling the whole fruit with a single pair of absorption and reduced scattering coefficients also presents a significant simplification. In reality the individual elements (skin, flesh and core) will have different parameters, which would require separate consideration in a more complex model.

To assess the light distribution measured by an imbedded probe like that used in these experiments the photon trajectory was checked at each interaction for coincidence with a probe location. The direction of the trajectory at these points was also assessed to enable modelling of both an omni-directional collecting fibre and a fibre with a total acceptance angle of 34°, each with a collecting cross-section 400 μm in diameter. The 34° case directly simulated the experimental measurement, whereas the omni-directional case assessed total light level at that

location. Collecting the light from all directions is interesting for theoretical comparison and will have better statistics since an estimated 46 times more light is collected by an omni-directional probe than the 34° probe.

4.2.3.3 Results and discussion

The experimental results as shown in Fig. 4-13 are backed up by the Monte Carlo simulations (Fig. 4-14), which show similar behaviour when modelling a peeled mandarin. The four upper curves in Fig. 4-14 illustrate the light level curves as measured by an omni-directional probe, which collects light from any direction and thus contains less statistical noise than the lower curves. The lower curves only show light meeting the 34° acceptance angle criteria, emulating our forward collecting experimental probe. The omni-directional probe is important from a theoretical point of view as well as providing greater statistical precision. Considering isotropic light, the lower curves are generated from only $\sim\frac{1}{46}$th of the light reaching each location and could be expected to have almost 7 times higher noise levels ($\sqrt{46}$). It is also observed that the slopes of the two curves are identical beyond 12 mm and are thereafter approximately a factor of 46 apart.

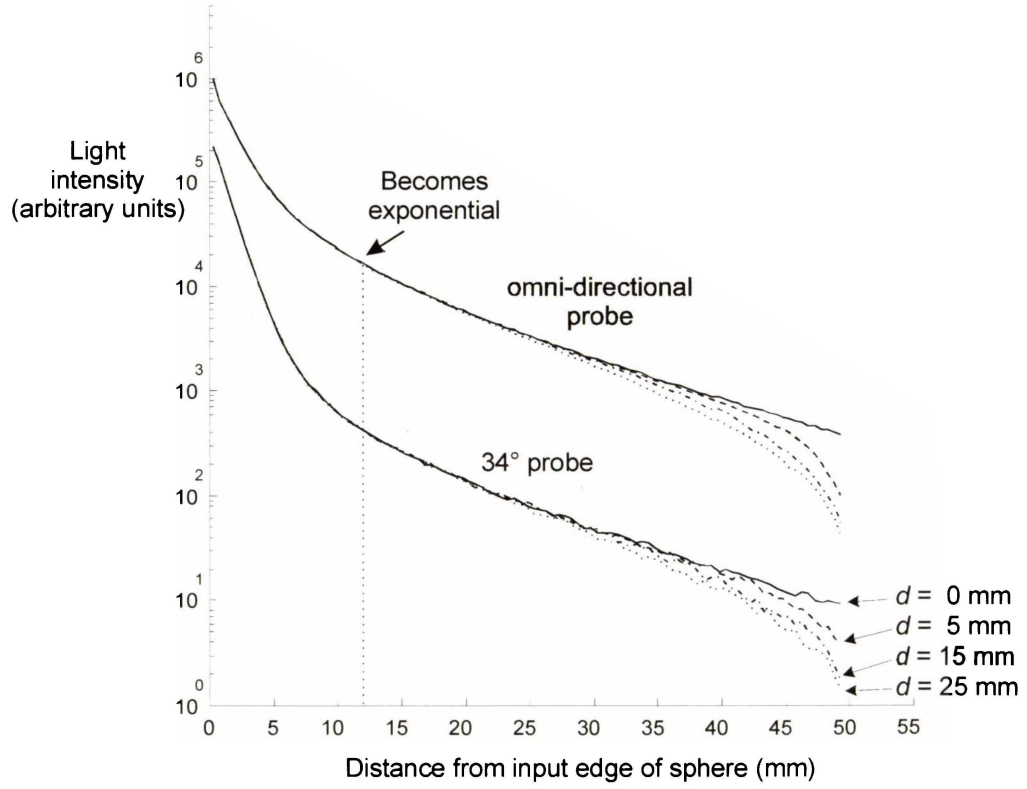

Fig. 4-14: Monte Carlo simulation of light levels in a 50 mm model with reflecting boundary partially removed.

Since the overall shape of the light level curves from the theoretical model and experimental result are similar, some confidence can be given to the whole fruit absorption and scattering constants used in the model. The mandarin whole fruit absorption coefficient could conceivably be higher, say 0.003 mm^{-1} as has been measured for peaches (Cubeddu et al., 2001). When a model with higher absorbance is simulated, a steeper decay for the exponential region is recorded, and the distance from the source to where the power-law drop off becomes exponential, reduces in turn decreasing the similarity to the experimentally measured shape. The reduction of light in the model far from the source as skin is removed appears to be similar to the result recorded in Fig. 4-13 and supports the theory that the reflective properties of the skin cause the elevated light levels in the mandarin.

In the proximal region of the fruit model (i.e., < 18 mm) where light is anisotropic, the forward flux drops off more rapidly than the total light so the two sets of curves in Fig. 4-14 are different. However, in the distal region the light becomes very nearly isotropic, and the two curve sets are similar apart from the scaling and consequent noise enhancing effects of the probe that only detects over a 34° cone. Wilson et al. (1984) found similar results, and published a detailed study of the forward, backwards and sideways fluxes as measured by a fibre probe in mammalian tissue.

