# Non-invasive detection of internal defects in fruit by using visible-shortwave NIR spectroscopy

***Prepared by***

Bed Prasad Khatiwada

M.Sc. (Ag.), Tribhuvan University, Nepal

A thesis submitted for the degree of Doctor of Philosophy

***To***

School of Medical and Applied Sciences
Central Queensland University, Rockhampton, Australia

August 2015

# STATUTORY DECLARATION

I hereby certify that "Non-invasive detection of internal defects in fruit by using visible-shortwave NIR spectroscopy" is a presentation of my original research work and all the sources used were acknowledged by means of complete reference. The material has not been submitted, either in whole or in part, for a degree at this or any other universities.

I acknowledge and certify that I have complied with the rules, requirements, procedures and policy of the Central Queensland University, Australia relating to research higher degree award. I authorize Central Queensland University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Name: Bed Prasad Khatiwada

Signature:

Date: 31st August 2015

# DEDICATION

This work is dedicated to my mother Lila Devi Khatiwada and late father Ghanashyam Khatiwada, for their patience, hard work and dedication to ensure the education of their child despite the hard times of their life.

Bed Prasad Khatiwada, 2015

# ABSTRACT

Non-invasive detection of three internal disorders of fruit of commercial relevance to Queensland horticulture was considered: (i) diffuse browning of apple fruit; (ii) gelling defect of mandarin fruit; and (iii) translucency of pineapple fruit. Visible - short wave near infrared spectroscopy (vis-SWNIRS) is in commercial use for non-invasive field and in line assessment of fruit dry matter and soluble solids content of mango and apple. Some claims exist for commercially available instrumentation for sorting of fruit internal defects, but no assessment of such systems exists in the scientific literature. Four vis-SWNIRS instruments were trialled, varying in optical geometry: (i) the Integrated spectronics's 'Nirvana' handheld instrument, operating with an interactance optical geometry; (ii) a purpose built unit employing a 300W halogen illumination source in a partial transmittance geometry, 'IDD0'; (iii) the MAF Roda Insight2 unit, employing a 150W halogen lamp and operated in a full transmission geometry, and (iv) the MAF Roda IDD2 unit, employing four near infrared light emitting diodes and operated in a full transmission geometry.

A number of reference methods were assessed for scoring level of apple flesh browning, including visual assessment, image analysis (% cross section area affected), chromameter CIE Lab values ($L^*$ and $a^*$ value) and juice $Abs_{420nm}$, of which visual scoring on a 5 point scale was recommended. Chlorophyll fluorescence and acoustic resonant frequency was poorly related to extent of defect, and thus these non-invasive techniques are not recommended. Apple flesh browning was best assessed using visible-shortwave NIRS in a transmission optical geometry, with a typical PLSR model $R^2_{cv}$ = 0.83 and RMSECV = 0.63 (5 point visual scale). Of different binary (good and defect fruit) classification approaches trialled, the best result was achieved using PLS discriminant analysis (PLS-DA) method, followed by linear discriminant analysis. More than 95% of defect fruit were predicted as defect (true negative rate) at the expense of having 10-20% of good fruit falsely predicted as defect fruit (false negative rate), across six populations.

A number of reference methods were also assessed for scoring level of granulation in mandarin fruit, including visual assessment (5 point scale), chromameter CIE Lab (L and colour index values) and % juice recovery, of which visual score and % juice recovery were recommended. Mandarin granulation, indexed by either visual score

or % juice recovery, was best non-invasively assessed using vis-SWNIRS in a transmission optical geometry, with a typical PLSR model $R^2_{cv}$ = 0.74 and RMSEP = 3.6 (% juice recovery). PLS-DAwas able to predict well for good fruit with up to 87 and 100% of good fruit as good fruit (true positive rate) using the IDD0 and MAF Roda Insight2 units, respectively. Defect fruit were wrongly predicted as good fruit (false positive) using both the machine with best result (97 % true negative rate) obtained with PLS-DA using IDD0 unit.

Translucency in pineapple, indexed by either a (5 point) visual score or image analysis was assessed using vis-SWNIR spectroscopy. Typical PLSR calibration results for models developed using the range 700-1000 nm results were modest ($R^2_{cv}$ = 0.58, RMSECV = 0.55 on 5 point scale), and prediction results were poor ($R^2_{p}$ = 0.41, RMSEP = 0.93. For binary classification, PLS-DA was able to predict 98.7% of good fruit as good (true positive rate) while only 34% of defect fruit were predicted defect (true negative rate) based on visual translucent score.

Sorting involves a trade-off between yield and quality. The use of a receiver operating characteristics curve (ROC) and a sorting optimisation curve (SOC) was explored for the comparison of binary classifiers and the optimisation of sorting set point. The need to adjust the sorting set point to maintain a desired quality specification (e.g. % of defect fruit in accepted class) as population mean and spread (SD) for the defect varies is explained. Internal defects of fruit under consideration are well detected and sorted for based under transmission optical geometry with visual defect score as a reference parameter.

# ACKNOWLEDGEMENTS

The journey of this doctoral degree is full of experiences of working with Australian people, culture and their inspiration, thank you everybody for being part of this journey. My sincere thanks and gratitude is to my principal supervisor Prof. Kerry B. Walsh for his knowledge sharing, and providing guidance during research and writing. Thanks for encouragement, enthusiasm and endurance. Thanks also to Dr Phul P. Subedi, my co-supervisor, for his support in laboratory and field works and guiding me to minimize human error during experimentation and in setting a research exercise. I also acknowledge his social help in making me settled in a new country.

I would like to thank Australian Government for providing me an International Postgraduate Research Scholarship (IPRS) and Australian Postgraduate Award (APA), which made me possible to pursue this degree. Help from my friend Umesh Acharya to link me with Non-Invasive Assessment Group (NIAG) paved my path to this opportunity, thank you Umesh. Thanks also to Larry Coleman and Clinton Hayes for your instrumentation and other technical help during the journey. Support from my seniors, colleagues and friends Louis Cunha Junior, Vineela Challagulla, Resham Gautam, Geeta Kafle, Prabin Rijal, Paul Thomas, Kishor Dahal, Surendra Osti, Thakur Bhattarai, Surya Bhattarai, Tanka Prasai, Deepak Poudel and Roshan Subediat different instances is highly appreciated. Great thanks to Ingrid Kennedy from Academic Learning Centre, Central Queensland University for her time on thesis formatting. Libbie Blanchard, Senior Librarian is also acknowledged for her support on Endnote referencing. It would be an injustice if I forget to acknowledge to Ken Mason, Rosie Savio and Robert Vedelago, knowledge-rich farmers of Queensland, Australia for fruit samples and insights into the research issues. Tropical Pines P/L is also acknowledged for fruit samples and MAF RODA Agrobotics, Montauban, France for instrumentation support. I also appreciate the support of administrative and technical staffs of the university. I am highly indebted to my brothers Uddab Khatiwada and Deepak Khatiwada and sister Muiyaand families in Nepal for their continuous encouragement, support and patience. Most importantly, I acknowledge the whole hearted wish and support of my son Swornim and wife Shobha during this journey.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

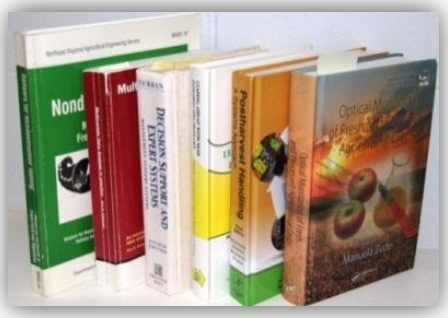
| | |
|---|---|
| Abs | Absorbance |
| ACC | Accuracy |
| ADCC | Analogue to digital converter counts |
| ASTM | American society of testing and materials |
| AUC | Area under curve |
| CA | Controlled atmosphere |
| CCD | Charge coupled device |
| CI | Confidence interval |
| CIE | International commission on illumination |
| $CO_2$ | Carbon dioxide |
| CT | Computed tomography |
| $D^2A$ | Second derivative of absorbance |
| DA | Difference in absorbance |
| DA | Discriminant analysis |
| DAFF | Department of Agriculture, Forest and Fisheries |
| DET | Detection error trade off |
| DM | Dry matter |
| EMSC | Extended multiplicative scatter correction |
| FNR | False negative rate |
| FPR | False positive rate |
| FT-NIR | Fourier transform near infrared |
| Fv | Variable fluorescence |
| Fm | Maximal fluorescence |
| FWHM | Full width at half maximum |

| | |
|---|---|
| GCMS | Gas chromatography and mass spectrometry |
| GVP | Gross value of production |
| HACCP | Hazard analysis and critical control point |
| HPLC | High performance liquid chromatography |
| Hz | Hertz |
| IB | Internal browning |
| InGaAs | Indium Gallium Arsenide |
| IDD | Internal defect detection |
| kHz | kilo Hertz |
| kNN | k-nearest neighbourhood |
| kPa | kilo pascal |
| LDA | Linear discriminant analysis |
| LED | Light emitting diode |
| LS-SVM | Least square support vector machine |
| LW-NIR | Long wave near infrared spectroscopy (1100-2500 nm) |
| LV | Latent variable |
| MD | Mahalanobis distance |
| MLR | Multiple linear regression |
| MLoR | Multiple logistic regression |
| mHz | mega Hertz |
| MMS | Miniature monolithic spectrometer |
| MRI | Magnetic resonance imaging |
| MSC | Multiplicative scatter correction |
| MVA | Multivariate analysis |
| NIRS | Near infrared spectroscopy |

| | |
|---|---|
| NMR | Nuclear magnetic resonance |
| $O_2$ | Oxygen |
| OSC | Orthogonal scatter correction |
| Pbs | Lead sulphide |
| PC | Principal component; synonymous to factor in multivariate analysis |
| PCA | Principal component analysis |
| PCR | Principal component regression |
| PDA | Photo diode array |
| PLSR | Partial least square regression |
| PTFE | Poly- tetrafluorethylene |
| PS | Photosystem |
| R | Correlation coefficient |
| $R^2$ | Coefficient of determination |
| RMSEC | Root mean square error of calibration |
| RMSECV | Root mean square error of cross validation |
| RMSEP | Root mean square error of prediction |
| RPD | Ratio of standard error of prediction to standard deviation |
| ROC | Receiver operator characteristics |
| SD | Standard deviation |
| SDR | Ratio of standard deviation to standard error of prediction |
| SEC | Standard error of calibration |
| SEP | Standard error of prediction |
| SG | Savitzky Golay |
| Si | Silicon |
| SIMCA | Soft independent modelling of class analogy |
| SNV | Standard normal variate |

| | |
|---|---|
| SOC | Sorting optimisation curve |
| SSC | Soluble solid content |
| SVM | Support vector machine |
| SW-NIR | Short wave near infrared spectroscopy (700-1100 nm) |
| T | Transmittance |
| TA | Titratable acidity |
| TF | Trifluoroacetic acid |
| TNR | True negative rate |
| TPR | True positive rate |
| TRS | Time resolved reflectance spectroscopy |
| TSS | Total soluble solids |
| UV | Ultra violet |
| Vis-SWNIR | Visible- shortwave near infrared spectroscopy (500-1100 nm) |
| W | Watts |
| Ws | Watts-second |

# Chapter 1. Introduction




## 1.1 Thesis theme

The fruit supply chain involves a range of actors including input suppliers, growers, packers, transporters and consumers, all working to provide quality fruit to consumers. Premium pricing requires consistent and high quality for both external (e.g. size, shape and colour) and internal attributes (e.g. dry matter, soluble solids). Internal defects deeply seated in the flesh can remain unnoticed until consumption. These types of defects are considered 'major defects' by the supply chain, e.g. retailer specifications of <2% of fruit. Sorting is required to ensure dispatched lots meet the tolerance limit for defects. There are reports of detection and sorting of internal defects based on X-ray imaging, magnetic resonance imaging (MRI), nuclear magnetic resonance (NMR), acoustic sensors however these methods have not achieved commercial application due to cost or inline applicability. NIR spectroscopy is a more promising technique for some internal defects. After the first spectroscopic assessment of internal defects within apple was made by Francis et al., (1965), this area of research remained dormant until 2000 when Upchurch et al., (1997), Clark et al., (2003) and McGlone et al., (2005) reported on the scope of internal defect detection in apple with NIR spectroscopy. The technology has achieved some commercial adoption, although there are no scientific literature reports on the efficacy of these commercial units. This thesis explores the use of such technologies, particularly the near infrared spectroscopy, in context of internal defects of commodities relevant to Queensland horticulture. Further, the sorting operation is a statistical 'fog', with uncertainties on measurement accuracy and on type I and II errors given the level of incidence of the disorder in a given population. The sorting technology operator therefore needs a decision support aid to guide setting of thresholds on the sorting operation.

## 1.2 Literature Review

This review consist five sections (i) fruit quality (ii) fruit internal defects (iii) physiological basis of defects (iv) technologies for detection of internal defects and (v) review of technologies.

### 1.2.1 Introduction to fruit quality

The measurement of ‘quality’ involves the measure of ‘excellence’ of a product. For fruit, this excellence can refer to its appearance (size, shape or colour), nutritive value, sensory attributes and/or internal composition. The consumer may ‘buy with their eyes’ (i.e. base purchase decision on external appearance) but repeat purchasing is determined by the eating experience, and thus by internal attributes. There is also an increasing emphasis by market regulators on food safety, requiring a strong focus throughout supply chain (Harker et al., 2008; Walsh, 2014).

Fruit quality can be viewed from the perspective of any actor in the ‘value chain’, but the ‘ultimate’ perspective is that of the consumer who purchases and consumes the product (Mowat & Collins, 2000). For fruit production, the value chain involves a network of participants, including suppliers of agricultural inputs (fertilisers, agrochemicals and seedlings), the grower, the packer, the transporter, the trader, the retailer, and ultimately the consumer (Walsh, 2014). The goal of each participant should be to add value at each step. The measure of success of a value chain is the creation of value to all of its members. Such success requires strict attention to the quality of the product, with each step in the chain maintaining this attention. For fruit, this also requires knowledge of the effect of growing condition on final product quality.

Quality product is achieved by consistent commitment to given standards that lead to uniformity of product, ultimately to satisfy consumers' need. The quality is ensured by setting a standard, e.g. dry matter (%DM) or soluble solids content (%SSC), depending on the mutual needs of participants in the fruit supply chain. The agreed specifications should be quantifiable and replicable. Produce standards are in place for all commodities entering major retail chains of developed economies (Walsh, 2014). The development of such standards will

also occur in emerging economies, firstly in service to export markets and secondly to ensure food safety and eating quality to the consumers (Abbott, 1999).

Historically, postharvest fruit grading on size, shape and colour was undertaken through manual inspection and sorting (Bollen & Prussia, 2009; Walsh, 2014). These practices depended on cheap labour, and suffered from lack of uniformity in sorting. An evolution of automated grading has occurred, progressing from simple diverging belts grading on fruit size to electronic platforms using load cells for fruit weight and machine vision (based on charge coupled device, CCD, cameras), combined with neural network routines and fuzzy logic to differentiate defects from acceptable features such as stalks (Bee & Honeywood, 2004; Bollen & Prussia, 2009). However, these technologies are limited to the assessment of 'external' fruit attributes.

A range of non-invasive techniques hold promise for grading of fruit on internal quality attributes. For example, near infrared spectroscopy (NIRS) is used in grading of fruit for the quality parameters of %SSC and DM of apple and mango fruit (Subedi & Walsh, 2011; Subedi et al., 2007). Application of NIRS to grading of fruit on firmness and acidity has been proposed (Mendoza et al., 2014). Other technologies also hold promise, e.g. X-ray imaging (Haff et al., 2006; Mehta et al., 2013), nuclear magnetic resonance (NMR) (Hernández-Sánchez et al., 2007; Hills & Clark, 2003), magnetic resonance imaging (MRI) (Herremans et al., 2014b), hyperspectral imaging (Lee et al., 2014; Lu & Ariana, 2013; Wu & Sun, 2013), acoustic methods including ultrasound (Mizrach, 2000, 2008; Mizrach & Flitsanov, 1999), and the electronic nose (Boeker, 2014; Kauer & White, 2009; Loutfi et al., 2015; Vagin & Winquist, 2015). However, commercial uptake of these technologies has been limited by cost, safety, accuracy, and compatibility for online application. This thesis will focus on the non-invasive detection of fruit defects of relevance to Queensland horticulture.

### 1.2.2 Fruit internal defects

#### *1.2.2.1 A definition*

The term 'internal defect' refers to an internal character that is deemed unacceptable by any member of the value chain, but principally the consumer. Attributes such as SSC and DM can be considered 'positive' attributes, used in sorting for fruit with an improved eating

experience and ideally an improved market value or acceptance, while defects such as internal rot represent a 'negative', resulting in consumer rejection. Fruit are metabolically active after harvest and may develop certain internal defects during storage or movement through the value chain. Although largely unnoticed until consumption, the presence of such defects will affect repeat purchase decisions and market share (Gamble et al., 2010). If such defects result in product rejection, not only are the costs incurred for production and harvest wasted, but also the costs incurred in sorting, packing, storage, postharvest treatments and transport. Re-establishment of brand reputation is difficult, with marketing studies indicating a delay on return purchase of fruit of 4 – 6 weeks after a consumer has a bad eating experience.

For example, a study conducted in Australia on the effect of apple eating quality on consumer decision reported that a bad eating experience by a consumer typically resulted in change in purchasing decisions. It was reported that more than half (58%) of respondents shifted to another cultivar or purchased fewer fruit (31%), while 24% of respondents shifted to other types of fruit, and 17% did not buy apple for few weeks (Batt & Sadler, 1999). Sale volumes of apples were noted to be influenced more by fruit quality than by price Harker (2001); Harker et al., (2003, 2008). Internal disorders, such as internal browning in apple, are considered major defects by retailers. Retailers or standards bodies introduce produce specifications for their fresh fruit suppliers for their retail chain and set a limit on the presence of such disorders at < 2% of the consignment for 'major' defect and applies to any fruit commodities (Woolworths, 2015) . Thus, there is incentive for development and use of non-invasive measurement techniques for sorting to remove fruit with internal defects.

The issue of internal quality of fruit has received considerable scientific attention. Approximately 33,250 articles are associated with the key words "fruit internal quality" over the period 2000-2015 (www.sciencedirect.com; data retrieved on 18.12.2014). Approximately 20% of this number was captured by the descriptor 'internal defect', and 10% were associated with the descriptor 'internal defects of apple' (Fig. 1.1).

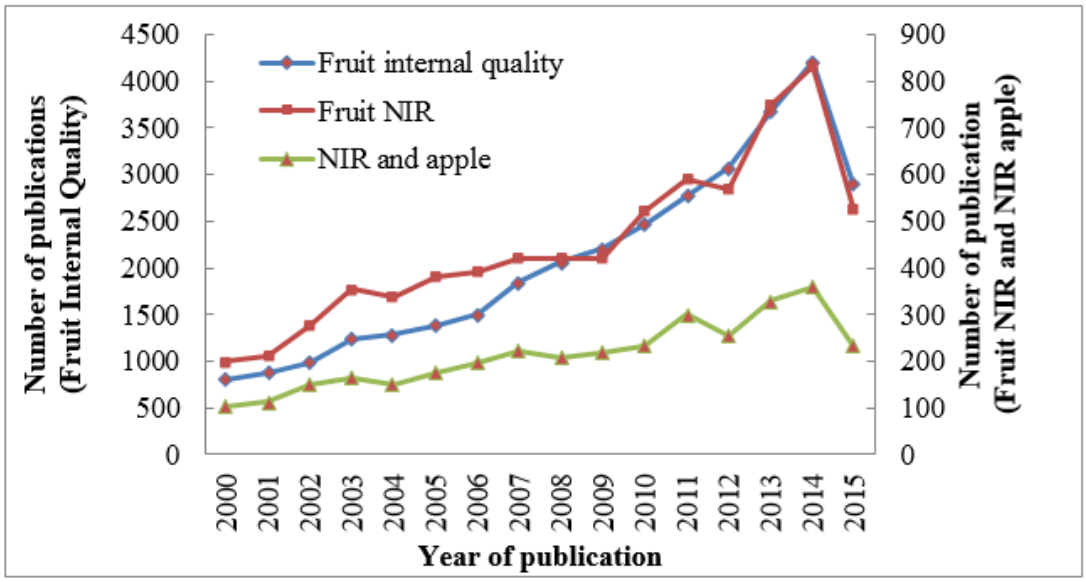



**Figure 1.1.** Number of refereed publications associated with key words 'fruit internal quality', 'Fruit NIR' and 'NIR and apple'. (Data sourced from Science Direct on May 15, 2015. Note use of a different scale for keywords 'fruit internal quality).

#### *1.2.2.2 Types of internal defects*

Many types of internal defects can occur in fruit (Table 1.1). Water-soaked tissues (e.g. watercore in apple), hard or spongy flesh (e.g. lumpiness and spongy tissues in mango) flesh browning (e.g. in apple, pear, pineapple and avocado), mealy flesh (e.g. mealiness in apple and stone fruit), dry juice sacs (e.g. in mandarin) and formation of cracks and voids are a few of the common physiological disorders of fruit (Wongs-Aree & Noichinda, 2014) (Table 1.1). These defects may be the result of preharvest or postharvest conditions and may appear before or after harvest. Pre-disposing factors that occur pre-harvest include growing conditions such as soil, temperature, aspect, humidity, growing degree days, nutritional deficiency or excess, water relations and harvest time (Benkeblia et al., 2011; Bergman et al., 2012; Galvis-Sánchez et al., 2004; Hannah, 2007; Hatoum et al., 2014b; James & Jobling, 2009; Jobling & James, 2008; Moggia et al., 2015; Paull & Reyes, 1996). Postharvest factors include storage environment conditions such as temperature, humidity and gas concentrations (Castro et al., 2008; Castro et al., 2007; Eksteen & Truter, 1987; Galvis-Sánchez et al., 2004; Hatoum et al., 2014a; Hatoum et al., 2014b; Ho et al., 2013).

Some common internal defects affecting fruit relevant to Queensland are discussed below (Table 1.1) .

1 **Table 1.1.** Major internal disorders of fruits of a range of fruits grown in Queensland.

| Commodity/ Major internal disorders | Predisposing factors /Causes | Symptoms | References |
|---|---|---|---|
| Apple | | | |
| Bitter pit | Low moisture in soil reducing the ability of tree to uptake and transport calcium from the soil. Excessive nitrogen and hot dry weather during fruit growth | Bruise like spot in skin, disorder starts internally and may lead to bitter flavour as it worsens. | de Freitas et al. (2013); de Freitas et al. (2010); Miqueloto et al. (2014); de Freitas et al. (2015) |
| Internal browning | Nutrition, maturity and $CO_2$ concentration in controlled atmosphere storage | Development of a brown to dark brown spot limited to cortex or vascular bundle, sometimes extending through entire flesh. | James and Jobling (2009); Holderbaum et al. (2010); Moggia et al. (2015); |
| Water core | Low night and high day temperature during maturity stages, low calcium and high nitrogen, over maturity and lean crop with large size fruit. | Translucent flesh with high sugar sorbitol content. | Yamada and Kobayashi (1999); Beaudry (2014); Dart and Newman (2005 ) |
| Avocado | | | |
| Chilling injury | Storage at 0-2 $^{o}$C for a week or more | Skin pitting, scald development and blackening . | Florissen et al. (1996); Pesis et al. (2002) |
| Internal flesh greying | Storage of fruit at 3-5 $^{o}$C for more than two weeks | Gray pulp and vascular browning | Woolf et al. (2005); Gudenschwager et al. (2013) |

| | | | |
|---|---|---|---|
| Banana | | | |
| Under peel discolouration | Bunches exposed to a temperature less than 12 $^{o}$C | Peel discolouration on ripening and develop grey or yellow colour | Huang et al. (2013); Pongprasert et al. (2011) |
| Citrus | | | |
| Granulation | Nutrition imbalance, sandy soils, humid climates, fast growth of trees, large fruit | Vesicle shrivelling and gel formation | Munshi et al. (1978); (Sharma et al., 2006; Singh & Singh, 1981a, 1981b; Wang et al., 2014) |
| Section drying | Nutrition, irrigation | Collapse of juice sacs or granulation also called dehydration disorder. | Peiris et al. (1998) |
| Gelling defect | Nutrition, irrigation, | Dry, chewy and tasteless sensation from fruit consumption. Juiciness decreased. | Subedi (2007) |
| Macadamia nut | | | |
| Shrivelled kernel | Nutrition | Low oil content and hard after drying | Wall (2013); Srichamnong and Srzednicki (2015) |
| Immature kernel | Nutrition | Excessive browning during roasting | Srichamnong and Srzednicki (2015) |
| Mango | | | |
| Spongy tissue | High temperature, convective heat and postharvest exposure to sunlight. | Patches of flesh fails to ripen and forms a lump | Shivashankara and Mathai (1999) |
| Chilling injury | cultivar, temperature and duration of exposure. | Skin discolouration, uneven ripening, poor colour and flavour | Chongchatuporn et al. (2013) |
| Internal flesh breakdown | late harvest | Flesh breakdown and development of internal cavities between seed and peduncle. | Raymond et al. (1998) |

| | | | |
|---|---|---|---|
| Jelly seed | Ca deficiency | Disintegration of flesh around seed into a jelly-like mass. | Raymond et al. (1998) |
| Soft-nose | Ca deficiency | Softening of tissue at apex. Flesh appears over-ripe and may discolour and become spongy. | Burdon et al. (1991) |
| Pear | | | |
| Core breakdown | Delayed harvesting, over maturity, cold storage at high $CO_2$ concentration | Brown and soft breakdown of core, may have vascular browning | Choi et al. (2015); Lammertyn et al. (2003); Verlinden et al. (2002); Franck et al. (2003) |
| Internal browning | Late harvest, CA storage | | Franck et al. (2007); Wang and Sugar (2013); Yan et al. (2013) |
| Watery breakdown | Late precooling and short storage | Soft and watery appearance of tissues, leakage of juice out of fruit | Wang and Sugar (2013) |
| Pineapple | | | |
| Internal browning | Low temperature (below 5 $^{o}$C) storage | Development of browning streaks in the pulp and expands to core | Lu et al. (2011); Pusittigul et al. (2012); Selvarajah et al. (2001); Soares et al. (2005) |
| Translucency | Small crown, low night temperature during fruit growth, followed by high temperature during fruit maturation | Low porosity and water soaked appearance, may develop overripe flavour | Chen and Paull (2000)<br>Paull and Reyes (1996) |

Apple fruit can display internal flesh browning, water core or chilling injury with associated development of off flavours during storage (Hatoum et al., 2014b). Water core is a condition associated with pre-harvest conditions of low transpiration rate, and is not a common disorder in Queensland. Internal flesh browning develops during storage, with certain pre-harvest conditions pre-disposing fruit to development of the disorder. Growers in the Stanthorpe region of Queensland report that the variety Pink Lady™ is prone to internal browning during controlled atmosphere (CA) storage, with loss up to 80% of fruit in some seasons for fruit from some paddocks and under certain nutrition management (Personal communication Rosie Savio, Savio farm, Stanthorpe). Indeed, in 2003, 35 containers of Pink Lady™ apples exported to UK were rejected on arrival due to internal browning (Jobling & James, 2008). This entrained loss of $35,000 per container, plus long term damage to brand reputation and loss of consumer confidence.

Avocado fruit, especially of the Hass cultivar, suffer from chilling injury e.g. when fruit developed under shaded canopy conditions are stored at 0 $^{o}$C for several days (Ferguson et al., 1999a). However, this disorder can be minimised by good temperature management during transport and storage.

Banana fruit are affected by chilling, with development of dark brown streaks in the flesh and browning of peel and flesh in severe cases (Pongprasert et al., 2011) . Severity is high if temperatures fall below 13 $^{o}$C for more than 24 h. Impact bruising in banana may also cause flesh browning without any external skin symptoms (Bugaud et al., 2014). Again, these disorders can be minimised by good temperature management during transport and storage, and care in fruit handling.

Macadamia kernels can develop internal discolouration leading to brown centres or rancidity during storage. Internal browning in macadamia is the result of presence of reducing sugars in the centre of kernel while an external browning developing during storage is associated with movement of sugars to the surface (Wall, 2013). Kernels with higher oil content are prone to rancidity. High moisture content in kernels during storage initiates rancidity, resulting from the decomposition of peroxides which, in turn, are the oxidative product of unsaturated fats and includes aldehydes, hydrocarbons and ketones (Wall, 2013). Current assessment requires physical inspection of kernels extracted from nuts. Unfortunately the thick exocarp and

mesocarp and woody endocarp of the fruit render non-invasive assessment techniques difficult to apply.

Mandarin and other citrus fruit, particularly limes, lemons and grapefruit, suffer from chilling injury, granulation and puffiness (Ladaniya, 2008c, 2008b). Of these disorders, granulation in Imperial mandarin is the largest concern in Queensland. Hofman (2011) reported the incidence of granulation in the Burnett region of Central Queensland, Australia, as affecting >25% of fruit on average over four years, and reaching incidence levels of up to 33% in some seasons.

Mango fruit can suffer ricey disorder, lumpy tissue, spongy tissue and black flesh due to preharvest reasons, while chilling injury, $CO_2$ injury and black flesh are disorders caused by postharvest conditions (Shivashankar, 2014). Other important issues in mango include stone (seed) weevil infestation (Thomas et al., 1995; Verghese et al., 2005) and soft nose (Young, 1957), internal breakdown and jelly seed (Burdon et al., 1991; Raymond et al., 1998). However, seed weevil does not occur in Queensland and the other disorders do not cause large economic losses (Meurant et al., 1999).

Pineapples can develop internal browning, internal rot or translucency. The growing or storage environment conditions that pre-dispose fruit to these disorders are not well defined (Haff et al., 2006; Marrero & Kader, 2006; Paull & Reyes, 1996; Zhang et al., 2013).

### *1.2.2.3 Horticultural production in Queensland*

Fruit production in Queensland is valued at $1169 million in Gross Value of Production (GVP) (five year average for the ten fruit commodities; Table 1.2). Of these commodities, banana fruit has the highest value, followed by avocado and mandarin (DAFF, 2013).

**Table 1.2.** Gross value of production (GVP) of different fruit crops grown in Queensland, Australia (in $ million).

| Commodities | 2008-09 | 2009-10 | 2010-11 | 2011-12 | 2012-2013* | 5 year average |
|---|---|---|---|---|---|---|
| Apple | 79 | 81 | 143 | 78 | 95 | 95 |
| Avocado | 60 | 80 | 170 | 145 | 140 | 119 |
| Banana | 390 | 448 | 283 | 415 | 500 | 407 |
| Macadamia | 16 | 29 | 35 | 53 | 52 | 37 |
| Mandarin | 64 | 76 | 89 | 71 | 64 | 73 |
| Mango | 83 | 72 | 55 | 50 | 70 | 66 |
| Pineapple | 88 | 70 | 50 | 68 | 83 | 72 |
| Strawberries | 87 | 145 | 74 | 150 | 125 | 116 |
| Table grapes | 24 | 36 | 32 | 18 | 50 | 32 |
| Other fruit and nuts | 126 | 257 | 129 | 92 | 200 | 161 |
| Total | 1017 | 1294 | 1062 | 1139 | 1334 | 1169 |

*forecast (Source: 2013)

No formal estimate exists for the loss in production value due to internal disorders, and losses will vary with commodity, variety, and season. However, from the discussion above, the loss associated with apple flesh browning, mandarin dryness defect and pineapple translucency are significant, and these defects are candidates for non-invasive assessment/sorting.

Based on economic value to Queensland (Table 1.2), the level of incidence of internal defects and potential application of technology for non-invasive detection of such defects, attention is given in this thesis to apple browning, citrus granulation defect and pineapple translucency.

### 1.2.3 Physiology and biochemical basis of internal defects

Fruit are living biological entities, even after harvest, and accordingly demand management to retain physiological condition. The growing and storage environment and inherent nature of the variety/fruit determines the incidence of defects, both internal and external. Fruit exposure to inappropriate environments (light, temperature, gaseous concentration, relative humidity etc.) will affect physiology of the fruit, potentially creating a defect. For example, frost (freezing) can induce cell damage and later development of 'dry' areas within citrus fruit. Similarly, nutrition (deficit or excess), water management (deficit or excess) and gaseous environment (e.g. $O_2$ and $CO_2$ levels) can also have significant impact on the development of disorders.

In this section, a review is presented on the basis of internal browning in apple, dryness defect in citrus and translucency in pineapple fruit.

#### *1.2.3.1 Apple Internal Browning*

Apple is a member of the Rosaceae family, characterised by presence of flower parts in 5 or its multiple. The fruit are pome, in which the inner core is the true fruit (ovary wall), with the endocarp lignified to a stony layer surrounding the seed. The outer part of the fruit (the hypanthium) is formed by enlargement of sepal, petal and stamen tissue surrounding the ovaries in this inferior fruiting structure. As such the fruit contains accessory (non-ovary) tissue.

Several types of internal browning have been recognised. Radial flesh browning is characterised by appearance of browning discolouration advancing along vascular bundles (Fig. 1.2A), while diffuse browning involved discolouration of the inter-vascular areas (Fig. 1.2B) (James & Jobling, 2009). Bulge browning refers to browning of an area under a surface lump (Fig. 1.2C). This type of defect can be detected based on the external feature of the surface deformation. An internal breakdown disorder termed 'brownheart' is associated with $CO_2$ injury and prolonged storage (Fig. 1.2D).

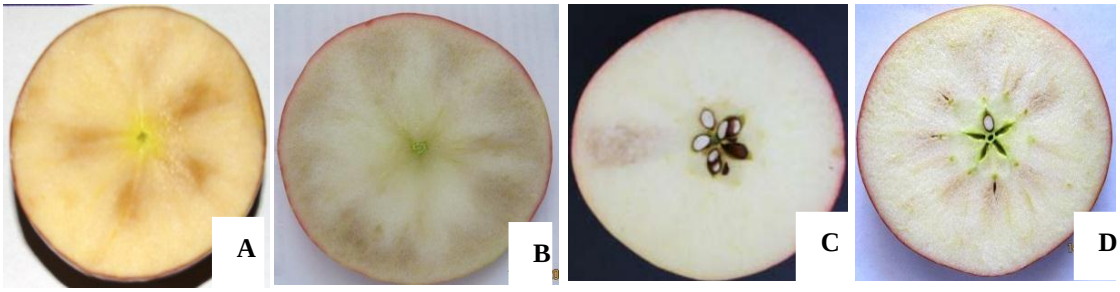



**Figure 1.2.** Types of internal flesh browning in apple A. radial. B. diffuse C. bulge and D. $CO_2$ injury. (Source: Bergman et al., 2012; James and Jobling (2009))

Apple fruit are rich in polyphenols, including procyanidin, catechin, epicatechin, chlorogenic acid, coumaroylquinic acid, and phloridzin, with the first four phenolic compounds predominating (Song et al., 2007). These polyphenols are responsible for flavour and colour development in apple but are also responsible for development of internal browning (Holderbaum et al., 2010; Lattaznzio, 2003; Macheix et al., 1990). Varieties that contain less of these compounds and polyphenol oxidase activities show less browning of cut surfaces, and less internal browning as a defect (Holderbaum et al., 2010). The browning symptom is thought to result from the oxidation of polyphenols to brown coloured pigments, namely quinone or its insoluble polymer, melanin as shown in Fig. 1.3 (Nicolas et al., 1994). These polyphenolic products have a phenolic ring with OH groups which has an absorption peak around 847 nm.

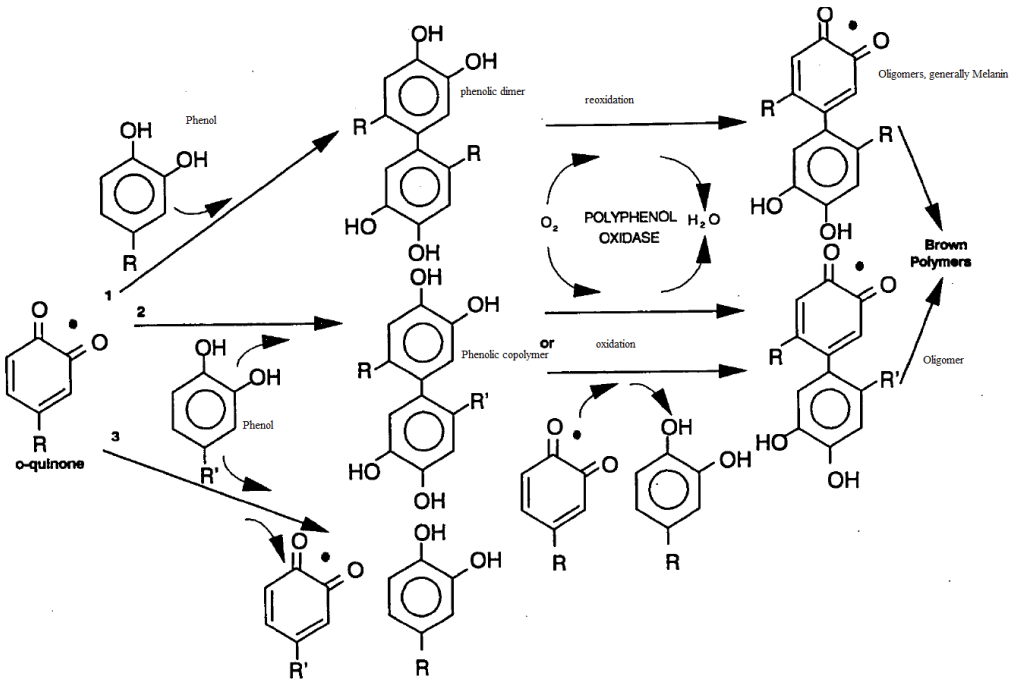



**Figure 1.3.** Proposed kinetic mechanism for enzymatic browning reaction in apple which leads to formation of brown polymer, melanin. Source: (Nicolas et al., 1994 )

The same process results in tissue surface browning following cutting of the fruit and exposure to air. Apple varieties that display a low propensity to brown on cutting (e.g. Granny Smith) possess a low concentration of precursor's catechin and chlorogenic acid and also display a low propensity to develop internal browning during prolonged controlled atmosphere (CA) storage, relative to varieties that contain higher amounts of these polyphenols (Holderbaum et al., 2014; Holderbaum et al., 2010; Huque et al., 2013; Ma et al., 2015). A loss of malic acid was reported by Vandendriessche et al. (2013) in apple tissues with internal browning.

For apples, the minimum level of physiological function and maximum storage life, as indexed by respiration, transpiration and enzymatic metabolism, is achieved through reduced oxygen ($O_2$) concentration and temperature, and increased humidity and partial pressure of carbon dioxide ($CO_2$) to minimize aerobic respiration while avoiding fermentation (Ho et al., 2013; Lumpkin et al., 2015; Nock & Watkins, 2013). Apples, and pome fruit in general, are generally stored at around 2-5 $^{o}$C with relative humidity above 70% in CA storage having gaseous concentration of oxygen and $CO_2$, respectively, of 2% and 5-6% v/v. Higher partial pressures of $CO_2$ and lower partial pressures of $O_2$ can result in anaerobic respiration, and to disruption of

cell membranes thereby leading to a conducive environment for oxidative enzymatic browning. In practice, following controlled atmosphere storage of apple, fruit internal tissues may become hypoxic, leading to membrane dysfunction. This phenomenon appears to be more prevalent in modern varieties (e.g. Pink Lady$^{TM}$) which have been bred for increased fruit 'crispiness', and indirectly for a reduced intercellular volume (Herremans et al., 2014a).

Membrane dysfunction is associated with de-compartmentalisation of oxidative enzymes, leading to oxidation of polyphenols to quinone or related polymer compounds, and resulting in browning of flesh. The discoloured tissues areas are characterised by collapse of vascular and cortical cells in radial and diffuse flesh browning, respectively while $CO_2$ injury is associated with formation of cavities in the cortical tissue (Fig. 1.4) (James & Jobling, 2009).

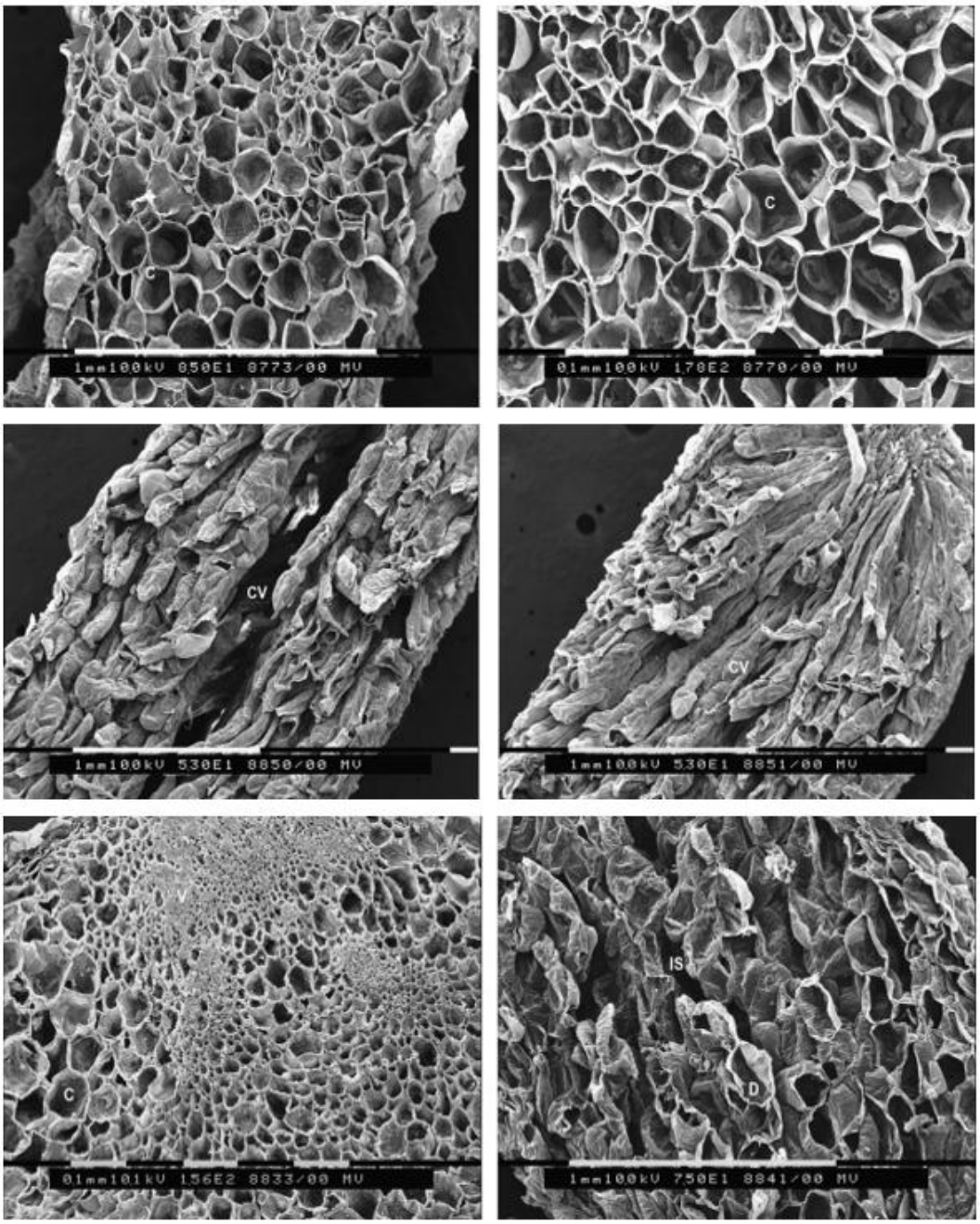



**Figure 1.4.** Scanning electron microscope (SEM) micrographs showing ultrastructure of apple vascular tissue (left) and cortex tissue (right) for sound fruit (upper row) and fruit with diffuse flesh browning (middle row)and with $CO_2$ injury (bottom row) V = vascular tissues, C = cortex cells, D = collapsed tissues, cv = cavity and IS = intercellular spaces. Source: James and Jobling (2009).