The experimental data observed in Fig. 4-13 shows that the initial steep drop off in the first half of the fruit reduces to exponential form 10 mm into the fruit (x = 13 mm). This is comparable to that observed in the modelled curves, both changing to exponential form around 12 mm into the model. For a semi-infinite tissue model with the same scattering and absorption parameters used in our mandarin model we would expect to reach the isotropic region at the $\frac{1}{e}$ depth δ, for which equation 1 gives a distance 18.2 mm. The difference is that we have imposed a reflecting boundary on the model rather than letting the photons escape, which causes significant deviation from what would be expected for the semi-infinite case.

When the skin is removed from the mandarin there will still be some internal reflection at least in part due to the refractive index mismatch from the flesh to air. The simulations remove the reflective boundary completely (i.e., 0% reflectance),

allowing the photons to easily escape. This may explain the rolling drop off observed in the simulated light level curves (Fig. 4-14) at the furthest distance from the source, which is not seen in the experimental data (Fig. 4-13). Repeating the simulations leaving 33% reflectance did partly reduce the roll off, lifting the light intensity at the edge of the model.

Our mandarin model shows the same external response as the real fruit that were measured, but the detail within the fruit may well be more complex. If the relative influence of the different regions in a complex fruit on the external spectroscopic measurements is to be assessed theoretically the inner detail may be important. To improve the simulation of internal light level changes for mandarins, the significant thickness and attenuation of the skin and the influence of the core need to be taken into account by a more complex model. The scattering and absorption constants for each region would need to be estimated and if internal interfaces are to be simulated. Refractive index estimates will also be required. It should be emphasized that the large value for δ means that boundary conditions cannot be ignored for modelling mandarins.

4.3 Summary

A probe has been developed to measure the light levels inside a fruit, which enables measurements close to the illuminated surface. The probe makes only a small puncture hole so it has only a minimal effect on the light levels being measured. The technique has considerable advantages over alternative procedures in the literature where large sections of the fruit are cut away or optical boundaries are otherwise altered.

The light levels inside mandarin, apple, kiwifruit and nashi have been measured directly and with a spatial resolution better than 1.5 mm. This has not been reported previously. Experimental rates of light extinction are reported for the different fruit types at 808 nm.

We have been able to confirm that the thick skin of the mandarin is highly attenuating. The core and segments of the mandarin lead to perturbations in the light level curves. The skin of the mandarin appears to create a significant internal reflection leading to elevated light level in the flesh of the mandarin.

For apples, our depth of penetration measurements at different wavelengths are quite different to those of Lammertyn et al. (2000) and suggest very poor transmission above 1350 nm. Admittedly their results are in terms of a maximum penetration depth for influencing diffuse reflectance measurements. However, if the light cannot be transmitted to a certain depth then it most certainly cannot get back. Indeed, following that logic it may be reasonable to halve our measured penetration depths and use them as the equivalent penetration depths for diffuse reflectance measurements. This gives a 1% depth of 17.5 mm at 713 nm and considerably less than 1 mm at 1400-1580 nm region. Diffuse reflectance measurements in this region of low light penetration will only hold optical information relating to the very near surface of the apple. For NIR transmission where the objective is to sample the internal tissue, the greatest signal to noise ratio will be achieved by using the wavelengths that are not strongly absorbed by water, specifically those in the 'diagnostic window' at 700-900 nm.

Single constants describing scattering and absorption for the whole fruit similar to those obtained by Cubeddu et al. (2001) allow the computation of an overall light distribution shape for a homogeneous substance, which has the measured μ_a and μ_s' coefficients. We have shown that a Monte Carlo simulation, using such constants, can approximate the experimental measurements made by the probe. The simulations and calculations showed that the boundary conditions cannot be ignored for modelling mandarins, and therefore measurements of light levels in mandarins should avoid altering the boundaries.

5

Summary and Conclusions

5.1 Overview

This thesis has advanced the field of NIR spectroscopy for the assessment of internal fruit properties. Significant contributions include:

- The analysis of moving as opposed to static samples of fruit.
- comparison of model performances at different angles.
- presentation of findings which demonstrate overfitting factors to interpret NIR spectra leads to higher mean variance and less confidence in the means.
- The development of a probe which renders obsolete the common practice of slicing to determine the light level in various portions of a fruit
- simulated and actual measurements which showed that light within fruits such as mandarin is internally reflected from the skin, while core and segments modify light level patterns,
- The measurement of penetration depths in fruit for different wavelengths.

This study has provided useful information for implementing practical NIR fruit graders. The study of reflectance and transmission spectra acquired continuously from moving fruit gives a better understanding of the timings and practical implications of on-line grading. Visible-NIR light levels in fruit have been measured directly and validated with Monte Carlo photon tracing simulations. By measuring or computing where the light goes inside the fruit we are better able to

understand which regions of the fruit are being sampled with different wavelengths and therefore how grading equipment might be improved or designed for both constituent assessment and defect detection.

5.2 Detail

For the NIR assessment of kiwifruit dry matter three optical modes at different angles are possible. It was found that transmission mode was better than interactance mode, which was better than reflectance mode. For soluble solids content transmission was similar in performance to interactance mode, and both were better than reflectance. The best prediction on the full set of ~780 kiwifruit gave an RMSEP for DM of 0.45 % (R_P^2 = 0.90) using 180° transmission.

The longitudinal variation in DM across a kiwifruit only slightly increased (< 1% DM) towards the stem end. It was found that fruit quality could be best predicted using 90° transmission if the fruit were presented to the detector in horizontal and vertical orientations. A model for randomly oriented fruit was built, and the optimal orientation was at 150° using transmission, which gave an RMSEP for DM of 0.58 % (R_P^2 = 0.86). This is very promising for a full transmission grader implementation; it means that the grader can be trained with fruit at different orientations to produce a robust model that can predict the fruit properties no matter what orientation is presented.

NIR spectroscopy was used to assess acidity in mandarins. The results suggest that the acid is not being estimated as a direct consequence of the acid absorption in the spectra. Rather the acid content is inferred from the absorption effects of other components (such as chlorophyll) in the fruit, which are strongly correlated to acid content via maturity and other relationships. Using only wavelengths that respond to skin colour and chlorophyll content, almost the same prediction error (RMSEP = 0.177 ± 0.010 %, R_P^2 = 0.69) is achieved as that for the full wavelength range using 5 PLS factors. Interactance mode performed slightly better, with acid RMSEP = 0.148 ± 0.004 % with R_P^2 = 0.78. Transmission mode was best able to determine SSC giving RMSEP = 0.39 ± 0.02 °Brix (R_P^2 = 0.89). It was found that varying the angle of transmission had little effect on the ability to predict from the spectra, which suggests that mandarins could be assessed at different fruit orientations without detriment.

By continuously measuring transmission or reflectance spectra from apples moving at different speeds (up to 3 fruit/s for reflectance) the optimal positions for assessment were ascertained. At the best position, transmission achieved RMSEP 0.54 ± 0.07 °Brix ($R_P^2 = 0.65$) for SSC and 0.66 ± 0.08 % ($R_P^2 = 0.80$) for DM. For reflectance the predictions could be improved by averaging selected adjacent spectra. The lowest RMSEP resulting from this averaging was 0.69 ± 0.08 °Brix ($R_P^2 = 0.64$) for SSC and 0.80 ± 0.08 % ($R_P^2 = 0.72$) for DM. These results are better than those reported in other studies of moving fruit where an RMSEP value has been given. In spite of the grading speed being more than 10 times slower, the transmission measurements did not give significantly better results for apples than the reflectance mode. Because the stationary results showed improvements when moving from reflectance to transmission mode, further equipment refinement (such as brighter light source while shielding fruit from unwanted wavelengths, larger dynamic range/sensitivity for the detector, superior shutter control, reduced vibration, custom designed spectrometer optimising wavelengths) may yet show this trend for moving fruit, but with the existing system there seems no advantage in using transmission for DM or SSC assessment of moving apples.

A fibre optic probe has been developed and used to measure the light distributions in fruit with minimal effect on the light being measured. Measurements of the depth of light penetration on apples confirm the presence of a diagnostic window in the range 700-900 nm for which the 1% penetration depth was more than 25 mm, reaching a peak of 35 mm at 713 nm. At higher wavelengths the penetration is greatly reduced due to the increased effect of water absorption and reaches less than 1 mm in the range 1400-1580 nm. This range is often used for improved information contend due to the narrower absorption bands and yet it is not capable of sampling flesh greater than 1 mm into the fruit at this level of sensitivity. The fact that the diagnostic window coincides with the most sensitive region of relatively cheap silicon detectors (300-1050 nm) makes this wavelength region attractive for NIR assessment of fruit quality.

The light extinction curves for different 'intact' fruit have been measured at 808 nm. The presence of an initial diffusion region followed by an exponential decay is a result that has been predicted but not experimentally shown for intact fruit until now. The exponential decay ranged from 0.014 mm^{-1} ($\delta = 24$ mm) over

one side of the mandarin to 0.241 mm^{-1} (δ = 4.2 mm) over the inner pericarp of a kiwifruit. The fact that the light has significant interaction with the deep tissue supports the argument that fruit grading using transmission mode will be more successful for quality selection and internal disorder detection than reflectance.

The light distribution over an equatorial slice of a mandarin and an apple has been measured and the influence of the internal components investigated. A further study showed the mandarin peel to be internally reflective. Removing pieces of skin from the mandarin reduced the light level in the fruit in the corresponding side of the fruit. This result emphasises how important it is not to alter the optical boundaries (for example by cutting away sections of the fruit) while studying the light inside fruit. Cutting away sections or removing the skin can significantly alter the internal reflections and/or backscattering properties.

Single constants describing scattering and absorption for the whole fruit allow the computation of an overall light distribution shape for a homogeneous substance, which has the measured μ_a and μ_s' coefficients. We have shown that a Monte Carlo simulation, using such constants, can approximate the experimental measurements made by the probe. In the case of a mandarin the experimental result was adequately imitated using $\mu_a = 0.001$ mm^{-1}, $\mu_s' = 1$ mm^{-1}, and a 90% reflective boundary. The simulations and calculations showed that well-defined boundary conditions are necessary for modelling mandarins.

5.3 Future outlook

Significant challenges still exist for implementing on-line NIR grading of fruit. There is potential to improve existing systems for automatically oriented fruit and spectrum acquisition timing and locations on moving fruit. Future graders are likely to move towards transmission grading as shown by the trend in Japan. This move will facilitate internal defect detection, which will be a crucial selling point of the technology with the potential to save whole lines of an affected crop. When the systems are proven reliable there will be a rapid uptake of the technology.

Appendix

A.1 Monte Carlo Software: Overview and Structure

The simulated light distributions within fruit reported in this thesis were generated from custom written software. A copy of the code is included on the CD that is attached to this thesis. The program was written in Matlab (The Mathworks Inc., Natick, MA, USA). Therefore, the user would need a working installation of Matlab (version 5.0 or later) to run the simulations. The function library "MC Modelling/matlabscripts/" should be added to the Matlab path.

The overall structure of the software is given in Fig. A-1. The user specifies model input using the shell interface function, which utilises the core simulation code and a function library to generate the light distribution.