### *1.2.3.2 Dryness defect/granulation in citrus*

Mandarin is a member of Rutaceae family, in which the fruit is a hesperidium. Anatomically, this fruit type consists of flavedo (skin and embedded oil glands), albedo (a thin layer of white tissue between skin and carpel), endocarp (with elongated multicellular structures projecting into the central locules forming juice vesicle) and seeds in the fruit axis.

Various forms of granulation are reported in citrus fruit. Peiris et al. (1998) noted the use of terms such as section drying, gelling, granulation, crystallization, vesicle collapse and core dryness to describe these disorders. They divided these disorders into two broad categories, namely granulation and dehydration. In dehydration, collapse of the vesicles follows the shrinking of juice vesicles due to loss of fluids within the juice sac. One type of dehydration defect is associated with frost damage, in which juice vesicles are ruptured, and later lose their water content (e.g. after storage) (Subedi, 2007). The defect is characterised by decreased fruit density. Granulation starts with 'hardening' of juice vesicles, without loss of water, progressing to a collapse of the inner cells within the juice sacs and formation of an empty cavity (Fig. 1.5).

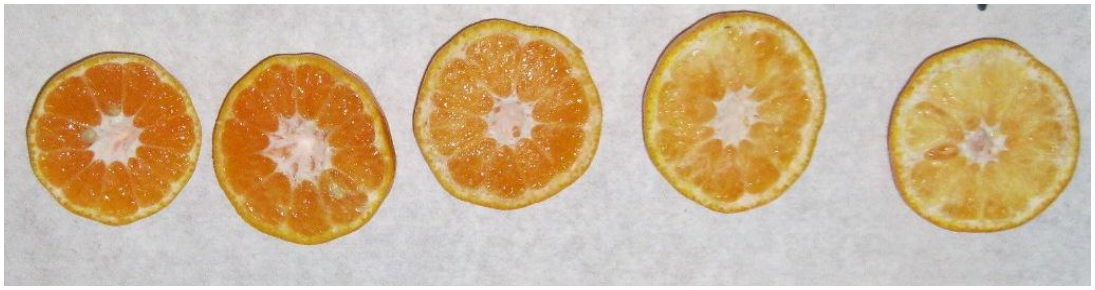


**Figure 1.5.** Granulation in 'Imperial' mandarin, from juicy (left) to dry (right**).**

The incidence of granulation in mandarin in Central Queensland cultivation has been associated with warm winters and high rainfall in spring or late summer, when there is high competition of resources between flowers, developing fruit and expanding flush. Hofman (2011) reviewed a range of hypotheses for the cause of gelling defect of Imperial mandarin, including sink competition between plant organs, accelerated senescence, temperature sensitivity, moisture stress, inefficient transport of

carbohydrates or too rapid growth, concluding that the sink competition hypothesis is the most favoured hypothesis. However, Hofman (2011) proposed a hypothesis based on results of trials on nutrition, thinning, plant growth regular use and irrigation (Fig. 1.6). This hypothesis postulates that juice cell water potential plays a role in the development of granulation, suggesting that the thickening of the juice cell walls is a protective mechanism to reduce water loss from the juice sacs. The 'sink competition' hypothesis suggesting the higher degree of competition between fruit, roots and leaves received wider attention. This is supported with corresponding low soluble solids (carbohydrates) in the granulated fruit and incidence of granulation in tree with high vegetative growth Kaur et al. (1991). However, this hypothesis didn't explain about cell wall thickening and gelling in vesicles (Hofman, 2011). Other theories consider granulation as normal maturation process (Chen et al., 2005) or temperature and/or moisture stress (Burns & Achor, 1989 ) or inefficient transport of metabolites into the fruit (Chakrawar & Singh, 1978 ).

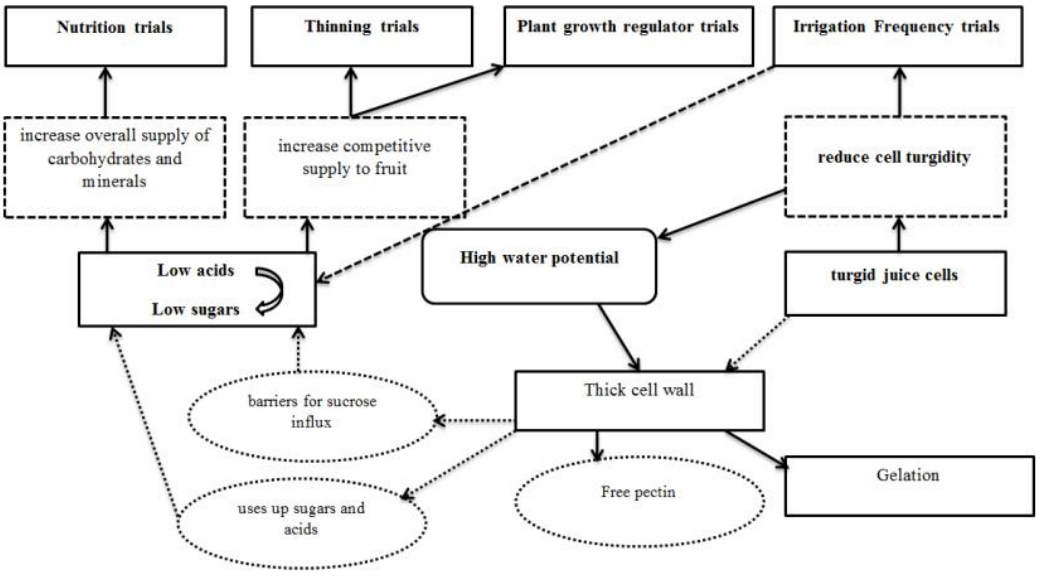



**Figure 1.6.** The water potential hypothesis for the development of granulation in mandarin. Source: adopted from Hofman (2011).

#### *1.2.3.3 Pineapple translucency and browning*

Pineapple is a member of Bromeliaceae family, in which a multiple fruit is fruit is formed by fusion of multiple fruitlets and accessory (bract) tissue. The fruit is infertile. Given this anatomical complexity, there can be a large variation in parameters such as TSS around the fruit. A water soaked appearance of the flesh of pineapple fruit is termed translucency (Haff et al., 2006) (Fig. 1.7). In this disorder, intercellular air spaces are filled with liquid, accompanied by loss of tissue gas permeability (Chen & Paull, 2000). High summer temperature and high rainfall during the ripening period followed by some duration of dry conditions pre-dispose fruit to this disorder (Paull & Reyes, 1996). The disorder involves a sugar induced solute-potential gradient between symplast and apoplast, involving a high rate of sucrose accumulation and increased sucrose synthase and cell wall invertase activity in affected fruit (Chen & Paull, 2000). Total sugar content and the sugar: acid ratio of translucent fruit is higher than that of normal flesh fruit. High levels of translucency result in poor oxygen diffusion into the fruit, and consequent anaerobic conditions resulting in \bacterial fungal infection and unpleasant flavour. The disorder occurs erratically, and is colloquially considered to occur following pre harvest rain. Translucency in pineapple is mentioned by Department of Agriculture and Fisheries, Queensland as an issue if it exceeds 50% in the fruit (www.daf.qld.gov.au).


**Figure 1.7.** Flesh translucency in pineapple: normal (1) to most translucent (5). (Source: Edwards Don, University of California, Davis)

### 1.2.4 Technology for assessment of internal defects in fruit

#### *1.2.4.1 Detection technologies*

All participants in a fruit value chain can take steps to minimise the incidence of defects, with some steps having more influence (e.g. selection of variety is a large determinant on apple internal browning incidence). A quality management system will aim to minimize the development of physiological disorders incidence in farm or storage or to identify and remove defect fruit before dispatch to retail. Small sample size destructive sampling can occur anywhere from in-field to retail distribution centre or in store. However, when for a retailer specification of no more than 2% of fruit to be affected by a defect, a typical destructive sample size of 30 fruit is statistically unsound. For screening of large numbers of fruit, or indeed, or all fruit, a non-invasive technology is required. There are limited options for effective implementation points of such technologies in the value chain. Such points occur on packing lines near the production area, or in re-pack centres following storage.

Recent advances in technology provide opportunities for detection of internal defects based on the physical and/or chemical differences between sound and defect fruit. Acoustic, X-ray imaging, Nuclear Magnetic Resonance (NMR) and Magnetic

Resonance Imaging (MRI), hyperspectral imaging and Near Infra-red Spectroscopy (NIRS) techniques have been proposed for in-line detection of internal defects, while other technologies (e.g. fluorescence; production of certain volatiles) have been discussed in terms of in-field or at-line assessment (Table 1.3). Some techniques have been used to non-invasively monitor the development of a disorder in fruit. For example, the use of chlorophyll fluorescence technique to assess the development of pithy brown core in pear, as caused by changes in partial pressure of $O_2$ in controlled atmosphere (CA) storage, has been considered by a number of authors (Lumpkin et al., 2015; Mattheis et al., 2013; Mattheis & Rudell, 2011; Pasquariello et al., 2013; Pedreschi et al., 2008; Wang & Sugar, 2013; Yan et al., 2013; Zhang et al., 2015). The use of the technologies presented in Table 1.3 for detection of specific internal defects is presented in Table 1.4.

A seminal review of technologies appropriate for non-invasive assessment of fruit was published by Abbott (1999), with subsequent reviews focussing on specific technologies such as NIRS by Nicolaï et al. (2007). Non-invasive quality assessment techniques have been well reviewed in recent years (Chen & Opara, 2013; Chen et al., 2013; Clark et al., 1997; Lin & Ying, 2009; Loutfi et al., 2015; Nicolaï et al., 2007; Ruiz-Altisent et al., 2010; Wu & Sun, 2013; Zhang et al., 2014). Commercial uptake of these technologies has been modest, as for commercial application a technique must be both economic and robust in detection of the defect. A brief description of these technologies is presented here.

Flotation of fruit in solutions of varied density has been used for sorting for internal defects involving change in fruit density, though with high inaccuracies (Clark et al., 2007; Moscetti et al., 2014). The technique is compromised by entrapment of air (e.g. around bracts with pineapple fruit) and the effect of the solution may on the quality of fruit, e.g. incidence of mould.

X-ray imaging and computer tomography have been used for detection of differences in density (and thus coefficient of absorption of X-ray) within fruit (Jha & Matsuoka, 2000). In practice, the density differences of tissues within fruit are modest, such that normal transmission X-ray imaging yields relatively little information. However, air spaces can be readily detected. X-ray CT provides 3D imaging depicting much greater detail, but at greater cost and too-low speed to be practical to fruit sorting.

**Table 1.3. Summary of methods for detection of internal defect in fruit.**

| Methods | Attributes for study/detection | Comments | References |
|---|---|---|---|
| X-ray transmission | Attributes with density different to normal fruit tissue, e.g. decay and insect infestation, voids | Well suited for detection of voids, health hazards, has been commercially used for detections of potato hollow heart | Herremans et al. (2014b) |
| Magnetic resonance and resonance imaging (MRI) | Tissues varying in water mobility and water content e.g. water core, ripeness, core breakdown, pits, seeds, voids, bruises, dryness, chilling and freezing injuries | Relaxation times are often used to describe the biological state of tissues and are interpreted as the ratio of bound water to free water. | Abbott (1999); Gonzalez et al. (2001) |
| Thermal Imaging | Tissues varying in temperature e.g. bruises | Up to 100% of apple bruises were detected using thermal imaging of the fruit surface | Baranowski et al. (2012); Varith et al. (2003) |
| NIR hyperspectral imaging | Tissues varying in O-H, C-H, content or scattering properties, e.g. bitter pit in apple, maturity in banana, internal defects | A discriminant calibration model was developed and successfully validated in different apple varieties. Cannot differentiate between bitter pit and corky tissues. | Ariana and Lu (2010); Rajkumar et al. (2012) |

| | | | |
|---|---|---|---|
| NIR Spectroscopy | | | |
| -transmission geometry | Major internal defects like apple browning, pear core breakdown, jujube insect infestation and cherry insect infestation, translucency in mangosteen | Detect the defects due to difference in light absorptivity and scattering of the fruit | Nicolaï et al. (2007); Nicolaï et al. (2009), Teerachaichayut et al. (2011); Terdwongworakul et al. (2012) |
| -interactance geometry | attributes to 1-2 cm depth like DM, SSC etc. | Differences in light absorption and scattering by the tissues | Upchurch et al. (1997); Nicolaï et al. (2009) |
| -reflectance geometry | Attributes in the surface e.g., skin colour, presence of the bruises and moulds | Differences in absorption of different colour in visible region | Lorente et al. (2015) |

1

2

**Table 1.4.** Summary of fruit defects and detection techniques used in detection of a range of fruit defects.

| Defects | Fruit Type | Detection | References |
|---|---|---|---|
| Internal Browning | Apple | X-ray micro CT | Herremans et al. (2013) |
| | | X-ray imaging | Herremans et al. (2014b) |
| | | NIRS | Upchurch et al. (1997) |
| | | MRI (sensors) | Chayaprasert and Stroshine (2005) |
| | | MRI | Gonzalez et al. (2001) |
| | | FT- NIR | Li (2011) |
| | | NIRS | McGlone et al. (2005) |
| | | NIRS | Clark et al. (2003) |
| | Pear | NIRS | Fu et al. (2007) |
| | | Thermography/ MRI | Baranowski et al. (2009) |
| | | MRI | Clark et al. (1998) |
| Internal Breakdown | Apple | NIRS | Upchurch et al. (1997) |
| Core breakdown | Pear | MRI | Clark et al. (1998) |
| Woolly breakdown | Nectarines | MRI | Clark et al. (1998) |
| Chilling Injury | Persimmon | MRI | Clark et al. (1998) |
| | Pear | NIRS | Han et al. (2006) |
| Mealiness | Apple | Ultrasonics | Bechar et al. (2005) |
| Pericarp hardening | Mangosteen | NIRS | Teerachaichayut et al. (2011) |

| | | | |
|---|---|---|---|
| Translucency | Mangosteen | NIRS | Teerachaichayut et al. (2007) Terdwongworakul et al. (2012) |
| Translucency | Pineapple | MRI | Clark et al. (1998) |
| Insect damage | Cherry | NIRS | Xing and Guyer (2008b) Xing and Guyer (2008a) |
| | Jujube | NIRS | Wang et al. (2011) Wang et al. (2010) |
| | Citrus | Gas chromatography | Kendra et al. (2011) |
| Gelling defect | Mandarin | NIRS | Subedi (2007) |
| Section Drying | Mandarin | NIRS/ X-ray CT | Peiris et al. (1998) |

Chlorophyll fluorescence is a commonly used physiological method for assessment of photosynthetic activity and physiological status. Fluorescence is induced by direct excitation of chlorophyll molecules of photosystem II (PSII) by light and their immediate relaxation. Stresses such as chilling or high temperature injuries and low $O_2$ can reduce PSII function thereby lowering the photochemical efficiency. A change in photosynthetic activity may be associated with the overall 'health' of a tissue. Thus the technique has been utilised with stored apples to detect some physiological disorders and associated internal defects on the basis that PSII is interrupted and chlorophyll fluorescence will be high in stressed fruits (Lumpkin et al., 2015; Mattheis et al., 2013; Mattheis & Rudell, 2011; Pasquariello et al., 2013; Pedreschi et al., 2008; Wang & Sugar, 2013; Yan et al., 2013; Zhang et al., 2015). Equipment intended for controlled atmosphere (CA) storage room management is available (e.g. Harvest Watch; Bessling P/L).

Nuclear magnetic resonance (NMR) and magnetic resonance imaging (MRI) take advantage of the property of certain nuclei (such as 1H or 13C) to absorb electromagnetic radiation at a characteristic frequency when placed in a strong magnetic field, and weakly emit a radio frequency when 'relaxing' (Zhang, 2012). The spin-spin relaxation time (T2) and spin- lattice relaxation time (T1) refers to the mode of relaxation process which involves loss of energy.

Tissues within a fruit that vary in water content and distribution can be differentiated based on NMR T2 relaxation times. Hills and Clark (2003) presented a summary of use of MRI in assessment of fruit quality attributes and in measurement of physiological changes in fruit and vegetables. Vandendriessche et al. (2013) reported the use of NMR in detection of internal flesh browning in Braeburn apple. There are number of reports of use of NMR in defects detection in pear (Franck et al., 2007; Hernández-Sánchez et al., 2007). Other potential applications include detection of woolly breakdown (Sonego et al., 1995), insect damage and freezing injuries in fruit (Clark et al., 1997). This technique is not in commercial use for detection of the internal defects in fruit owing to its cost and processing. However, Aspect Imaging (www.aspectimaging.com) claims development of an in-line real time application suitable for fruit quality measurement and sorting with speed of 10-12 fruit per second.

Ultrasound technology (sonography) involves use of high-frequency sound waves wave for non-invasive assessment (Mizrach, 2008). Vibrations in the order of 20 Hz to 20 kHz are audible to humans, vibration below this range is termed infrasound, and above this range, ultrasound. In sonography, a pulse of ultrasound is propagated into the sample from a transducer, with reflected sound (echoes) from the sample detected and displayed as an image, typically of the acoustic impedance of a two-dimensional cross-section of tissue. The frequency used in an application is a compromise between spatial resolution and imaging depth. Higher frequencies allow greater resolution but with a higher attenuation coefficient the depth of penetration of the sound wave into the sample is limited. For human and animal tissues, the frequency range of 1 to 18 mHz is normally used.

This technique has been widely used in human and animal medicine but it does not work well in plant tissue due to prevalence of air-water interfaces. These interfaces reflect sound waves, limiting effective visualisation depth, even with use of lower frequencies. Other limitations include the need for a bridging medium between the transducer and the sample. In human applications a gel is used, and in industrial applications a stream of water has been used. There are number of reports of use of ultrasonic technology for fruit quality assessment including fruit defects (Bechar et al., 2005; Mizrach & Flitsanov, 1999).

Acoustic methods utilise lower frequencies than ultrasound to index the stiffness of the fruit rather than to create an image. Typically an impact or an electromechanical device is used to produce acoustic vibration, with assessment of the frequencies best transmitted through the sample, the 'resonant' frequencies. The stiffness index is calculated from the resonant frequency, the mass of the sample and an index that describes the shape of the sample. A number of researchers have reported correlation between the stiffness index and firmness of the fruit, or presence of internal defects such as voids or bruises, using acoustic methods (Jamal, 2012). For example, the Aweta unit (AFS, Aweta, Nootdrop, The Netherlands) creates vibration using the impact of an electro-magnetically activated plunger and is used for firmness assessment of avocado (Arpaia, 2007 ), apple and tomato (Ramos- Garcia et al., 2005 ). A microphone records the frequencies of transmitted sound to detect the dominant frequency. Subedi and Walsh (2009) reported use of a device that also

induces vibration by impacting the sample, but with measurement of transmission time of the vibration across the tissue to assess the fruit firmness. In another approach, vibration is applied to the sample using a sinusoidal wave vibration from 10 Hz to 2 kHz in a sweep mode, with a detector on the opposite side of the sample used to detect the dominant frequency (Applied Vibro-Acoustics http://www.ava.co.jp/kataroguEnglish.pdf). This technique is claimed to be useful in assessment of fruit texture (Ramos- Garcia et al., 2005 ) and trialled in kiwifruit firmness measurement by Muramatsu et al. (1997). These techniques have seen very limited use in pack houses due to the need for contact to the sample, and to interference from other vibrations in the packing house.

Near infra-red spectroscopy was developed for assessment of quality of dry agricultural produce, beginning with grain and (dry) forage materials in the 1980s, with subsequent extension to a wide range of industrial, pharmaceutical and medical applications. The technique is based on the absorption of frequencies matching that of the stretching or vibration of chemical bonds such as C-H and N-H. This technique is focus for this research study and is reviewed in greater detail in a subsequent section.

#### *1.2.4.2 Case study: citrus*

Various non-invasive measures have been used for quality evaluation of citrus. Zheng et al. (2010) reported use of Vis-SWNIR reflectance spectroscopy for predicting oleocellosis (a skin disorder in citrus fruit leading to development of surface blemishes), with the oleocellosis disorder associated with lower absorbance over the range 400-1000 nm ($R_{cv}$ = 0.98, RMSECV = 0.0015 and $R_p$ = 0.97, RMSEP = 0.035). Lorente et al. (2015) explored use of laser light backscatter imaging method for the early fungal (*Penicillium digitatum*) decay detection of citrus. An average correlation coefficient of determination of 0.99 was reported, given inputs at 532, 660, 785, 830 and 1060 nm. Kawano et al. (1993) introduced SWNIRS for evaluation of the sugar content of Satsuma mandarin.

In summary, various technologies are in use for detection of internal defects in different fruit species (Table 1.4), though very few of these have achieved commercial use. For example, the potential health hazards and cost of X-ray imaging

limit the adoption of this technology, while cost and speed of assessment are limits to adoption of magnetic resonance imaging.

### *1.2.4.3 Detection of internal defects of fruit using NIR*

There is a substantial scientific literature on the field of NIR and fruit quality. For example, in the 2010-13 period there were 900 papers associated with the terms 'fruit internal quality, of which approximately 10% were associated with 'apple', and 5% with 'apple internal defect'(Fig. 1.8).

NIR spectroscopy is so far the only method for non-invasive assessment of internal quality in commercial use (e.g. packline manufacturers Greefa, www. greefa.nl/UK, MAF-Roda, www.maf.com; and Compac-Taste Technologies, www.taste-technologies.com), primarily in context of sorting on fruit SSC or DM content. Within recent years the Taste Technology website has specifically mentioned the detection of water core and browning in apple and internal browning in kiwifruit while Greefa claims technology available to detect water core and apple browning with its 'Combisort' (http://www.greefa.nl/UK/products-grading-machines-combisort.htm) and 'Geo sort' (http://www.greefa.nl/ UK/ products-grading-machines-geosort.htm) products. However, there is no characterisation of these systems or validation of their efficacy in the scientific literature.

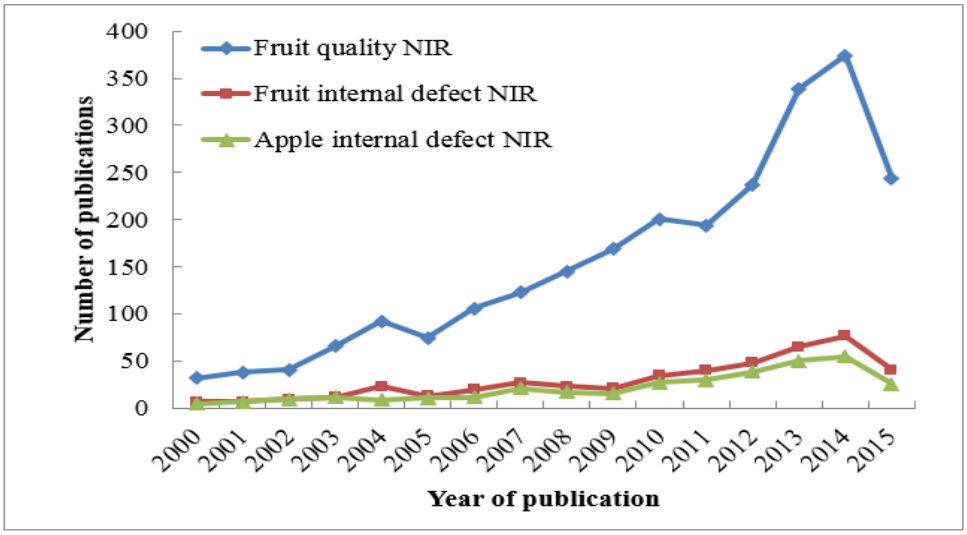



**Figure 1.8.** Number of refereed publications associated with key words 'fruit quality NIR', 'fruit internal defect NIR' and 'apple internal defect NIR'. Data sourced from Science Direct on May 15, 2015. Note use of a different scale for keywords 'fruit quality and NIR'.

### *1.2.4.4 NIR spectroscopy*

*Electromagnetic spectrum*

The electromagnetic spectrum (Fig. 1.9) encompasses all frequencies of electromagnetic radiation. Human eyes are sensitive to a small portion of this spectrum, 400-700 nm but it is this portion that most of the energy of sunlight at earth's surface is present.

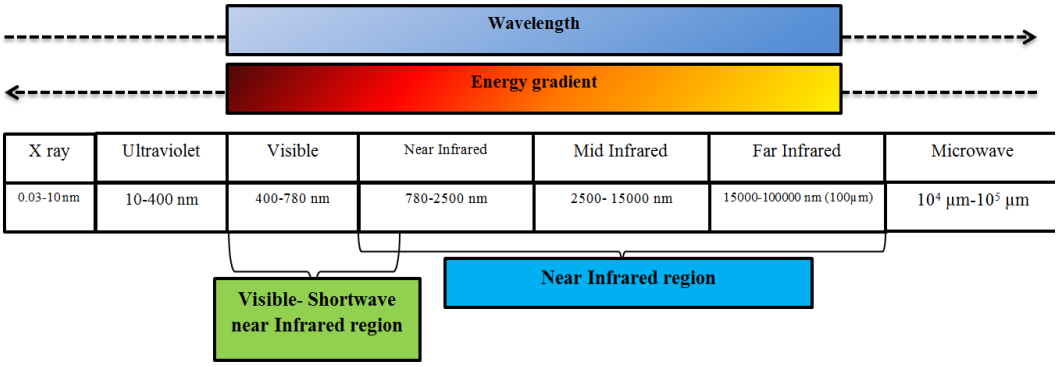



**Figure 1.9.** A diagram of the electromagnetic spectrum, showing various properties across the range of frequencies and wavelengths.

The presence of energy beyond the red part of the spectrum was discovered by Friedrich Wilhelm Herschel in 1800 A.D. The American Society of Testing and Materials (ASTM) have defined the NIR region of the electromagnetic spectrum as the wavelength range of 780– 2526 nm. In practice this region is divided by detector function, with Silicon based (charged coupled devices, CCD or photodiode arrays PDA) detectors sensitive from 450 to 1100 nm, and lead sulphide (PbS) detectors sensitive from 1100 to 2500 nm. More recently InGaAs detectors have become available, typically functioning from 800 to 2200 nm. Thus the term NIR is generally taken to infer the range 1100 to 2500 nm, while the term short wave near-infrared (SWNIR) refers to operation within the Herschel region of the spectrum, from 780-1100 nm. This wavelength region of the electromagnetic spectrum is of particular importance as water O-H bonds are strong absorbers. As fruit contains above 80% water, absorption around 1850-1920 nm and 1400-1450 nm, associated with the first and second overtone bands of the O-H will be very strong such that the effective path length is a few millimetres. The absorption coefficients of the third and fourth

overtone stretch feature, at approximately 1150 and 930 nm, are lower, allowing effective penetration of biological tissue, albeit at the expense of peak broadening and thus overlapping of spectra features (Lin & Ying, 2009). The three absorption peaks of 760 nm, 840 nm and 970 nm are associated with water or OH group (Clark et al., 2003).

*Principles of Spectrophotometry*

When electromagnetic radiation penetrates an object, it subject to scattering and absorption processes. Incident radiation can be either reflected, absorbed or transmitted. The proportion of each phenomenon depends on the chemical composition and physical parameters of the sample. The scattering properties of light are a physical phenomenon and are dependent upon the changes in refractive index within the sample while absorption of light depends on the chemical composition of the product (Fu et al., 2007; Nicolaï et al., 2007).

Radiation of certain frequencies will be absorbed by particular chemical bonds present in the biological materials (e.g. an O-H bond within sugar or water), causing stretching of the bond. The absorption difference is governed by the magnitude of dipole change during the displacement of atoms and its anharmonicity during vibration (Pasquini, 2003). High anharmonicity and displacement is observed in presence of bonds with hydrogen atoms with heavier elements like oxygen, carbon, nitrogen and sulphur in O-H, C-H, N-H and S-H bonds. The response to incident radiation depends on the match of radiation energy with the difference in energy level of the various states of the bond. This will lead to absorption or non-absorption of radiation frequencies at particular wavelength. For example, aromatic hydrocarbon (Ar-OH) of phenolic compounds shows high absorbance at 750 nm and 1000 nm (Fu et al., 2007). The series of absorption features for the mix of bonds present in a sample results in the absorption spectrum of the sample under consideration, and its spectral ‘signature’.

*Acquisition Modes – optical geometry*

Different optical geometries should be used for acquisition of spectra depending on the nature of the sample under consideration. Typical geometries are reflectance, interactance and transmission (Fig. 1.10). Some of the incident radiation reaching an

object will be reflected. Reflected light will consist of specular reflection and external diffuse reflection. Specular reflections are mirror-like, and the light carries no information on the sample it was reflected from. Diffuse reflection represents light that has penetrated to some depth into the sample and re-emerged and thus carries information about the sample. For an optical geometry based on use of diffuse reflection, the detector is typically set at an angle of $45^{o}$ relative to the light source – object axis, to minimise detection of specular reflection (Fig. 1.10A). Full transmission geometry is suggested for parameters deeply seated in the sample like internal defects (Fig. 1.10C), while the information on skin colour can be obtained better by a reflectance geometry. Transmittance is the fraction of incident light (electromagnetic radiation) at a specified wavelength that passes through a sample. Issues associated with internal structure such as the presence of a large seed or difficulties in detecting low levels of transmitted light can be addressed using a partial transmittance optical geometry (Fig. 1.10C). An interactance geometry is a special case of a partial transmission geometry in which the light source and detector are typically set parallel to each other in a way that light due to specular reflection cannot directly enter the detector but the re-emitted light from sample is detected (Fig. 1.10C).

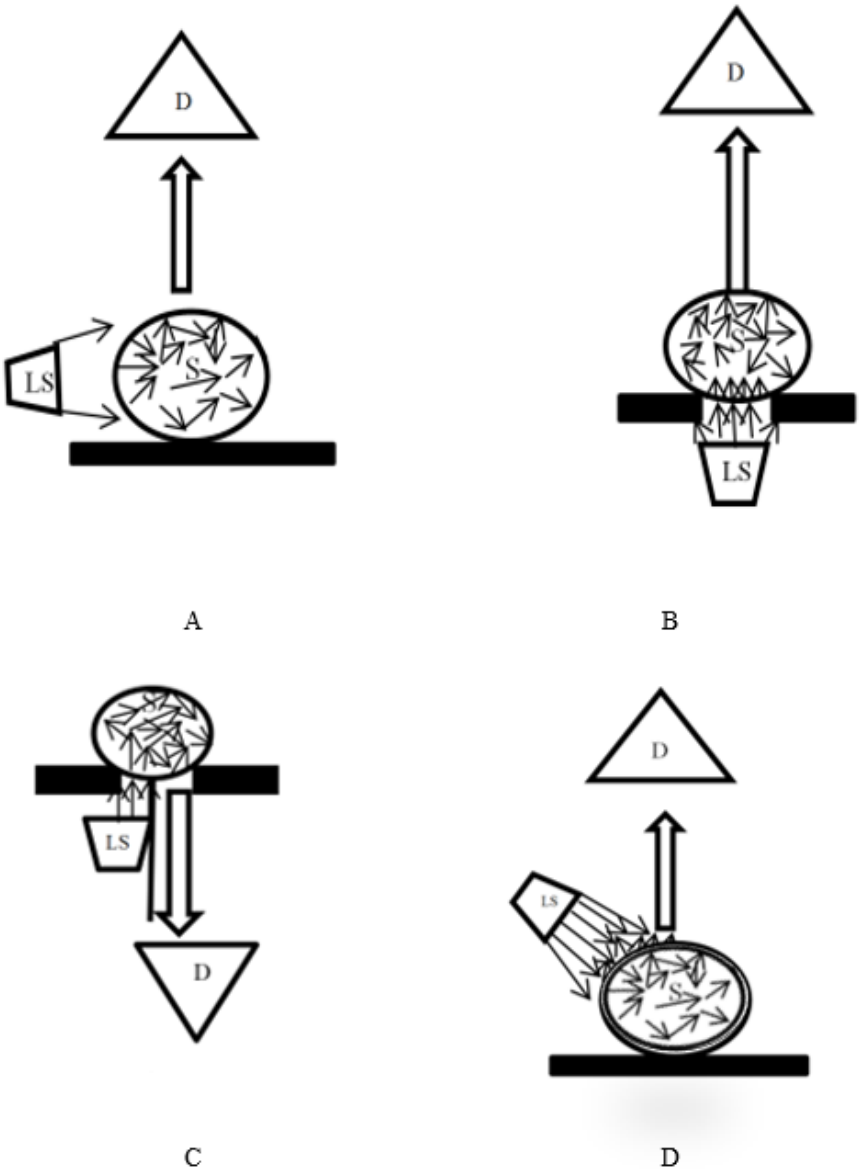



**Figure 1.10.** Optical geometries for the acquisition of spectra from samples, involving a light source (LS), sample (S), and detector (D) in A. Partial transmittance B. Full Transmission and C. Interactance and D. Reflectance optical geometries (modified from Nicolai et al., 2007).

### *1.2.4.5 Chemometrics*

#### *1.2.4.5.1 Pre-treatment techniques*

Chemometrics aims at extracting hidden information from chemical systems by data-analysis. The NIR spectrum of a fruit contains information on multiple variables, including chemical variables such as water content and sugars, physical variables such as light scattering, and non-sample variables such as instrumental noise, and effect of external temperature and light (Nicolaï et al., 2007). Chemometrics

approaches can also be taken to deal with skewed data sets, using data de-noising and smoothing pre-treatment procedures. A range of spectral pre-treatments methods are available to remove the influence of variables such as scattering. Chemometrics can thus give important qualitative and quantitative information to best describe the parameter(s) under consideration.

Two broad categories of pre-treatments methods are employed in spectra pre-processing in NIR spectroscopy namely scatter correction and derivatization (Rinnan et al., 2009). Scatter correction includes standard normal variate (SNV), multiplicative scatter correction (MSC), Inverse MSC, extended MSC (EMSC), extended inverse MSC, de-trending, and normalization while derivatization consists Norris-Williams (NW) derivatives (gap segment) and Savitzky-Golay(SG) polynomial derivatives with different data points. Other treatments methods include size and weight correction (by dividing the spectra features for particular sized/weight fruit with average spectra). Mean centering and auto scale are some other pre-treatment techniques.

Some of commonly used pre-processing techniques are briefly discussed here:

*Derivatives*

Derivatization is undertaken done to overcome effect of baseline shift and to separate peaks of overlapping bands. Calculation of a derivative involves use of a 'window' of wavelength data. For example, calculation of Savitzky Golay (Savitzky and Golay, 1964) derivative requires the setting of the number of data points used for polynomial fitting as well as the polynomial order. A higher number of data points used in derivatization will led to creation of smooth spectrum but potentially at the loss of useful information. Derivatization can result in magnification of noise and can complicate interpretation of results in terms of absorbance features.

*Standard Normal Variate (SNV)*

This transformation removes variation caused by differential scattering related to particle sizes (Rinnan et al., 2009). The SNV involved auto scaling of each spectrum individually by deducting the mean spectrum not by entire sample set, which in

effect, removes the scatter effect from spectra. SNV does not use any average spectrum in the set.

*Multiplicative Scatter Correction (MSC)*

MSC is a scatter correction method by use of least squares regression of a given spectrum to the ideal spectrum (commonly the averaged spectrum). This involves the use of slope and intercept of the best line of fit to correct the sample spectra (Martens & Naes, 1989; Rinnan et al., 2009). MSC pre-treated spectra, in many cases, look similar to SNV treated spectra. Many chemometricians prefer SNV over MSC as MSC is set dependent, and thus if sample membership changes, the mean spectrum will change and a recalculation is required (Dhanoa et al., 1994).

*Spectral truncation*

Spectral truncation involves optimisation of the wavelength range used in analysis, to involve the assessment of wavelengths carrying information related to the attribute of interest, and avoiding inclusion of irrelevant wavelengths that contribute only noise to the analysis (Huang et al., 2010). For example, wavelength range of 729-975 nm was recommended for estimation of SSC and DM in fruit (Golic et al., 2003 ).

1.2.4.5.2 *Quantitative predictions*

Partial least square regression (PLSR) is the most commonly employed algorithm for prediction of a continuous variable in a sample from spectral data, although other techniques are also used (e.g. multiple linear regression, MLR, support vector machines, SVM).

PLSR is a statistical tool used to regress a large set of variables and combines features of multiple linear regression and principal component analysis (Herve, 2003 ). PLSR involves reduction of the spectral data set to a few latent variables (LVs) or principal components (PCs) that are then regressed on the attribute values. This regression technique utilises the reference values in calculation of PCs. The end result is a model with a weighting on absorbance at each wavelength in the chosen region.

PLSR model development typically involves the following steps (Conzen, 2003 ):

i. Sample selection for model development: Sample selection should represent the range of variation that could be observed in given population. Sample size can vary from 30 to thousands but at least 10 samples per principal component and more samples than the independent variables is a generic rule. Many literature reports are based on over-fitted models, wherein apparently good calibration results are achieved by use of too many principal components (Al-Najjar & Pai, 2014; Andreassen, 2015; Subramanian & Simon, 2013).
ii. Prediction set: A test set different from the calibration set, representing future samples, should be used. Many studies fail to use an independent test set, but rather go to some length to select a test set that is matched to the calibration set, with all samples drawn from a common population.
iii. Spectra acquisition from samples
iv. Spectral pre-treatment: Appropriate spectra pre-treatments are applied to the spectra before making a regression model, e.g. standard normal variate with or without derivatization.
v. Wavelength selection: An appropriate wavelength region should be selected, e.g. based on spectroscopy knowledge, univariate correlation across the available wavelength range with attribute level, or a wavelength range optimisation procedure using multiple range settings in the multivariate regression method of choice.
vi. Regression: Pre-treated spectra are regressed against the reference values, with cross validation in which a number of samples are sequentially left out from the calibration set as a validation set, and the resulting model is tested on the validation set. Cross validation set selection will influence the regression outcome, with the 'leave one out' procedure being least onerous.
vii. Outlier removal: A fault in reference assessment or spectral acquisition or instrumental noise for a sample will increase error in the model if included in the calibration modelling. These outliers should be removed from the data set, but ideally only where the cause is known. Indiscriminate removal of data that does not fit the model will produce apparently good calibration statistics but a model that is not robust in use with future populations.

viii. Testing the model: The model should be tested using an independent test set.

The most commonly used statistics in multivariate data analysis are:

Correlation coefficient (r or R) and Correlation Coefficient of determination ($r^2$ or $R^2$). Calibration statistics may be distinguished from cross validation or prediction statistics using the notation r and R, respectively, or by use of subscripts c, cv and p (e.g. $R^2_{cv}$).

***RMSEC*** (Root Mean Square Error of Calibration), *RMSECV* (Root Mean Square Error of Cross Validation) and ***RMSEP*** (Root Mean Square Error of Prediction) are estimates of error in calibration, cross validation and prediction, respectively. Prediction can involve a bias, an average increase or decrease in set value, and so a bias corrected RMSEP (SEP or RMSEPbc) is also a useful statistic.

$$RMSECV = \sqrt{\frac{\sum (p-a)^2}{n-1}}$$

$$RMSEP = \sqrt{\frac{\sum (p-a)^2}{n-1}}$$

$$SEP = \sqrt{\frac{\sum (p-a-b)^2}{n-1}}$$

$$bias = \frac{\sum p-a}{n}$$

where p = predicted value, a = actual value, n = number of samples, b = bias

There also exists a relationship among $R^2$, RMSEP and SD. The derivation of this relationship is as follows:

$R^2$ = ratio of the total variance to the explained variance

$R^2$ (coefficient of determination) is the ratio of total variance to explained variance, expressed mathematically as:

$$R^2 = \frac{1 - SSR}{SST}$$

where, SSR is the sum of squared residuals, and SST is the sum of squared total variance.

$$R^2 = 1 - \frac{\sum (p-a)^2}{\sum (a-\mu)^2} \qquad \text{Equation 1}$$

where μ is the mean value of analyte.

We also know:

$$RMSEP^2 = \frac{\sum (p-a)^2}{n} \qquad \text{Equation 2}$$

$$SD_y = \frac{\sum (a-\mu)^2}{n} \qquad \text{Equation 3}$$

Substituting equation 1 with values equation 2 and 3:

$$R^2 = 1 - \frac{n(RMSEP)^2}{n(SD_y{}^2)}$$

$$R^2 = 1 - \frac{RMSEP^2}{SD_y{}^2} \qquad \text{Equation4}$$

There exists a relationship between RMSEP, bias and SEP as shown below:

$$RMSEP^2 = Bias^2 + SEP^2$$

SDR: This is ratio between standard deviation (SD) and RMSECV/RMSEP of an parameter of a set of samples either used in calibration or prediction.

$$SDR = \frac{SD_y}{RMSECV} \text{ or } SDR = \frac{SD_y}{RMSEP}$$

RPD: This is the measure of ratio of bias corrected standard error of prediction (SEP) to sample standard deviation. RPD for calibration and prediction results are reported as RPDc and RPDp.

$$RPD = \frac{SD_y}{SEC} \text{ or } RPD = \frac{SD_y}{SEP}$$

*1.2.4.5.3 Discriminant predictions*

A range of methods are used for discrimination of groups based on spectral data, including principal component analysis (PCA), linear discriminant analysis (LDA), PLS-discriminant analysis (PLSDA), PCA linear discriminant analysis, support vector machine classification (SVM), k nearest neighbourhood (kNN), soft independent modelling of class analogy (SIMCA) and logistic regression.

Principal Component Analysis (PCA): This procedure converts a set of spectral data into linearly uncorrelated variables named Principal Components, where the first principal component accounts for the highest possible variances in the spectral data. The PCs can be used to describe the difference of a given sample to that of an existing set.

Linear Discriminant Analysis (LDA): LDA is a pattern recognition technique involving supervised classification involving pre-specified classes. This classification tool involves a reduction in the dimension of variation in the given sample while retaining discriminatory information, and uses a linear combination of features that enables classification to given classes or groups (Zhang et al., 2006). For example, Vanoli et al., (2014) report use of linear discriminant (canonical) analysis for classification of Braeburn apples with internal defects based on time resolved reflectance spectroscopy estimates of absorbance and scattering at 670 nm and 740-1040 nm. A classification accuracy of 89.7% for healthy fruit, 42.7% for brown core fruit and 75% for brown pulp fruit was reported.

K Nearest Neighbours algorithm (kNN): kNN is one of the non-parametric tests used for classification. In this method the number of nearest neighbours in a given feature

space is calculated for the sample with reference to the training set (Gongal et al., 2015). In the example Fig 1.11, the unknown (?) has 2 neighbours of class 1 (red triangles) and 3 neighbours of class 2 (purple squares). For a k value of 2, the unknown has more neighbours in class 1 than 2 and the sample will be assigned to class 1. If we assign the k value of 3, this unknown sample will fall in class 2. Higher values of k are recommended to reduce noise in classification.

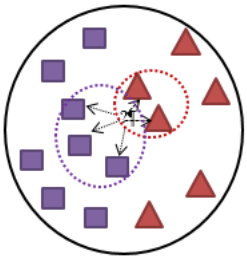


**Figure 1.11.** An explanation of k-NN decision based on nearest neighbours numbers
Source: Mucherino et al. (2009)

Partial Least square – discriminant analysis (PLS–DA): Partial least square regression discriminant analysis (PLS-DA) uses the PLS linear classification technique in combination the discrimination ability of a classification technique. In essence, a PLS regression is carried out in which groups are assigned unit value (e.g. 0 and 1). The model is used to predict the class of the unknown, using a threshold value (e.g. 0.5) to differentiate samples to classes.

Support vector machine discriminant analysis (SVM-DA): This is a supervised classification technique using pattern recognition. Samples are arranged in a different feature space separated by a hyperplane for classification (Khanmohammadi et al., 2014). Samples lying on the borderline are named support vectors and are used in designing rules for classifying rest of samples. This classification method is considered suitable for heterogeneous classes with overlapping features. For example, the technique has been used in discriminating animal feed contaminated with toxic material and for separation of plastic types in recycling operations, based on NIR spectral information (Abe, 2010).

Soft independent model of class analogy (SIMCA): Soft independent model of class analogy (SIMCA) is based on use of a principal components model, with

discrimination based on PCA model scores of the unknown and training sample. SIMCA is considered more suitable where classes do not overlap.

Logistic regression: This is a regression algorithm for ordinal variables (class or groups) which calculates the probability of a sample belonging to a given class/category based on independent variables (Bielza et al., 2003; Kleinbaum & Mitchel, 2002; Lammertyn et al., 2000). Logistic regression can be binomial (only two class/categories) or multinomial (more than two categories) (Schein & Ungar, 2007 ). For example, Yang et al. (2012) used this technique with a training set 1378 samples, reporting a classification accuracy of more than 95% for discrimination of mature and immature fruit, and leaves of blueberry based on multispectral imaging using UV-Vis-NIR regions.

#### *1.2.4.6 Sorting of fruit*

The terms 'sorting' and 'grading' are interchangeable terms that refer to the segregation of material into quality categories, be that a binary sorting operation for removal of fruit that do not meet a specification, or assignment of consumable fruit into several subgroups based on a set of specifications (Bollen & Prussia, 2014). In the case of fruit with internal defect, the primary requirement is to separate affected from sound fruit, that is, consumer acceptable from non-acceptable fruit. If unsorted fruit is marketed, the consumer becomes a discriminant classifier, accepting or rejecting fruit. As discussed above, various technologies may be employed for non-invasive detection of internal defect. However, any sorting operation will have measurement error, resulting in misclassifications. If the classifier (e.g. defect level) is a continuous value, a threshold value must be set to act as a boundary between classes. As the threshold value is varied, the number of 'true' and 'false' positive identifications will vary. There are four possible outcomes from a binary classification, a positive identification that is either true (TP) or false (FN), or a negative identification that is either true (TN) or false (FP). Several evaluation indices have been used based on these values (Table 1.5).

Practical operation requires the operator to manage a compromise between classification accuracy and yield of the operation. There are several tools used in other fields that hold promise for use in fruit classification.

**Table 1.5.** Definition of terms and indices used in binary classification.

| Terms | Nomenclature | Description |
|---|---|---|
| Total positive | P | Total number of actual good samples in the given population |
| Total negative | N | Total number of actual defect samples in the given population |
| True positive | TP | actual good sample predicted as good by model |
| True negative | TN | actual defect sample predicted defect by model |
| False negative | FN | actual good sample predicted as defect, Type II error |
| False positive | FP | actual defect sample predicted good, Type I error |
| Prevalence | | P/(P+N) |
| False discovery rate | FDR | measure of total defect out of total predicted good, FP/(FP+TP) |
| False omission rate | FOR | measure of total good out of total predicted defect, FN/(FN+TN) |
| False positive rate (Hit rate) | FPR | false positive divided by condition negative, FP/N |
| False negative rate (Miss rate) | FNR | false negative divided by condition positive, FN/P |
| True positive rate (Sensitivity or recall), positive predictive value | TPR, PPV | true positive divided by condition positive, TP/P |
| True negative rate (Specificity) or negative predictive value | TNR, NPV | true negative divided by condition negative, TN /N |
| Accuracy | ACC | true positive plus true negative divided by total population, TP+TN/P+N |
| Positive likelihood ratio | LR+ | TPR/FPR |
| Negative likelihood ratio | LR- | FNR/ TNR |
| Diagnostic Odd ratio | DOR | LR+/LR- |

*Receiver operating characteristic (ROC) curve*

A 'receiver operator characteristic' or ROC (also known as a Receiver Operating Curve) can be used to compare the performance of binary classifiers. This curve is a plot of the true positive rate (sensitivity, d-prime or recall) against the false positive rate (fall-out or 1 – Specificity) at various threshold settings (Santos-Pereira & Pires, 2005). The ROC plot was developed by the British army in WWII for comparison of operators interpreting radar signals, with true positives being incoming German planes and false positives associated with flocks of birds or other events. The plot was later used extensively in psychology test interpretation and more recently for medical diagnostic tests based on pathological, biomarker or imaging results (Hanley & McNeil, 1982).

The ROC curve is equivalent to a plot of the cumulative distribution of the detection probability (y axis) against the cumulative probability of the false alarm probability (x axis), which can be calculated if the probability distributions for TP and FP are known. The best result for a binary classifier is 100% TPR and 0% FPR (perfect classification, i.e. a point at 0,100 on the ROC). A random assignment results in a point on the diagonal line (line of no-discrimination). A result below the line has discriminatory power, but the predictions must be reversed (shifting the result above the line of no-discrimination). Several summary statistics are used to describe the overall performance of a ROC curve (e.g. intercept of a line at $90^{o}$ to both ROC curve and no-discrimination line, known as Youdens J statistic; area between ROC curve and no-discrimination line; area under ROC curve, known as AUC or A' or c-statistic; or d', the distance between the means of distribution for the native and sorted populations, divided by their respective standard deviations, assuming normal distributions). The AUC is equivalent to the probability that a classifier will rank a positive event higher than a negative event (Fawcett, 2006; Mason & Graham, 2002 ) and is related to Mann-Whitney U and the Wilcoxon test of ranks (Fawcett, 2006). However such summary statistics lose information on the shape of the ROC curve, i.e. the pattern of trade-offs.

A body of work exists around classifier training techniques to optimise AUC, particularly for the case where the training set is large, so that computation time becomes significant. Calders and Jaroszewicz (2007) describe techniques which use

differentiable algorithms to estimate AUC as 'soft-AUC', and propose a technique based on a polynomial approximation of the AUC. Other techniques have been proposed, e.g. a support vector machine learner (Joachims, 2005). Calders and Jaroszewicz (2007) also advocate a method that assigns more than a simple binary score for points close to the threshold border, such that points with a small difference (predicted value – threshold value) are weighted lower than those with large difference, as occurs in calculation of mean squared error.

As well as use for comparison of classifiers, a ROC curve can be used in the choice of the threshold value for a given classifier. This threshold value may be chosen to maximise both sensitivity and specificity, i.e. the upper left point of the curve, however in practice the optimal threshold value depends on the application demands for sensitivity and specificity. In practical terms, there are times in a sorting operation when the FNR (presence of defect items in the accepted category) is critical to the value chain, even at the expense of a high miss rate (FN/P), while at other times the operator may desire to maximise pack-out (TP+FN) at the expense of a higher false discovery rate FP/(FP+TP).

Ooms et al. (2010) note that the ROC is well suited to the diagnosis of disease, for which the impact of mis-classification is large, but they suggest it is less relevant to the sorting of materials. As an alternative, the detection error trade off (DET) graph gives more attention to the area of interest of the ROC curve (i.e. region of minimal FPR and maximal TPR, top left region of plot). The DET plots missed detections (false negative rate, FNR) against false alarms (false positive rate, FPR) using x and y scales transformed by the quantile function of the normal distribution (the inverse of the cumulative normal distribution).

Ooms et al. (2010) also describe compromise between yield and a quality factor to produce a 'sorting optimisation curve' (SOC) to guide the choice of the threshold value used in the classifier. In this approach, yield refers to (TP+FP)/(P+N), while quality refers to TPR with an assigned value (e.g. price) to sorted objects based on the proportion of defect items present in the lot.

Given the apparent relevance of the use of ROC and its derivatives, there are surprisingly very few reports of application to fruit sorting as discussed by Bollen

and Prussia (2009, 2014). An ROC curve was developed for the application of detection of the apple bruises based on spectral information over the range 380-1000 nm by (Luo et al., 2012). Reflectance difference between 808 and 760, 832 and 772, 834 and 762 and 788 and 742 nm was recommended for bruise detection in cultivars Fuji, Jonagold, Orin and Sinano Gold, respectively, based on the AUC character. Siedliska et al. (2014) reported use of the AUC of a ROC curve for comparison of binary classifiers for bruise detection based on the hyperspectral imaging data. The binary classifier used was reflectance at 718-890 nm and 1017-1118 nm where reflectance of non-bruised (normal) fruit was higher than bruised one. The accuracy of bruise detection was as high as 95% and 90% for calibration and prediction sets respectively.

*Example data*

A simple explanation of the ROC curve is presented for the case of sorting of defect fruit. Good fruit may be predicted as good (True positive, TP) or defect (False negative, FN). Similarly defect fruit may predicted defect (true negative, TN) or false positive (FP). Let us assume an arbitrary population of fruit comprising 100 good and 100 defect fruit, noninvasively assessed for the degree of defect on a 1 to 5 (nil to severe defect). Let score 1 and 2 represent acceptable fruit, while 3 to 5 are unacceptable. Given the measurement technique has error, it can be seen that setting the sorting threshold at 2 will result in a level of errors. Increasing the threshold will result in more Type 1 errors (FP), while decreasing the threshold will result in more Type II errors (FN). An example data set is given in Table 1.6. In this data set, for a threshold value of 3.0, the distribution of true positive, false negative, false positive, true negative are 88, 12, 19 and 81, respectively

Sensitivity (S, also called true positive rate, TPR), a measure of how well the good fruit were predicted, is calculated as TP/P. In this example, Sensitivity = 88/100 = 0.88. Other parameters are documented in Table 1.6.

**Table 1.6.** Example data set of a sorting operation involving prediction of defect level and a sorting threshold value varied between 0.5 and 5, with calculations of various descriptive parameters based on these data.

| | Threshold value | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Descriptive parameters | 0.5 | 1 | 1.5 | 2 | 2.5 | 3 | 3.5 | 4 | 4.5 | 5 |
| True positive (TP) | 10 | 35 | 52 | 72 | 84 | 88 | 92 | 95 | 97 | 100 |
| True negative (TN) | 100 | 100 | 100 | 100 | 98 | 81 | 65 | 35 | 22 | 0 |
| False negative (FN) | 90 | 65 | 48 | 28 | 16 | 12 | 8 | 5 | 3 | 0 |
| False positive (FP) | 0 | 0 | 0 | 0 | 2 | 19 | 35 | 65 | 78 | 100 |
| Prevalence | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| False discovery rate (FDR) | 0 | 0 | 0 | 0 | 0.02 | 0.18 | 0.28 | 0.41 | 0.45 | 0.5 |
| False omission rate (FOR) | 0.47 | 0.39 | 0.32 | 0.22 | 0.14 | 0.13 | 0.11 | 0.13 | 0.12 | NA |
| False positive rate (FPR)/(Hit rate) | 0 | 0 | 0 | 0 | 0.02 | 0.19 | 0.35 | 0.65 | 0.78 | 1 |
| False negative rate (FNR)/(Miss rate) | 0.9 | 0.65 | 0.48 | 0.28 | 0.16 | 0.12 | 0.08 | 0.05 | 0.03 | 0 |
| True positive rate (Sensitivity or recall), TPR | 0.1 | 0.35 | 0.52 | 0.72 | 0.84 | 0.88 | 0.92 | 0.95 | 0.97 | 1 |
| True negative rate (Specificity) TNR | 1 | 1 | 1 | 1 | 0.98 | 0.81 | 0.65 | 0.35 | 0.22 | 0 |
| Accuracy (ACC) | 0.55 | 0.68 | 0.76 | 0.86 | 0.91 | 0.85 | 0.79 | 0.65 | 0.60 | 0.5 |
| Positive likelihood ratio (LR+) | NA | NA | NA | NA | 42.00 | 4.63 | 2.63 | 1.46 | 1.24 | 1 |
| Negative likelihood ratio (LR-) | 0.9 | 0.65 | 0.48 | 0.28 | 0.16 | 0.15 | 0.12 | 0.14 | 0.14 | NA |
| Diagnostic Odd ratio (DOR) | 0 | 0 | 0 | 0 | 257.3 | 31.3 | 21.4 | 10.2 | 9.1 | NA |

Specificity (or true negative rate, TNR) is measured as TN/N. Sensitivity (or true positive rate) is measured as TP/P. Adjustment of the threshold value will alter classification accuracy, both in terms of sensitivity and specificity. In this example, for a threshold of 3, Specificity = 81/100 = 0.81 and Sensitivity is 0.88 (Table 1.6).

The Receiver Operating Characteristic (ROC) curve is obtained using the 1-Specificity and sensitivity values for a range of threshold values (Fig. 1.12). For a specificity of 0.81, false positive rate (1 - Specificity) = (1-0.81) = 0.19. In this example, the threshold is varied from 0.5 to 5 for fruit scored on a 1-5 defect scale

(Table 1.6). The ideal ROC would track to the top left (ie. to 0,1) of the ROC (Fig. 1.12).

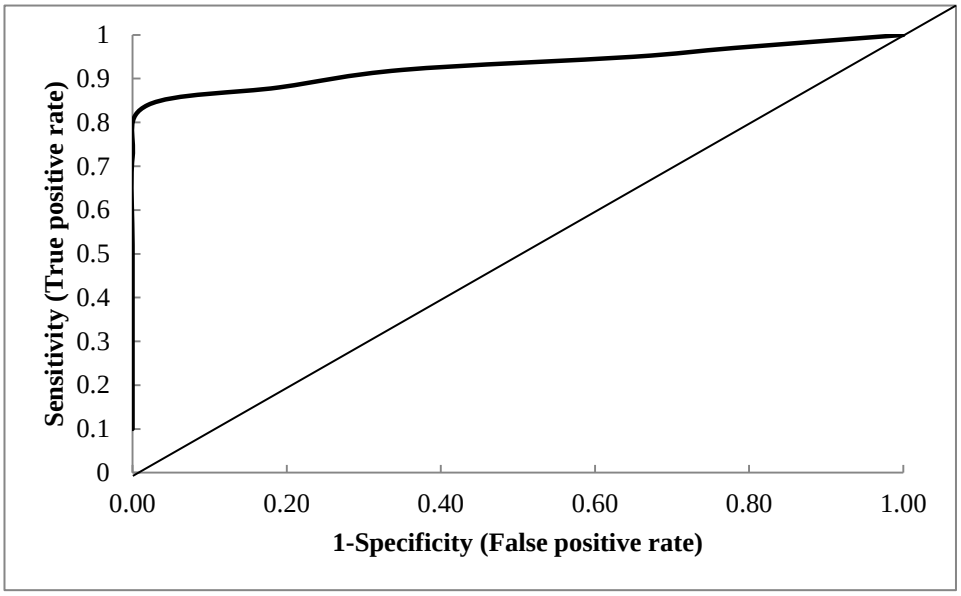



**Figure 1.12.** Receiver operating characteristic (ROC) curve for an arbitrary population plotting the true positive rate and false positive rate.

A discriminant function may also involve a threshold value, being the value that separates the two classes during prediction and yields a classification.

A Sorting optimisation curve (SOC) is intended to aid the decision in sorting operation. The operator of a sorting operation on defect fruit must compromise between yield (TP+FP/P) and false discovery rate (FP/FP+TP), with control exerted through selection of a threshold value. Thus, yield and false discovery rate can be plotted against Threshold value (for the example data set, Fig. 1.13). In this example, a threshold value of approx. 2.5 is required to achieve a false discovery rate of <2%, and this setting involves a yield of 43%. However, other parameters may also impact operator behaviour, e.g. pricing on premium and defect lines. The design of a SOC to include such factors will be considered later.

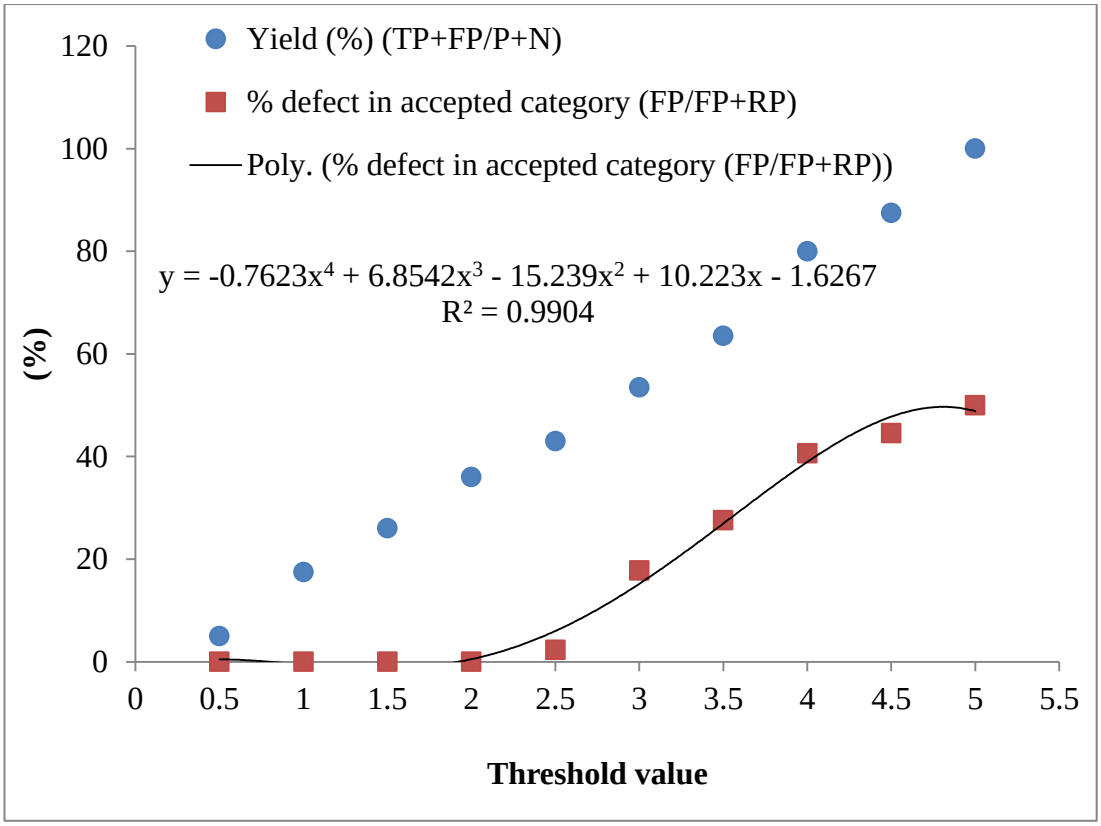



**Figure 1.13.** Sorting optimisation plot of relevant characters (yield and false reject rate) at a range of threshold values for visual score of cut surface.

Alternately, the Detection Error Trade-off (DET) curve can be used, in which the missed detection rate (false negative rate) is plotted against the false alarm (false positive rate) using x and y scales transformed by the quantile function of the normal distribution (the inverse of the cumulative normal distribution) or by logarithm base 10. This presentation expands the false reject scale in the area of interest, in this case around a value of 2%.

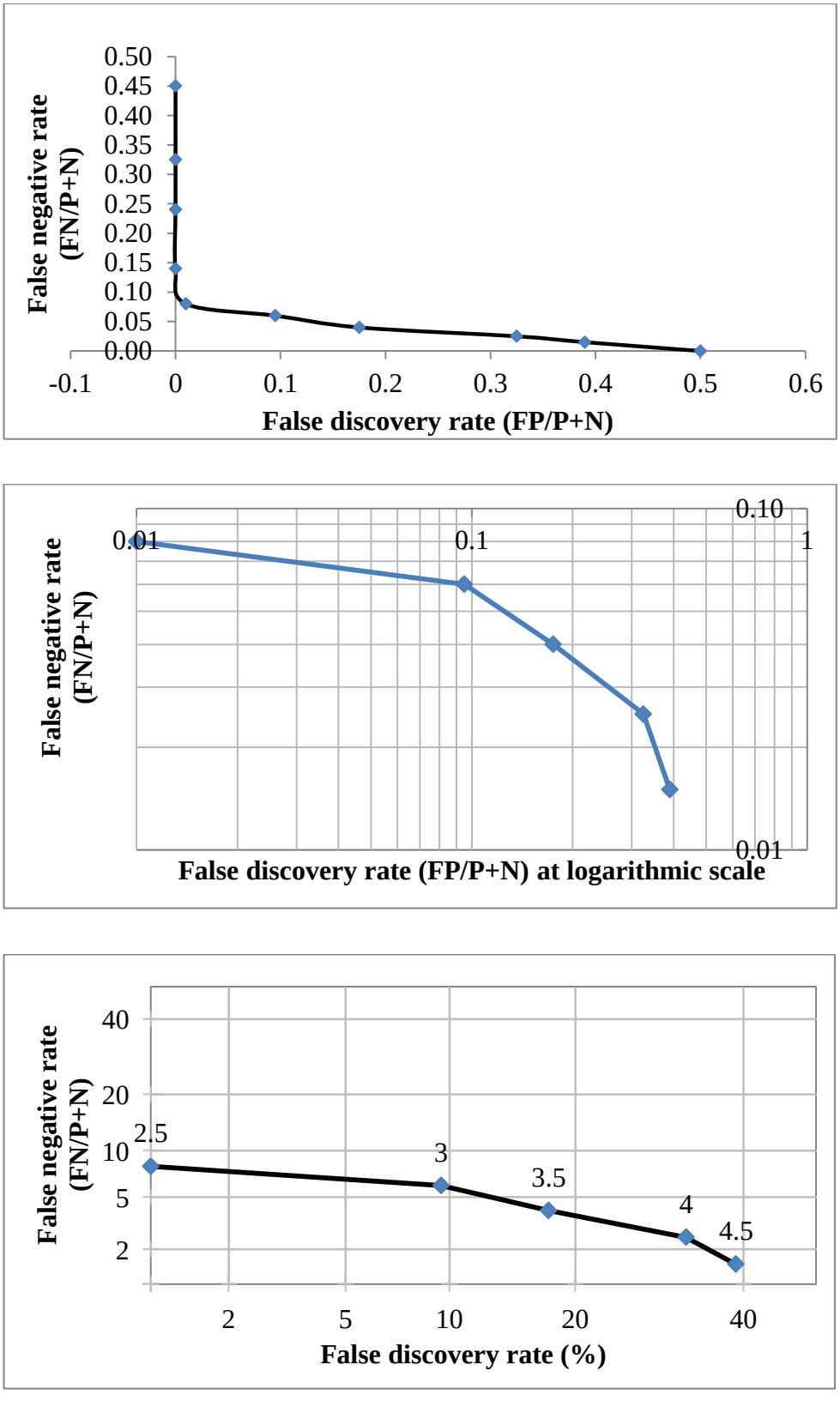



**Figure 1.14.** Detection Error Trade-off (DET) curve: plot of missed detection rate (false detection rate) against the false alarm (false positive) rate: top panel, raw scale; middle panel, log-log plot; bottom panel, quantile (probit) function plot.

## 1.3 Conclusion

A number of internal defects within fruit result in consumer dissatisfaction and thus in consignment rejection by retailers. Internal browning of apple, granulation of mandarin and translucency of pineapple are disorders of economic importance to Queensland. There is some understanding of the physiological basis of these disorders, but little ability to control their incidence by management practices. With variable occurrence comes the need for sorting to remove affected fruit.

Instrumentation for fruit internal defect detection and sorting must be rapid, accurate, non-invasive and cost effective for commercial adoption. The use of several methods have been explored over the past 50 years, including methods based on X-ray transmission, NMR, MRI, acoustics, and ultrasound. Complexity, cost, speed and reliability have limited adoption of these methods. Two methods have seen limited adoption: chlorophyll fluorescence for monitoring physiological status of fruit in storage, and NIR spectroscopy. NIR spectroscopy has been applied to the non-invasive measurement of soluble sugars, dry matter content and some other attributes of fruit. The technique usually involves a tedious process of calibration model development and seasonal updates for increasing model robustness, a process which limits widespread adoption.

Non-invasive detection and sorting of fruit on internal defects is yet limited in terms of commercial use. This thesis will explore use of visible-short wave near infrared spectroscopy to non-invasively detect the disorders of internal browning of apple, granulation of mandarin and translucency of pineapple, with attention to approaches that increase classification accuracy and that optimizing the sorting operation and provide guidance to the operator of the sorting operation.

# Chapter 2.
# Non-invasive detection of internal flesh browning of apple using visible-short wave near infrared spectroscopy[1]

## Abstract

Certain cultivars of apple are prone to an internal flesh browning defect in regular air storage and/or following extended controlled atmosphere storage. A number of (destructive) reference methods were assessed for scoring the severity of this defect in a fruit, including visual assessment, image analysis (% cross section area affected), chromameter CIE Lab values and juice $Abs_{420nm}$, of which visual scoring on a 5 point scale and a colour index based on CIE Lab were recommended. Fruit chlorophyll fluorescence and acoustic resonant frequency were not consistently related with presence of defect, and these non-invasive methods were discounted for this application. Non-invasive detection of this disorder using three instruments operating in the visible-shortwave NIR but varying in optical geometry (interactance, partial transmission and full transmission) was attempted. Quantitative prediction of defect level was best assessed using visible-shortwave NIRS in a transmission optical geometry, with a typical partial least squares (PLS) regression model $R^2_p$ = 0.83 and RMSEP = 0.63 (5 point defect score). Of 12 binary (good and defect fruit) classification approaches trialled, the best result was achieved using the PLS discriminant analysis (PLS-DA) method, followed by linear discriminant analysis and support vector machine classification. Classification accuracy ((TP+TN)/(P+N)) on an independent validation population of > 95% was achieved, with a false discovery rate (FP/(TP+FP)) of <2% using linear discriminant analysis, PLS-discriminant analysis, a support vector machine approach and a logistic regression. A simple two wavelength based discriminator was also demonstrated to support promising sorting results.

**Keywords:** model, colour index, score, optical geometry, classification

[1] This chapter is in preparation for submission to Postharvest Biology and Technology, a journal of the Elsevier group. Part of the work was presented at the International Horticultural Congress, 2014 Brisbane, Australia, as an oral paper ("Internal defect detection in fruit") and as a peer reviewed paper in the conference proceedings.

## 2.1 Introduction

Several types of internal browning are recognised in apple (*Malus domestica* Borkh.) and other pome fruit, including radial, diffuse, bulge or internal breakdown (Bergman et al., 2012; James & Jobling, 2009). The browning symptom is considered to result from membrane disruption, with consequent oxidation of polyphenols otherwise localised to the vacuole to the brown compound, quinone, or its insoluble polymer, melanin (Hatoum et al., 2014a). The browning disorder developing during ambient storage or following controlled atmosphere (CA) storage is associated with varieties with lower intercellular space content but incidence severity is influenced by both pre and postharvest factors, and their interaction (Benkeblia et al., 2011; Bergman et al., 2012; Castro et al., 2008; Castro et al., 2007; Felicetti et al., 2011; Franck et al., 2007; Hatoum et al., 2014b; Lau, 1998; Wang & Sugar, 2013). The incidence of the disorder can therefore be erratic (James & Jobling, 2009). Pre-harvest factors include nutrition (chiefly calcium and nitrogen), irrigation, harvest maturity, growing degree days and days after full bloom (Moggia et al., 2015). The main postharvest factor associated with diffuse internal browning is high carbon dioxide concentration, referred to as $CO_2$ phyto-toxicity (Ferguson et al., 1999b; Hatoum et al., 2014b), with probability of incidence increasing for fruit in controlled atmosphere storage beyond six months (Castro et al., 2007).

Internal browning is considered a ‘major defect’ by retailers, with consignments subject to rejection if more than 2% of fruit display the disorder (Woolworths, 2015). This market pressure creates demand for a technology capable of detection of the disorder in fruit, allowing for sorting to remove defect fruit.

Chlorophyll fluorescence has been widely adopted for assessment of stress in plants. The technology has been commercialised for measurement of stress of fruit in controlled atmosphere storage, particularly apples (Neuwald et al., 2007; Prange et al., 2002). It is possible that fruit developing internal browning may display increased levels of fluorescence. For example, Prange et al. (2002) report a decrease from 0.8 to 0.63 in photochemical efficiency (Fv/Fm) of apple fruit in storage as $O_2$ concentration decreased from 3.5 to 0.5% over 120 hours. Similarly, Fv/Fm of Braeburn apple stored for two weeks at 0 and 30 kPa $CO_2$ was 0.51 and 0.44, respectively (Neuwald et al., 2007). Indeed, Saquet and Streif (2002) reported use of

chlorophyll fluorescence measurements for early detection of high $CO_2$ and low $O_2$ stress in 'Conference' pear fruit. A marked decrease in chlorophyll fluorescence was noted after two weeks of storage in conditions which were associated with the development of internal tissue browning at a later stage, following prolonged CA storage (Saquet & Streif, 2002). However, the arrest in photosystem II might be observed due to any adversities in change in fruit physiology and does not necessarily imply an internal disorder.

Of technologies with potential for detection of this disorder, visible-short wave near infrared spectroscopy (SWNIRS) shows more promise than acoustic, X-ray or nuclear magnetic imaging techniques for adoption into online sorting. The first report of detection of apple water core and internal breakdown using light transmittance was produced by Francis et al. (1965), with the topic then left unreported for 30 years, until the work of Clark et al., (2003), Upchurch et al., (1997) and McGlone et al., (2005) and then others, as summarised in Table 2.1.

The work of Francis et al. (1965) was based on Richard Delicious apples, with absorbance measured in a full transmission optical geometry using a single beam Biospec spectrophotometer (high intensity grating and Dumont 6911 photomultiplier tube) at three wavelengths (740, 805 and 840 nm). The absorbance difference of 740 -805 nm was negatively correlated (R = -0.81) with severity of watercore while the absorbance difference between 840 - 740 nm was correlated with severity internal breakdown (R = 0.91), unless internal browning was present.

**Table 2.1.** Overview of reports of use of NIR spectroscopy for non-invasive assessment of internal browning in apple.

| Population (n) | Reference method | Wavelength (nm) | Detector | Geometry | $R^2$ | Classification (%) TPR, TNR | References |
|---|---|---|---|---|---|---|---|
| 100 | score | 740, 840 | Biospect | Transmission | 0.82 | 98, 93 | Francis et al. (1965) |
| 556 | score | 450-1050 | SiPDA | Interactance | 0.71 | 94, 88 | Upchurch et al. (1997) |
| 240 | % area | 697- 861 | SiPDA | Transmission | 0.91 | NA | Clark et al. (2003) |
| 117 | % area | 650-950 | SiPDA | Transmission | 0.9 | NA | McGlone et al. (2005) |
| 512 | juice $A_{420}$ | 950-2300 | InGaAs | Interactance | 0.87 | NA | Li (2011) |
| 120 | visual | 780 | | TRS | | 90, 71 | Vanoli et al. (2014) |

PDA: Photodiode Array, SiDA: Silicon Diode Array, InGaAs: Indium Gallium Arsenide, True Positive Rate (TPR) refers to the % of good fruit predicted as good and True Negative Rate (TNR) refers to the % of defect fruit predicted as defect.

Upchurch et al. (1997) reported use of a full transmittance optical geometry and a spectrometer with entrance slit of 500 µm, 150 grooves/mm ruled grating and a silicon diode array detector, supporting a spectral resolution of 12 nm, for detection of internal browning in apple. Spectra were acquired using an exposure time of 5 s (a time impractical for on-line grading). The study was based on 556 Delicious apples stored at 0 $^{o}$C for 6 months to develop internal breakdown. A set of 136 apples were used for calibration set development while the remaining (420) apples were used for prediction purposes. A visual score of 0 (no browning) to 3 (extreme browning) was assigned after cutting each fruit perpendicular to stem-calyx axis. Defect fruit were reported to absorb more light below 750 nm (particularly 720-750 nm) and less above 750 nm, relative to healthy apples. A ratio between transmittance at 720 and 810 nm was selected as a classifier for discrimination between defective and good apples. The ratio decreased as the browning intensity increased and was correlated with the degree of internal breakdown ($R^2$ = 0.71). Error rates of 6.3% good apples misclassified as defect (false negative) and 12% defect fruit misclassified as good (false positive) were reported. The presence of bruises contributed to instances of false negatives.

Clark et al. (2003) reported use of SWNIRS to detect internal browning in Braeburn apple fruit that had been selected to represent a range of severities of the disorder using clinical magnetic resonance imaging (MRI). The light source and detector were located at right angles to one another relative to the fruit (partial transmittance geometry), with spectra collected over the range 300-1140 nm using a Zeiss MMS1 spectrometer. Errors in assessment were partly attributed to asymmetric distribution of browning within fruit. Regression based on simple wavelength ratios, multiple linear regression and partial least squares regression were considered for segregation of good and bad fruit. Changes in spectral features between 680 and 900 nm were observed as browning intensity increased. Absorbance was higher in defect fruit over the range 680-825 nm and lower at wavelengths above 825 nm. The best PLSR model prediction result reported was obtained using a model developed utilizing absorbance spectra over the range 697-861 nm ($R^2_p$ = 0.91, RMSEP = 7.9% of cut surface affected), obtained by averaging spectra from opposite sides of the fruit.

McGlone et al. (2005) considered detection of brownheart disorder of Braeburn apples using two detector types with integration times short enough to allow on-line grading. SWNIR transmittance spectra were acquired over the 650-950 nm range, with fruit moving at a speed of 500 mm per second (5 fruit per second). A set of 117 fruit kept for 4 weeks in cold storage was used, with severity of browning ranging from 0-60% of the surface area of an equatorial cut of the fruit. A 'large aperture spectrometer' was recommended for online measurement, with a PLSR model result achieving $R^2_p$ = 0.9, RMSEP = 4%. Multiple measurements of a given fruit in different orientations was recommended for more accuracy, and for application on line, a good seal between fruit and the conveyor cup was recommended to block stray light to the detector.

Li (2011) reported assessment of internal browning in Fuji apples (n = 512) using Fourier Transform (FT) NIR (Bruker Optics, Germany, 800-2500 nm, InGaAs detector, 50 W light source) in an interactance optical geometry. Four spectra were acquired per fruit and were averaged for further analysis. Fruit were stored in controlled atmosphere (CA) storage (at 8% $O_2$ and 5% $CO_2$ for 20, 40, 60, 80, 100, 120, 140 and 160 days. Of these fruit, 300 fruit were used for calibration and 212 fruit for validation purposes. The best PLSR model ($R^2$ = 0.87) was obtained using the wavelength ranges of 950-1440 nm, 1480-1890 nm and 1960-2300 nm. The degree of apple browning was analysed using 4 mL of juice mixed with 0.05 g of polyvinyl polypyrrolidone (PVPP) and 6 mL of 95% ethanol for measurement at wavelength of 420 nm using a UV-visible spectrophotometer (UV 2600 Shanghai, China).

Working with another member of the pome fruit family, pear (*Pyrus communis* L.), Han et al. (2006) reported use of a discriminant technique for detection of brown core using spectra collected over the range 651-1282 nm with a transmission optical geometry at. A perfect classification result was obtained with discriminant analysis (DA) using Mahalanobis distance (MD). Using the absorbance difference between 713 and 743 nm as a classifier, 5.3% of good pears were classified as defect while only 4.3% of defect fruit were classified as good.

Fu et al. (2007) compared the use of two instruments for detection of internal browning in pear. Transmission spectra were acquired using a 50 W tungsten

halogen lamp and a Si based Ocean Optics spectrometer operating over the range of 400-1028 nm. Spectra were collected from fruit in two fruit orientations, with stem-calyx axis vertical or horizontal. Diffuse reflectance spectra were also acquired with an FT-NIR spectrometer (Bruker Optics Corporation, Germany) operating with two detectors (silicon 670-1100 nm and InGaAs 800-2630 nm). Only 57 pears were used for the exercise, which were stored for 5 months at 4 $^{o}$C to allow development of brown heart. Samples were cut perpendicular to stem-calyx axis for visual scoring on a three point scale (no browning, slight browning and severe browning). Discriminant Analysis (DA) was used to discriminate classes using the Mahalanobis distance. Of the 57 fruit, 37 (24 with brownheart and 13 good) were used for calibration and the remaining 20 were used for prediction purposes. Better results were obtained with the transmission geometry instrumentation. Defect pear fruit demonstrated a higher absorbance below 750 nm and a lower absorbance above 750 nm. Increased absorbance below 750 nm is consistent with absorption in the visible wavelengths, presumably by phenolics associated with the browning symptoms. In calibration, an accuracy of 89% was noted, for validation accuracy decreased to 81%.

Other detection technologies continue to benefit from development. Chayaprasert and Stroshine (2005) reported use of magnetic resonance imaging (MRI) for detection of browning in intact apple in an online sorting conveyor belt at speed of lower than 150 mm/s achieving the classification accuracy (TP+TN/P+N) of 88%. Gonzalez et al. (2001) also reported the use of MRI in detection of internal browning in Fuji apples with difference in longitudinal (T1) and transverse (T2) relaxation time and proton density between the normal, moderate and severe browning fruit. Vanoli et al. (2014) reported the use of time resolved reflectance spectroscopy (TRS) estimated absorption and scattering coefficients at 780 nm for separation of internal browning in apple fruit. They reported correct classification of 90% of good fruit and 71% of defect fruits.

Thus a number of reports indicate that non-invasive sorting of fruit on internal browning is possible. For the light transmission studies, two approaches have been used – the estimation of a continuous variable associated with the level of defect (e.g. the output of a partial least square (PLS) regression prediction) and a binary classification to defect and sound fruit (principally LDA). Other discriminant

techniques of promise to this application include PLS–discriminant analysis (PLS-DA), support vector machine (SVM) classification, soft independent modelling of class analogy (SIMCA), and multiple logistic regression (MLR) (Yang et al., 2012).

The published reports of detection of internal browning generally recommend a transmission geometry and the use of wavelengths in the vis-SWNIR region, but there is discrepancy between the reports on the best range to use, and on the algorithm to use (PLS regression, simple ratio, discriminant analysis etc.). The comparison of published studies is difficult, as results depend on instrumentation, population distribution and the reference method used to assess level of defect. Further, all of the reports mentioned above are based on relatively small sample sizes, basically limited to a single population of fruit divided into training and validation sets. As such, these studies have failed to consider the range in variation of fruit optical properties occurring between populations (of different growing conditions etc.).

Perhaps it should not be a surprise that commercial adoption of such technology appears limited, an indication that application is not as straightforward as some of the scientific literature suggests. Indeed, there are no reports in the scientific literature of use of commercially available equipment targeted to detection of apple fruit with internal diffuse browning. In the current exercise, results for detection of internal browning in apple are compared for spectra collected with instruments using an interactance, partial transmittance and full transmission geometries, including the instrumentation of fruit grading equipment manufacturer, MAF Roda, with consideration of reference method, wavelength region and algorithm.

## 2.2 Materials and Methods

### 2.2.1 Fruit

Apple (*Malus domestica* Borkh.) cv Pink Lady$^{TM}$ fruit were commercially harvested in Stanthorpe, south Queensland, Australia, in April 2013 and stored for six months in controlled atmosphere storage with 1-2% $O_2$ and 4-5% $CO_2$ at temperature of 2 $^{o}$C and relative humidity more than 85%. Fruit were sampled randomly from various bins, and represents different harvest dates in the harvesting season, transported to Rockhampton, Queensland at 25 $^{o}$C, and then stored for another two months at 4 $^{o}$C

in a refrigerator with relative humidity of 85%. Fruit were allowed to equilibrate to room temperature of 25 $^{o}$C for 6 h before spectra were acquired. The room and fruit temperature during spectra acquisition was at 25 $^{o}$C.

The four independent lots (populations) of fruit contained a total of 296 fruit (93 good and 203 defect). Population 1, consisting of 90 fruit (31 good and 59 defect), was used for calibration model development. Population 2 and 3 contained 60 (12 good and 48 defect) and 77 (28 good and 49 defect) fruit, respectively, and were used as independent prediction sets. Population 4, containing 69 fruit (22 good and 47 defect fruit), was used for trials on sound frequency, resonant frequency, DA index measurement, chlorophyll fluorescence and absorbance at 420 nm of extracted juice. Population 1 and 2 (a total of 150 fruit) was subset into five groups of 30 fruit for a study of the impact of fruit temperature on model prediction. Each lot of the fruit represents fruit from same harvest season but differ in storage time of 3 weeks in refrigerator at 4 $^{o}$C before experiment.

### 2.2.2 Chlorophyll fluorescence

Chlorophyll fluorescence was assessed for each fruit after dark adaptation (Fig. 2.1). Fruit were wrapped with aluminium foil for 24 h prior to measurement of photosynthetic efficiency using an OptiSciences 30p (Bioscientific, Australia) at a modulation intensity of 4. This unit employs a PIN photodiode with a 700 - 750 nm band pass filter. One reading was taken per fruit within 2 s of removing the aluminium foil.


**Figure 2.1.** Measurement of chlorophyll fluorescence of apple fruit.

### 2.2.3 SWNIRS

Four spectra were acquired of each fruit (detector facing each of four equidistant locations around the equator of the fruit) using each of three instruments, involving different optical geometries and detectors (for an expanded description, see Appendix 1).

Nirvana (Integrated Spectronics, Sydney, Australia); a handheld instrument using an interactance optical geometry (Greensill & Walsh, 2000) with a 32 W halogen lamp and a MMS-1 photodiode array spectrometer (302-1150 nm);

IDD0 (Fig. 2); an instrument developed 'in-house' employing a partial transmission optical geometry with a 300 W halogen lamp and a MMS1 photodioide array spectrometer (302-1150 nm) operated with an integration time of 1000 ms;

Insight-2 (MAF Roda, Montabaun, France); an online sorting instrument using a full transmission geometry with a 150 W halogen lamp and an Avantes spectrometer (600-973 nm) operated with an integration time of 9 ms.

IDD2 (MAF Roda, Montabaun, France); an online sorting instrument employing a full transmission optical geometry with pulsed light emitting diodes (LEDs) of peak emission wavelengths 700, 780, 810 and 880 nm and a single photodiode detector.

All units achieved repeatability (SD of absorbance of 20 repeated measurements of a white tile) of around 1 mAbs unit (Appendix 1).