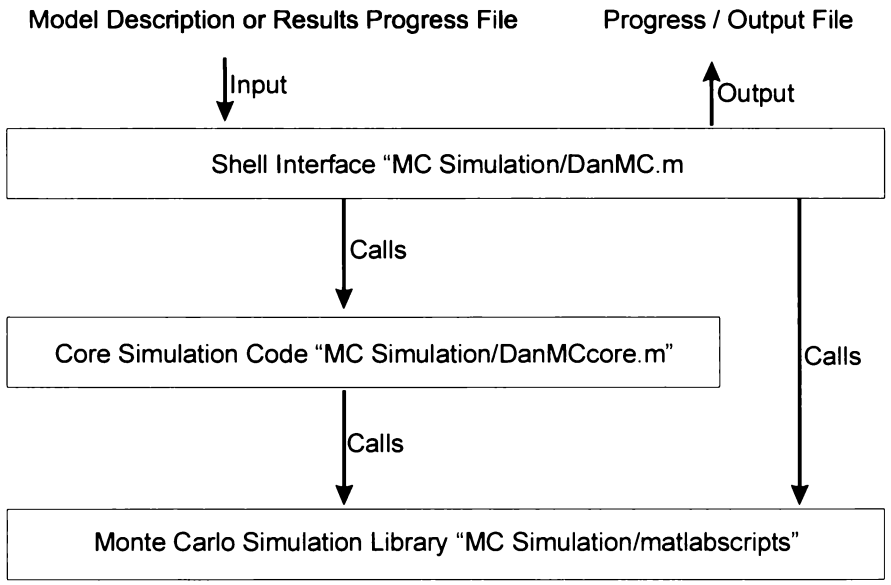

Fig. A-1: Structure of the MC simulation software.

The model to be simulated is specified in a text file which must conform to the format demonstrated in "MC Modelling/DefaultDMCModel.txt". The parameters specified must use compatible units. The spreadsheet: "MC Modelling/Extra/Diffusion equations.xls" may help a new user understand the unit conversions. If a model file is not specified default values are used as specified in "MC Modelling/DefaultDMCModel.txt". To launch the simulation use: DanMC('Model File Name.txt')

As the simulation is being run (which can take many hours) cumulative output is stored in Matlab arrays and saved at regular intervals as results files (extension *.dmc). This provides progress indication and prevents loss of data in the event of a failure. The final output file name will be of the form: "MC MachineName 9999999 dd-mon-yyyy hhmm.dmc" where MachineName is the name of the machine (computer) the simulation was run on, 999999 is a code corresponding to the model parameters used, and the date and time the simulation completed is appended at the end. If a simulation is interrupted before completion it can be resumed using: DanMC('Progress file name.dmc').

When a simulation has been completed it can be examined using "examineDMC2('resultfilename.dmc',1)" leaving off the last parameter prevents graphs being displayed. Simulation result files can also be combined (which enables sharing a calculation of a model on multiple computers) by using the function combineDMC('file1.dmc','file2.dmc'['joined.dmc']).

The directory "MC Modelling/ Mandarin Simulation" contains the models and results files for the simulations performed on mandarins, including the pealing experiment reported in this thesis.

A.2 Kiwifruit Trail Sampling Regime

The kiwifruit trail described in section 3.3 uses a large amount of fruit which are difficult to classify. In order to help visualise the sets used in each of the result graphs and tables figure A-2 maps out the described sets and the corresponding result graphs.

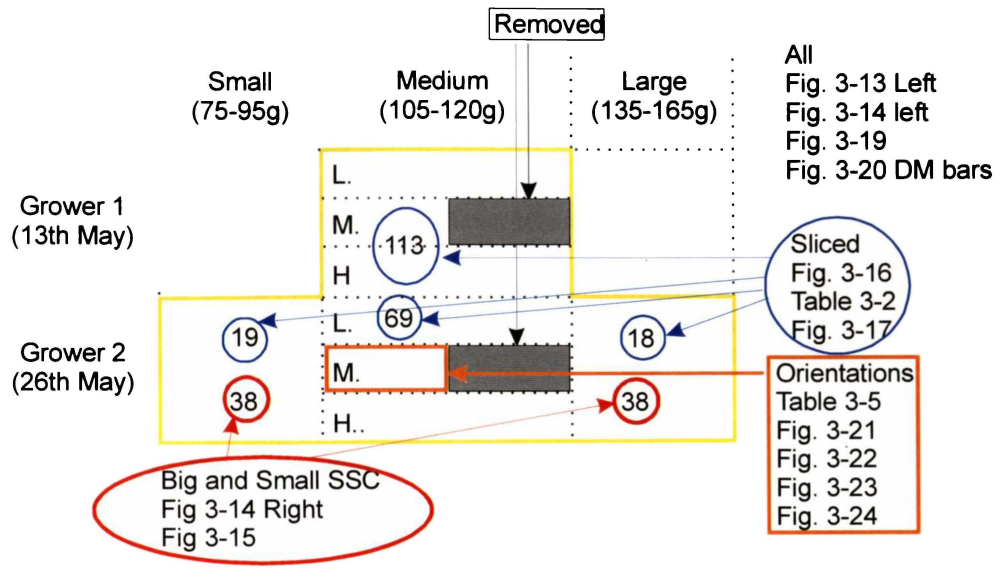

Fig. A-2: Map of the kiwifruit trial sample sets (L=Low density, M=Medium density, H = High density).

References

Anderson, R.R., Beck, H., Bruggemann, U., Farinell, W., Jacques, S.L., Parrish, J.A., 1989. Pulsed photothermal radiometry in turbid media: internal reflection of backscattered radiation strongly influences optical dosimetry. Applied Optics, 28 (12), 2256-2267.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Applied Spectroscopy, 43 (5), 722-777.

Bartholomew, E.T., Sinclair, W.B., 1941. Unequal distribution of soluble solids in pulp of citrus fruits. Plant Physiology, 16, 293-312.

Beever, D.J., Hopkirk, G., 1990. Fruit development and fruit physiology. In: Warrington, I.J. and Weston, G.C. (Eds.), Kiwifruit science and management, Ray Richards Publisher, New Zealand, pp. 97-126.

Bellon, V., Sévila, F., 1993. Optimization of a non-destructive system for on-line infrared measurement of fruit internal quality. In: IV international symposium on fruit, nut, and vegetable production engineering, March 22-26, Valencia-Zaragoza, Spain, 162-163.