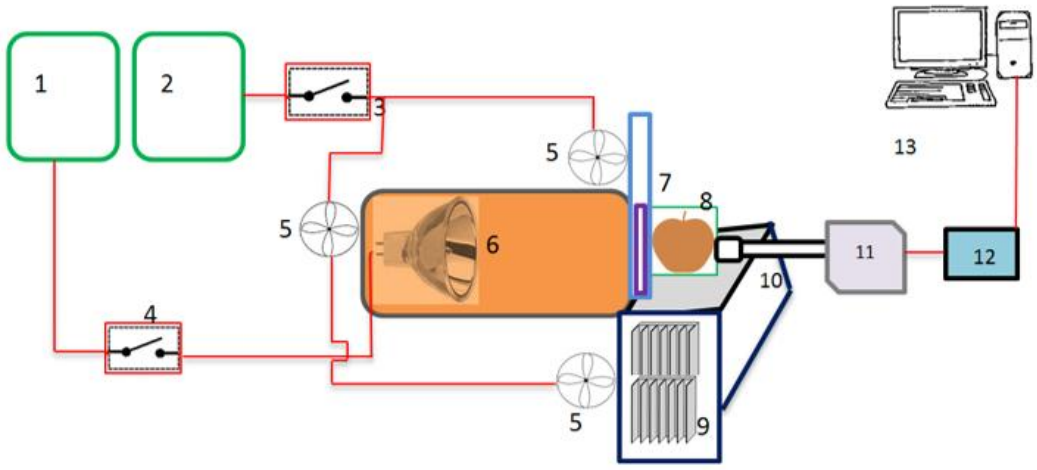



**Figure 2.2.** Schematic diagram of IDD0 instrumentation, featuring:

Lamp power supply (1): step-down transformer model MM 115/250EN (Peter Martin Sales P/L, Newstead, Qld, Australia), input 240VAC 50Hz, output: 115V-250V, Fan power supply (2): adjustable transformer; input: 240VAC, 50Hz, output 12VDC 0.8A to 30A, Power switches (3,4) on fans (5) and lamp (6) (300W, 120V, Quartzline lamp, General Electric Company, Nela Park, Cleveland, OH, 44112), Adjustable shutter (7) with aperture to suit fruit (8) of different sizes. on aluminium block heat sink (9), Lens assembly (10) with SMA connection to fibre optic of Zeiss MMS 1 spectrometer (11) (employing Hamamatsu S3904-256Q diode array); dimension 70x50x40 mm ), supported by a 16 bit analog to digital converter (tec5, Germany) (12) and computer (13) for data acquisition and processing using an in-house developed LabView based program, 'Hortical'.

The IDD0 (Table. 2.2) was developed for this project, and was based on the Zeiss MMS-1 spectrometer, operated with a 16 bit analogue to digital converter (tec5, Germany). This unit has a mean pixel pitch of 3.3 nm, a wavelength resolution (FWHM) of approximately 10 nm, a signal to noise of 1 milli absorbance and a well depth of 1013 counts/watt seconds and operates over the spectral range 310-1100 nm. Thus the unit has high signal to noise but modest sensitivity. To achieve a 180$^{o}$ geometry (lamp-fruit-detector angle) involving transmission of light through a whole fruit, a high intensity lamp (300 W Quartzline) and a long integration time were employed. The use of the 300 W lamp entailed heating problems of both fruit and lamp housing, so the unit was designed with a series of fans and a failsafe circuit to

cut power if lamp holder temperature exceeded 100 $^{o}$C. A shutter was incorporated to allow shielding of the fruit until measurement.

Spectrometer analogue to digital conversion counts (ADCC) of over 67% of detector maximum were achieved for citrus fruit using this lighting, but for apple, reasonable detector ADCC and integration time was achieved only with a 90$^{o}$ illumination-sample-detector angle.

### 2.2.4 Reference measurements

The extent of the browning disorder within apple fruit was assessed using several reference methods:

#### *2.2.4.1 Visual browning score*

A panel of 6 members was trained in the assignment of a score for the severity of defect as visible in a transverse equatorial cut of the fruit, using either a 5 (Fig. 2.3) or 10 point scale, with scorers assisted by a reference pictorial chart. The average value of all scorers was used. The 10 point visual scale was initially used, to mimic practice observed in industry (personal communication, Philippe Plagnol, MAF Roda, France). A 5 point scale (Jobling, 2005) was later trialled. In the 5 point scale, score 1 is symptomless and score 2 is associated with a faint 'off white' colouration that can be considered acceptable to consumers. Scores 3-5 are associated with increasingly distinct symptoms, but all categories can be considered unacceptable to consumers. Therefore fruit of score 1 and 2 are considered 'acceptable' fruit, while those of 3, 4 and 5 can be considered as 'defect' fruit (Fig 2.3). For the 10 point scale, scores 4-10 are associated with defect fruit.

**Figure 2.3.** Cut surfaces of apple with diffuse internal browning symptoms in order of increasing browning intensity from left to right, with visual score ratings of 1 to 5, respectively.

### *2.2.4.2 Area of defect*

The cut apple (equatorial transverse cut), as used in visual scoring, was imaged using a Canon DS 126211 camera with a 17-85 mm lens at a fixed focal distance. Images were processed using ImageJ software, with conversion of the RGB colour image into grayscale and use of a threshold setting to differentiate defect tissues and masking of the core area (Fig. 2.4). Affected tissue area was expressed as a percentage of total fruit cross-sectional area, with subtraction of the core area.


**Figure 2.4.** Image processing using ImageJ software for quantification of affected area as a % of total cross-sectional area.

### *2.2.4.3 CIE colour space readings*

The cut apple (equatorial transverse cut), as used in visual scoring, was also used for the measurement of surface colour (CIE Lab colour space) a Chromameter CR 400 (Konica Minolta; 2 degree observer, D65 illuminant), with averaging of five randomly located measurements per fruit. The measurements was taken following the cut and completed within 15 seconds for each fruit so as to avoid further browning. The Chromameter was calibrated using a standard calibration procedure. CIE Lab values were used to calculate the Colour index (CI), as described by Magwaza et al. (2014):

$$Color\ Index = \left(\frac{1000\ a}{L\ b}\right)$$

#### *2.2.4.4 Absorbance of juice aliquot at 420 nm*

A measurement of absorbance of extracted juice was attempted, after the method of Song et al. (2007). Tissue was sampled for this measurement at the point where Chromameter readings were taken. Apple flesh (10 g FW) was chopped into fine pieces, blended with 12.5 ml of water using an ultra turrax and centrifuged for 5 min at 3000 rpm. An aliquot (5 mL) of juice extract was then mixed with 7.5 mL of ethyl alcohol and centrifuged for another 10 min at 3000 rpm before measurement of supernatant absorbance at 420 nm using an UV-Visible spectrophotometer (Cary 50 Bio, Ontario, Canada).

#### *2.2.4.5 Weight, SSC and DM*

Following weighing of each fruit, a 20 mm diameter core was extracted to 10 mm depth from the point of spectral assessment. The core was halved radially, with one half used for dry matter (DM) and the other for soluble solids content (SSC) measurement. For DM assessment, tissue was dried in a fan forced oven at 65 $^{o}$C for 48 h. For SSC, juice was extracted using a garlic press and measured using a refractometer (RFM 320, Belingham and Stanley Ltd).

### 2.2.5 Data analysis and Chemometrics

Data analysis was undertaken using Excel (Microsoft, USA), the multivariate data analysis software The Unscrambler 10.3 (Camo, Norway) and Matlab (Mathworks Inc., USA) with PLS Toolbox (Eigenvector, USA). IDD0 and Nirvana spectral data was collected in Absorbance units. Insight2 spectral data was output in units of % transmittance (T) and was converted to Absorbance (log 1/T). Absorbance data of all instruments was subject to a number of pre-treatments including mean centring, standard normal variate (SNV), multiple scatter correction (MSC), and Savitzky Golay second derivatization using a window of 9 points (SG-9). Spectral data used for the analysis was restricted to 500-975 nm for IDD0 and Nirvana instrumentation while for the Insight2 unit the entire available range (600-973 nm) was used.

Predictive quantitative models were developed using the calibration data set using partial least square regression (PLSR). PLSR models were assessed using the criteria of correlation coefficient of determination ($R^2_{cv}$), root mean square error of cross validation (RMSECV) and number of principal components (PCs), while predictive

performance was assessed based on coefficient of determination of prediction ($R^2_p$), root mean square error of prediction (RMSEP) and bias.

The binary classification algorithms of linear discriminant analysis (LDA), PLS–discriminant analysis (PLS-DA), support vector machine (SVM) classification, soft independent modelling of class analogy (SIMCA), and multiple logistic regression (MLoR) were assessed based on the true positive rate (TPR = TP/P) and true negative rate (TNR = TN/N) and accuracy (TP+TN/P+N).

## 2.3 Results and Discussion

### 2.3.1 Population description

Of the four populations assessed, population 2 had the highest mean visual (5 point) scores (Table 2.2). The mean CIE a and colour index values of the three populations was not proportionate to the mean visual scores, consistent with a poor relationship between these variables (Table 2.2).

**Table 2.2.** Mean and standard deviation of reference parameters for calibration and prediction sets.

| Parameters | Calibration Pop 1 | Prediction 1 Pop 2 | Prediction 2 Pop 3 | Preliminary Pop 4 |
|---|---|---|---|---|
| Number of fruit | 90 | 60 | 77 | 69 |
| Score (5) | 3.21 ± 1.53 | 3.81 ± 1.39 | 3.36 ± 1.75 | 2.8 ± 1.60 |
| Colour Index | 0.37 ± 1.29 | 1.05 ± 1.4 | 1.38 ± 1.38 | |
| CIE a* value | 0.63 ± 2.2 | 1.75± 2.32 | 2.45 ± 2.48 | |

The five sets of fruit used in the exercise on assessment of influence of temperature on prediction performance (section 1.3.8) each contained 30 fruit with mean and SD of 2.96 ± 1.49, 3.28 ± 1.57, 3.14 ± 1.48, 3.36 ± 1.51 and 3.85 ± 1.38, respectively.

### 2.3.2 Reference methodologies

Consumer assessment of the diffuse browning defect of apple involves subjective visual inspection of a cut surface of the fruit. However there are inexacitudes involved in classifying this disorder, e.g. how should a fruit with a diffuse overall browning be scored compared to a fruit with a patchy but darker areas. In an attempt

to quantify the visual assessment, five and ten point scoring systems were compared. The average RMSED (root mean of squares of error of differences) of assessment by six different operators relative to an 'average assessor' using the 5 point scale was 0.54 (n = 6), while for the 10 point scale it was 0.79 (adjusted to a 5 point scale). For a repeated assessment using the 5 point scale by one assessor only the RMSED was 0.28, Appendix 2). These errors are likely to be associated with the ability of a human observer to remember the descriptions of multiple levels. Further, the use of scales with an odd number of classes was recommended by Rensis Likert in 1932, based on social science research involving participant agreement or disagreement with a declaration statement (Likert, 1933).

Thus the 5 point scale is recommended over the 10 point scale for human scoring of this disorder, given the error of assessment and the preference for an odd number of classes. The lower limit benchmark (RMSED) against which non-invasive assessment techniques should be judged is 0.28, being the error of the reference method.

For a set of 69 fruit, the two visual score ratings (5 and 10 point scales) were better related ($R^2 = 0.85$) to each other than any other reference parameters, although visual score was better indexed by % area (image analysis) than juice $Abs_{420nm}$ (Table 2.3). The image analysis procedure required manual intervention to adjust the threshold level on each image, to achieve segmentation of areas comparable to that see by human eye. This was despite imaging of the cut fruit under conditions of constant illumination. Potentially results could be improved by use of a more diffuse source of illumination, but at this point the image analysis method cannot be recommended over simple visual 5 point scoring.

**Table 2.3.** Correlation coefficient of determination ($R^2$) between attributes in context of assessment of extent of the internal diffuse browning in apple fruit. Data of population 4 (n = 69). Attribute % area refers to extent of browning on a cut fruit surface, while visual 5 and 10 relate to a visual score of the same surfaces using a 5 and 10 point scale, respectively. $A_{420nm}$ refers to a measurement of extracted juice.

| Parameters | % area | Fv/Fm | $Abs_{420}$ nm | Visual 1-5 | Visual 1-10 |
|---|---|---|---|---|---|
| % area | 1 | | | | |
| Fv/Fm | 0.26 | 1 | | | |
| $Abs_{420}$ nm | 0.045 | 0.0013 | 1 | | |
| Visual 1-5 | 0.64 | 0.20 | 0.013 | 1 | |
| Visual 1-10 | 0.65 | 0.19 | 0.0005 | **0.85** | 1 |

Note: Values greater than 0.75 are shown in bold.

The poor relationship between visual score and juice $A_{420}$ nm (Table 2.3) could be due to either differences in the sample volume assessed by the two methods or method artefact. Visual score involves consideration of an equatorial plane, juice absorbance involves juice extracted from a sample of the fruit. As diffuse browning is not uniformly distributed through the fruit, the difference in the two measures may reflect sampling differences. Alternately, the procedure used for the juice absorbance may be flawed in that it did not employ PVPP as recommended by (Li, 2011). PVPP is commonly used in extraction protocols precipitate polyphenols and tannins, and to prevent browning reactions through reduction of polyphenols. PVPP was therefore avoided as it could remove the coloured compounds in the diffuse browning affected tissue. Nonetheless, future work should consider use of this additive.

In a second exercise involving 150 fruit (Pop 1 and 2), visual score was better correlated to CIE a ($R^2$ = 0.78) than L or the colour index, and poorly related to SSC or DM (Table 2.4). Again, the visual 5 point score is recommended as an easy to undertake, though destructive, reference method for indicting the extent of diffuse browning in apple fruit.

**Table 2.4.** Correlation coefficient of determination ($R^2$) between various destructive measures of Pink Lady$^{TM}$ apple fruit (Populations 1 and 2; n = 150). Values above 0.75 are in bold.

| Parameters | Score (5) | CIE L* | CIE a* | CIE b* | Colour Index | DM | SSC |
|---|---|---|---|---|---|---|---|
| Score (5) | 1 | | | | | | |
| CIE L | 0.68 | 1 | | | | | |
| CIE a | **0.78** | **0.90** | 1 | | | | |
| CIE b | 0.40 | 0.56 | 0.52 | 1 | | | |
| Colour Index | 0.70 | 0.71 | **0.77** | 0.43 | 1 | | |
| DM | 0.11 | 0.14 | 0.12 | 0.07 | 0.19 | 1 | |
| SSC | 0.23 | 0.30 | 0.28 | 0.18 | 0.38 | 0.61 | 1 |

SSC was weakly related to DM ($R^2$ = 0.61). McGlone et al. (2003) reported a correlation coefficient of determination ($R^2$) of 0.93 between harvest time DM and post storage SSC for Royal Gala apple. For Pink Lady, the relation between DM and SSC for mixed population of good and defect fruit was modest ($R^2$ = 0.61), with a poorer relationship between DM and SSC for defect fruit ($R^2$ = 0.47) than sound fruit ($R^2$ = 0.76) (data not shown).

### 2.3.3 Chlorophyll fluorescence

Chlorophyll fluorescence involves emission from the de-excitation of excited chlorophyll molecules, with values of photochemical efficiency (Fv/Fm), a measure of the efficiency of the light reactions of photosynthesis, decreasing with increasing stress. Fv/Fm of sound fruit was 0.57, a level consistent with previous reports of healthy apple fruit (Neuwald et al., 2007). Fv/Fm was not significantly ($P > 0.05$) different between fruit of visual score ratings of 1 to 4, but severely affected fruit (score 5) demonstrated a decreased value of 0.47 (Fig. 5).

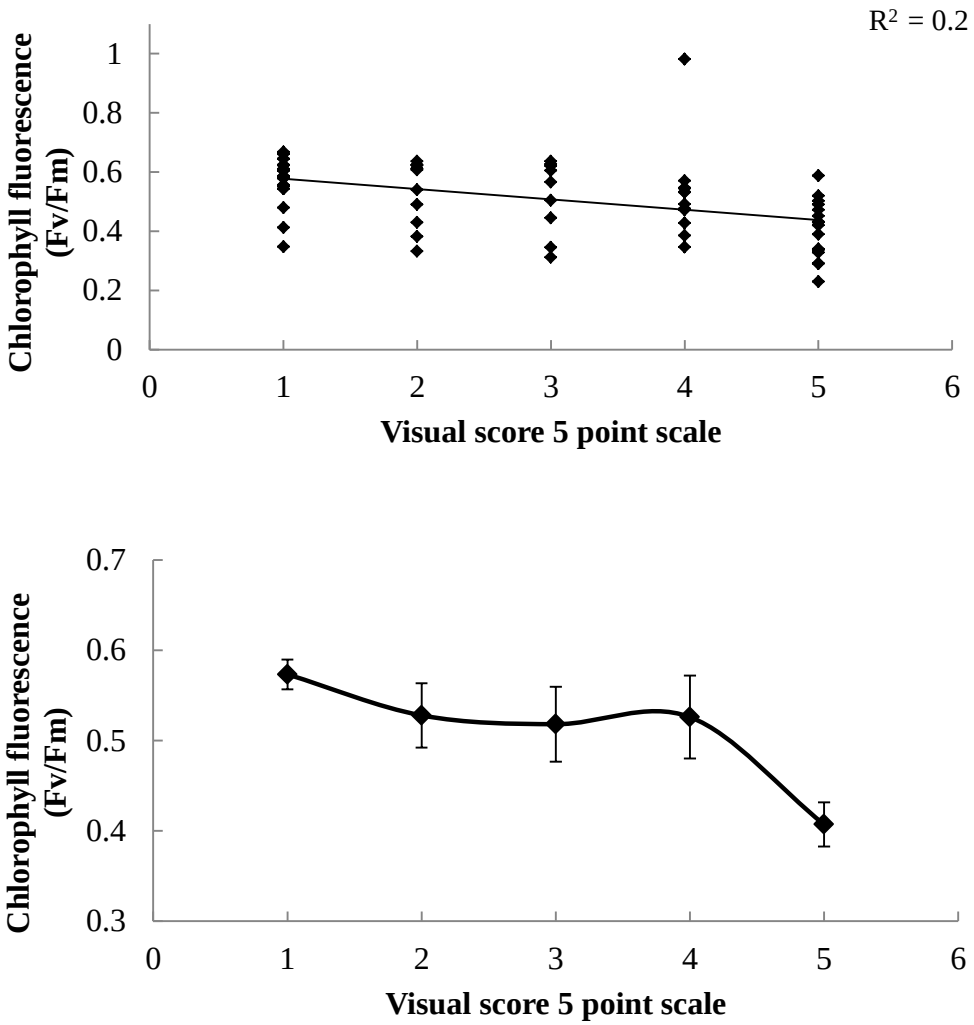



**Figure 2.5.** Photochemical efficiency (Fv/Fm) of photosynthesis in apple fruit relative to degree of internal browning in population 4 (n = 69). Top panel displays raw values, bottom panel shows mean and associated SE.

Thus the chlorophyll fluorescence technique may have value for detection of severe defect levels in 'sentinel' fruit within CA storage consignments, using equipment such as that offered by HarvestWatch or Bessling P/L but with a form of carousel to allow monitoring of a larger sample of fruit. Potentially the technique could be implemented for in-line grading, however detection of only severely affected fruit is not a practical outcome in terms of sorting to remove defect (score 3-5) fruit.

### 2.3.4 Visible – SWNIR spectral features

Fruit absorbance spectra of fruit collected using all instruments were characterised by features associated with anthocyanin at around 550 nm, chlorophyll at around 665 nm and water at 730, 840 and 950 nm (Fig. 2.6). The Nirvana unit employs a gold coated reference, while poly-tetrafluorethylene (PFTE) references were used with other two units, accounting for difference in absorbance values in the visible range.

Average absorbance values were higher for defect relative to sound fruit at wavelengths less than 830 nm for the Insight2 spectra, and less than 730 nm for the IDD0 spectra (Fig. 2.6). This observation is consistent with previous reports. Clark et al. (2003) reported higher absorbance for apple with internal browning over the range 600-750 nm, and weaker absorbance beyond 850 nm. McGlone et al. (2005) noted higher absorbance in the red to near red region of spectrum. For pear, higher absorbance was noted over the range 640-860 nm for fruit with brown core (Han et al., 2006). The higher absorbance values in this region may represent absorption of light by the polyphenols associated with browning.

The strongest correlation between absorbance at a single wavelength and defect intensity was achieved at around 620 and 710 nm for the Nirvana, IDD0 and InSight2 instruments, respectively (Fig. 2.6).

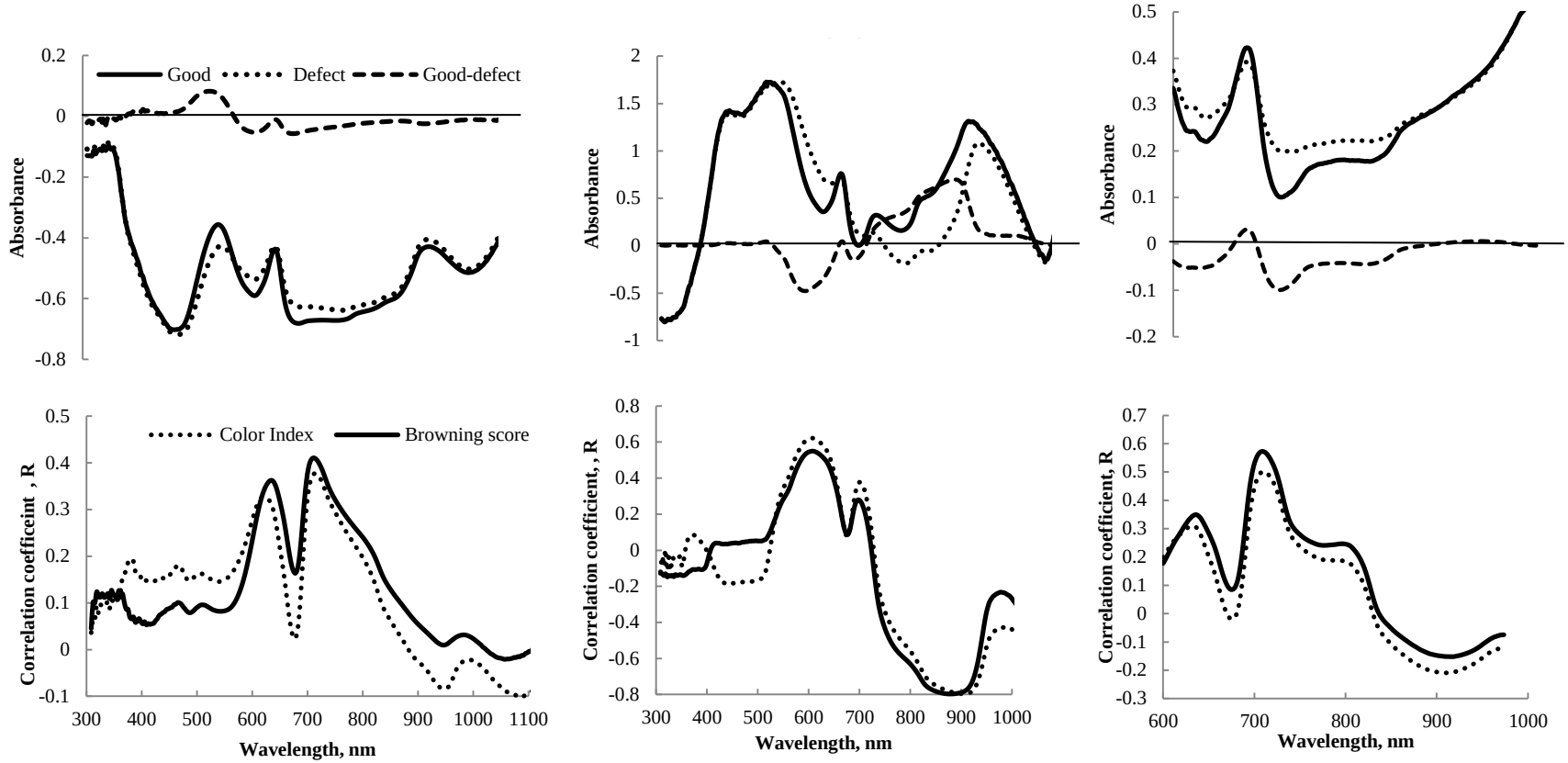



**Figure 2.6.** Mean absorbance spectra for good and defect fruit and their difference (score 1 and 5, respectively, from populations 1 and 2) and univariate correlation coefficient (R) of internal defect parameters (visual score and colour index) with absorbance at each wavelength.

### 2.3.5 Partial Least Square Regression

Partial least square regression (PLSR) models were developed using the reference values of visual score, colour index (Table 2.5). Similar results were obtained for Lab colour space based parameters (a* and Colour Index) across the three instruments, with slightly better results for Lab a* value (data not shown). Better calibration results were obtained using the Lab colour space based parameters than the visual score value, but the reverse held for prediction results.

Initially, PLSR models were developed using the full wavelength range available from a given instrument. Smooth regression coefficients were obtained for the region 500-975 nm for the units based on a MMS 1 spectrometer (Nirvana and IDD0), and for the entire available range of 600-973 nm for Avantes spectrometer (Insight2 instrumentation) (Fig 2.7). Thus, these ranges were selected for PLS regression models. For the IDD0 and InSight2 units, the PLSR models gave strong weighting to absorbance at 670, 710 and 900 nm, and 680, 710, 800 and 835 nm, respectively.

For the reference parameter of visual score and spectra obtained using the IDD0 and InSight2 units, better calibration and prediction performance was obtained for models based on raw absorbance spectra, followed by use of SNV and MSC pre-treatments (Table 2.5). A correlation coefficient of determination ($R^2$) of 0.89 was achieved for IDD0 SNV or MSC treated absorbance spectra. However, model predictive performance was higher for models developed using second derivative spectra. Poor predictive performance was obtained using spectra acquired with the Nirvana unit.

Overall, of the three instruments, best results were obtained using the IDD0 unit (e.g. $R^2_{cv}$ = 0.83, RMSECV = 0.62, for visual score model), with comparable results from the InSight2 unit but poorer results from the Nirvana unit (e.g. $R^2_p$ = 0.62, RMSEP = 0.99, visual score). The poorer performance of Nirvana unit is expected given its interactance geometry, i.e. assessment of only a localised area of the fruit. IDD2 was operated in a $90^o$ optical geometry, but scattering of light within the fruit apparently resulted in an adequate volume of fruit explored by detected light, supporting models comparable or superior to that achieved with the full transmission InSight2 system.

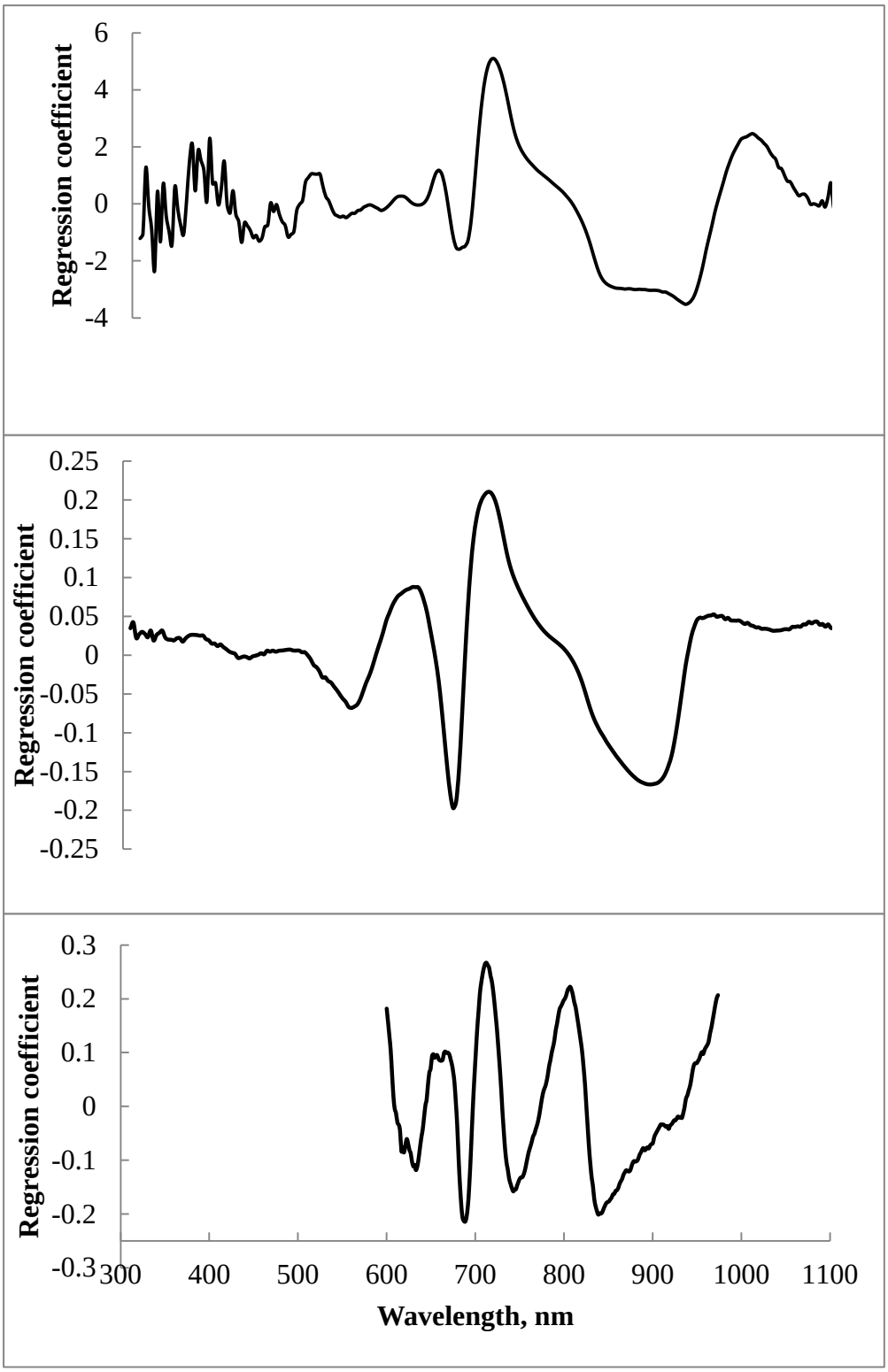



**Figure 2.7.** PLS Regression coefficients for Nirvana (top), IDD0 (middle) and Insight2 (bottom) for visual score models based on absorbance data.

**Table 2.5.** Partial least square regression results for spectra from three instruments, for raw absorbance spectra and for spectra treated with several pre-processing methods for Nirvana, IDD0 and Insight2 instrumentation. Calibration, prediction set 1 and prediction set 2 consist of 90, 60 and 77 fruit respectively.

**A. Nirvana**

| **Pre treatments /Parameters** | **Calibration statistics** | | | **Prediction statistics on set 1** | | | **Prediction statistics on set 2** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias | $R^2_p$ | RMSEP | Bias |
| **Raw Abs** | | | | | | | | | |
| Score | 0.79 | 0.7 | 14 | 0.57 | 1.01 | -0.1 | 0.62 | 0.99 | 0.14 |
| Colour Index | 0.86 | 0.45 | 14 | 0.8 | 0.62 | 0.15 | 0.7 | 0.83 | -0.12 |
| **Abs SNV** | | | | | | | | | |
| Score | 0.77 | 0.71 | 14 | 0.52 | 0.4 | -0.11 | 0.53 | 1.3 | 0.24 |
| Colour Index | 0.85 | 0.47 | 14 | 0.77 | 1.15 | -0.48 | 0.71 | 0.82 | -0.1 |
| **Abs MSC** | | | | | | | | | |
| Score | 0.76 | 0.73 | 11 | 0.53 | 1.08 | -0.03 | 0.6 | 0.99 | 0.2 |
| Colour Index | 0.83 | 0.51 | 11 | 0.75 | 0.73 | 0.2 | 0.66 | 0.88 | -0.1 |
| **d2A** | | | | | | | | | |
| Score | 0.77 | 0.71 | 10 | 0.6 | 0.96 | 0.01 | 0.64 | 0.95 | 0.28 |
| Colour Index | 0.84 | 0.49 | 10 | 0.74 | 0.69 | 0.12 | 0.71 | 0.82 | -0.11 |
| **Abs SNV d2A** | | | | | | | | | |
| Score | 0.76 | 0.74 | 10 | 0.56 | 1.01 | -0.08 | 0.59 | 1.00 | 0.2 |
| Colour Index | 0.83 | 0.52 | 10 | 0.71 | 0.71 | 0.02 | 0.66 | 0.9 | -0.19 |

**B . IDD0**

| **Pre-treatments /parameters** | **Calibration statistics** | | | **Prediction statistics on set 1** | | | **Prediction statistics on set 2** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias | $R^2_p$ | RMSEP | Bias |
| **Raw Abs** | | | | | | | | | |
| Score | 0.83 | 0.62 | 4 | 0.70 | 0.83 | 0.24 | 0.87 | 1.08 | 0.88 |
| Colour Index | 0.87 | 0.46 | 4 | 0.73 | 0.72 | 0.04 | 0.75 | 0.73 | -0.16 |
| **Abs SNV** | | | | | | | | | |
| Score | 0.78 | 0.71 | 4 | 0.76 | 0.7 | 0.16 | 0.89 | 0.82 | 0.56 |
| Colour Index | 0.87 | 0.46 | 3 | 0.71 | 0.75 | 0.03 | 0.77 | 0.61 | -0.53 |
| **Abs MSC** | | | | | | | | | |
| Score | 0.77 | 0.73 | 3 | 0.75 | 0.71 | 0.13 | 0.89 | 0.9 | 0.66 |
| Colour Index | 0.88 | 0.45 | 3 | 0.74 | 0.71 | 0.03 | 0.77 | 0.68 | -0.06 |
| **d2A** | | | | | | | | | |
| Score | 0.75 | 0.75 | 3 | 0.73 | 0.73 | 0.06 | 0.84 | 0.8 | 0.02 |
| Colour Index | 0.86 | 0.49 | 4 | 0.68 | 0.79 | -0.08 | 0.77 | 0.82 | -0.47 |
| **Abs SNV d2A** | | | | | | | | | |
| Score | 0.76 | 0.75 | 3 | 0.76 | 0.68 | 0.03 | 0.86 | 0.79 | 0.02 |
| Colour Index | 0.87 | 0.47 | 5 | 0.7 | 0.76 | -0.06 | 0.77 | 0.67 | -0.1 |

**C. Insight2**

| Pre-treatments /parameters | Calibration statistics | | | Prediction results on set 1 | | | Prediction results on set 2 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias | $R^2_p$ | RMSEP | Bias |
| **Raw Abs** | | | | | | | | | |
| Score | 0.83 | 0.63 | 7 | 0.75 | 0.81 | 0.26 | 0.84 | 1.12 | 0.87 |
| Colour Index | 0.85 | 0.50 | 7 | 0.68 | 0.81 | 0.11 | 0.65 | 0.89 | 0.23 |
| **Abs SNV** | | | | | | | | | |
| Score | 0.84 | 0.61 | 6 | 0.72 | 0.88 | 0.25 | 0.82 | 1.18 | 0.86 |
| Colour Index | 0.84 | 0.52 | 6 | 0.68 | 0.82 | 0.13 | 0.64 | 0.99 | -0.13 |
| **Abs MSC** | | | | | | | | | |
| Score | 0.83 | 0.62 | 5 | 0.73 | 0.84 | 0.23 | 0.79 | 1.49 | 1.10 |
| Colour Index | 0.84 | 0.51 | 5 | 0.68 | 0.82 | 0.12 | 0.70 | 0.92 | -0.16 |
| **d2A** | | | | | | | | | |
| Score | 0.78 | 0.71 | 5 | 0.75 | 0.69 | -0.08 | 0.84 | 0.92 | 0.56 |
| Colour Index | 0.79 | 0.58 | 8 | 0.58 | 0.90 | -0.1 | 0.61 | 1.08 | -0.66 |
| **Abs SNV d2A** | | | | | | | | | |
| Score | 0.80 | 0.68 | 7 | 0.74 | 0.72 | -0.03 | 0.87 | 0.73 | 0.39 |
| Colour Index | 0.77 | 0.61 | 7 | 0.56 | 0.94 | -0.18 | 0.69 | 1.02 | -0.68 |

### 2.3.6 PLS model robustness to temperature across population

Fruit consist of approximately 80% water. The IR/NIR spectrum of water is affected by temperature due to an effect on the extent of hydrogen bonding. This issue has been well explored in context of prediction of SSC and DM in intact fruit using SWNIRS (Acharya et al., 2014), with the simplest accommodating measure being the inclusion of samples of a range of temperatures into the training sets. Given that the spectral information relevant to internal browning is in the red-far red region, spectral based measures of internal browning may be free from interference associated with temperature change. It can be inferred that the brix model is temperature dependent but can be made more robust by including spectra from range of temperatures in calibration model while the model for internal browning is independent of temperature.

PLSR models developed on internal browning score using fruit at 10, 25 or 35 $^{o}$C or combined possessed near identical b coefficients (Fig. 2.8). The effect of sample temperature on prediction of apple diffuse browning was considered for the case of a PLSR model on diffuse browning developed using 120 absorbance spectra of 30 intact fruit (Pop 5) at 10 $^{o}$C, with prediction undertaken for the same fruit at 10, 25 and 35 $^{o}$C, and for further lots of fruit at these three temperatures. There was no bias associated with prediction of fruit at temperatures different to that of the calibration samples (Table 2.6), consistent with use of wavelength regions in the model (e.g. 570, 620, 670, 720, 900) that were insensitive to temperature. The b coefficients for the combined temperature model were similar to those models developed using fruit at a single temperature. This similarity is consistent with model insensitivity to temperature.

**Table 2.6** Prediction performance for a PLSR model on apple internal browning visual score (calibration $R^2_{cv}$ = 0.87, RMSECV = 0.53, PCs = 8) developed using IDD0 absorbance spectra (500-975 nm) collected of intact fruit at 10 $^oC$. Populations are subsets of population 1 and 2, with 30 fruit in each population.

Calibration set: Pop 5, 10 $^oC$: $R^2_{cv}$ = 0.87, RMSECV = 0.52, PCs = 4

| **Populations** | **Temp ($^oC$)** | **$R^2_p$** | **RMSEP** | **Bias** |
|---|---|---|---|---|
| Pop 5 | 10 | 0.88 | 0.50 | 0.00 |
| SD = 1.49 | 25 | 0.86 | 0.58 | 0.14 |
| | 35 | 0.88 | 0.52 | 0.14 |
| Pop 6 | 10 | 0.81 | 0.65 | -0.09 |
| SD = 1.57 | 25 | 0.84 | 0.59 | 0.07 |
| | 35 | 0.83 | 0.60 | 0.03 |
| Pop 7 | 10 | 0.59 | 1.41 | 0.25 |
| SD = 1.48 | 25 | 0.66 | 1.07 | 0.17 |
| | 35 | 0.57 | 1.28 | 0.27 |
| Pop 8 | 10 | 0.46 | 1.18 | 0.04 |
| SD = 1.51 | 25 | 0.53 | 0.99 | 0.06 |
| | 35 | 0.54 | 0.93 | 0.07 |
| Pop 9 | 10 | 0.64 | 1.49 | 0.64 |
| SD = 1.38 | 25 | 0.69 | 1.15 | 0.51 |
| | 35 | 0.73 | 0.98 | 0.47 |

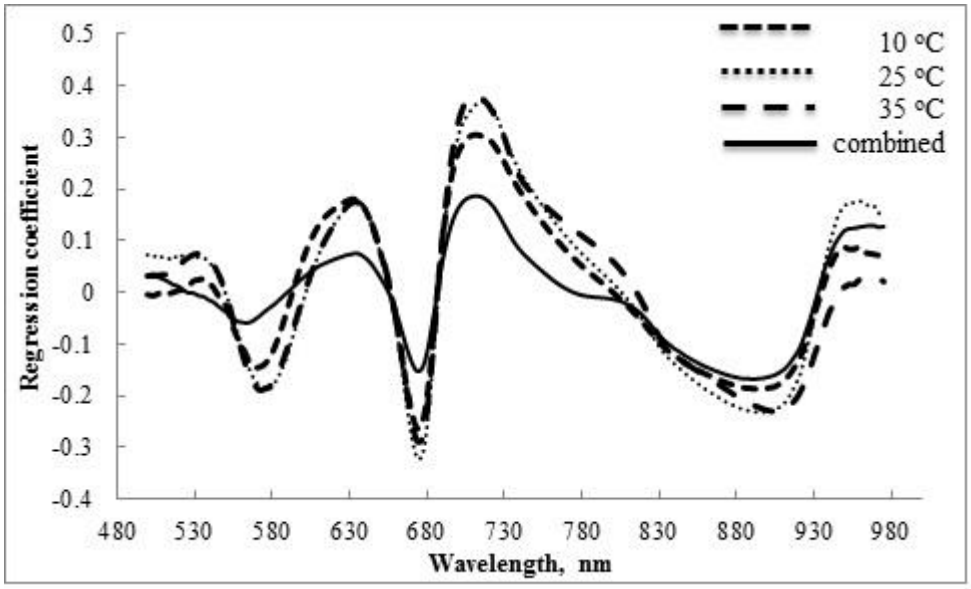



**Figure 2.8.** PLSR b-coefficients for browning score models developed using spectra of fruit at three different temperatures (IDD0 instrumentation).

### 2.3.7 Principal component analysis

Principal component analysis (PCA) was undertaken using MSC treated absorbance data from the three instruments. A principal component plot using the first three PCs (explaining >95% of variation in the three cases) demonstrated separation of consumer acceptable fruit (visual score of 1 or 2) from defect fruit (visual score of 3 to 5 (Fig. 2.9).

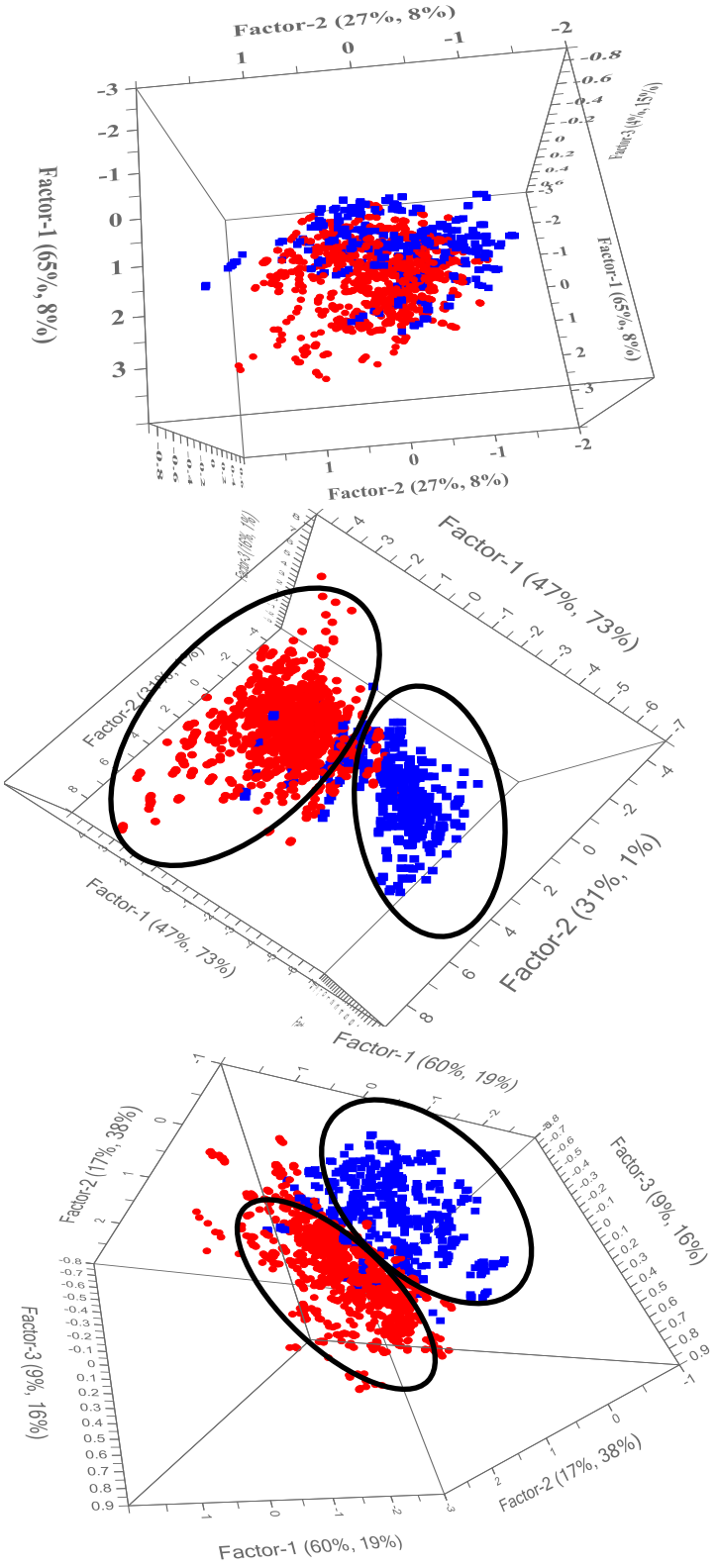



**Figure 2.9.** Score plots for principal components 1, 2 and 3 from a principal component analysis of MSC treated absorbance spectra (500-975 nm) using an interactance (A), partial transmission (B) and full transmission (C) optical geometries for 150 fruit (populations 1 and 2). Values for defect fruit displayed as circular dots while those of good fruit is displayed as square dots.