Bellon, V., Vigneau, J.L., Leclercq, M., 1993. Feasibility and performances of a new, multiplexed, fast and low-cost fibre-optic NIR spectrometer for the on-line measurement of sugar in fruits. Applied Spectroscopy, 47 (7), 1079-1083.

Bellon-Maurel, V., Steinmetz, V., Dusserre-Bresson, L., Jacques, J.C., Rannou, G., 1997. Real-time NIR sensors to sort fruit and vegetables according to their sugar content. In: International symposium on fruit nut and vegetable production engineering 5, Sept. 3-10 1997, Davis, USA.

Birth G.S., Hecht, H.G., 1987. In: William, P. and Norris, K. (Eds.), Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists Inc., USA, pp. 1-15.

Birth G.S., Dull, G.G., Magee, J.B., Chan, H.T., Cavaletto, C.G., 1984. An optical method for estimating papaya maturity. Journal of American Society Horticultural Science, 109 (1), 62-66.

Birth, G.D., Olsen, K.L., 1964. Proceedings of the American Society Horticultural Science, 85, p. 74.

Birth, G.S., 1978. The light scattering properties of food. Journal of Food Science, 43, 916-925.

Bittner, D.R., Norris, K. H., 1976. Optical properties of selected fruit vs maturity, Quality detection in foods. American Society of Agricultural Engineers, 57-59.

Budiastra, I.W., Ikeda, Y., Nishizu, T., 1998. Optical methods for quality evaluation of fruits (Part1) - Optical properties of selected fruits using the Kubelka-Munk theory and their relationships with fruit maturity and sugar content. Journal of the Japanese Society of Agricultural Machinery, 60 (2), 117-128.

Chen, P., Nattuvetty, V.R, 1980. Light transmittance through a region of an intact fruit. Transactions of the American Society of Agricultural Engineers, 23 (2), 519-522.

Chuma, Y., Kawano, S., Sein, K., 1976. Optical properties of fruits to serve the automatic selection in the packinghouse line (2) Light reflectance of Satsuma orange, Journal of the Japanese society of agricultural machinery, 37 (4), 587-592.

Chuma, Y., Sein, K., Iwamoto, M., 1974. Optical properties of fruits and their utilization in automatic sorting facility (1) suitable wave length to detect the peel colour and maturity of citrus unshiu, Journal of the Japanese society of agricultural machinery, 35 (4), 416-423.

Clark, C., McGlone, V.A., Jordan, R.B., 2001. Use of transmission near infra-red (NIR) to detect brownheart in 'Braeburn' apple. In: Proceedings of the Australasian postharvest conference, Sept. 23-27, Adelaide, Australia (submitted).

Cubeddu, R., D'Andrea, C., Pifferi, A., Taroni, P., Torricelli, A., Valentini, G., 2001. Non-destructive quantification of chemical and physical properties of fruits by time-resolved reflectance spectroscopy in the wavelength range 650-1000 nm. Applied Optics, 40 (2), 538-543.

Davies, A.M.C., 2000. William Herschel and the discovery of near infrared energy. NIR news, 11 (2), 3-5.

Dull, G.G, Leffler, R.G., Birth, G.S., Smittle, D.A., 1992. Instrument for nondestructive measurement of soluble solids in Honeydew melons. Transactions of the American Society of Agricultural Engineers, 35 (2), 735-737.

Dull, G.G., Birth, G.S., Smittle, D.A., Leffler, R.G., 1989. Journal of Food Science, 54, p. 393.

Dull, G.G., Leffler, R.G., Birth, G.S., Zaltzman, A., Schmilovitch, Z., Russell, R.B., 1991. The near infrared determination of moisture in whole dates. HortScience, 26 (10), 1303-1305.

Earp, R.W., 1990. Export and marketing of New Zealand kiwifruit. In: Warrington, I.J. and Weston, G.C. (Eds.), Kiwifruit science and management. Ray Richards Publisher, New Zealand, pp. 485-510.

Eilert, A.J., Sweat, J.A., Wetzel, D.L., 2000. Parabolic concentration of diffusely transmitted near infrared radiation in an acousto-optic tunable filter spectrometer. Journal Near Infrared Spectroscopy, 8, 239-250.

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

Ferguson A., 1990. The genus Actinidia. In: Warrington, I., and Weston, G.C. (Eds.), Kiwifruit science and management. Ray Richards Publisher, New Zealand, pp. 15-35.

Flock, S.T., Paterson, M.S., Wilson, B.C., Wyman, D.R., 1989. Monte Carlo modelling of light propagation in highly scattering tissues I: Model predictions and comparison with diffusion theory. IEEE Transactions on Biomedical Engineering, 36 (12), 1162-1168.

Fraser, D.G, Jordan, R.B., Künnemeyer, R., McGlone, V.A., Near infrared light distribution in a mandarin. Journal of Postharvest Biology and Technology (Submitted July 2001).

Fraser, D.G., Künnemeyer, R., Jordan R.B., McGlone V.A., 2001b. A new method for mapping visible-near infrared light levels in fruit. In: 10th International conference on near-infrared spectroscopy, 10-15th June 2001, Korea (in press).

Fraser, D.G., McGlone V.A., Jordan R.B., Künnemeyer, R., 2001a. NIR (Near Infra-Red) light penetration into an apple. Postharvest Biology and Technology, 22 (3), 191-194.

Freund, I., 1992. Surface reflections and boundary conditions for diffusive photon transport. Physical review A, 45 (12), 8854-8858.

Fujiwara, T., Honjo, T., 1995. Determination of sugar and acid contents in fruit juice of Satsuma mandarin by near infrared spectroscopy, Nippon-Shokuhin-Kogyo-Gakkaishi. Journal of the Japanese Society for Food Science and Technology, 42, 109-117.

Geladi, P., MacDougall, D., Martens, H., 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. Applied Spectroscopy, 39 (3), p. 491-501.