### 2.3.8 Classification of defect and good fruit

Discrimination of sound fruit from defect based on spectra collected using the three instrumentations was attempted using several classification algorithms. Relatively poor prediction results were achieved for the Nirvana instrument compared to that for the other two instruments for prediction set 1, and use of this instrument was discontinued (Table 2.6).

Comparable results were obtained with PLS DA, LDA and PCA LDA classification algorithms (Table 2.7). The PLS DA classification method yielded fairly consistent classification accuracy for all three instruments, while PLS based classification gave best results for the IDD0 instrumentation.

Prediction set accuracy of more than 95% and false discovery rate of < 2% was achieved using IDD0 spectra with either SVM classification, logistic regression or PLS-DA using 4 PCs. The specifications were achieved using InSight2 spectra using LDA, PCA-PLDA and PLS-DA using 6 PCs (Table 2.7).

**Table 2.7.** Classification of good and defect fruit in calibration and in prediction of independent sets using 12 classification algorithms, using absorbance spectra from two instruments, against a reference assessment of visual score. TPR is true positive rate, TNR is true negative rate. Accuracy (Accu.) is mean of TPR and TNR, false discovery rate (FDR), FP/(TP+FP). False discovery rates <2% associated with accuracy rate of >95% are shown in bold. Units are in percentage.

**A. IDD0**

| | Calibration set (n=90) | | | | Prediction set 1 (n=60) | | | | Prediction set 2 (n=77) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Algorithms | TPR | TNR | Accu. | FDR | TPR | TNR | Accu. | FDR | TPR | TNR | Accu. | FDR |
| PLS-DA | 83 | 98.3 | 90.7 | 2.0 | 83.3 | 99.1 | 91.2 | 1.1 | 83.9 | 93.9 | 88.9 | 6.8 |
| LDA, Linear | 87.6 | 98.6 | 93.1 | 1.6 | 80.5 | 99.5 | 90.0 | 0.6 | 91 | 86.7 | 88.9 | 12.8 |
| LDA MD | 95.2 | 97.6 | 96.4 | 2.5 | 76.3 | 98 | 87.2 | 2.6 | 92.9 | 93.9 | 93.4 | 6.2 |
| PCA LDA | 78.5 | 97.2 | 87.9 | 3.4 | 98.6 | 97 | 97.8 | 3.0 | 98.2 | 83.7 | 91.0 | 14.2 |
| PCA LDA MD | 85.5 | 93.9 | 89.7 | 6.7 | 89.6 | 99.5 | 94.6 | 0.6 | 96.4 | 95.9 | 96.2 | 4.1 |
| k-NN | 94.3 | 98.3 | 96.3 | 1.8 | 93 | 94.6 | 93.8 | 5.5 | 83 | 97.5 | 90.3 | 2.9 |
| SIMCA | 94.6 | 55.5 | 75.1 | 32.0 | 82.6 | 77.8 | 80.2 | 21.2 | 100 | 3.57 | 51.8 | 49.1 |
| SVM. Linear | 77.7 | 98.2 | 88.0 | 2.3 | 97.9 | 99.6 | **98.8** | **0.4** | 92.9 | 100 | **96.5** | **0.0** |
| Logistic regression | 86.8 | 94.2 | 90.5 | 6.3 | 99.3 | 100 | **99.7** | **0.0** | 70.5 | 100 | 85.3 | 0.0 |
| PLS DA(1-5) 4 PCs | 76.3 | 98 | 87.2 | 2.6 | 89.6 | 96 | 92.8 | 4.3 | 97.3 | 91.84 | 94.6 | 7.7 |
| PLS DA (5) 4PCs | 84.9 | 92.4 | 88.7 | 8.2 | 97.2 | 99.8 | **98.5** | **0.2** | 96.4 | 100 | **98.2** | **0.0** |
| PLS DA (2) 4PCs | 79.8 | 94.2 | 87.0 | 6.8 | 97.9 | 97.9 | 97.9 | 2.1 | 100 | 82.7 | 91.4 | 14.7 |

**B. Insight2**

| | Calibration set (n=90) | | | | Prediction set 1 (n=60) | | | | Prediction set 2 (n=77) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TPR | TNR | Accu. | FDR | TPR | TNR | Accu. | FDR | TPR | TNR | Accu. | FDR |
| PLS-DA | 79.4 | 99.6 | 89.5 | 0.5 | 100 | 97.6 | 98.8 | 2.3 | 87.9 | 97.8 | 92.9 | 2.4 |
| LDA, Linear | 91.6 | 99.1 | 95.35 | 1.0 | 80.5 | 98.8 | 89.65 | 1.5 | 91 | 98.5 | 94.8 | 1.6 |
| LDA MD | 99.2 | 100 | 99.6 | 0.0 | 33.3 | 100 | 66.65 | 0.0 | 0 | 100 | 50.0 | 0.0 |
| PCA LDA Linear | 82.2 | 97.7 | 89.95 | 2.7 | 85.2 | 99.8 | 92.5 | 0.2 | 89.3 | 100 | 94.7 | 0.0 |
| PCA LDA MD | 91.4 | 90.2 | 90.8 | 9.7 | 88.1 | 92.2 | 90.15 | 8.1 | 96.4 | 80.1 | 88.3 | 17.1 |
| k-NN | 96.5 | 98.0 | 97.25 | 2.0 | 73.6 | 92.8 | 83.2 | 8.9 | 77.7 | 95.4 | 86.6 | 5.6 |
| SIMCA | 98.4 | 43.5 | 70.95 | 36.5 | 95.1 | 59 | 77.05 | 30.1 | 96.4 | 84.2 | 90.3 | 14.1 |
| SVM. Linear | 99.2 | 99.8 | 99.5 | 0.2 | 84.7 | 95.6 | 90.15 | 4.9 | 75.9 | 100 | 88.0 | 0.0 |
| Logistic regression | 86 | 94.9 | 90.45 | 5.6 | 84.7 | 99.5 | **92.1** | **0.6** | 67.9 | 100 | 84.0 | 0.0 |
| PLS-DA(1-5) 6PCs | 87.1 | 88.9 | 88 | 11.3 | 89.6 | 93.7 | 91.65 | 6.6 | 91 | 93.8 | 92.4 | 6.4 |
| PLS (1-5) 6 PCs | 90 | 95.6 | 92.8 | 4.7 | 82.6 | 99.3 | **90.95** | **0.8** | 86.6 | 100 | **93.3** | **0.0** |
| PLS-DA (1-2) 6 PCs | 88.2 | 96.6 | 92.4 | 7 | 85.4 | 99.5 | **92.45** | **0.6** | 91 | 100 | **95.5** | **0.0** |

### 2.3.9 Classification based on two wavelengths

To enable wider adoption of defect sorting technology across the fruit industry, cheaper instrumentation is required. Use of models based on only a few (two) wavelengths and a single detector rather than use of a full spectrometer allows for a decrease in the cost of instrumentation. In this exercise, the coefficient of determination ($R^2$) was compared for the relationship between several two wavelength indices and internal defect detection (Table 2.7).

The indices trialled included those recommended by Francis et al. (1965) and Upchurch et al. (1997), an index based on the wavelength available in the IDD2 unit, and indices chosen based on the wavelength of maximum absorbance difference of defect and good fruit recorded in this study normalised by that of minimum difference (Abs 525-1040 nm for Nirvana, 600-482 nm for IDD0 and 714-900nm for Insight2). An $R^2$ of 0.74 was achieved with the two wavelength index of 720 and 810 nm for IDD0 data, and 714 and 900 for InSight2 data (Table 2.8; Fig. 2.10). As expected given the partial transmission optical geometry, results for Nirvana data were poor.

**Table 2.8.** Regression coefficient of determination ($R^2$) for two wavelength difference model against visual score for three instruments (Nirvana, IDD0 and Insight2).

| | **Instrumentation** | | |
|---|---|---|---|
| **Two wavelength based index** | **Nirvana** | **IDD0** | **Insight2** |
| A740 - 805 nm (Francis et al., 1965) | 0.37 | 0.62 | 0.15 |
| A720 - 810 nm (Upchurch et al., 1997) | 0.44 | 0.74 | 0.64 |
| A 840 - 740 nm (Francis et al., 1965) | 0.30 | 0.72 | 0.66 |
| A 552- 1040 nm, 600 – 482nm , 714 – 900 | 0.39 | 0.39 | 0.74 |
| A 880 / 780 nm (IDD2) | 0.39 | 0.69 | 0.67 |

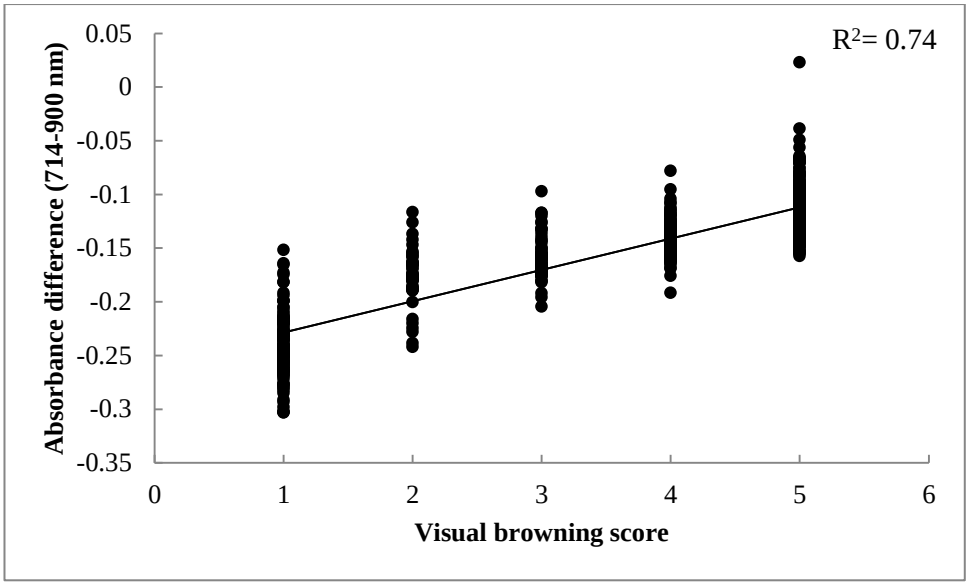



**Figure 2.10.** Scatter plot between the absorbance difference between 714 and 900 nm and visual score (5 point scale), InSight2 data.

### 2.3.10 Sorting operation based on IDD2

This section begins a discussion that is further developed in Chapter 5 'Sorting Optimisation'. For the techniques that return a continuous value in prediction of defect level in fruit, a threshold value must be set to act in discrimination of accepted and rejected groups. For the IDD2 instrumentation, a value can be placed on absorbance (A880/780) ratio index. Using a value of 76 as the threshold, sorting accuracy of between 70 to 83% were recorded across three populations with false discovery rate (FP/FP+TP) of less than 10% (Table 2.9).

**Table 2.9.** Classification of fruit as defect or good using the IDD2 instrumentation based on A 880/780 nm and assessed based on visual score for three population sets at threshold of 76. Values are in percentages. TPR = true positive rate (TP/P), TNR = true negative rate (TN/N), Accuracy = (TP+TN/P+N) and FDR = false discovery rate (FP/(FP+TP)).

| Population | TPR | TNR | Accuracy | FDR |
|---|---|---|---|---|
| Population 1 (n = 90) | 35.48 | 98.31 | 76.66 | 8.33 |
| Population 2 (n = 60) | 41.66 | 100.00 | 70.83 | 0.00 |
| Population 3 (n = 77) | 64.28 | 93.87 | 83.11 | 8.69 |

Altering the threshold value on the sorting operation will impact these results (Fig. 2.11). For example, at a threshold of 76, 94% of defect fruit are identified as defects (TNR), while only 64% of non-defect fruit are correctly identified (TPR), while at a threshold of 104, TNR is 28% while TPR is increased to 100%.

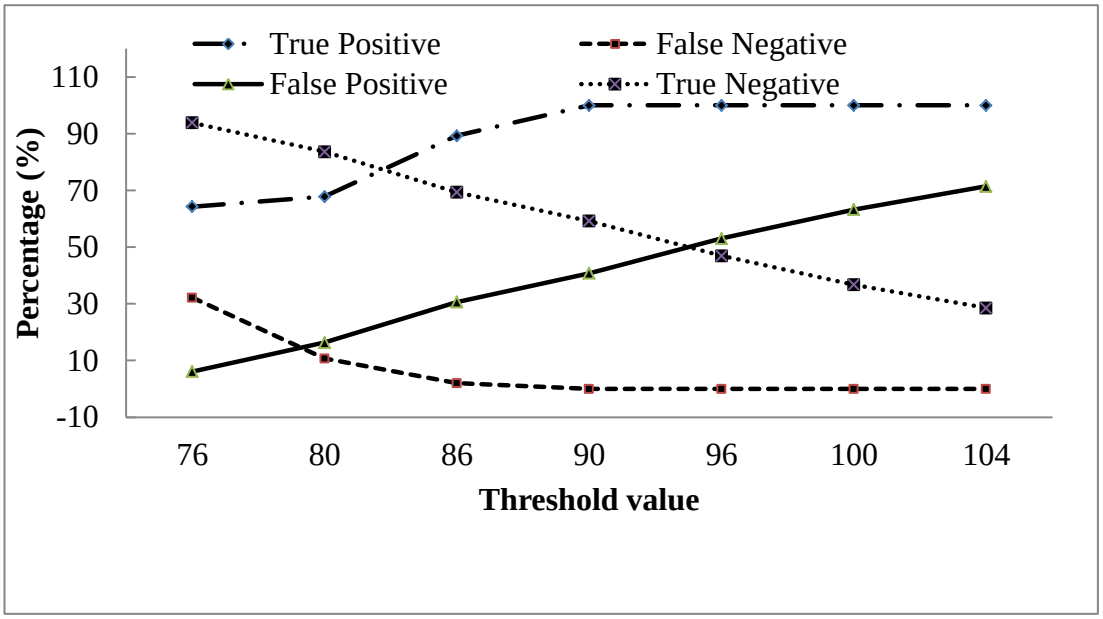



**Figure 2.11.** Classification of fruit into different categories as dependant on sorting threshold value: IDD2 data population 3 involving 77 fruit (combined population).

Formats for presentation of this data (in terms of individual score bands) are explored in Figures 2.12 to 2.14 (Fig. 2.12, 2.13). Decreasing the threshold value increases the proportion of score 1 (good) fruit in the accepted population, but at the expense of overall yield (TP/P). Cumulative distribution plots convey similar information (Fig 2.14).

While useful in a general descriptive sense, these plots are not recommended for use by a packline operator to aid the decision on selection of threshold value. This decision requires compromise between yield (TP/P), false discovery rate (FP/(TP+FP)) and knowledge of pricing. As indicated, these issues are explored further in Chapter 5.

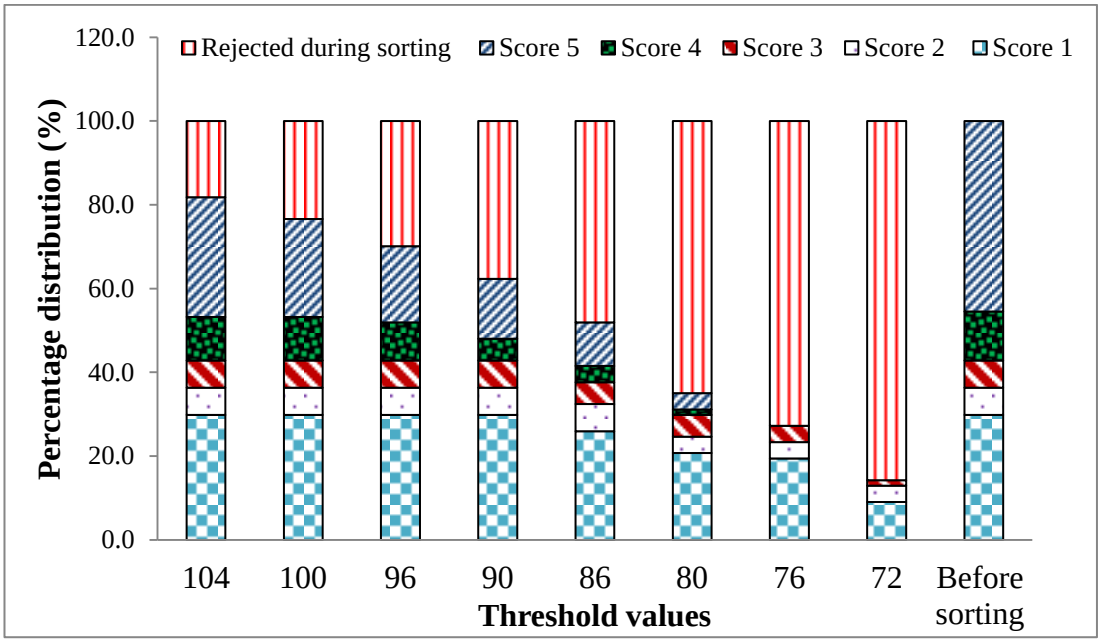



**Figure 2.12.** Proportion of fruit from different quality category and rejection based on the threshold value.

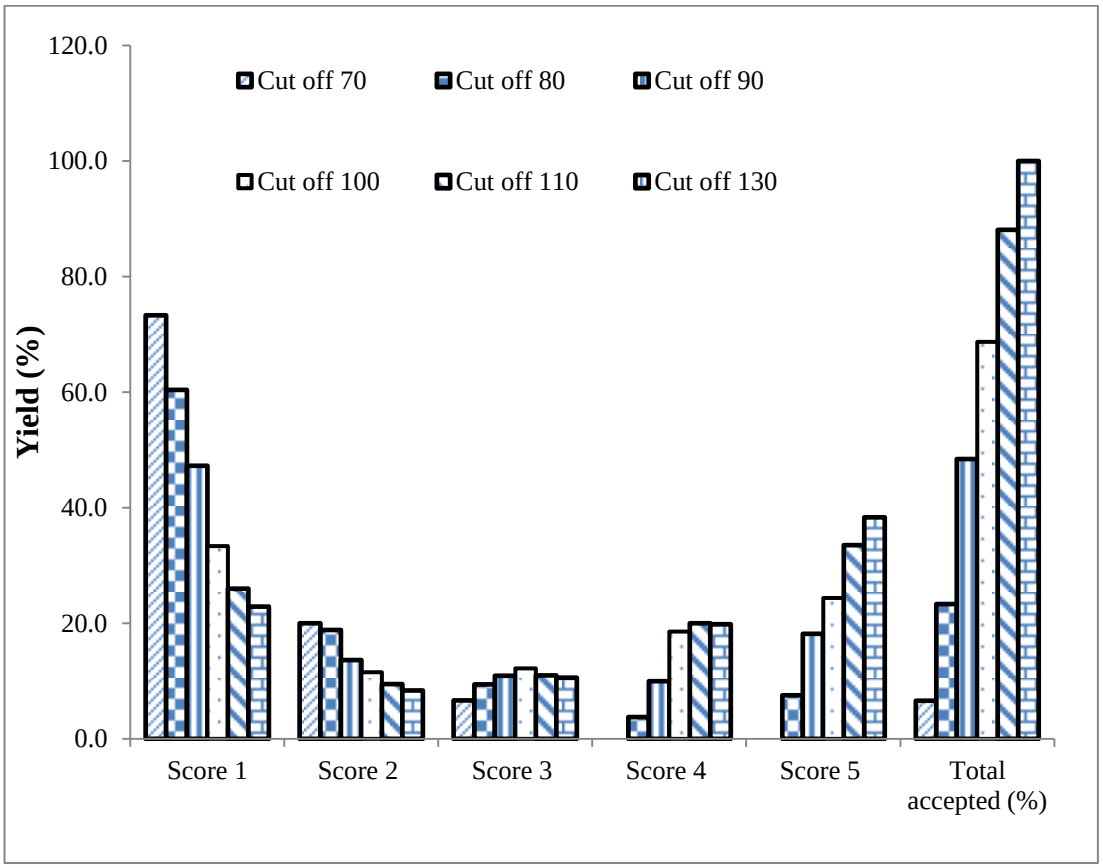



**Figure 2.13.** Distribution of fruit in new sorted population based on different cut off settings with respective yield.

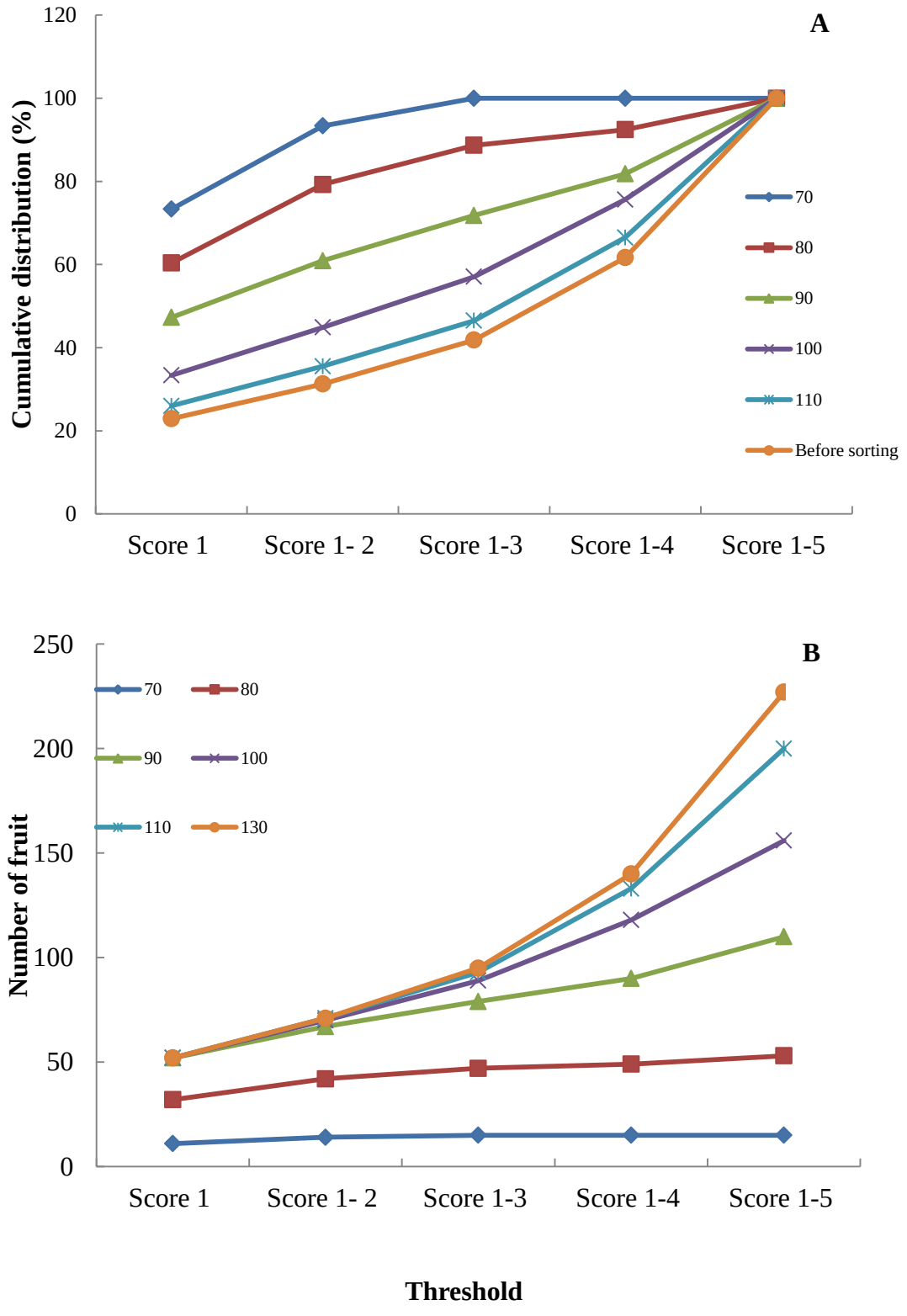



**Figure 2.14.** Distribution of defect for a population of apple before and after sorting to various threshold values A. cumulative distribution in each threshold in percentages and B. distribution of fruit in each threshold in number of fruit. (Threshold value is ratio of B/A at two wavelengths; IDD2 unit).

## 2.4 Conclusion

Internal diffuse browning disorder is an economically important disorder induced during storage in some apple cultivars, particularly Pink Lady TM apples. This susceptibility is related to the physiology of fruit tissue where tissue undergoes stress under low oxygen in CA storage, leading to increase in reactive oxygen which causes membrane integrity loss, finally visible as oxidation of phenolic compounds into brown colored polymers. This issue is further exacerbated by recent breeding programs focussed on improving crispiness in the fruit, which has resulted in decreased air space volume in fruit tissue. Further, although this issue is evident after storage, it is influenced by various pre harvest factors and so there is a large variation in incidence of the disorder between orchards and seasons. Multivariate approaches to understanding of the causes of the disorder may help management of the level of incidence, but sorting technologies to remove defect fruit from affected consignments will remain necessary for the foreseeable future.

The absorbance of defect fruit was generally higher than that of control fruit. Previous work (e.g. Clark et al., 2003; McGlone et al., 2005) on the utility of near infrared spectroscopy using partial or full transmission geometry over the wavelength range 500-975 nm or selected wavelengths from this region for the detection of the internal flesh browning was validated, with a discriminant analysis classification method achieving a sorting accuracy as high as 98%, independent of fruit temperature. Every detection technique has errors, however, with the proportion of type 1 and type 2 errors changing with defect level and distribution within a population. Thus further effort is required to provide operators of a sorting process with management tools, extending that produced in chapter 5.

Another application areas may be in use with fruit in storage, for monitoring of ‘sentinel’ fruit to detect the onset of the disorder, and so guide the decision to out-turn fruit, and to provide information on conditions that pre-dispose fruit to development of the defect.

The chlorophyll fluorescence measures of fruit also association with level of defect, and further work on diagnosis of fruit with the disorder using this measurement technique is recommended.

# Chapter 3.
# Detection and characterisation of granulation in mandarin by using NIR spectroscopy [1]


## Abstract

Granulation in 'Imperial' mandarin was attributed to cell wall (cellulose, hemicelluloses) proliferation in the hypodermis of the juice sacs, leading to increased light scattering (L value of cut surface) and reduced juice recovery although constant water content except in extreme granulation. Spectroscopic instrumentation operating in full transmission optical geometry was compared for the non-invasive assessment of the disorder, using visual score, luminosity (L) and juice recovery as reference attributes. Classification based on wavelength ratio based algorithm and other classification algorithm yielded classification accuracy as high as 98 % for good fruit while for defect fruit, the classification accuracy was lower. High prediction accuracy (92%) was reported for IDD0 instrumentation using PLS-DA classification method .

**Keywords:** sorting, regression, classification, score, spectra

[1] Acknowledgements: The UQ EM centre is acknowledged for the EM sectioning and electron microscopy imaging and Prof S. Mansfield of University of British Colombia, Canada for the cell wall characterisation analysis. Fruit were provided by Ken Mason , Yeppoon based farmer and Central Fruit Packers, Mundubbera, Queensland, Australia.

## 3.1 Introduction

Citrus fruit can develop a range of physiological disorders including fruit cracking, sunburn, puffiness, rind breakdown, chilling injury and internal dryness (Munshi et al., 1978; Peiris et al., 1998; Subedi, 2007). The descriptor of internal dryness and the associated terms granulation, section drying and gelling have been used somewhat interchangeably by several authors. Peiris et al. (1998) categorized internal dryness into two broad categories, namely dehydration and granulation (Fig. 3.1). Dehydration involves shrinkage of the tissue followed by complete collapse of the affected vesicles due to loss of vesicle contents, e.g. following frost damage. In contrast, granulation begins with the hardening of the affected vesicles following gradual collapse of the inner cells resulting in an empty cavity (Peiris et al., 1998). Juice recovery rate is decreased proportionate to the extent of the disorder (e.g. 40% v/w for normal fruit to 5% in defect fruit), but water content and TSS is constant, except in severely granulated fruit, in which levels of these attributes are decreased (Subedi, 2007). Affected fruit become unfit for fresh consumption due to a chewy, dry and tasteless mouth-feel.

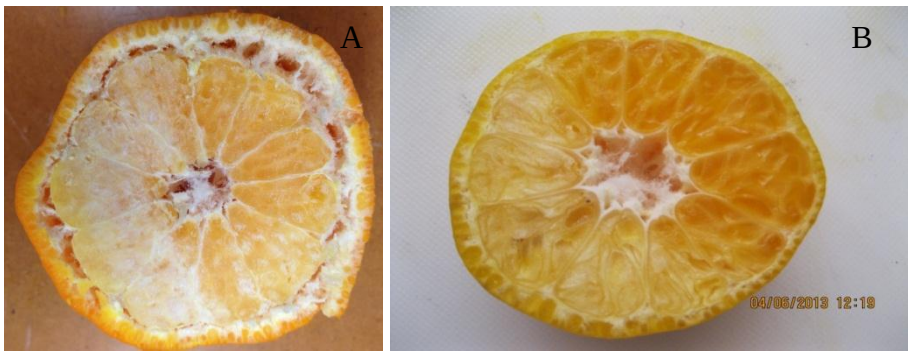



**Figure 3.1.** Internal defects of citrus: A. dehydration defect (white areas) following freezing injury B. granulation defect in 'Imperial' mandarin.

Approximately 28,000 ha of land in Queensland is under citrus production, with 13% of this (3,590 ha) associated with Imperial mandarin production (Citrus Australia website, accessed 28/4/15). 'Imperial', a brightly coloured and easy-to-peel cultivar, is prone to the granulation disorder. An erratic incidence of this disorder is reported in Queensland, with variation among soil types, nutrition, irrigation, rootstock and orchard locations (Hofman, 2011). There can be difference in incidence of this defect within a single tree. Hofman (2011) reported the incidence of granulation is

associated with the very early fruit development stages with competition between fruitlets, flowers and flush for the nutrients and water and nitrogen. It was observed that granulation was decreased when winter nitrogen application was followed by an additional spring application.

Affected fruit cannot be recognised on visual external appearance. Severely dry fruit can be detected by hand feel based on firmness. Given the difficulty in recognising affected fruit and that unpredictable nature of incidence of the disorder, there is a clear need for development of a non-invasive sorting technology.

Difference in fruit optical properties (extent of light scattering) and water content offer promise for the application of non-invasive detection technologies. Peiris et al. (1998) reported use of SWNIR absorption spectrometry and X-ray computed tomography for detection of the tangerine tissue drying disorder. NIR absorption spectra (500-1000 nm) were acquired using a 75 W tungsten halogen lamp as the light source and a Si CCD based spectrometer (Ocean Optics, SD 1000-TR). A multiple linear regression was undertaken, with a model based on second derivative absorbance values at 768 nm and 960 nm yielding a correlation coefficient of determination ($R^2$) of 0.77 on a two point scale of visual granulation scores. The 960 nm absorbance feature of fruit is well known to be associated with the second overtone of O-H stretching, and thus to water content (Acharya et al., 2014). This result is consistent with detection of a dehydration defect, but this measurement may not be appropriate for detection of granulation defect within a fruit given the lack association of granulation level to water content (Subedi, 2007).

Subedi (2007) explored the use of SWNIR (500-1000 nm) spectroscopy to assess granulation of intact Imperial mandarin. Defect level was scored by visual assessment and by a chromameter (luminosity, L*) reading of the cut surface of the fruit. For fruit without peel, luminosity was reasonably well modelled ($R^2$ = 0.84). However, for whole fruit, calibration $R^2$ was decreased to 0.74 and validation of the model using an independent population was poor (Subedi, 2007).

There does not appear to have been a successful commercial implementation of a non-invasive technology for in-line assessment of ‘granulation’ of citrus or indeed other fruit. The work of Peiris et al. (1998) with dehydration defect, but applicability

to detection of granulation defect remains to be demonstrated. Further work is therefore warranted to explore appropriate techniques to non-invasively detect the granulation disorder.

## 3.2 Materials and methods

### 3.2.1 Anatomical and chemical characterisation

Representative fruit samples were chosen for chemical and anatomical characterisation. The procedure of Cunnigham and Walsh (2002) was followed for analysis of the sugar composition of labile carbohydrates in the insoluble (cell wall) fraction. Normal and 'gelled' juice sacs were dissected from the fruit segments and the sac contents were dissected from 'gelled' juice sacs. A sample of known weight (approximately 0.5 g) was washed repeatedly with 70% ethanol, and then digested at 110 $^{o}$C for 1 hour in 2.0 ml trifluoroacetic acid (TFA) within a pressure vessel. The TFA was then evaporated with forced air assistance on a water bath at 40 ºC, and the residue dissolved into 2.0 ml sterile $dH_2O$. This solution was filtered to pass a 5.0 µm and then a 0.4 µm filter. The resulting solution was then analysed for sugar content by HPLC using a Supelcosil LC-$NH_2$, 5 µm, 250 x 4.6 mm ID amino-propyl bonded phase column (Sigma Aldrich 5-8338).

For anatomical characterisation, normal and granulation affected tissue was fixed in a 3% glutaraldehyde in 10% sucrose solution, dehydrated through an alcohol/acetone series, and embedded in LR white resin before sectioning to 2 µm thickness. Fresh, hand cut sections and resin embedded sections were examined using a Nikon epifluorescence microscope, as unstained sections and following staining with toluidine blue, safranine, phloroglucinol, and aniline blue. Fresh material was also rapidly frozen by plunging into liquid nitrogen, and either used directly in a crude freeze fracture preparation, or freeze dried, cut with a sharp razor blade and gold sputter coated before examination using a JEOL scanning electron microscope (model JSM-6360), operating at 15 kV.

### 3.2.2 Fruit

Fruit of the mandarin variety 'Imperial' were sourced from two commercial farms in Central Queensland, Australia during 2014 and 2015, and stored at 10 $^{o}$C. Fruit were harvested from areas which the farm manager reported as having a high incidence of

granulation. Population 1 represents a fruit from 2014 harvest and consisted of 125 (76 good and 49 defect fruit) while the population 2 from 2015 harvest, includes a total of 310 fruit (234 good and 76 defect samples).Each fruit was marked at two locations on the equator of the fruit, opposite to each other. Fruit were harvested at commercial harvest maturity. Fruit were stored for a week at 10 $^{o}$C before assessment. Few fruit observed air gap between skin and flesh however, most of the fruit didn't observe the air gap.

### 3.2.3 Instrumentation and fruit measurements

#### *3.2.3.1 Visible-SWNIR and IDD2*

Full transmission measurements were acquired using IDD0, InSight2 and IDD2 instruments (for detail, see Appendix 1). Briefly, the IDD0 was developed in-house and utilised a 300 W tungsten halogen lamp and a MMS 1 Zeiss spectrometer (300-1100 nm with interval of 3.3 nm) and was operated using an integration time of 400 ms. The Insight-2 (MAF Roda, France) unit utilises a 150 W tungsten halogen lamp and an Avantes spectrometer (600-973 nm) and was operated using an integration time of 4 ms. The two units were characterised by a repeatability of less than 2 mA in the 600-900 nm range. The IDD2 unit (MAF Oceania) is based on sequential operation of LEDs at four peak wavelengths (700, 810, 780 and 880 nm) with transmitted light detection by a single photodiode detector. For IDD0, and InSight2 instruments, four spectra were averaged from the each of the two marked positions per fruit.

For the IDD0 and InSight2 instruments, dark and white reference measurements were acquired at the initiation of each run, with integration time set to achieve an analogue to digital conversion count with fruit samples of >50% of saturation level (32,000 for IDD0, 64,000 for InSight2). Example spectra (raw analogue to digital conversion count, ADCC) are displayed in Fig. 3.2.

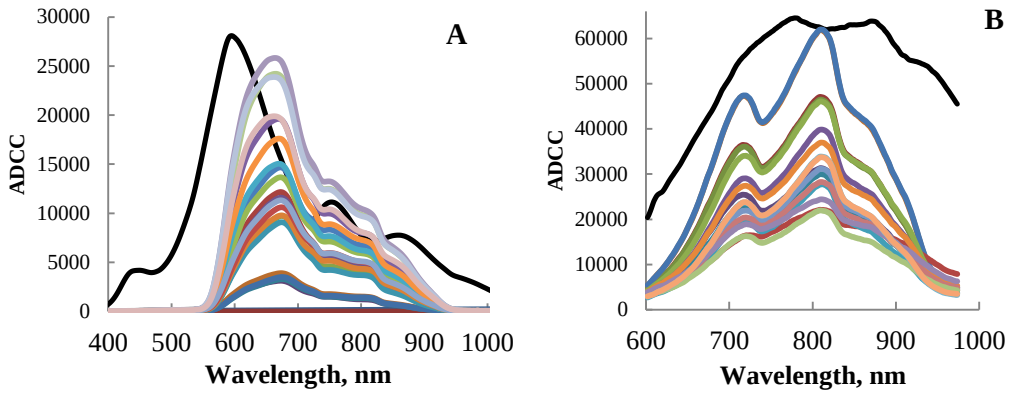



**Figure 3.2.** White reference and sample spectra from IDD0 (A) and Insight2 (B) units, with spectra acquired using an integration time of 400 ms and 2 ms, respectively, with samples stationary and moving, respectively. Black solid line represents white reference.

### 3.2.4 Other non-invasive measures

Each fruit was wrapped with aluminium foil for 24 h prior to measurement of photosynthetic efficiency (chlorophyll fluorescence) using an OptiSciences 30p (Bioscientific, Australia) at a modulation intensity of 9. This unit employs a PIN photodiode with a 700 -750 nm band pass filter. Four readings were taken per fruit within 4 s of removal from the aluminium foil, in a darkened room. Values for a given fruit were averaged.

The Difference in Absorbance (DA) index, calculated as difference in absorbance at 670 nm and 720 nm, was recorded for each marked spot using the DA meter (TR Turoni srl, Italy). Values for a given fruit were averaged.

Resonant frequency was measured using Aweta resonant frequency device (AFS, Aweta, Nootdrop, The Netherlands).

### *3.2.4.1 Destructive reference methods*

*Visual granulation score*

Each fruit were cut transversely at the equator of the fruit and the cut surface photographed using a Canon PC1474 12.1 megapixel digital camera.

The cut surface image was visually scored for the extent of granulation score based on a five point scale, aided by reference images (Fig. 3.3). This subjective score depends on the area affected by the defect and the degree of the defect (whiteness of the tissue). The fruit having visual score 1 to 3 are considered acceptable while those scoring 4 and 5 are considered unacceptable or defect.

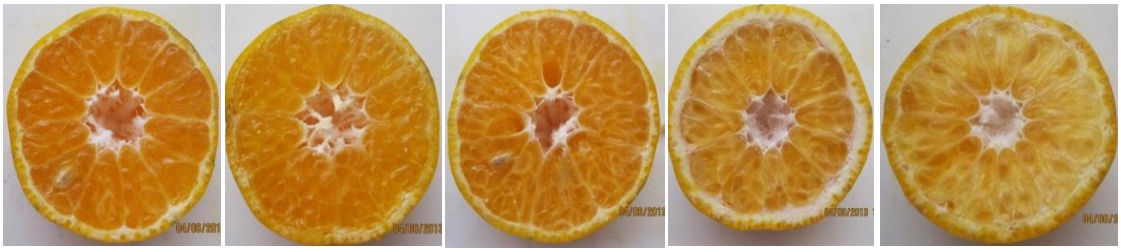


**Figure 3.3.** Visual score (1 to 5 scale) in granulation in 'Imperial' mandarin.

*CIE colour space values*

The fruit CIE Lab colour space was meassured at four locations on the cut surface of each fruit using a Chromameter CR 400 (Konica Minolta; 2 degree observer). Values for a given fruit were averaged. The Chromameter was calibrated using a standard calibration procedure before each lot of Lab measurements.

*Other measurements*

Diameter and weight of each fruit was recorded. Juice was extracted from both hemispheres of the cut fruit using a manual juice extractor. Juice recovery was calculated as total juice weight divided by total fruit weight including peel weight. Total Soluble Solids (TSS) was measured using a refractometer (RFM320, Bellingham and Stanley Limited).

### 3.2.5 Data analysis and Chemometrics

Chemometric analysis was undertaken using The Unscrambler 10.3 (Camo Inc. Oslo, Norway) and Matlab (Mathworks Inc.) software. Principal component analysis (PCA), partial least square regression (PLSR) and multiple linear regression (MLR) were utilised. The classification algorithms of linear discriminant analysis (LDA), partial least square discriminant analysis (PLS DA), support vector machine (SVM) classification and soft independent modelling of class analogy (SIMCA) and multiple logistic regression were trialled for classification of good and defect fruit. General computation and graphs were performed by Microsoft Excel software. The test of significance for chemical components of normal, moderate and severe defect fruit was performed by using Genstat 16.0 (VSN International, UK), using one way analysis of variance (ANOVA) without blocking.

## 3.3 Results and discussion

### 3.3.2 Characterisation of defect

#### *3.3.2.1 Chemical characterisation*

The granulation defect is associated with a decrease in juiciness of fruit although the level of water content is unchanged, except in extreme defect. A population comprising 50 fruit with various degree of granulation shows a dry matter content of 10.02, 10.07 and 12.61% respectively for normal, moderate granulated and severely granulated fruit. Thus the water content of the juice sacs is gelled, raising the possibility of involvement of a galactomann (plant gum). Cell wall chemical characterisation was undertaken to investigate this possibility.

Mandarin juice sac cell wall material was composed of cellulose, fucose, arabinose, rhamnose, galactose, xylose, mannose and lignin (Table 3.1). The primary change in granulated material was a higher glucose content (17% w/w). This result is consistent with an increase in cellulose and hemicellulose, i.e. cell wall, rather than galactomannan, in granulated tissue. Increased hemicellulose content could result in the gelling of water, resulting in decreased juiciness. Future studies might attempt to fraction the cell wall components (e.g. using a mild TFA digest to solubilise hemicelluloses but not crystalline cellulose). The difference of these insoluble cell

wall material for good, moderate defect and severe defect fruit are highly significant at 5% level of significance except fucose and manose.