Glodschmidt E.E., Koch, K.E., 1996. Citrus. In: Zamski, E., Schoffer, A.A. (Eds.), Photoassimilate distribution in plants and crops: source-sink relationships. Marcel Dekker Inc., New York, pp. 797-823.

Graaff, R., Koelink, M.H., de Mul, F.F., Zijlstra, W.G., Dassel, A.C.M., Aarnoudse, J.G., 1993. Condensed Monte Carlo simulations for the description of light transport. Applied Optics, 32 (4), 426-434.

Greensill, C.V., Newman, D.D., 1999. An investigation into the determination of the maturity of pawpaws (Carica papaya) from NIR transmission spectra. Journal Near Infrared Spectroscopy, 7, 109-116.

Greensill, C.V., Walsh, K.B., 2000. A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit. Measurement Science and Technology, 11, 1674-1684.

Groenhuis, R.A.J., Ferwerda, H.A., Ten Bosch, J.J., 1983. Scattering and absorption of turbid materials determined from reflection measurements. 1: theory. Applied Optics, 22 (16), 2456-2462.

Guthrie, J., Walsh, K., 1999. Influence of environmental and instrumental variables on the non-invasive prediction of brix in pineapple using near infrared spectroscopy. Australian journal of experimental agriculture, 39 (1), 73-80.

Hallett, I., MacRae, E.A., Wegrzyn, T.F., 1992. Changes in kiwifruit cell wall ultrastructure and cell packing during postharvest ripening. International Journal Plant Science, 153 (1), 49-60.

Harris, C.M., Couey, H.M., Harvey, J.M., 1972. Effect of harvest date, storage period, and ripening time on the quality of Chinese gooseberries. Marketing research report 940, Agricultural Research Services, U.S. Department of Agriculture, Washington, D.C., U.S.A., p. 10.

Harty A.R., 1995. Satsuma mandarin production in New Zealand. In: HortResearch Satsuma and mandarin workshop, November 1995, New Zealand.

Henyey L.C., Greenstein, J.L., 1941. Diffuse radiation in the galaxy. Astrophysics Journal, 93, 70-83.

Holler, F., Burns, D.H., Callis, J.B., 1989. Direct use of second derivatives in curve-fitting procedures. Applied Spectroscopy, 43 (5), 877-883.

Hother, K., Herold, B., Geyer, M., 1995. Grenzen der Erkennung von Qualitätsfehlern im Apfelgewebe bei Messung der spektralen Reflexion. Gartenbauwissenschaft, 60 (4), 162-166 (in German).

Hwang, I., Noh, S., 2000. Preliminary study for the development of an algorithm for the on-line analysis of the sugar content of intact fruits using near infrared spectroscopy. In: Davies, A.M.C., Giangiacomo, R. (Eds.), 9th International conference on near-infrared spectroscopy, June 1999, Verona, Italy, 379-384.

Jacques, S.L., 1989. Simple theory, measurements, and rules of thumb for dosimetry during photodynamic therapy. In: Proceedings of the SPIE, photodynamic therapy: mechanisms, 1989, 100-108.

Jacques, S.L., 1998a. A minimal Monte Carlo program for simulating steady-state light distributions in tissues. http://omlc.ogi.edu/news/dec98/mc321/index.html (accessed 23 February 2001).

Jacques, S.L., 1998b. Light distributions from point, line, and plane sources for photochemical reactions and fluorescence in turbid biological tissues. Photochemistry and Photobiology, 67 (1), 23-32.

Jacques, S.L., 1998c. Diffuse reflectance from a semiinfinite medium, http://omlc.ogi.edu/news/may99/rd/index.html (accessed September 2001).

Jordan, R.B. et. al., 1996. Kiwifruit density as an Indicator of consumer acceptability, Confidential Client Report No 96/44, HortResearch, Filed March 1996.

Jordan, R.B., Walton, E.F., Klages, K.U., Seelye, R.J., 2000. Postharvest fruit density as an indicator of dry matter and ripened soluble solids of kiwifruit. Postharvest Biology and Technology, 163-173.

Kawano, S., 1994. Non-destructive NIR quality evaluation of fruit and vegetables in Japan. NIR News, 5 (6), 10-12.

Kawano, S., 1998. New application of nondestructive methods for quality evaluation of fruits and vegetables in Japan. Journal Japanese Society Horticultural Science, 67 (6), 1176-1179.

Kawano, S., Fujiwara, T., Iwamoto M.J., 1993. Nondestructive determination of sugar content in Satsuma mandarin using near infrared (NIR) transmittance. Journal Japanese Society Horticultural Science, 62 (2), 465-470.

Kawano, S., Watanabe, H., Iwamoto, M., 1992. Determination of sugar content in intact peaches by near infrared spectroscopy with fiber optics in interactance mode. Journal Japanese Society of Horticultural Science, 61 (2), 445-451.

Kays, S.J., Dull, G.G., Leffler, R.G., 2000. Challenges and opportunities in the use of near infrared for the analysis of intact, high moisture plant products. In: Davies, A.M.C., Giangiacomo, R. (Eds.), 9th International conference on near-infrared spectroscopy, June 1999, Verona, Italy, 841-847.

Kerr J.P., Aitken, A.G., 2001. Horticulture facts and figures 2000. Martech Consulting Group Ltd., New Zealand.

Koch, K.E., Avigne, W.T., 1990. Postphloem, nonvascular transfer in citrus: kinetics, metabolism, and sugar gradients. Plant Physiology, 93, 859-866.

Kortüm, G., 1969. Reflectance spectroscopy: principles, methods, applications. Springer-Verlag, Berlin-Heidelberg, Germany.

Kramer, R., 1998. Chemometric techniques for quantitative analysis. Marcel Dekker Inc., New York, USA.

Lammertyn, J., Nicolaï, B., Ooms, K., DeSedt, V., De Baerdemaeker, J., 1998. Non-destructive measurement of acidity, soluble solids, and firmness of Jonagold apples using NIR-spectroscopy., Transactions of the American Society of Agricultural Engineers, 41 (4), 1089-1094.

Lammertyn, J., Peirs, A., De Baerdemaeker, J., Nicolaï, B., 2000. Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. Postharvest Biology and Technology, 18, 121-132.

Lammertyn, J., Peirs, A., De Baerdemaeker, J., Nicolaï, B., 2000. Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. Postharvest Biology and Technology, 18, 121-132.

Lide, D.R. (Ed.), 1998. Handbook of chemistry and physics 79th edition. CRC Press Ltd., USA, p. 8-62 & p. 8-81.

Martens, H., Næs, T., 1989. Multivariate calibration. John Wiley and Sons, New York, USA.

Martinsen, P., 1999. PhD Thesis: quantitative near-infrared imaging spectroscopy of fruit. The University of Auckland, New Zealand.

Massie D.R., Norris, K.H., 1965. Spectral reflectance and transmittance properties of grain in the visible and near infrared. Transactions of the American Society of Agricultural Engineers, 8 (4), 598-600.

McClure, W.F., 1994. Near-infrared spectroscopy. In: Wilson, R.H. (Ed), Spectroscopic techniques for food analysis. VCH Publishers Inc., New York, USA, pp. 15-57.

McGlone, V.A., Kawano, S., 1998. Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. Journal of Postharvest Biology and Technology, 13, 131-141.

McGlone, V.A., 2001. Technology Development Group, HortResearch, Hamilton, New Zealand. Personal Communication.

McGlone, V.A., Jordan, R.B., Martinsen, P.J., 2001b. NIR estimation at harvest of pre- and post-storage quality indices of apples. Journal of Postharvest Biology and Technology (accepted).

McGlone, V.A., Jordan, R.B., O'Donnell, K., 1997a. Internal report (97/74): NIR technology for fruit inspection, HortResearch, New Zealand.

McGlone, V.A., Jordan, R.B., Schaare, P., 1997b. Anomalous firmness changes in cool-stored kiwifruit. Journal of Postharvest Biology and Technology, 12, 147-156.

McGlone, V.A., Jordan, R.B., Seelye, R., Martinsen, P., 2001a. Comparing density and NIR analysis for the measurement of kiwifruit dry matter and soluble solids content. Journal of Postharvest Biology and Technology (accepted).

Miyamoto, K., Kawauchi, M., Fukuda, T., 1998. Classification of high acid fruits by PLS using the near infrared transmittance (NIT) spectra of intact Satsuma mandarins. Journal Near Infrared Spectroscopy, 6 (1-4), 267-271.

Miyamoto, K., Kitano, Y., 1995. Non-destructive determination of sugar content in Satsuma mandarin fruit by near infrared transmittance spectroscopy. Journal of Near Infrared Spectroscopy, 3, 227-237.

Mohsenin, N.N., 1986. Physical properties of plant and animal materials: structure, physical characteristics and mechanical properties 2nd edition. Gordon and Breach Science Publishers, New York, USA.

Moons, E., Dardenne, P., 2000. Determination of internal apple quality by non-destructive visible and near infrared spectroscopy. In: Davies, A.M.C., Giangiacomo, R. (Eds.), 9th International conference on near-infrared spectroscopy, June 1999, Verona, Italy, 785-789.

Moons, E., Dardenne, P., Dubois, A., Sindic, M., 1997. Nondestructive visible and NIR spectroscopy for determination of internal quality in apple. In: Proceedings from the sensors for non-destructive testing international conference and , 122-132.

Morton, J., 1987. Mandarin Orange. In: Morton, J.F. (Ed.), Fruits of warm climates. Miami, FL, USA. pp. 142-145.

Mowat, A.D., Poole P.R., 1997a. Use of visible-near infrared diffuse reflectance spectroscopy to discriminate between kiwifruit with properties altered by preharvest treatments. Journal Near Infrared Spectroscopy, 5, 113-122.

Mowat, A.D., Poole, P.R., 1997b. Non-destructive discrimination of persimmon fruit quality using visible-near infrared reflectance spectrophotometry, Acta Horticulturae, 436, 159-163.

Norris, K.H., Hart, J.R., 1965. Direct spectrophotometric determination of moisture content of grain and seeds. Principles and methods of measuring moisture content in liquids and solids, 4, 19-25.

O'Rourke, D., 2000. World kiwifruit review-2000. Belrose Inc., Pullman, USA.

Osborne, S.D., Künnemeyer, R., Jordan, R.B., 1999. A low-cost system for the grading of kiwifruit. Journal of Near Infrared Spectroscopy, 7 (1), 7-15.

Osborne, S.D., Jordan, R.B., Künnemeyer, R., 1997. Method of wavelength selection for partial least squares. Analyst, 122, 1531-1537.

Osbourne, B.G, Fearn, T., Hindle, P.H., 1993. Practical NIR spectroscopy: with applications in food and beverage analysis 2nd edition. Longman Scientific and Technical, Essex, England.

Peiris, K.H.S., Dull, G.G., Leffler, R.G., Kays, S.J., 1999a. Spatial variability 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 (1), 114-118.

Peiris, K.H.S., Dull, G.G., Leffler, R.G., Kays, S.J., 1998a. Near-infrared spectrometric method for nondestructive determination of soluble solids content in peaches. Journal American Society Horticultural Science, 123 (5), 898-905.

Peiris, K.H.S., Dull, G.G., Leffler, R.G., Kays, S.J., 1999b. Rapid, nondestructive method for determination of processed soluble solids in intact unprocessed tomato fruit using near infrared spectroscopy. In: Bieche, B.J. (Ed.), Proceedings 6th international ISHS Symposium on the processing tomato, 413-418.