The soluble (juice) fraction contained glucose, fructose, sucrose and fucose, with a notable decrease in sucrose content only in fruit with severe symptoms (Table 3.2). These results are highly significant at 5% level of significance for glucose while results are significant at 1% level of significance for rest of parameters.

**Table 3.1.** Chemical characterisation (insoluble fraction) of the normal, moderate defect and severe defect 'Imperial' mandarin based on tissue analysis using acid hydrolysis method (average for 5 samples). Unit in mg / 100 mg of dry weight of insoluble fraction.

| | | **Fucose** | **Arabinose** | **Rhamnose** | **Galactose** | **Glucose** | **Xylose** | **Manose** | **Lignin** |
|---|---|---|---|---|---|---|---|---|---|
| Normal | Mean | 0.68 | 7.88 | 1.52 | 3.77 | 17.62 | 4.55 | 4.38 | 3.07 |
| | SD | 0.03 | 0.74 | 0.06 | 0.04 | 0.37 | 0.04 | 0.05 | 0.83 |
| Moderate | Mean | 0.74 | 4.37 | 1.70 | 3.28 | 22.72 | 5.14 | 4.82 | 5.80 |
| | SD | 0.08 | 0.45 | 0.27 | 0.46 | 3.86 | 0.27 | 0.28 | 1.24 |
| Severe | Mean | 0.80 | 4.66 | 1.31 | 2.91 | 20.59 | 5.41 | 5.35 | 5.35 |
| | SD | 0.12 | 0.47 | 0.20 | 0.23 | 0.79 | 0.74 | 1.06 | 0.54 |

**Table 3.2.** Chemical characterisation (soluble fraction) of the normal, moderate defect and severe defect 'Imperial' mandarin based on tissue analysis using acid hydrolysis method (n=5). Unit in mg / 100 mg of dry weight of soluble fraction.

| **Sample** | **Statistics** | **Glucose** | **Fructose** | **Sucrose** | **Fucose** | **Total** | **Gluc:Fruc** |
|---|---|---|---|---|---|---|---|
| Normal | Mean | 14.97 | 50.31 | 40.84 | 0.40 | 106.5 | 0.29 |
| | SD | 3.76 | 3.38 | 6.39 | 0.06 | 9.42 | 0.07 |
| Moderate | Mean | 11.6 | 61.23 | 21.34 | 0.70 | 94.9 | 0.19 |
| | SD | 1.42 | 3.91 | 3.21 | 0.15 | 5.13 | 0.03 |
| Severe | Mean | 9.76 | 79.21 | 12.92 | 1.15 | 103.1 | 0.12 |
| | SD | 1.60 | 6.75 | 3.75 | 0.42 | 5.31 | 0.02 |

#### *3.3.2.2 Anatomical characterisation*

Based on both light and scanning electron microscopy observations (Figs. 3.4, 3.5), juice sacs from control fruit were characterised by large internal cells, enveloped by an epidermal and hypodermal layers, approximately three to five cells deep (Fig. 3.4 A,B). These layers were composed of smaller cells, relative to the internal cells. In granulated fruit these hypodermal layers were increased. In extremely affected fruit (fruit with no extractable juice), the cell wall thickness of juice cells was also increased (Fig. 3.6). These cell walls appeared to consist of an expansion of the primary cell wall only (based on colour of toluidine blue staining).

Thus, anatomical observations are consistent with granulation involving a proliferation of juice sac hypodermal layers, such that juice is less easily released from the juice sac. Only in severely granulated fruit were juice sacs collapsed, consistent with loss of water. This description is different to that for citrus drying defects reported by Albrigo et al. (1980), in which cell wall thickening and cell collapse was reported as a general symptom. Cell division and maturation within juice sacs occurs early in the growth of a fruit (Hofman, 2011; Peiris et al., 1998). The mature juice sac cells are highly specialized for storage and would presumably not be able to de-differentiate and divide. It is unlikely that the cell proliferation associated with the granulation could occur late in fruit development. Thus the disorder must be ‘set’ early in fruit development through the effect of some assembly of environmental conditions. Future studies could usefully consider the extent of cell proliferation in the juice sacs during fruit development as a function of the final extent of the disorder in mature fruit.

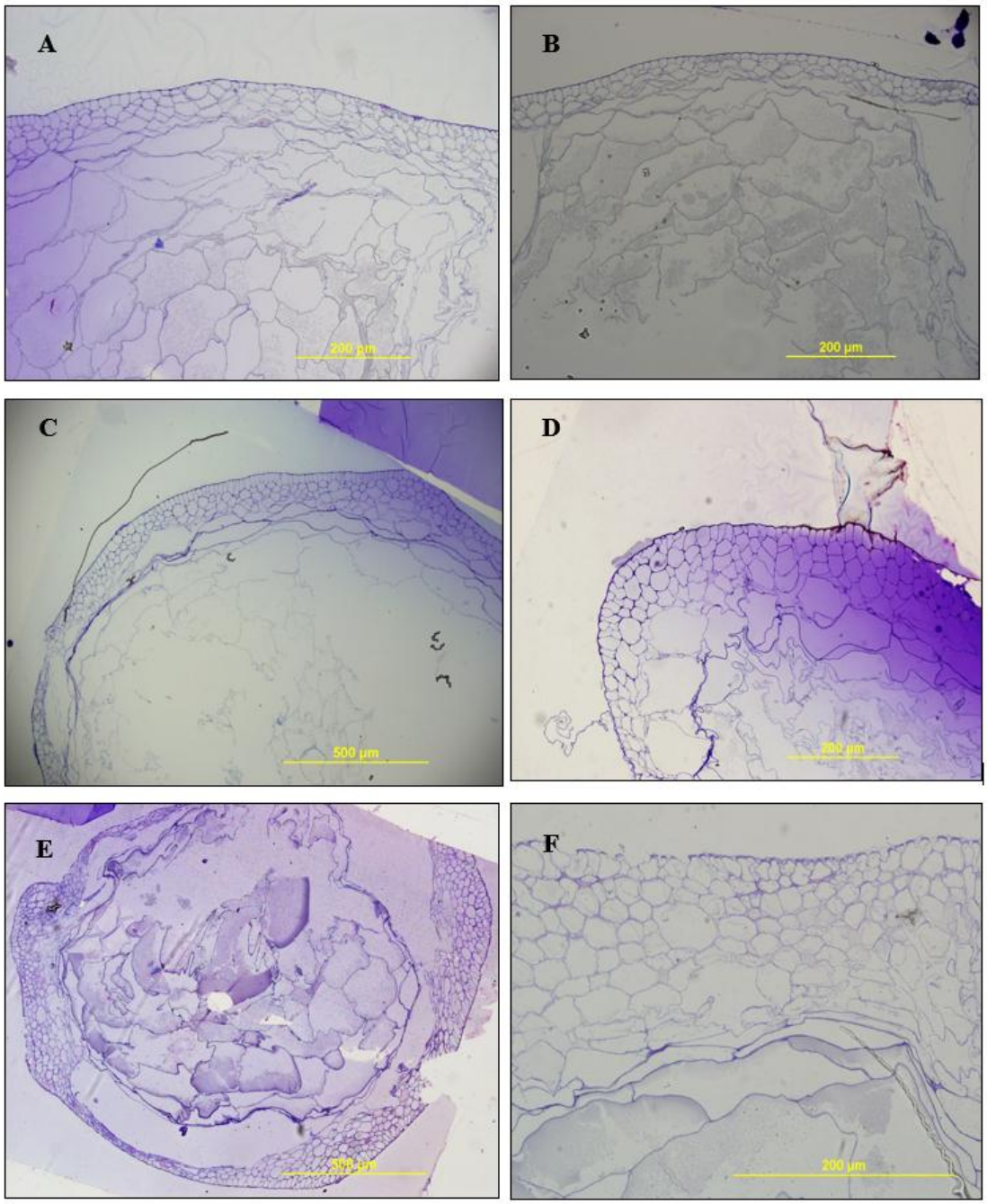



**Figure 3.4.** Light microscopy of toluidine blue stained sections of resin embedded juice sacs, for (A, B) control fruit; (C, D), fruit with moderately severe symptoms of granulation, and (E, F) severely affected fruit. The number of cell layers in the hypodermal tissue increased in drynness affected fruit. However, there was no apparent difference in cell wall thickness.

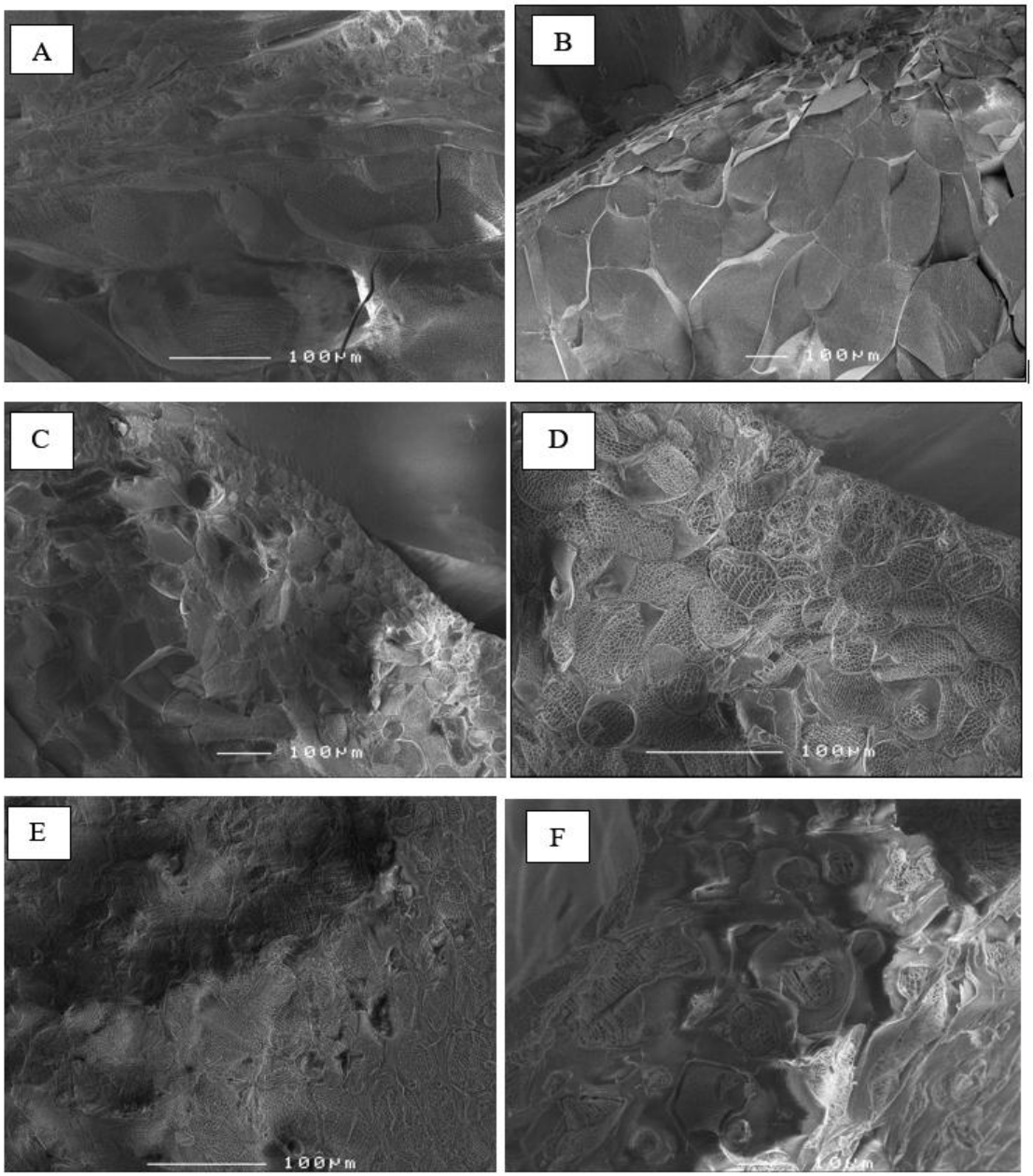



**Figure 3.5.** Scanning electron microscopy of the periphery of cryo-section frozen juice sacs, for (A, B) control fruit; (C, D) fruit with moderately severe symptoms of granulation; and (E,F) extremely affected fruit (no extractable juice). Collapse of juice sacs is evident in severely affected fruit.

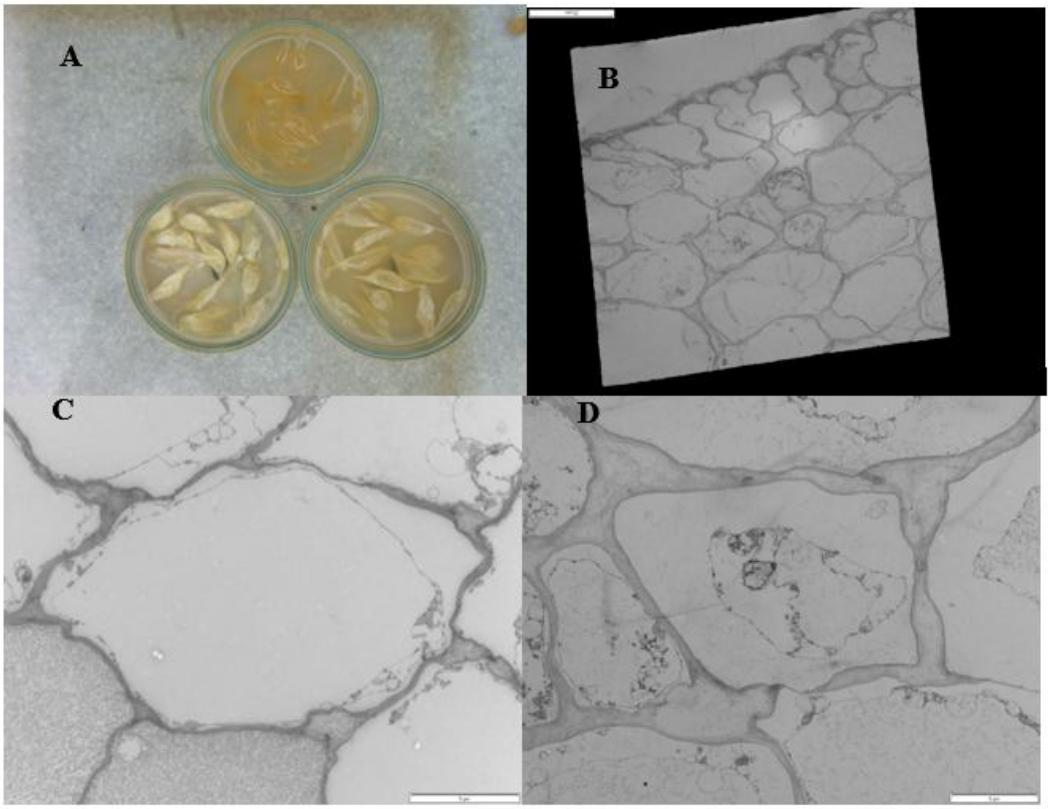



**Figure 3.6.** Transmission electron microscopy of section resin embedded juice sacs Photograph of (A) juice sacs used in microscopy work, (B) normal fruit, (C) moderate defect and (D) and severe defect.

### 3.3.3 Fruit sample and population structure

Statistics on the populations of fruit used in terms of reference quality parameters is presented in Table 1. Unfortunately, different fruit populations were used with each instrument, due to instrument availability windows.

**Table 3.3.** Population statistics on visual score (5 point scale) and juice recovery (% w/w, peel included) for calibration and prediction sets in season 1 and season 2. Data presented as mean ± SD.

| | IDD0 instrumentation | | Insight2 instrumentation | |
|---|---|---|---|---|
| Set | Calibration | Prediction | Calibration | Prediction |
| Sample | 75 | 50 | 200 | 110 |
| Visual score | 3.0 ± 1.4 | 2.8 ± 1.4 | 2.2 ± 1.3 | 2.9 ± 1.4 |
| Juice recovery | 21.5 ± 7.9 | 21.8 ± 7.9 | 28.4 ± 7.9 | 29.0 ± 7.4 |
| CIE L* value | | | 51.9 ± 4.2 | 49.7 ± 4.1 |

### 3.3.4 Non spectral measures as an index of granulation

As with any sorting operation, clarity on the attribute to be assessed is critical. Consumers object to the granulation defect on two grounds, a visual assessment (whiteness of the segment contents, as revealed when bitten) and on an eating quality basis (dry mouth feel). Visual score on a 5 point scale and cut surface luminosity, and % juice recovery were assessed in an attempt to provide relevant objective measurements of the defect. However the linearity of such measures with a physical (e.g. scattering) or chemical (e.g. water content) fruit attribute associated with granulation defect is not clear.

Relationships between reference attributes were only weakly correlated (e.g. correlation coefficient of determination between juice recovery and visual score of $R^2 = 0.57$, and with L value of cut surface, of $R^2 = 0.83$, Fig 3.7). A non-linearity is evident in the relationship between score and % juiciness, with visual score 5 fruit having low juiciness. There was effectively no relationship between DA meter reading, resonant frequency, chlorophyll fluorescence (Fv/Fm) or juice TSS and visual score (Table 3.4).

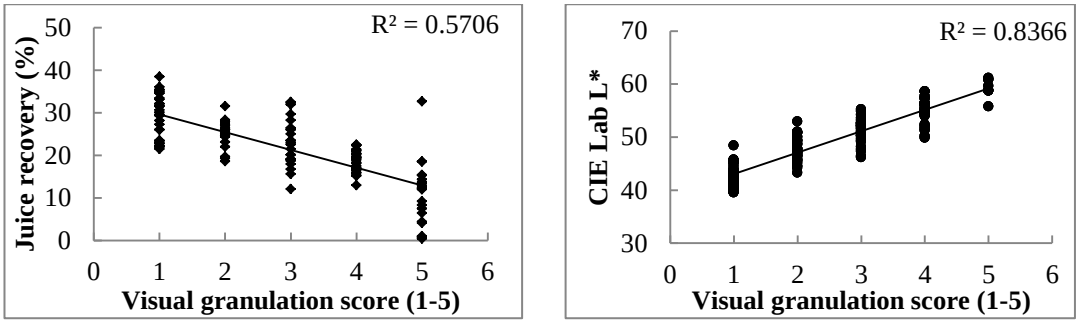



**Figure 3.7.** Scatter plot of juice recovery and surface Luminosity and visual score (Season 1, n=125).

Cut surface L* value is thus a potential quantitative reference method for assessing this disorder, to mimic the human observation, while % juiciness is a measure related to eating experience.

**Table 3.4.** Correlation coefficient of determination ($R^2$) between reference parameters

| | Score | Juice recovery | DA | Sound frequency | Chlorophyll Fluorescence | TSS |
|---|---|---|---|---|---|---|
| Score | 1 | | | | | |
| Juice recovery | - 0.57 | 1 | | | | |
| DA | 0.0017 | 0.0078 | 1 | | | |
| Sound Frequency | 0.0001 | 0.0227 | 0.0183 | 1 | | |
| Fv/Fm | 0.00004 | 0.0152 | 0.0092 | 0.0322 | 1 | |
| TSS | 0.0096 | 0.0107 | 0.0042 | 0.0035 | 0.16 | 1 |
| CIE Lab L* | 0.84 | 0.34 | | | | 0.0019 |

### 3.3.5 Spectral features – linear regressions

The average (IDD0) absorbance spectra of defect (visual score 5, n= 88) fruit was higher than that of good (visual score 1, n = 248) fruit across the wavelength range 500 - 980 nm, with a maximum difference at 578 nm (Fig. 3.8). This result is consistent with higher scattering of light in defect fruit, as manifest in the increased L value of the cut fruit surface.

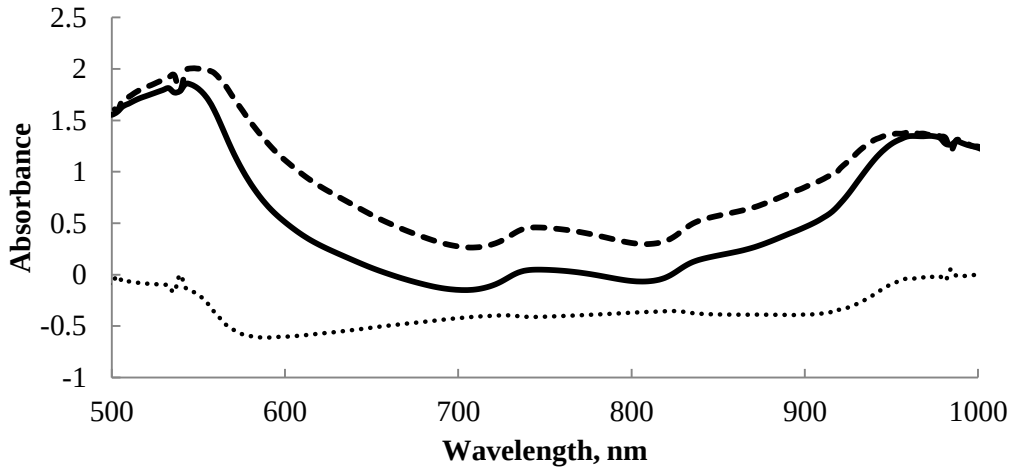



**Figure 3.8.** Average IDD0 absorbance spectra of good (n = 30 fruit; solid line) and defect (n = 22 fruit; dashed line) fruit and the difference spectra (good – defect; dotted line)

However the linear correlation between the absorbance at the single wavelength of 578 nm and the visual granulation score was relatively poor ($R^2$ = 0.3) and is

therefore not useful as an index for discriminating good and defect fruit (Fig. 3.9). The relationship between Abs578 and score demonstrated non-linearity, with higher absorbance associated with score 5 fruit.

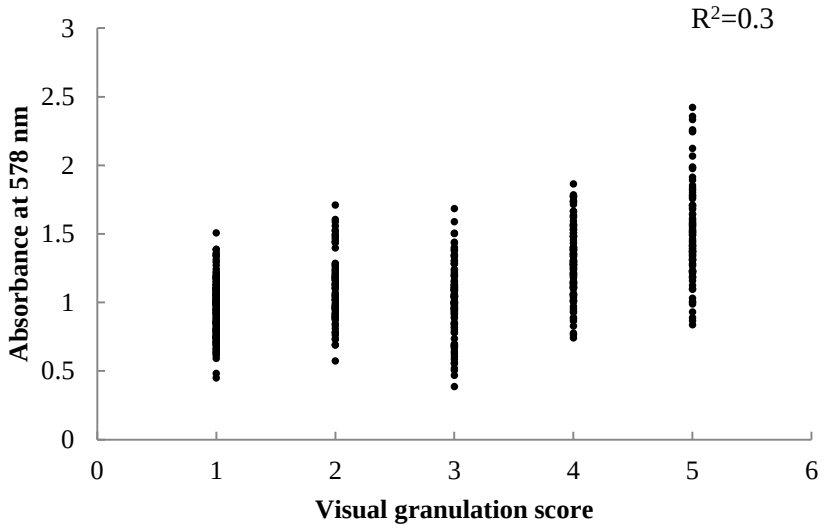



**Figure 3.9.** Scatter plot of apparent absorbance at 578 nm and visual score of cut surface.

This consideration of the relationship between defect attribute level and absorbance at a single wavelength was extended to all wavelengths (Fig. 3.10). A correlation coefficient R of approx. -0.75 existed between apparent absorbance at wavelengths between 600 and 920 nm and % juiciness (scale 1-5) (Fig. 3.10). The relationship between absorbance and score was similar, if slightly weaker. Water absorption features (e.g. as expected for the second overtone of the O-H stretch at 960 nm) were not weighted, consistent with spectral information relevant to this defect being related to scattering rather than water content.

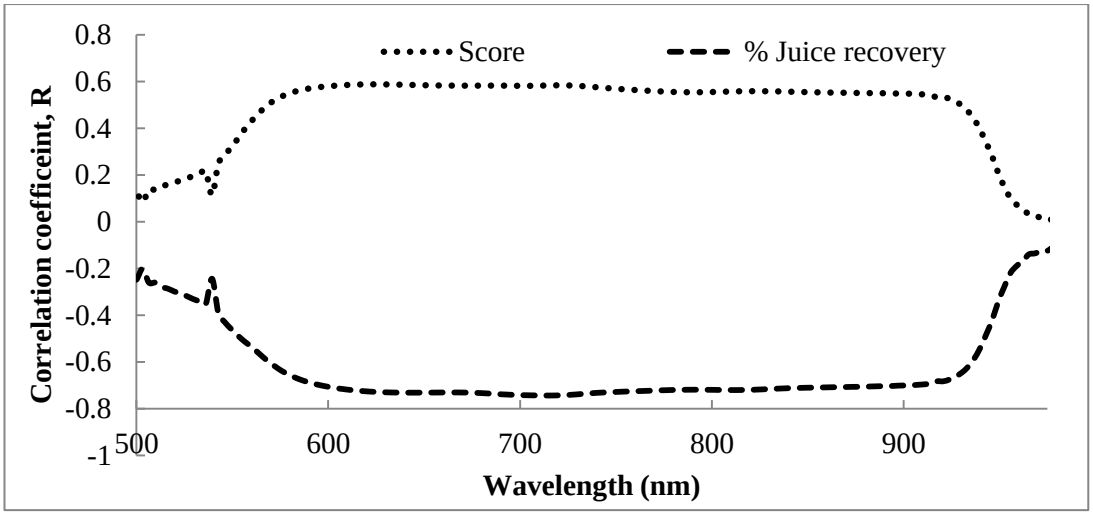



**Figure 3.10.** Univariate correlation coefficient (R)between absorbance at a given wavelength for internal quality parameters visual score and % juice recovery for mandarin (n=125) .

Multiple linear regression on score achieved a calibration $R^2_c$ = 0.41, RMSEC = 0.97 with use of absorbance at of wavelengths 621, 634, 667 and 790 nm. For % juiciness, a result of ($R^2_c$ = 0.6, RMSEC = 5.16) was achieved, using wavelengths 611, 617, 621, 732, 738, 832 and 903 nm.

The regression coefficients for a PLS regression model based on absorbance data and defect score using the entire available wavelength range were noisy at wavelengths below 550 and above 870 nm, suggesting little information was carried in these regions. The PLS model awarded high positive coefficient values to 550, 575, 661 and 765 nm, and high negative values to 560, 621, 713 and 869 nm.

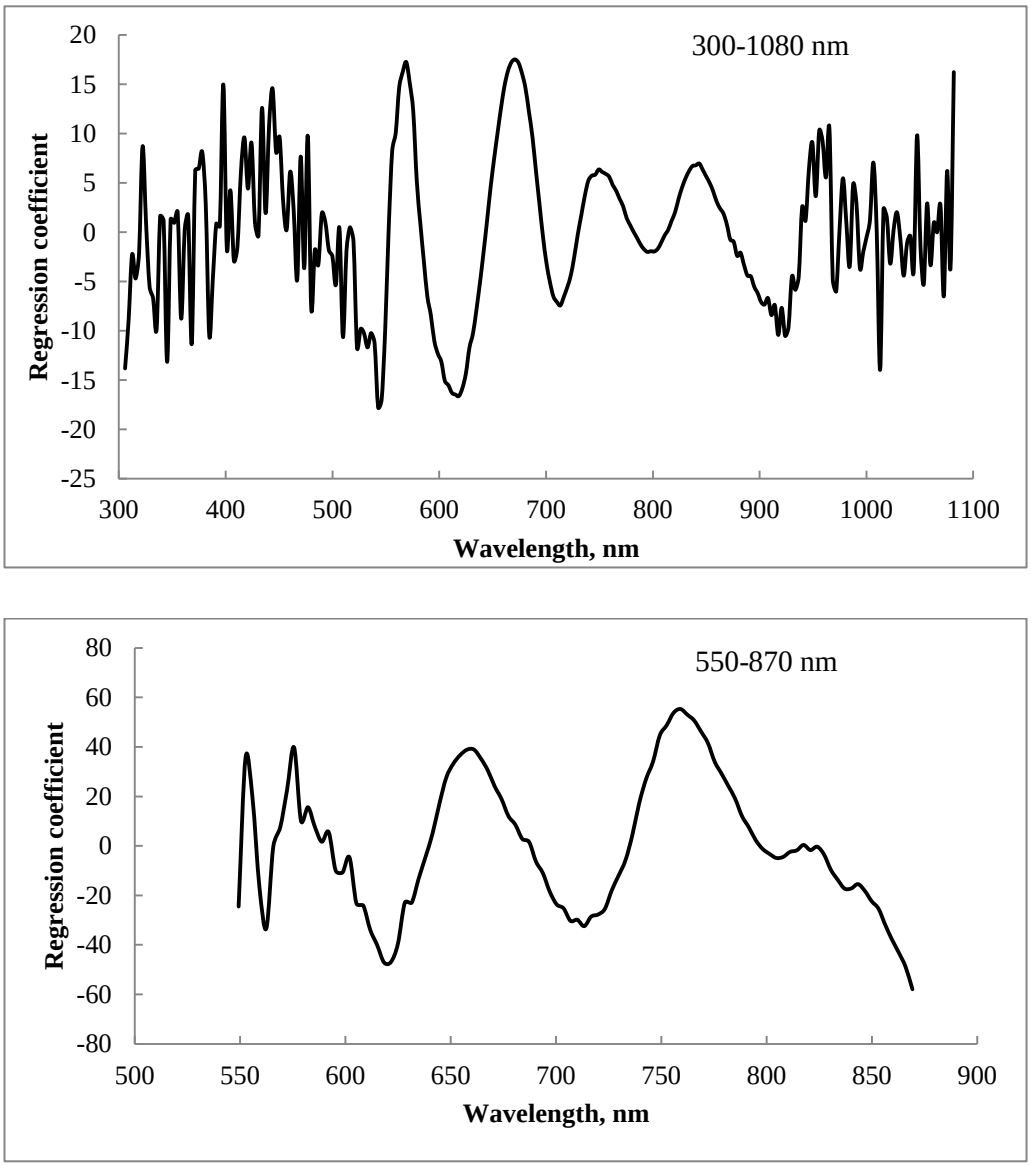



**Figure 3.11.** Regression coefficients for a PLSR model of visual score using absorbance values over the wavelength range of 300-1080 nm (top) and 550-870 nm (bottom).

The predictive performance of the PLSR model was improved by use of pre-processing treatments (IDD0 spectra, Table 3.5). The best result, in terms of prediction of an independent set, was achieved with standard normal variate and second derivative pre-treatments. Similar results were achieved for PLSR models based on spectra of the Insight2 unit (Table 3.6). However, while better results were obtained for the % juiciness model than the visual score model with the IDD0 unit, the reverse was true for the Insight2 unit (Table 3.6). This result could be due either

to the difference in wavelength ranges of the two instruments, or aspects of the two populations of fruit.

**Table 3.5.** Partial least square regression (PLSR) model performance for IDD0 spectra at 550-870 nm for (A) visual granulation score and (B) % juice recovery. (Season 1)

| Parameter | Calibration set (n = 75) | | | Prediction set (n = 50) | | |
|---|---|---|---|---|---|---|
| Visual score | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | bias |
| Abs | 0.6 | 0.87 | 10 | 0.33 | 1.35 | 0.6 |
| Abs SNV | 0.56 | 0.93 | 9 | 0.45 | 1.1 | 0.27 |
| Abs MSC | 0.51 | 0.98 | 10 | 0.49 | 1.04 | 0.19 |
| d2A | 0.56 | 0.93 | 7 | 0.49 | 1.04 | 0.27 |
| SNV d2A | 0.55 | 0.94 | 8 | **0.49** | **1.07** | **0.33** |
| MSC d2A | 0.54 | 0.95 | 7 | 0.47 | 1.06 | 0.28 |
| % juice | | | | | | |
| Abs | 0.76 | 3.64 | 9 | 0.43 | 6.49 | -1.36 |
| Abs SNV | 0.76 | 3.61 | 7 | 0.49 | 6.01 | -0.53 |
| Abs MSC | 0.65 | 4.23 | 11 | 0.6 | 5.22 | -0.71 |
| d2A | 0.72 | 3.97 | 7 | 0.66 | 5.04 | -0.75 |
| SNVd2A | 0.74 | 3.59 | 7 | **0.68** | **4.87** | **0.13** |
| MSC d2A | 0.66 | 4.37 | 7 | 0.66 | 4.86 | -0.34 |

**Table 3.6.** Calibration and prediction statistics based on the partial least square regression (PLSR) of SNV d2A spectra at 600-973 nm for Insight2 taking visual granulation score (1-5) as reference parameters.

| Instrumentation | Insight2 instrumentation | | | | | |
|---|---|---|---|---|---|---|
| | Calibration statistics (n = 200) | | | Prediction statistics (n = 110) | | |
| Parameter | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias |
| Granulation score | 0.63 | 0.8 | 7 | 0.62 | 1.05 | -0.57 |

### 3.3.6 Spectral features – discriminant analysis

MLR and PLSR are essentially linear regression techniques, although PLS can handle a degree of non-linearity in the data. However, the level of granulation defect in mandarin fruit as assessed by visual score, luminosity or % juiciness, does not necessarily link to a linear quantitative change in a physical or chemical attribute with an associated spectral feature, in the way that, e.g. water content is related to dry matter content, with water having clear absorbance features in the SWNIR. Therefore use of a discriminant technique rather than a regression technique is logical for assessment of this attribute.

Various algorithms were trialled for classification of fruit as defect or sound based on spectral information over the range 550-870 nm and 600-973 nm for the IDD0 and Insight 2 units respectively (Table 3.7). For IDD0 instrumentation, the highest prediction sorting accuracy was achieved with a PLS-DA routine, while for InSight2, best results were obtained using a PCA-LDA-MD or a PLS-DA (undertaken in Matlab).

**Table 3.7.** Results of several algorithms for classification of good and defect fruit based on visual score using IDD0 and Insight 2 instrumentation using raw absorbance spectra at 550-870 nm for IDD0 and 600-973 nm for Insight2 data. Units in percentage.

| | IDD0 | | | | | | Insight 2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Calibration statistics (n = 75) | | Prediction statistics (n = 50) | | | | Calibration statistics (n = 200) | | Prediction statistics (n = 110) | | | |
| Classification methods | TPR | TNR | TPR | TNR | Accuracy | FDR | TPR | TNR | TPR | TNR | Accuracy | FDR |
| PLS-DA (Unsb.) | 87.5 | 81.0 | 87.0 | 97.0 | 92.0 | 3.33 | 98.00 | 59.25 | 100 | 9.83 | 54.9 | 47.4 |
| PCA LDA Linear 8 PCs | 87.7 | 69.2 | 98.1 | 43.1 | 70.6 | 36.7 | 88.75 | 78.75 | 83.63 | 43.58 | 63.6 | 40.3 |
| PCA LDA MD 5 PCs | 82.9 | 75.5 | 96.2 | 50.0 | 73.1 | 34.2 | 98.75 | 37.5 | 97.3 | 47.2 | 72.2 | 35.2 |
| KNN (4 neighbours) | 94.9 | 81.6 | 85.4 | 42.0 | 63.7 | 40.4 | 98.3 | 86.9 | 92.2 | 34.0 | 63.1 | 41.7 |
| SIMCA | 97.8 | 48.6 | 89.2 | 43.75 | 66.5 | 38.6 | 98.8 | 30 | 100 | 2.77 | 51.4 | 49.3 |
| SVM. (Csvc Linear) | 98.0 | 40.1 | 89.6 | 25.00 | 57.3 | 45.6 | 98.8 | 37.5 | 100 | 4.16 | 52.0 | 48.9 |
| PLS-DA-matlab | 89.8 | 67.9 | 91.4 | 48.2 | 69.8 | 36.2 | 86.4 | 82.5 | 80.7 | 52.8 | 66.7 | 36.9 |

### 3.3.7 IDD2 two wavelength model

IDD 2 unit employs LEDs producing four wavelengths. Using a population of 160 fruit (a subset of Season 2 set), the use of various ratios of these wavelengths were assessed (data not shown), with the best result for detection of granulation defect obtained using Abs 810/Abs 700 nm. A total of 160 fruit were scanned using the IDD2 instrumentation. (Fig. 3.12).

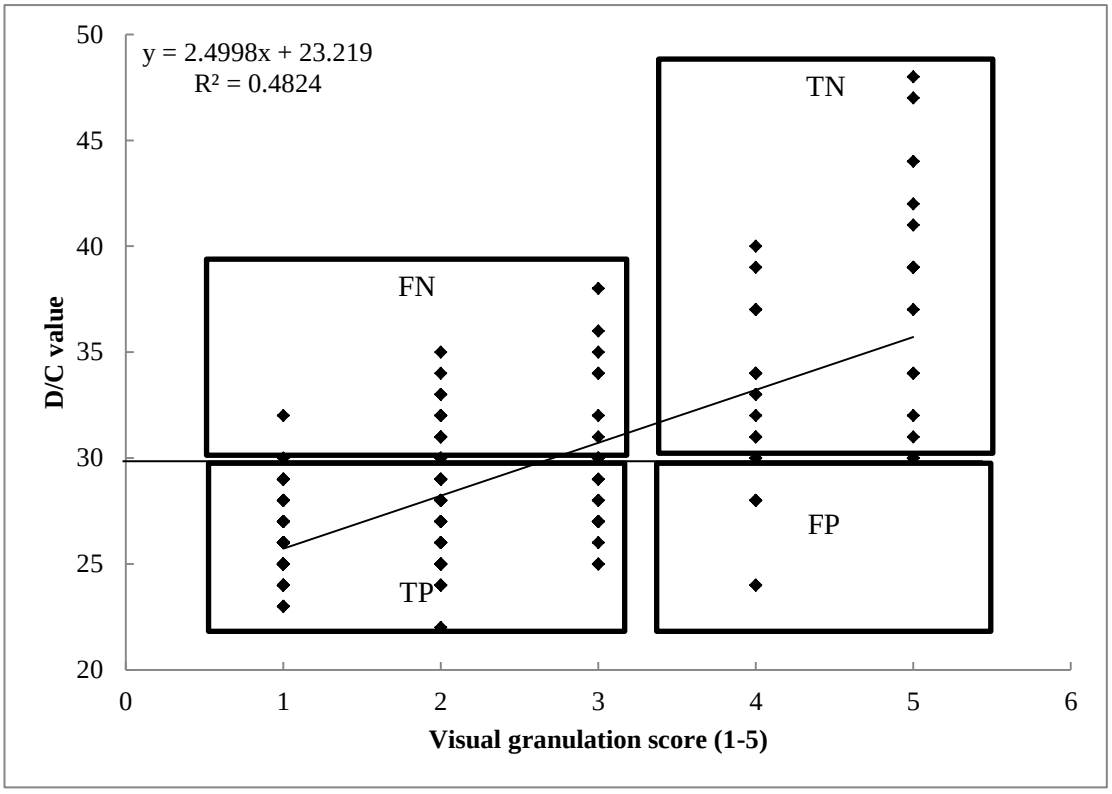



**Figure 3.12.** Scatter plot of ratio of absorbance at 810/700 nm and visual score, for a population involving 160 fruit. FN = false negative, TN = true negative, TP = true positive, FP = false positive.

### 3.3.8 IDD2 sorting optimisation

The impact of varying the threshold set on a sorting function on output classes was considered for the IDD2 results for a population involving 160 fruit (126 or 79% score 1-3 and 34 or 21% score 4 and 5) (Fig. 3. 13). The population distribution of defect level was shifted to lower values following removal of fruit above decreasing threshold values, although this was at the expense of the number of accepted fruit.

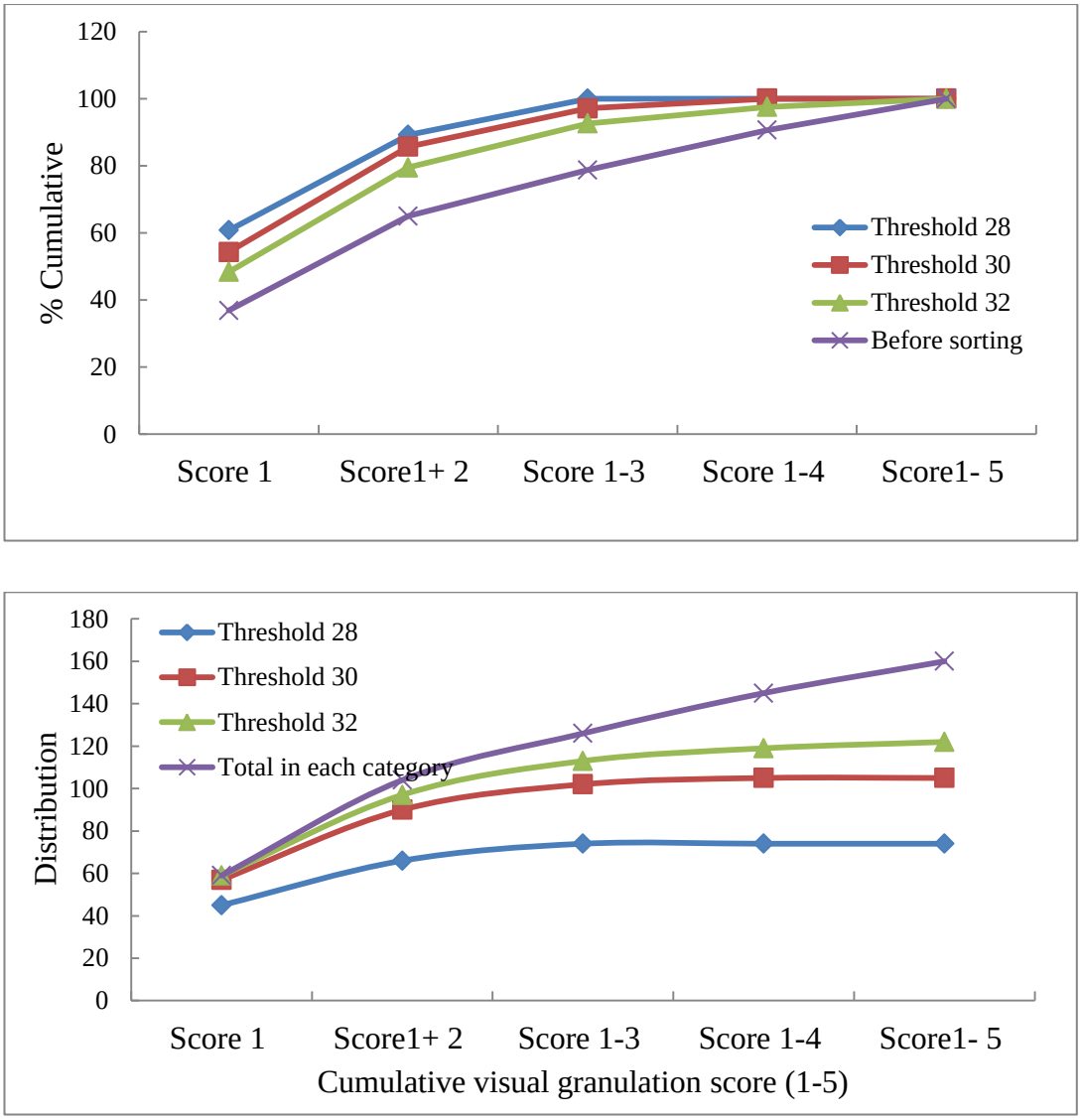



**Figure 3.13**. Cumulative distribution of mandarin defect for populations before and after sorting, to various threshold values on the IDD2 detection result. A. Distribution based on % B. Distribution by fruit number.