Peiris, K.H.S., Dull, R.G., Leffler, R.G., Kays, S.J., 1998b. Near-infrared (NIR) spectrometric technique for nondestructive determination of soluble solids content in processing tomatoes. Journal of American Society Horticultural Science, 123 (6), 1089-1093.

Prahl, S., 2001. Optical absorption of water. http://www.omlc.ogi.edu/spectra/water/ (accessed 16th September 2001).

Profio, A.E., 1979. Radiation shielding and dosimetry. John Wiley and Sons, New York, 168-199.

Roger, J.M., Bellon-Maurel, V., 2000. Using genetic algorithms to select wavelengths in near-infrared spectra: application to sugar content prediction in cherries. Applied Spectroscopy, 54 (9), 1313-1320.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 36 (8), p. 1627-1639.

Schaare, P.N., Fraser D.G., 2000. Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit (Actinidia chinensis). Journal of Postharvest Biology and Technology, 20, 175-184.

Seelye, R., Jordan, R.B., Schaare, P.N., Cate, L.R., 1997. A method of calculating the density of kiwifruit, NZ Patent 314439.

Segelstein, D.J., 1981. M.S. Thesis: the complex refractive index of water. University of Missouri-Kansas City, USA.

Skoog, D.A., Holler, F.J., West, D.M., 1990. Analytical chemistry an introduction. 5e, Saunders College Publishing, Philadelphia, USA.

Slaughter, D.C., 1995. Nondestructive determination of internal quality in peaches and nectarines. Transactions of the ASAE., 38 (2), p. 617-623.

Slaughter, D.C., Cavaletto, C.G., Gautz, L.D., Paull, R.E., 1999. Non-destructive determination of soluble solids in papayas using near infrared spectroscopy. Journal of Near Infrared Spectroscopy, 7 (4), p. 223-228.

Slaughter, D.C., Crisosto, C.H., 1996. Non-destructive internal quality assessment of kiwifruit using near infrared spectroscopy. In: 8th International diffuse reflectance conference.

Snee, R.D., 1976. Validation of regression models: Methods and examples. Technometrics, 19, 415-428.

Sohn, M., Cho, R., 2000. Non-destructive determination of free acid and malic acid content in apples using near infrared spectroscopy. In: Davies, A.M.C., Giangiacomo, R. (Eds.), 9th International conference on near-infrared spectroscopy, June 1999, Verona, Italy, 797-800.

Srahl, S., 2001. http://www.omlc.ogi.edu/spectra/water/ (accessed 15 September 2001).

Steinmetz, V., Biavati, E., 1998. Predicting the maturity of oranges with non destructive sensors. In: Proceedings second international symposium on sensors in horticulture, Acta Horticulturae, 421, p271-278.

Stone, M., 1974. Cross-validatory choice and assessment of statistical precision. Journal Royal Statistical Society B, 111-133.

Thomas, Corden, 1970. Tables of composition of Australian foods. Canberra, Australian Government Publishing Service, 15-17.

Throop, J.A., Aneshansley, D.J., Upchurch, B.L., Anger, B., 2001. Apple orientation on two conveyors: Performance and predictability based on fruit shape characteristics. Transactions of the American Society of Agricultural Engineers, 44 (1), 99-109.

Ting, S.V., 1969. Distribution of soluble components and quality factors in the edible portion of citrus fruits. American Society of Horticultural Science, 94, 515-519.

Tuchin, V., 2000. Tissue optics: light scattering methods and instruments for medical diagnosis. SPIE Bellingham, Washington, USA.

Upadhyaya, S.K., O'Brien, M.O., 1989. A 3-D light diffusion model for biological materials. In: Dodd & Grace (Eds.), Land and water use. Balkema, Rotterdam, pp. 2471-2477.

Upchurch, B.L, Throop, J.A, Aneshansley D.J., 1997. Detecting internal breakdown in apples using interactance measurements. Journal of Postharvest Biology and Technology, 10, 15-19.

Wang, L.H., Jacques, S.L., Zheng, L.Q., 1995. MCML Monte Carlo modeling of photon transport in multi-layered tissues. Computer Methods and Programs in Biomedicine, 47, 131-146.

Welford W.T., Winston, R., 1979. High collection nonimaging optics. Academic Press, New York, USA.

Williams, P., Norris, K., 1987. Qualitative applications of near-infrared reflectance spectroscopy. In: Williams, P., Norris, K. (Eds.), Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists Inc., Minnesota, USA, pp. 241-246.

Wilson B.C., Patterson, M.S., 1986. The physics of photodynamic therapy. Physics in Medicine and Biology, 31, 327-360.

Wilson, B.C., Adam, G., 1983. A Monte Carlo model for the absorption and flux distributions of light in tissue. American Association of Physics and Medicine, 10 (6), 824-830.

Wilson, B.C., Jacques, S.L., 1990. Optical reflectance and transmittance of tissues: principles and applications. IEEE Journal of Quantum electronics, 26 (12), 2186-2199.

Wilson, B.C., Jeeves, P., Lowe, D.M., Adam, G., 1984. Porphyrin localization and treatment of tumors. In: Doiron, D.R., Gromer, C.L (Eds.), Proceedings of the Clayton Foundation International Symposium on Porphyrin localization and treatment of tumors. April 24-28, Santa Barbara, California, Alan R. Liss Inc., New York, 115-132.

Wold, Herman, 1966. Nonlinear estimation by iterative least square procedures. In: David, F.N. (Ed.), Research papers in statistics. John Wiley & Sons, New York, USA, pp. 411-445.

Young, P.J., 1994. A reformulation of the partial least squares regression algorithm. Siam Journal Scientific Computing, 15 (1), 225-230.

Zhu, J.X., Pine, D.J., Weitz, D.A., 1991. Internal reflection of diffusive light in random media, Physical review A, 44 (6), 3948-3959.