The sorting results wer re-presented to emphasise features of interest to operator with change in threshold value (Fig. 3.14). For this population, a false discovery rate (FP/(FP+TP)) of 5% (a market tolerance point) was achieved at a threshold of 29, associated with a 67% yield. If a false discovery rate of 10% is acceptable, a threshold of 31 can be used, achieving a yield of 78%. (Fig. 3.14).

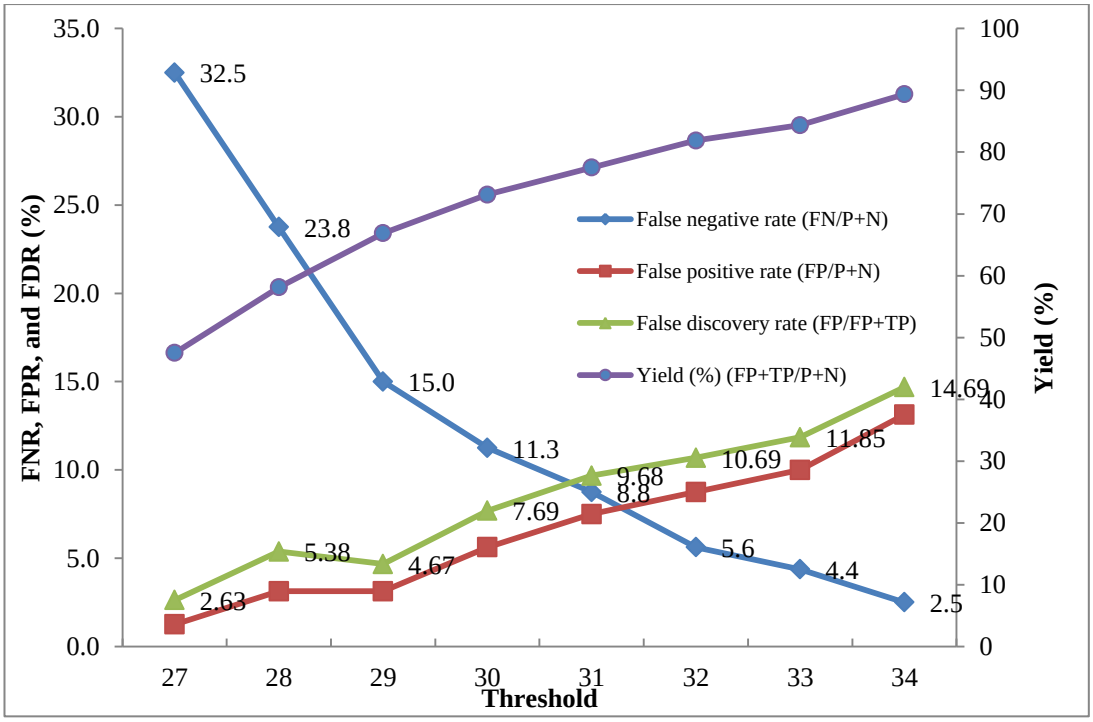



**Figure 3.14**. Effect of threshold sorting value on classification error for mandarin defect (with score 1 to 3 considered acceptable fruit, and score 4 and 5 deemed defect fruit) for population 2 (involving 160 fruit).

## 3.4 Conclusion

The granulation is an economically important defect in mandarin. The defect was characterised as involving proliferation of hypodermal cells within juice sacs, and consequent increase in cell water material, likely celluloses and hemicelluloses. This material is hypothesised to act in the gelling of cellular water, resulting in the observed decrease in % recovery of juice without decrease in water content, except in severe granulation where juice sacs collapsed. This renders the defect hard to detect based on water content differences. The increase in cell number is consistent with increased light scattering, and thus an effect on the luminosity of defect fruit. For non-invasive sorting of fruit using SWNIR spectroscopy, a discriminant analysis approach, such as PLS-DA, was recommended over linear regression techniques such as PLS and MLR. However, a simple two wavelength discriminator was also shown to be useful, with an accuracy of 92% demonstrated for the assessed population. Further trials are recommended to demonstrate robustness of detection for fruit of a range of growing conditions.

# Chapter 4.
# Non-invasive assessment of pineapple translucency using NIR transmission measurements [1]

## Abstract

Pineapple fruit with translucent flesh were not characterised by higher fruit density or TSS levels. Non-invasive assessment of this disorder in intact fruit was trialled using near infrared (300-1100 nm) transmission measurements using a total of 222 fruit that were cut and rated on a scale of 1 (no translucency) to 5 (extremely translucent). PLS regression at 500-1000 nm region yielded a prediction model ($R^2_p$ = 0.41, RMSEP = 0.93 and bias -0.36 units of score 1-5 on defect level) while linear discriminant analysis achieved classification accuracy (TP+TN/P+N) of 72% for prediction of independent population using IDD1 instrumentation whereas for IDD0 instrumentation, the classification accuracy of 66% was achieved. While this accuracy is not suitable for use by industry, the results are promising and suggestions are made for further work.

**Keywords:** score, defect, transmission geometry, classification

[1] I acknowledge the support of Col Scott from Tropical Pines P/L for fruit samples and market information on pineapple.

## 4.1 Introduction

Australia produced and marketed 89,099 t (tonnes) of pineapple (*Ananas comosus* (L.) in 2013-2014 season (ABS, 2015). In 2012-2013 season, production was 48,000 and 39,000 t for fresh and processing, valued $70 million (DAF, 2015). Queensland produced approximately 80 % of the national crop. The farm value of pineapple external quality is judged on size, shape, colour and absence of defects, while internal quality includes total soluble solids (TSS), dry matter, titratable acidity (TA), protein (bromelain) levels and internal defects (Haff et al., 2006; Hong et al., 2013).

Internal defects include internal browning, blackheart and translucency (Collins, 1968). Internal browning is associated with low temperature storage and affects flesh and core (Sukwanit & Teerachaichayut, 2013) and can be avoided with proper temperature management. Blackheart is a form of chilling injury associated with temperatures below 10 $^{o}$C in field or in storage (Zhou et al., 2003). Blackheart is considered as a major postharvest problem in Australia, often limiting refrigerated export with corresponding annual loss of US$ 1.3 million out of total production worth of US$ 30 million (Ko et al., 2006 ). Waxing, controlled atmosphere storage and heat treatment are recommended as effective methods for reducing blackheart incidence (Abdullah et al., 2010). Translucency is regarded as an issue to deal with when it exceeds 50% leading to decrease in palatability and posing transportation problem (www.daf.qld.gov.au).

Translucency is associated with growing condition, mainly pre-harvest rain and temperature (Chen & Paull, 2000). It is a common cause of reduced quality in Thailand (Pankasemsuk et al., 1996). The disorder is characterized by the presence of water soaked flesh, a symptom - due to accumulation of sugar and other metabolites in the intercellular spaces (Chen & Paull, 2000), similar to that occurring in mangosteen (*Garcinia mangostana* L.) and apple (*Malus domestica* Borkh.) (Herremans et al., 2014b; Pankasemsuk et al., 1996). The physiological basis of these disorders is based on sugar loss to the apoplast and consequent osmotic movement of water into intercellular spaces.

It is to be expected that the filling of apoplast spaces with water should result in changes in the optical properties of the fruit, e.g. less scattering and greater transmission. Short wave NIR spectroscopy (SWNIRS) appears suited to assessment of such a disorder, given its assessment of apparent absorption (i.e. both scattering and absorption) and the relatively low absorptivity by water, allowing for measurement of light passage through an intact fruit (Nicolaï et al., 2007; Subedi, 2007).

Teerachaichayut et al. (2007) reported prediction of internal translucency of intact mangosteen fruit using a transmission optical geometry and a wavelength range of 640-980 nm, taking the percentage weight of translucent flesh as the reference method with accuracy of classification (correctly classifying good and defect fruit out of the total sample) of 89%. The spectra were acquired using a 200W light source with an integration time of 78 ms. Using the same apparatus, Terdwongworakul et al. (2012) reported use of PLS regression based on pre-treated spectra (MSC, SNV, first and second derivatives, 665–955 nm) of intact mangosteen fruit and percentage area affected by translucency in images of the transversely cut fruit (as analysed using Image J software). The highest correlation coefficient of determination ($R^2 = 0.74$) was obtained with the stem to crown axis horizontal and at $135^o$ angle between the light source and detector. A similar type of disorder in apple (watercore) has been non-invasively detected using light (SWNIR) transmission, X-ray CT, MRI and thermography (Beaudry, 2014; Herremans et al., 2014b; Olsen et al., 1962; Wang et al., 1988).

Non-invasive assessment of several characters of pineapple fruit has been reported, using vis-SWNIRS, X-ray imaging and electronic noses (Chia et al., 2012; Sukwanit & Teerachaichayut, 2013; Torri et al., 2010). Guthrie et al. (1998) reported relatively poor results for assessment of TSS of intact pineapple using vis-SWNIRS, a result ascribed to the high variability of TSS and optical properties in this compound and accessory fruit, in which optical assessment involves passage of light through several tissue types (e.g. bract, vascular core). Assessment of TSS using vis-SWNIR reflectance spectroscopy and artificial neural networks was also reported by (Chia et al., 2012, 2013). Pathaveerat et al. (2008) reported the use of specific gravity and stiffness coefficient as estimated from an acoustic measurement for assessment of

pineapple fruit maturity. Fourier Transform (FT) NIR and FT-IR spectroscopy has been used for assessment of shelf life of fresh cut pineapple with corresponding change in SSC, TA and pH during 10 days of storage at three different temperatures (Egidio et al., 2009). Torri et al. (2010) reported the use of an electronic nose for monitoring of volatile compounds in relation to quality assessment.

Haff et al. (2006) utilized an X-ray imaging technique for detection of translucency in pineapple. Visual observation of the X-ray images was undertaken to separate affected fruit. Fruit were then cut and assigned a translucency score on a 1-5 scale. A true positive rate of 86% (predicted good/total good) and a true negative rate of 95% (predicted defect/total defect) was reported. More recently, Donis-González et al. (2014) reported use of X-ray computed tomography imaging the extent of translucency in intact fruit.

Sukwanit and Teerachaichayut (2013) reported use of SWNIRS for the non-invasive assessment of pineapple internal browning using a $90^{o}$ partial transmission optical geometry (Purespect HOS-200 spectrometer, Japan) to collect spectra over the range 665-955 nm. Spectra were collected from six locations on the fruit and averaged, but the reference method involved a medial slice only, with visual categorisation as sound or defect. PLS-DA was used, with a 0.5 threshold value resulting in a true positive rate of 89% and a false positive rate of 93%.

There has been no report of use of visible-SWNIRS to non-invasively detect translucency in pineapple. This exercise explores the use of SWNIR in partial transmission, as used by Sukwanit and Teerachaichayut (2013) in assessment of internal browning in pineapple fruit, for the application of assessment of pineapple translucency.

## 4.2 Materials and Methods

### 4.2.1 Samples and sample preparation

A total of 222 mature fruit (cultivar Hybrid Gold) of commercial harvest with shell color changes from green to yellow at the base of fruit were used, in two sets (season 1, n = 138; 106 normal and 32 translucent; season 2, n = 84; 59 normal and 25 translucent), sourced from a commercial pineapple farm in Yeppoon, Queensland. Season 2 fruit consists of four lots of fruit each collected and scanned at weekly

intervals. Fruit were stored at 15 $^{o}$C for up to a week before experimentation to let them to optimum ripening for consumption. On the day of assessment, fruit were allowed to equilibrate to room temperature for 6 h. Each fruit was marked at four equidistant locations around the equator of the fruit. The diameter of fruit at equator ranged from 63 to 92 mm.

### 4.2.2 Instrumentation and spectra acquisition

Four spectra per fruit were acquired, at each marked spot. Spectra were acquired using a MMS 1 spectrometer (silicon photodiode detector array operating over the range (308-1100 nm) at an integration time adjusted to achieve a signal >67% of saturation intensity. Integrations times as high as 1000 ms were required with less mature fruit. Two purpose built systems were used:

(i) a single 300W halogen lamp used in a partial transmission geometry (Fig. 4.1), with the stem - crown axis of the fruit horizontal, and a light source – sample – detector angle of 90$^{o}$.

(ii) eight 60 W halogen lamps (four on either side of the fruit) illuminating the fruit at 90$^{o}$ to the point of detection (Fig. 4.2), with the stem-crown axis of the fruit horizontal, a light source – sample – detector angle of 90$^{o}$ and the detector mounted under the fruit.

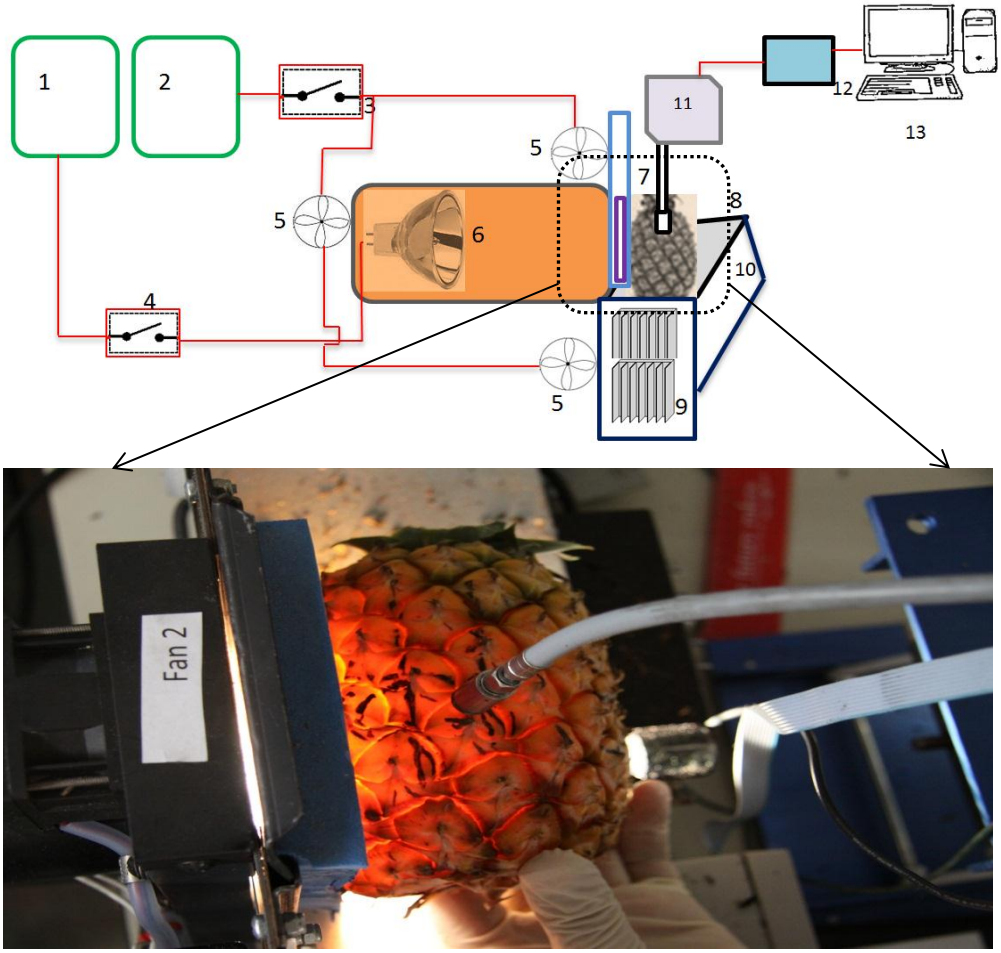



**Figure 4.1**. Schematic (top) and image (below) for a single 300 W light source unit, designated IDD0, used in acquisition of spectra of pineapple fruit.

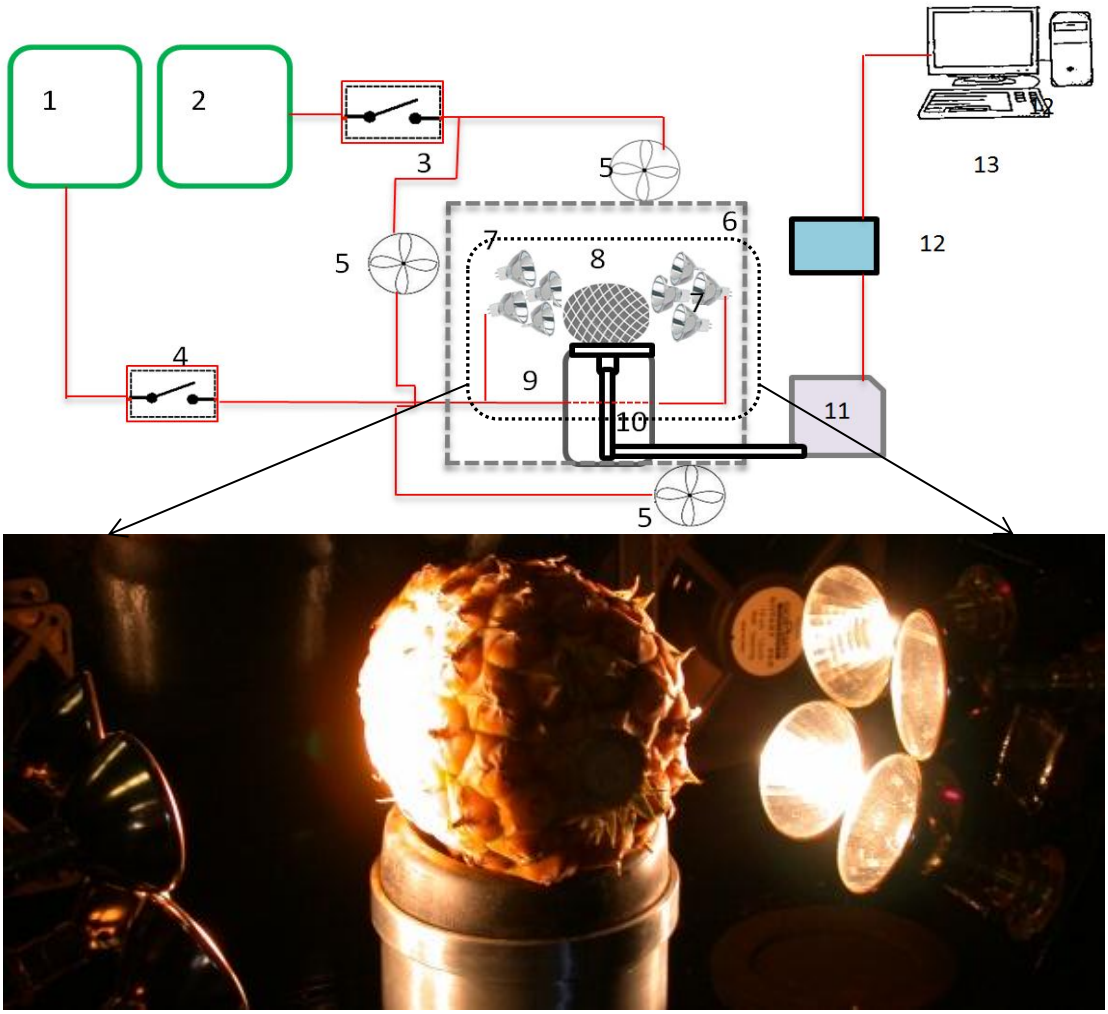



**Figure 4.2**. Schematic (top) and image (below) for an eight 60 W light source unit, designated IDD1, used in acquisition of spectra of pineapple fruit.

The IDD1 and IDD0 systems were used to acquire fruit spectra in season 1 and 2 respectively. These systems are comparable to that utilised by Teerachaichayut et al. (2007) and Sukwanit and Teerachaichayut (2013) for detection of translucency in mangosteen fruit and internal browning in pineapple fruit, respectively.

### 4.2.3 Reference assessment methods

Following spectral acquisition, each fruit was assessed for fruit density, TSS and extent of defect visible on the cut surface of a medial section of the fruit, as indexed by imaging and by visual assessment. Fruit density was measured by volume displacement method (Wanitchang et al., 2010) into 75% methanol (density = 0.792). Pineapple fruit can have a density lower than water but their density is always greater than methanol and thus fruit sink unaided.

Each fruit was then cut perpendicular to the axis of fruit at the middle of the fruit (i.e. through the marked areas) and the cut surface was imaged using a Canon DS 126211 camera with a 17-85 mm lens at a fixed focal distance. In preliminary work, fruit were repeatedly cut in 3 locations however scoring was uniform and the method was scaled back to a single cut assessment at middle of fruit for visual score assessment. Images were analysed for the percentage of area affected by translucency using ImageJ public domain software based on a user adjusted threshold on a grayscale value. The core of the fruit was masked and excluded from this analysis. Additionally, fruit were ranked based on visual assessment of images using a five point scale based on affected area and intensity of translucency (Fig. 4.3). TSS was measured of juice extracted from a 3 cm sided cube of flesh was taken from each marked spot at equator just below the skin, using a digital refractometer (RFM320, Bellingham and Stanley P/L).

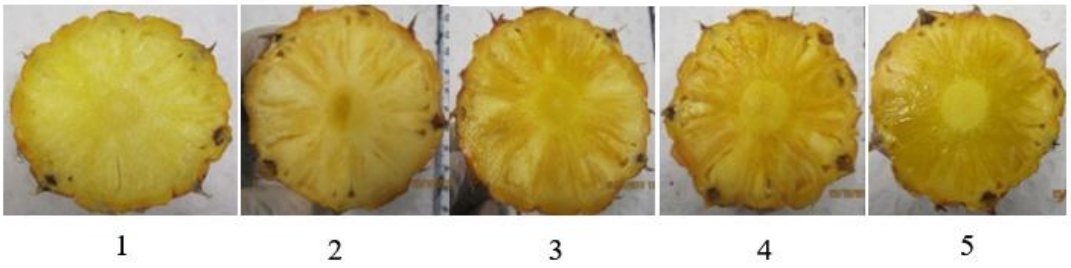



**Figure 4.3.** Cut medial surface of pineapple fruit ranked on a visual five point scale for extent of translucency. Fruit with score 1 and 2 are considered acceptable to consumers while score 3, 4 and 5 are considered as defect fruit.

### 4.2.4 Chemometrics

The multivariate data analysis software Unscrambler 13.2 (Camo, Oslo, Norway) and MS-Excel were used. Partial least square regression (PLSR) was performed using random cross validation with 10 samples per segment. The optimum number of principal components (factors) was selected based on the lowest root mean square of errors of cross validation (RMSECV). Model performance was described in terms of correlation coefficient of determination ($R^2_{cv}$) and RMSECV. Each model was validated using a prediction set and performance assessed in context of the correlation coefficient of determination in prediction ($R^2_p$) and the root mean square error of prediction (RMSEP). Attempts to reduce scattering effects and instrumental noise were made using various pre-processing treatments, including maximum

normalization, multiple scatter correction (MSC), standard normal variate (SNV), second derivatization using Savitzky Golay (9 points) or using gap segments with four samples per segment.

For classification of fruit into good and defect, discriminant analysis methods including linear discriminant analysis (LDA), partial least square discriminant analysis (PLS-DA), k nearest neighbourhood, support vector machine classification (SVM) were employed.

## 4.3 Results and Discussion

### 4.3.1 Population structure

Population mean and SD are presented for the calibration and prediction sets of the two populations in Table 4.1. Of the 84 fruit collected in season 2, the first 50 were used as a calibration set and the last 34 were used as validation set. Thus, second set had a notably higher level of defect (higher visual score and affected area), as these represents the fruit scanned as last two lots and were over ripen.

**Table 4.1.** Population statistics of different quality parameters of pineapple for calibration and validation sets.

| | Population 1 (n = 138) | | | | Population 2 (n = 84) | | | |
|---|---|---|---|---|---|---|---|---|
| | Calibration set (n = 90) | | Prediction set (n = 48) | | Calibration set (n = 50) | | Prediction set (n = 34) | |
| Parameters | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Translucency score | 1.75 | 0.87 | 1.92 | 1.07 | 1.72 | 1.56 | 2.67 | 2.06 |
| TSS | | | | | 12.25 | 1.86 | 12.87 | 1.27 |
| % area translucent | | | | | 23.0 | 20.1 | 32.1 | 23.5 |
| Fruit density | | | | | 0.94 | 0.01 | 0.94 | 0.02 |

The extent of the disorder within the fruit was quite variable, both radially and along the length of the fruit. Attempts were made to estimate an average value based on assessment of defect extent in transverse slices taken at multiple locations along the length of the fruit, but use of average values did not improve results (data not shown). An estimate of error on the visual estimation of defect level was attempted

by scoring the same set of images on two occasions. The RMSE of the repeated measurement of translucency score was 0.37 on a 5 point scale.

### 4.3.2 Overview of Vis- SWNIR spectra

Absorbance spectra of pineapple fruit were characterised by absorbance maxima at 430, 530 and 670 nm, presumably associated with carotenoids and chlorophyll (Fig 4.4 A). The chlorophyll associated peak was variable in intensity. Absorbance maxima were evident at 740, 840 and 960 nm as expected for water (O-H) features. SNV treatment decreased the spread of values somewhat and the second derivative treatment removed baseline differences, as expected (Fig 4.4, B and C).

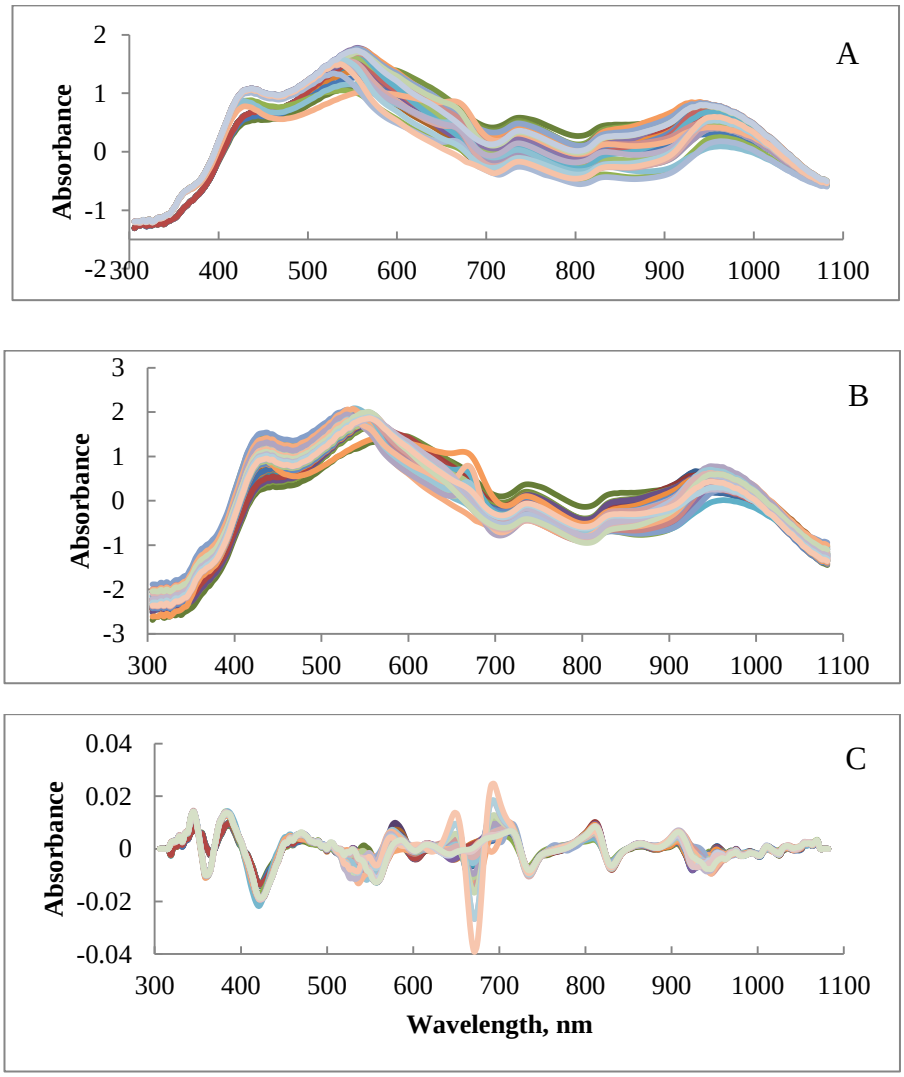



**Figure 4.4.** Vis-SWNIR spectra **A.** raw absorbance **B.** SNV pre-treated absorbance and **C.** second derivative absorbance.

### 4.3.3 Relation between reference parameters

The visual translucency score was relatively poorly indexed by the image analysis of % area affected ($R^2$ value of 0.65) (Table 4.2). The image analysis method depended on matching of an image gray scale level with areas of translucency density, a relationship that was apparently weak.

The visual translucency score was effectively not related to density (Table 4.2). The filling of apoplast spaces with solution in translucent fruit might be expected to increase fruit density, however, the pineapple fruit is complex, and contains locular air spaces associated with each fruitlet, and air may also be entrained under bracts.

The visual score was also not related to TSS (Table 4.2). The disorder is believed to be associated with the presence of sugars in the apoplast where sugars have largest share in TSS (Chen & Paull, 2000). Evidently, the mechanism of filling of the apoplast space with solution does not require a significantly higher overall level of sugar in the fruit.

**Table 4.2.** Correlation coefficient of determination ($R^2$) among the studied quality parameters from reference assessments from season 2 population (n = 84).

| | Score | % area translucent | Density | TSS |
|---|---|---|---|---|
| Score | 1 | | | |
| % area translucent | 0.65 | 1 | | |
| Density | 0.009 | 0.09 | 1 | |
| TSS | 0.049 | 0.056 | 0.09 | 1 |

Spectra of season 1 fruit demonstrated a stronger chlorophyll absorbance feature at 670 nm than fruit of season 2. The average absorbance of fruits with a visual score of 1 or 2 was higher than that of score 3 to 5 fruit between the wavelengths of 500 and 950 nm in Season 1 but not 2 (Fig. 4.5). Evidently the two populations are different in a number of ways despite both being harvested at commercial maturity. The season 2 fruit were more yellowish than season 1 on visual observation of skin color. The objective maturity assessment like assessment of soluble soilds content and acidity were not considered, as are not under our scope.

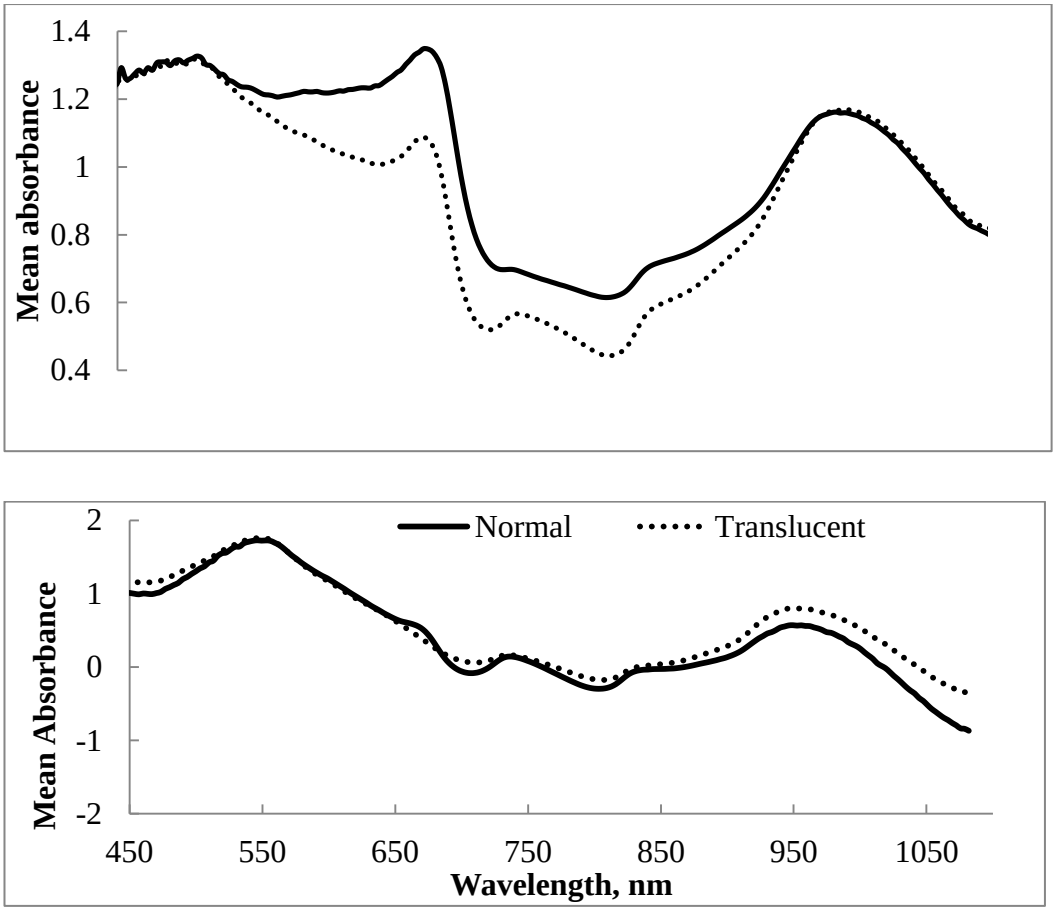



**Figure 4.5.** Mean spectra of acceptable and defect (translucent) fruit. Top panel: Season 1 fruit assessed with T1 unit; bottom panel: Season 2 fruit assessed with IDD0 unit.

### 4.3.4 Univariate linear regression and multiple linear regression

The correlation between absorbance at a single wavelength and visual score was strongest between 650 and 850, peaking at 710 nm for Season 1, but the weightings were quite different in Season 2 (Fig. 4.6). The 710 nm weighting is consistent with a difference in chlorophyll level in defect fruit in Season 1. The shape of the wavelength weights for R were similar for the attributes of visual score and % area affected, although the absolute levels of R were higher for the % area affected attribute (season 2 data, Fig. 4.6).

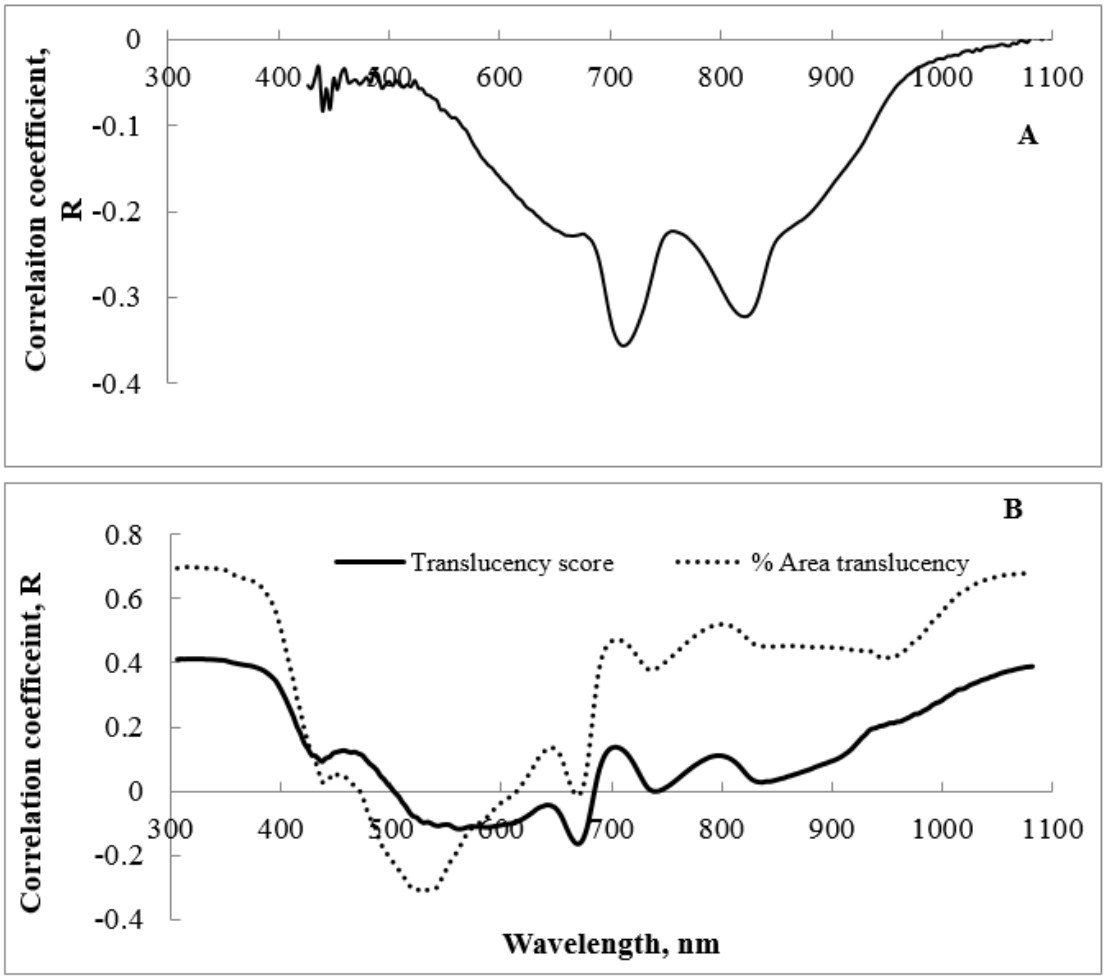



**Figure 4.6** Correlation coefficient (R) for visual score and % area affected by translucency for absorbance at each pixel over the range 300-1100 nm for Season 1 fruit assessed with IDD0 (A) and Season 2 fruit assessed with IDD 1 (B).

Multiple linear regression on visual score involving wavelengths 582, 644, 670 and 700 nm yielded a fairly poor statistics ($R^2_C$ = 0.35, RMSEC = 1.16).

### 4.3.5 Partial least square regression (PLSR)

Of the pre-treatment options trialled, similar results were obtained for raw absorbance and for SNV and second derivative data (Table 4.3). However, in prediction the lowest bias was achieved with models that involved a SNV pre-treatment of spectra, while the best overall result in prediction was achieved using absorbance data. This result is consistent with the expectation that raw absorbance spectra would support a better result than derivatised data based on the presence of both scatter and absorbance information in the raw (apparent) absorbance data.

**Table 4.3.** PLSR model calibration and prediction statistics for models for translucency score developed using a number of pre-processing treatments. Spectra from IDD1 and population 1 (n = 138; 700-1000 nm) and spectra from IDD0 and population 2 (n = 84; 490-980 nm).

| Spectra pre-treatments | Calibration statistics (n = 90) | | | Prediction statistics (n = 48) | | |
|---|---|---|---|---|---|---|
| Population 1 | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias |
| Abs | 0.58 | 0.55 | 10 | 0.41 | 0.93 | -0.36 |
| Abs MSC | 0.55 | 0.55 | 7 | 0.34 | 0.86 | 0.10 |
| Abs SNV | 0.56 | 0.55 | 7 | 0.39 | 0.83 | 0.009 |
| Abs d2A | 0.52 | 0.6 | 8 | 0.29 | 1.23 | 0.54 |
| Abs SNV d2A | 0.58 | 0.56 | 6 | 0.32 | 0.87 | -0.05 |
| Population 2 | Calibration statistics (n = 50) | | | Prediction statistics (n = 34) | | |
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias |
| Abs | 0.50 | 1.01 | 10 | 0.40 | 1.46 | 0.77 |
| Abs MSC | 0.42 | 1.09 | 7 | 0.29 | 1.34 | 0.30 |
| Abs SNV | 0.42 | 1.09 | 7 | 0.29 | 1.35 | 0.33 |
| Abs d2A | 0.47 | 1.05 | 5 | 0.28 | 1.55 | 0.81 |
| Abs SNV d2A | 0.42 | 1.1 | 6 | 0.33 | 1.31 | 0.35 |

(Note: Abs = Raw absorbance, SNV = standard normal variate, MSC = multiple scatter correction, SG = Savitzky Golay, GS = Gap segment, d2A = Second derivative of absorbance)

The weighting of the regression coefficients for a PLSR model developed using the range 490-980 nm was similar to that obtained for a model developed using the range 305-1080 nm (Table 4.4). This similarity and the relatively ‘smooth’ nature of the coefficient plot augers well for the stability of the model (Fig 4.7).

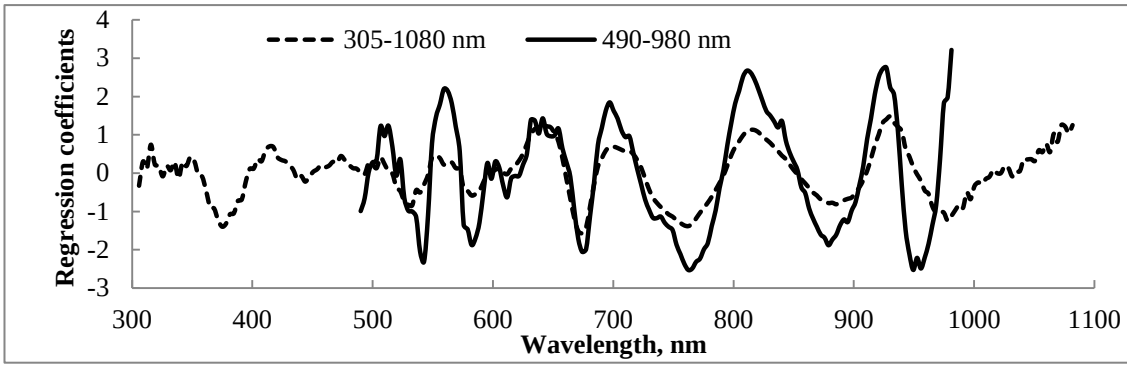



**Figure 4.7.** PLS regression coefficients for the parameter of translucency score for models developed using different wavelength regions (population 2).

**Table 4.4.** PLSR model statistics for population 1 (n = 138) and population 2 (n = 84). Wavelength regions were selected based on lack of noise in regression coefficients.

| Wavelength range (nm) | Calibration statistics (n = 90) | | | Prediction statistics (n = 48) | | |
|---|---|---|---|---|---|---|
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias |
| 426-1150 | 0.29 | 0.75 | 7 | 0.05 | 2.89 | 1.54 |
| 700-1000 | 0.58 | 0.56 | 6 | NA | 0.32 | 0.87 |

| Wavelength range (nm) | Calibration statistics (n = 50) | | | Prediction statistics (n = 34) | | |
|---|---|---|---|---|---|---|
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias |
| 305-1080 | 0.50 | 1.00 | 8 | 0.45 | 1.44 | 0.90 |
| 490-980 | 0.50 | 1.01 | 10 | 0.40 | 1.46 | 0.77 |

The PLSR model for % area affected by translucency was superior to that for visual score (Table 4.5). The character of % area affected may be a better reference attribute than visual score.

**Table 4.5.** PLS regression statistics based on absorbance data (population 2, n = 84) for the wavelength range 490-980 nm for the parameters of density and percentage area affected by translucency.

| | Calibration statistics (n = 50) | | | Prediction statistics (n = 34) | | |
|---|---|---|---|---|---|---|
| | $R^2_{cv}$ | RMSECV | PCs | $R^2_p$ | RMSEP | Bias |
| % area translucent | 0.57 | 13.21 | 8 | 0.26 | 20.46 | -0.76 |

A comparable study was undertaken by Terdwongworakul et al. (2012) for detection of mangosteen translucency using NIR transmittance measurements (n = 135). NIR transmission spectra (665–955 nm) were acquired with instrumentation based on a single 100 W tungsten halogen lamp, averaging four spectra for each fruit. Using PLS regression with pre-treated (MSC, SNV and derivatization), the highest correlation coefficient of determination ($R^2$) achieved was 0.74. The short integration time and the higher correlation relative to that of the current study is ascribed to fruit structure that is more suitable to spectroscopy. Mangosteen fruit have a smooth skin, a rind to 10 mm and a diameter of 50-80 mm.

### 4.3.6 Classification

Of the discriminant methods trialled, linear discriminant analysis based on Mahalanobis distance gave the best results, with attention to prediction true negative rate (TNR). Support vector machine classification (SVM-DA) failed in this context (TNR = 0), although a 100% true positive rate was achieved. Better results were achieved with the prediction set of population 1 (TNR = 69%) than 2 (TNR = 31%).

**Table 4.6.** Classification of intact pineapple based on the raw absorbance spectra (700-1000 nm for population 1 and 490-980 nm for population 2) based on visual translucency score (score 1 and 2 as normal fruit, score 3-5 as translucent defect fruit).

| | Population 1 (n = 138) | | | | | | Population 2 (n = 84) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Classification methods | Calibration set (n = 90) | | | Prediction set (n = 48) | | | Calibration set (n = 50) | | | Prediction set (n = 34) | | |
| | TPR | TNR | Accuracy | TPR | TNR | Accuracy | TPR | TNR | Accuracy | TPR | TNR | Accuracy |
| PLS-DA | 91.5 | 41.2 | 66.3 | 60.0 | 23.0 | 41.5 | 98.7 | 25.0 | 61.8 | 98.7 | 33.3 | 66.0 |
| LDA-MD | 80.2 | 68.4 | 74.3 | 74.3 | 69.2 | 71.7 | 72.5 | 52.5 | 62.5 | 34.2 | 31.7 | 33.0 |
| LDA-Linear | 74.6 | 73.7 | 74.1 | 54.2 | 69.2 | 61.7 | 57.5 | 62.5 | 60.0 | 42.1 | 20.0 | 31.0 |
| SVM | 94.7 | 64.3 | 79.5 | 91.9 | 27.3 | 59.6 | 83.0 | 79.0 | 81.2 | 38.9 | 59.4 | 49.2 |
| *k*-NN | 89.5 | 0 | 44.7 | 100.0 | 9.0 | 54.5 | 76.0 | 77.0 | 76.5 | 70.8 | 59.4 | 65.1 |

Classification was undertaken using partial least square discriminant analysis, with accuracy (true positive+true negative/total fruit) of 92% with informative wavelength of 673, 692, 723, 727, 743 and 762 nm. A similar result was reported by Benjakul et al. (2013) who employed a logistic regression method for discrimination of mangosteen translucent fruit based on electrical impedance measurements. These authors report an accuracy of discriminating the good and defect fruit (true positive + true negative/total fruit) as 87.3% while setting cut off point at 0.5. Again, the better results achieved with mangosteen than the pineapple work reported here is ascribed to a fruit structure that is more suitable to light transmission. Sukwanit and Teerachaichayut (2013) reported a true positive rate of 89% and false positive rate of 93% with use of SWNIR spectra for sorting of pineapple for internal browning. This result was based on very clear difference in spectra between good and defect fruit, which were not present for the translucent defect of the current study.

## 4.4 Conclusion

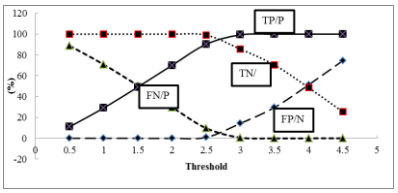
This work is the first reported for use of NIR spectroscopy for the detection of translucency in pineapple. The results achieved for non-invasive detection of translucency in pineapple were encouraging, but not at an accuracy suitable for use by the industry. This result, in comparison to literature reports of detection of translucency in mangosteen, is ascribed to the more complex anatomy of the pineapple fruit. Variation in bract tissues, internal locular spaces and rachis vasculature (i.e. fruit core) will all impact light transmission, 'diluting' the impact of the translucency disorder on light transmission. Also, the extent of the disorder within the fruit was quite variable, both radially and along the length of the fruit.

Future research could include refinement of a reference method, inclusion of more populations in an attempt to develop a more robust data set, use of more sensitive detectors to enable a shorter integration time, and use of a rotating sample stage to 'sample' more of the fruit (with a series of spectra taken as the fruit rotates and detection of light from various points along the length of fruit).

Alternatively, other detection technologies, such as electrical impedance or X-ray imaging, could be considered. Translucency involves extra water content in intercellular spaces which presumably affects electrical conductivity and certainly

should impact density and thus X-ray imaging. Likely problems for the former method include variation in electrical conductivity for reasons other than water content, e.g. salt content. A likely problem for the latter method is that overall fruit density is not related to defect level, due to variation in other air spaces within the fruit (e.g. locule and bract space). A micro-focus method would be required to assess tissue only.

# Chapter 5.
# Sorting optimisation for internal defects in fruit based on NIR spectroscopic measurements[1]




## Abstract

The sorting of defect from good fruit represents a binary classification, in which measurement error will cause both Type I and II errors. Typically the operator of a sorting operation has control of a threshold value on the classifier, compromising between yield (true positive rate) and adherence to a specification on proportion of defect fruit (false discovery rate). Adjustment of this threshold value is required in response to changes in measurement error, the severity of incidence of defect fruit in a given population (e.g. as proportion increases, more type II errors occur) and in terms of the pricing achieved for populations of varying defect incidence. A 'sorting optimisation' decision support guide was developed for use of a defect sorter based on a model of the sorting function, knowledge of the defect distribution in a population and pricing.

**Keywords:** binary classification, ROC, SOC, optimisation, threshold

[1] I acknowledge Mark Loeffen of Delytics P/L (Hamilton, New Zealand) for assistance with this Chapter and Mrs. Rosie Savio of Savio farm, Stanthorpe, Queensland, for advice on marketing of apple.

## 5.1 Introduction

Sorting of agricultural products is undertaken to maintain quality and to bring uniformity to the sorted population. Separation of a consignment to two or more quality classes is generally done to specifications agreed mutually between parties (e.g. packer and wholesaler) (Bollen & Prussia, 2009, 2014; Laofor & Peansupap, 2012). However, whether the sorting/grading operation is a manual or a machine task, class assignment errors are anticipated (Ladaniya, 2008a).

To achieve a binary classification for a measurement system producing continuous values, a threshold or cut off value (t) can be used, e.g. with the output values (v) higher than the threshold resulting in rejection of the product (Ooms et al., 2010).

v<t: Accept; v>t: Reject

There are four possible classes in such a classification: true positive, TP, true negative, TN, false positive FP and false negative, FN. The performance of a binary classifier is generally assessed in terms of accuracy, classification error for positive and negative classes, as defined below:

Accuracy = (TP+TN)/(P+N)

Classification error for positive item (good fruit category) FPR= FP/N

Classification error for negative item (defect fruit category), FNR= FN/P

where P and N denote number of actual positive and actual negatives.

The operator in control of a sorting operation normally has control of the threshold value, balancing a trade-off between the false negative rate (good fruit predicted as defect, FNR) and the false positive rate (defect fruit predicted as good, FPR). However, consignments may be rejected if a specification on the false discovery rate (defect fruit in accepted group; FP/(TP+FP)) is exceeded. Further, the economic outcome will also depend on the true positive rate (TPR, TP/P or yield) and return (price) achieved for batches with different levels of defect. The operators task in optimisation of the setting of this threshold value could be assisted by provision of tools that demonstrate these trade-offs.

A receiver operating characteristic (ROC) curve, the plot of true positive (TP/P, or Sensitivity) against false positive (FP/N, or 1 - Specificity) rates of classification (Gómez Sanchis et al., 2013; Luo et al., 2012; Ooms et al., 2010), is one such tool. This plot illustrates the trade-off between the yield of the sorting operation (TP/P) and the compliance to a criterion of relevance to the market (FP/N). However, the area of most interest in a ROC plot is the region of minimal FPR and maximal TPR, displayed in the top left region of the plot. As an alternative, the detection error trade off (DET) graph gives more attention to the area of interest of the ROC curve. The DET plots missed detections (false negative rate, FNR) against false alarms (false positive rate, FPR) using x and y scales transformed by the quantile function of the normal distribution (the inverse of the cumulative normal distribution).

A sorting optimization curve (SOC) utilises additional information to optimise the selection of the threshold value (Ooms et al., 2010). For example, the marketplace criterion on diffuse browning in apple is not FP/N per se, but rather FP/(TP+FP) (false discovery rate, or defect fruit in the accepted class). There may also be latitude given in this criterion (e.g. the market cannot accurately measure and so does not strictly enforce a 2% false discovery rate on apple browning), and ideally this latitude should be included in the SOC. The price differential on product with different levels of defect is another driver of operator behaviour that should be included in an effective SOC.

Given knowledge of measurement error (e.g. RMSEP for a SWNIRS based prediction of an internal attribute) and the distribute of attribute level in a population it should be possible to model the proportion of Type I and II errors that will be made in a sorting operation, and thus provide information to support a SOC that incorporates pricing (ie. a SOC based on value of sorting outcomes). In this section, the value of these various aids in guiding the operation of a sorting operation is discussed. Data of sorting of fruit for internal browning is used.

## 5.2 Materials and Methods

### 5.2.1 Data

Data of 227 apple fruit from six populations sorted using the IDD0 unit were considered for this exercise. The overall diameter at equator and weight of the fruit was 72.8 ± 4.1 mm and 154.3 ± 22.4 g (mean ± SD). An additional data set of 7 populations of fruit (each of approximately 200 fruit) from a pilot in-packhouse sorting operation (near Stanthorpe, Queensland) were also accessed. A visual browning score of cut fruit was used as the reference measurement. On this scale, scores 1 and 2 are associated with consumer acceptance while scores 3 – 5 represent unacceptable fruit. A partial least square regression model based on IDD0 absorbance data over the wavelength range 500-975 nm was used in prediction using The Unscrambler 13.1 (CAMO, Oslo, Norway).

Details of the fruit, instrumentation and spectra acquisition and manipulation are presented in Chapter 2.

### 5.2.2 Data analysis

Microsoft Excel 2010 was used, utilizing features of the Data Analysis ToolPak. The Detector Error Trade-off probit function plot was undertaken using the Excel +NORMSINV function, which returns the inverse of the standard normal cumulative distribution (i.e. mean of zero and standard deviation of 1).

## 5.3 Results and Discussion

### 5.3.1 Sampling statistics

The number of samples required to estimate the mean value of an attribute with a continuous value (such as severity of internal browning of apple) can be estimated from the relationship between sample uncertainty, probability and standard deviation (Fornasini, 2008).

$$n = \left(\frac{t\,SD}{e}\right)^2$$

where n is the number of samples required, SD is the standard deviation of the population, t is the t statistic (1.96 for n>30 and a 95% confidence interval) and e is

the measurement error. A preliminary sampling is required to estimate population SD.

Thus to estimate mean level of internal browning in a population of fruit with a 95% confidence interval (CI), a SD of 1.5 and an error of 0.24 (RMSED of operator repeatability on a 5 point visual score scale), the number of samples required is:

$$n = \left(\frac{1.96\ X\ 1.5}{0.24}\right)^2 = 150$$

However, the fresh fruit value chain specification on this defect is one of discrimination (accept/reject), not estimation of level of severity of defect. Diffuse browning in apple is described as a major defect by major retailers, with a specification of no more than 2% of fruit in a consignment to be affected (Woolworths,2015). The typical test imposed by retailers is the cutting and assessment of 30 fruit on receival of a consignment of fruit at the distribution centre. The probability of encountering a defect fruit if the defect is present in 2% of the population when n fruit are sampled is:

Let P = probability of finding one or more defect fruit when sampling n fruit from a population with p incidence of defect.

$$P = 1 - (1 - p)^n$$

where p is the actual proportion of defect fruit in the population and n is the number of fruit sampled

For x ≥1, p = 0.01 and n = 30

$$P = 1 - \ (1 - 0.01)^{30} = 0.2603$$

The same result can be estimated using the Excel binomial distribution function, as :

$$P = 1 - BINOMDIST(x - 1, n, p, TRUE)$$

Examples are provided in Figure 5.1 for the probability (P) of selecting at least one defect fruit when sampling 10, 30 and 300 fruit from a consignment, for various levels of incidence of defect in the consignment. For a consignment with a defect incidence of 2%, there is a 17.5, 45.5 and 99.98% chance of choosing a defect fruit when sampling 10, 30 and 300 fruit respectively (Fig. 5.1).

Obviously a sampling strategy using only 10 or 30 fruit would often be in error, failing to detect presence of defects, for low incidence rates. Thus a supplier with perfect knowledge of product defect levels might choose to provide the retailer with fruit of a slightly higher level of defect incidence than the specification level, given the probability of detection. In practice, the major retailers must recognise this sampling issue, as although the official specification is no more than 2% incidence, suppliers report retailer tolerance of 5% incidence of defect fruit (personal communication, Rosie Savio, Savio packing house, Stanthorpe, Queensland, Australia). At a 5% incidence level, a 30 fruit sample would include select a defect fruit in 77% of sampling events.

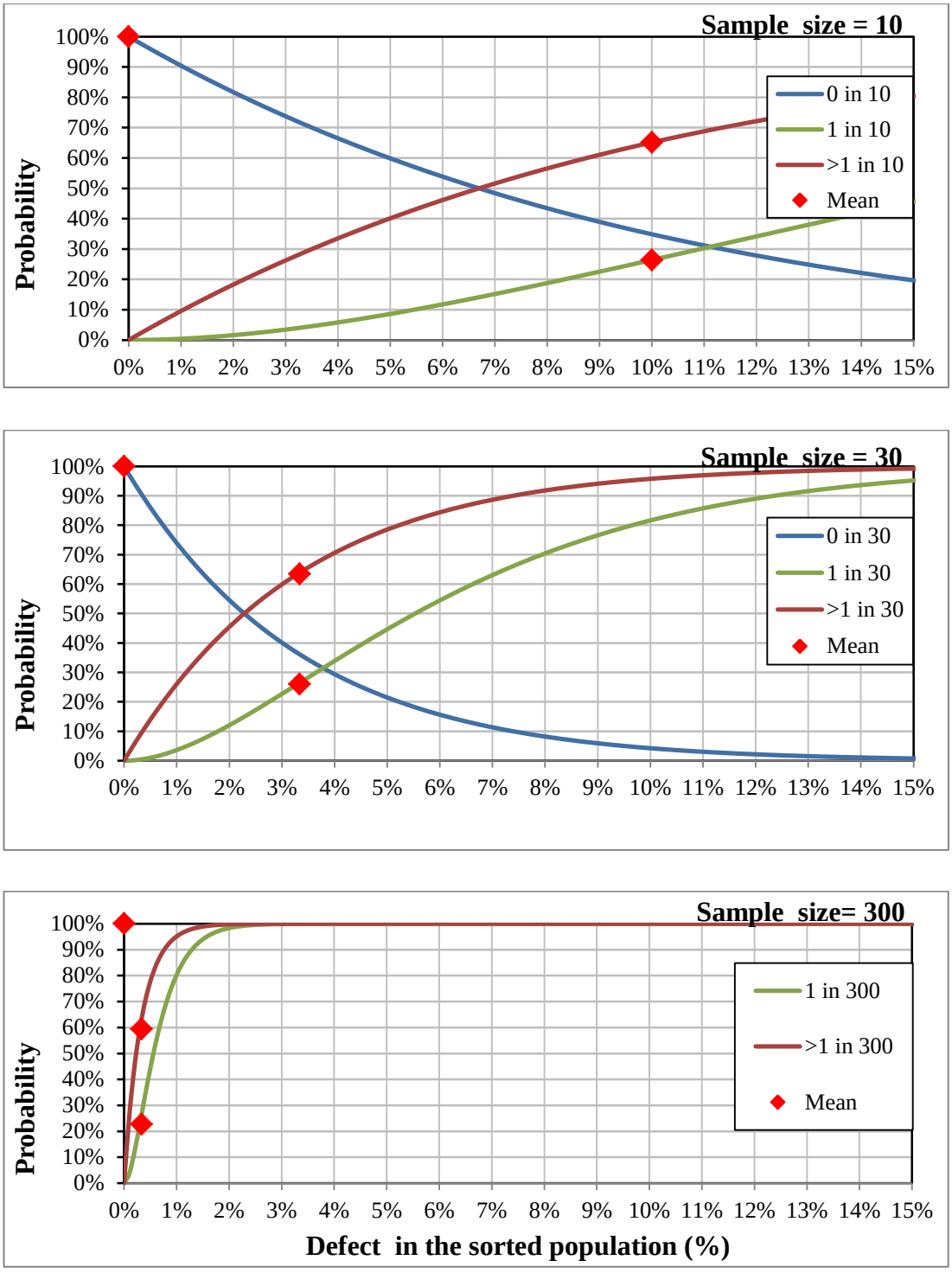



**Figure 5.1.** Probability of finding 0,1 or >1 defect fruit in a sample of (A) 10, (B) 30 and (C) 300 fruit from a consignment containing between 0 and 15% defect fruit.

### 5.3.2 Population Description

In total, 227 fruit (71 good and 156 defect) were considered. This exercise involved five populations each containing 30 fruit and one population with 227 fruit which is

77 fruit from a population in addition to fruit in all 1-5 pop combined. Population statistics are presented in Table 5.1.

**Table 5.1.** Average and variation of visual browning score (5 point scale) in each population.

| **Population** | **Pop 1** | **Pop 2** | **Pop 3** | **Pop 4** | **Pop 5** | **Pop 6** |
|---|---|---|---|---|---|---|
| # fruit | 30 | 30 | 30 | 30 | 30 | 77 |
| Mean | 2.96 | 3.28 | 3.14 | 3.36 | 3.85 | 3.32 |
| SD | 1.49 | 1.57 | 1.48 | 1.56 | 1.38 | 1.42 |

### 5.3.3 A sorting operation based on PLSR model output

Let us consider the sorting of an example batch of fruit (population 6) using SWNIRS for the level of diffuse browning on a five point scale, given variation of the threshold value on sorting (Fig. 5.2). For example, at a threshold of 1, the FPR (FP/N) is 0% but there was 70.8% rejection of actual good fruit (FNR, FN/P), with consequent economic loss. At a threshold value of 2.5, the % of defect fruit predicted as good (FP/N) was increased to 0.93% while only 9.7% of good fruit were falsely considered to be defect fruit (FN/P).

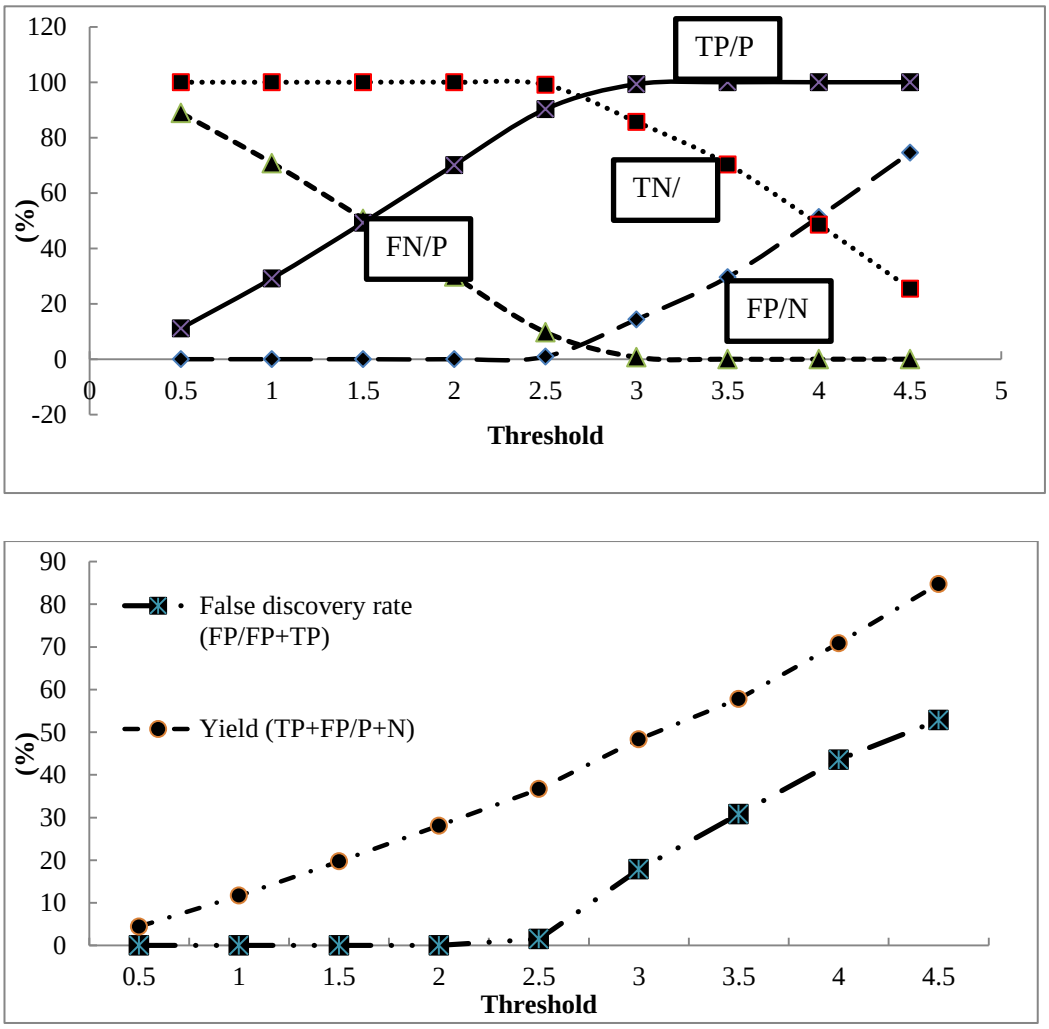



**Figure 5.2.** Classification of an apple fruit population as the threshold value for visual score is varied, in terms of % of good and defect fruit accepted and rejected respectively. % on y axis relates to descriptors associated with each line. Fruit with values < threshold are graded as acceptable fruit, while those > threshold value are rejected as defect. Prediction of defect made using absorbance data over the range 500 – 975 nm using IDD0 instrumentation for population 1.

The effect of change in threshold value on classification results for five populations of apple fruit is shown in Figure 5.3. For example, for population 1, a threshold value of 2.5 would achieve a FP/N rate of 0.93% and a FP/(TP+FN), or false discovery rate, of 1.5% at the expense of a FN/P rate (% of good fruit rejected as defect, with economic loss) of 9.7%.

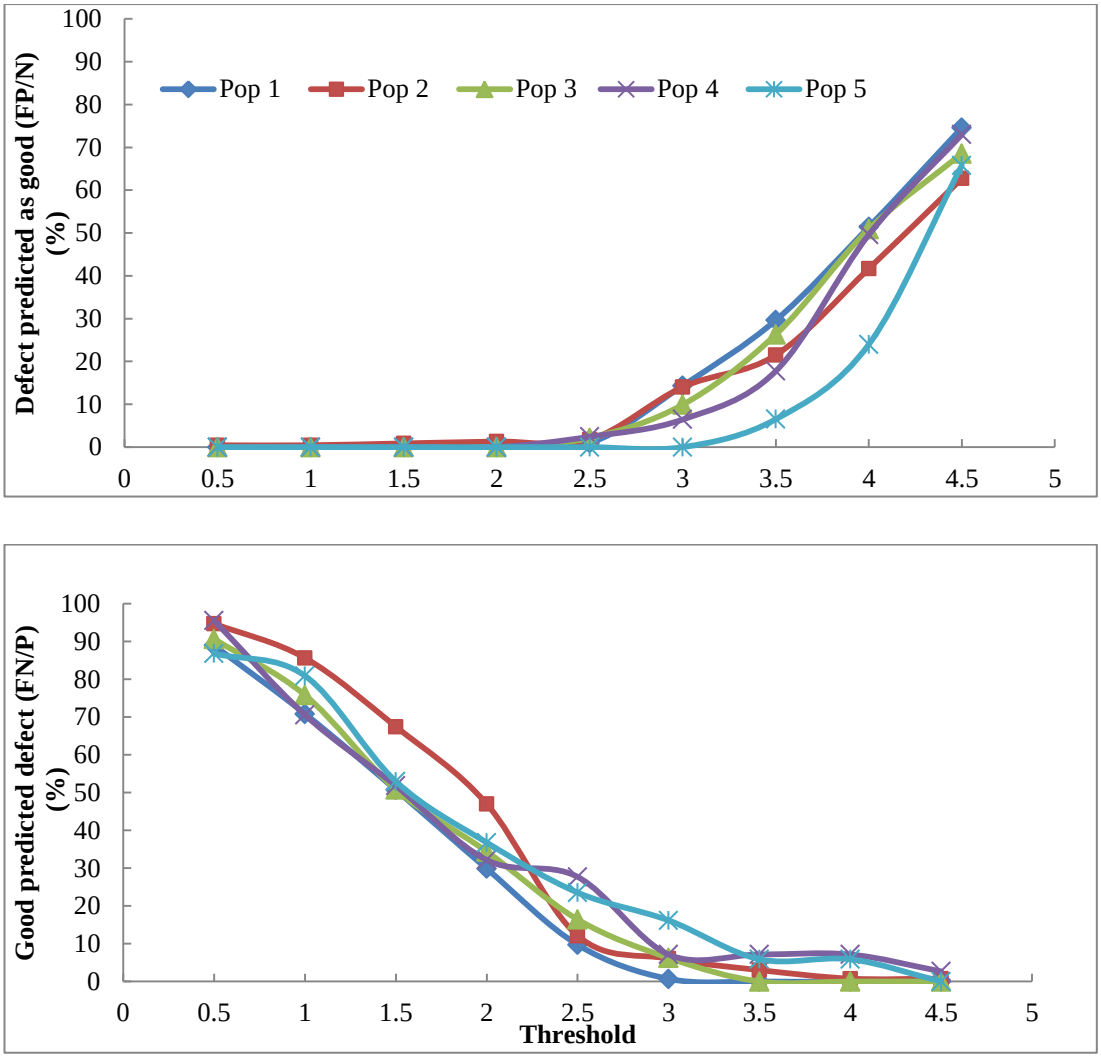



**Figure 5.3.** Classification error for apples with internal browning based on assessment using IDD0 instrumentation for populations 1-5. Top panel depicts the rate of defect fruit predicted as good fruit (FP/N) and bottom panel depicts the rate of good fruit predicted as defect (FN/P), as influenced by the threshold value used in sorting.

### 5.3.4 Describing the sorting operation with ROC and DET

Receiver Operating Characteristic (ROC) and Detection Error Trade-off (DET) curves are commonly used in describing sorting operations. The Receiver Operating Characteristic curve depicts the true positive and false positive rate. Such a curve can be generated for the apple internal browning example using sorting outcomes resulting from the use of different threshold values (e.g. for five populations, Fig. 5.4). Note that the area under the RUC curve is a measure of sorting effectiveness, suggesting sorting of population 1 was achieved more effectively than other populations (Fawcett, 2006). The ideal operating point would be (0,1), representing perfect operation, achieving 100% accuracy in sorting both the good and defect fruit.

In practice, the optimum operating point (and associated threshold value) is associated with the shoulder of the curve in the top left of the graph.

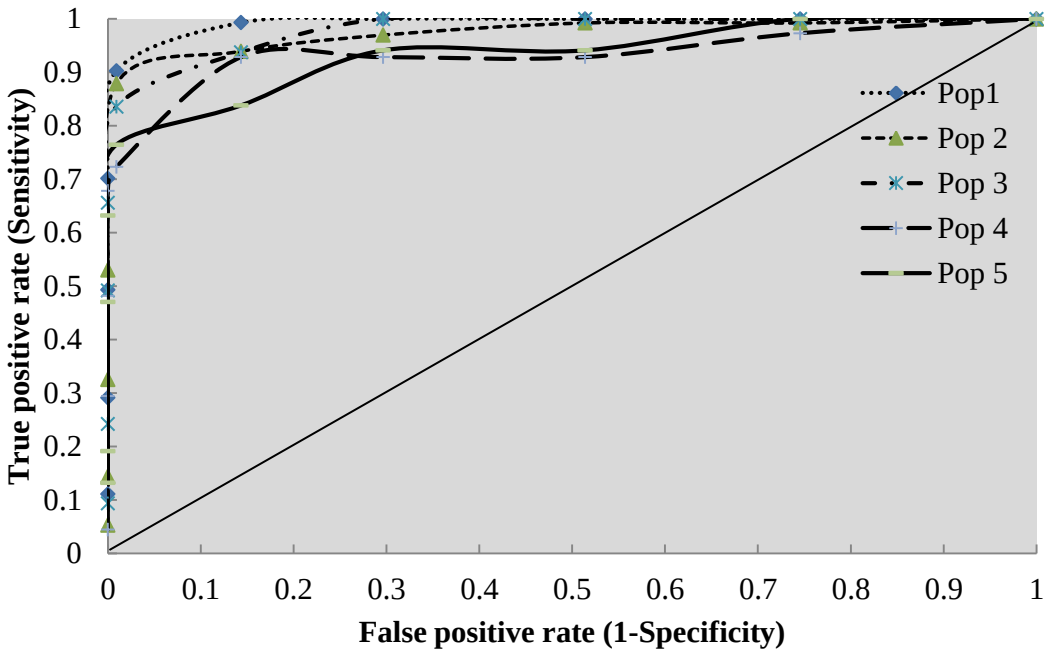



**Figure 5.4.** Receiver operating characteristics (ROC) curve for five population of apple showing true positive rate (TP/P or Sensitivity) and false positive rate (FP/N or 1 - Specificity) for data from IDD0 instrumentation.

The Detection Error Trade-off (DET) curve involves a plot of the missed detection rate (false reject rate, FN/(P+N)) against the false alarm (false accept rate, FP/(P+N)) using x and y scales transformed by the quantile function of the normal distribution (the inverse of the cumulative normal distribution) or by logarithm base 10. This presentation expands the false reject scale in the area of interest, in this case around a value of 2%.

However, while such presentations require development for each population assessed. Developments of some guidelines or 'rules of thumb'on how to chose threshold values therefore has merit. Further, attention to the marketplace criterion of false discovery rate (FP/(FP+TP)) is required.

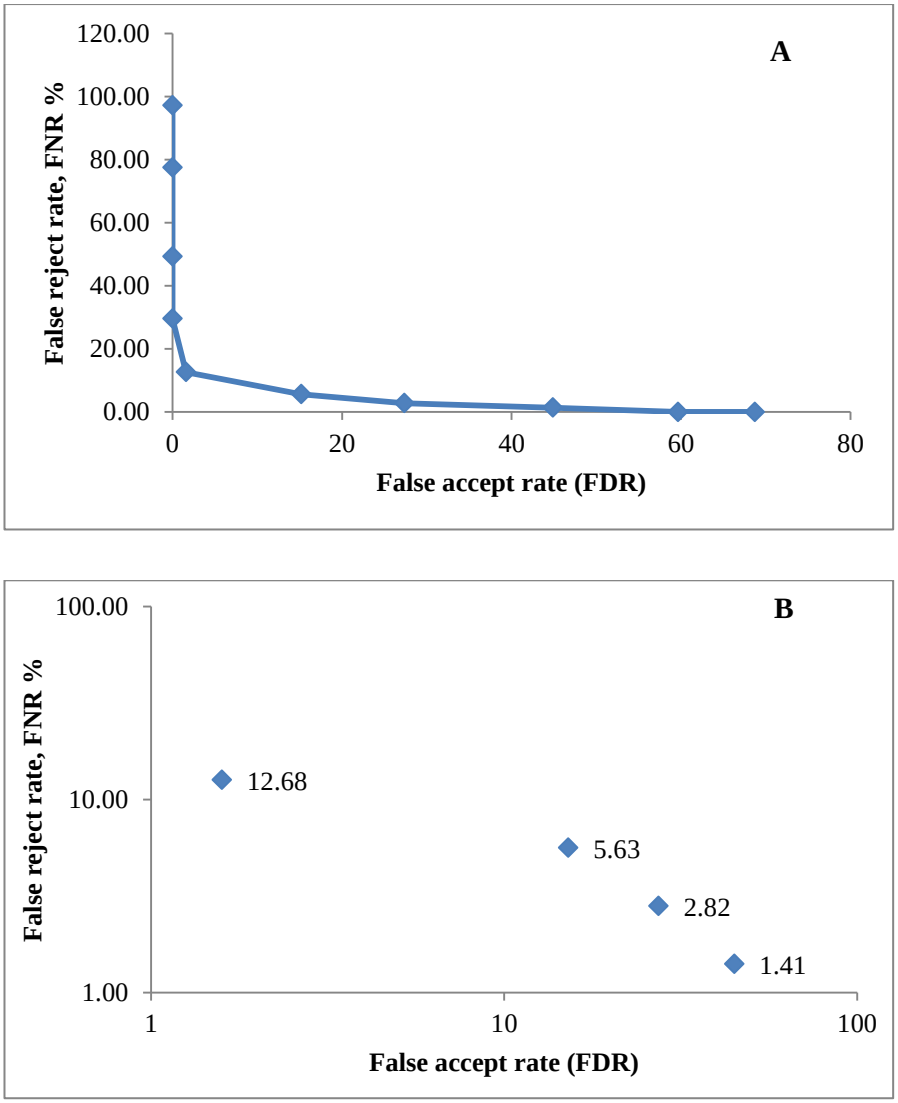



**Figure 5.5.** Detection error trade off (DET) curve – the plot of false negative against false positive rate for a range of sorting threshold values. A. Raw data. B. log-log plot plot.

### 5.3.5 Choosing a threshold value

For internal browning of apple, the retail specification is the percentage incidence of defect in the consignment (false discovery rate or total defect in accepted category; FP/(FP+TP)). If a set sorting threshold value is used across all populations, the false discovery rate will vary (e.g. for a threshold setting of 2.5, Fig, 5.5). Thus the threshold value required to achieve a False Discovery Rate of 2% will vary with population, depending on defect mean and range.

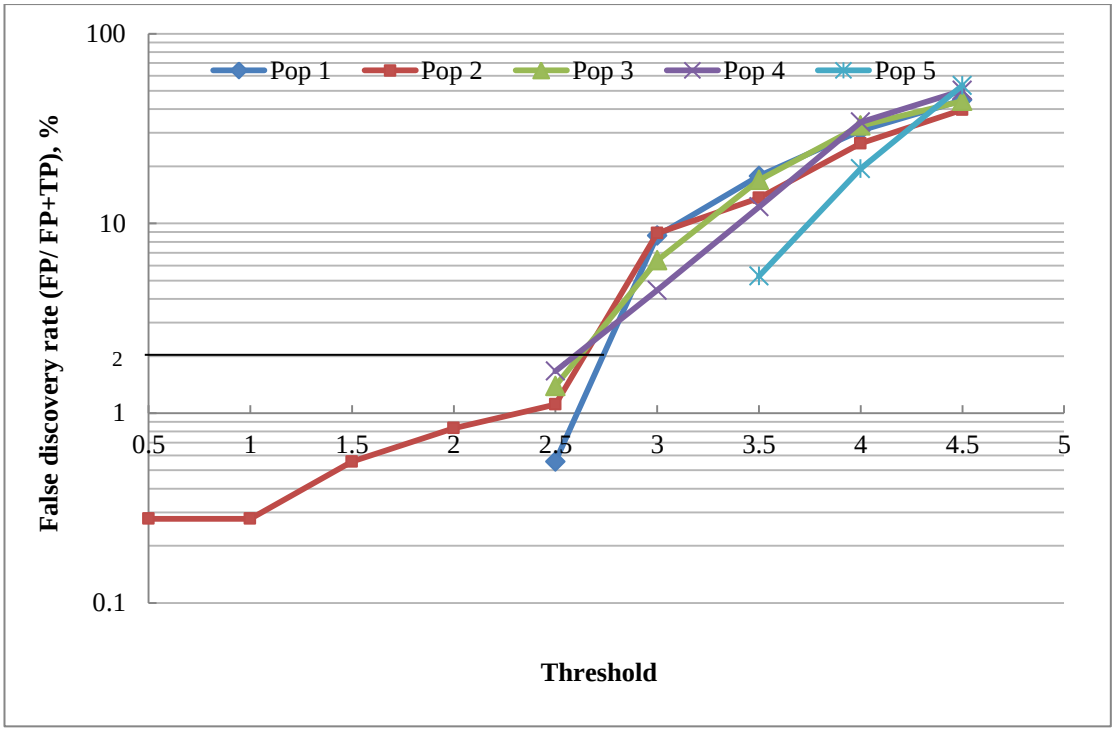



**Figure 5.6.** False discovery rate (FP/(FP+TP)) for five populations sorted using a range of threshold value. FDR displayed on a log scale.

The mean and range of defect level in the populations displayed in Fig. 5.4 did not differ greatly (Table 5.1), and the threshold value required to achieve a 2% FDR was similar for these populations. For another set of data involving 7 populations of fruit (each approximately n=200), the threshold value required to achieve a 2% FDR was noted to alter with change in population defect level distribution (mean score or % of unacceptable fruit in population). With addition of data from further populations an operating rule might be developed for setting of a threshold value given knowledge of defect incidence.

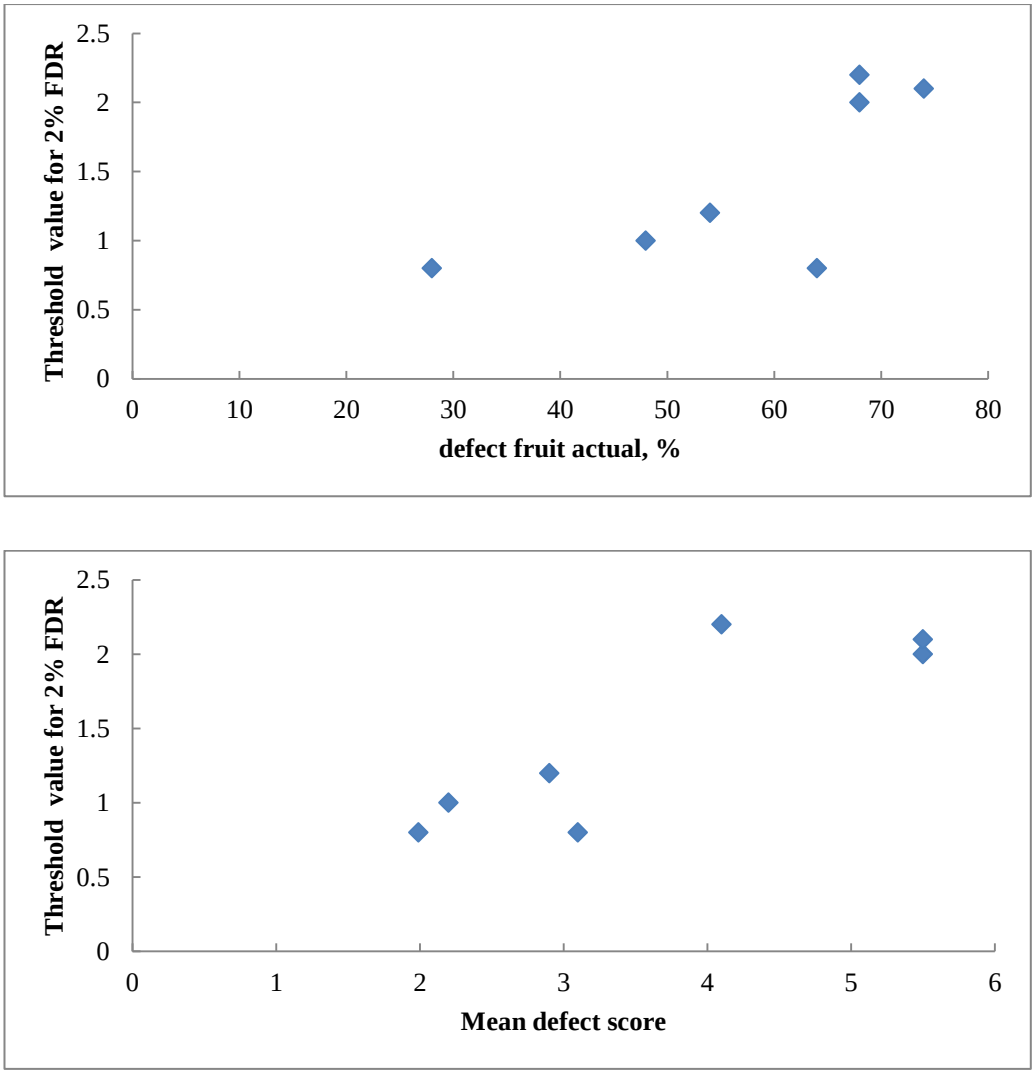



**Figure 5.7.** The threshold value required to achieve a 2% false discovery rate (FP/(FP+TP)) for seven populations, as a function of the % of defect fruit in the population (top panel) or mean defect score (bottom panel) for each population.

### 5.3.6 Estimation of defect distribution in the sorted population using a transfer function

Given knowledge of the accuracy of the sorting operation (i.e. RMSEP), the effect of error in the sorting operation on the distribution of defect levels in the end population should be able to be estimated using a 'transfer function'. This transfer function is proposed to have the shape of a normal distribution with a SD equal to the RMSEP of the sorting operation, with its mean set on the desired threshold level. The product of the transfer function and the defect distribution in the unsorted population will yield an estimate of the distribution of defect level in the sorted population.

For example, consider a population with some (normal) distribution of defect levels around a mean value x and with a standard deviation SD. If we set the sorting threshold at the value of x-SD, a perfect sorter would yield two populations, one with values <(x-SD), one with values >(x+SD), ie. composed of 16 and 84% of the original population respectively. However, the sorting accuracy is not perfect. Score 5 fruit are unlikely to be graded as score 1, but a score 3 fruit might be graded as score 2. In the example, let us assume a sorting error (RMSEP) of 0.5*SD. There will be the possibility of incorrectly grading good fruit with scores of up to x-SD-2*RMSEP as defect fruit, with a decreasing probability for more extreme values. This probability is estimated using a normal distribution with a standard deviation termed 'band width'.

An Excel based model was developed using the 'NORMDIST' function as the base of the transfer function, using the desired threshold value as the mean for this function, and setting a 'band width' as the standard deviation. A routine was added to compare the difference between the predicted defect levels in accepted and rejected classes with the actual levels in populations sorted using different threshold values, to allow for optimisation of the bandwidth setting at a given threshold setting. Data of population 6 (227 fruit) was used to illustrate model performance.

As the bandwidth of the transfer function is decreased, sorting accuracy increases. For example, a 'perfect' sorting result (i.e. no errors) is achieved using a bandwidth of 0.1 (Fig. 5.8), while a bandwidth of 1.39 results in considerable errors (Fig. 5.9). For this data set, model error was minimised at a bandwidth setting of 0.6 (Fig. 5.10).

The fit of the model was surprisingly good, indicating that the normal distribution used in the transfer function was a reasonable choice. Interestingly, the band width required to minimise the model output error to actual sorting results was less than the actual RMSEP of the sorting operation. In this example, the RMSEP was 1.39, the value used for the bandwidth in Figure 5.9. Model performance error was minimised using a bandwidth of 0.6 at a threshold of 2.5, but bandwidth optimum varied with threshold setting (Table 5.2). Thus the sorter was performing better than expected by the model in terms of adopting the RMSEP as the SD input of the transfer function.

Consideration of a number of populations is recommended to allow interpretation of optimal bandwidth and RMSEP. Likely factors include variation in the value of residuals with level of attribute.

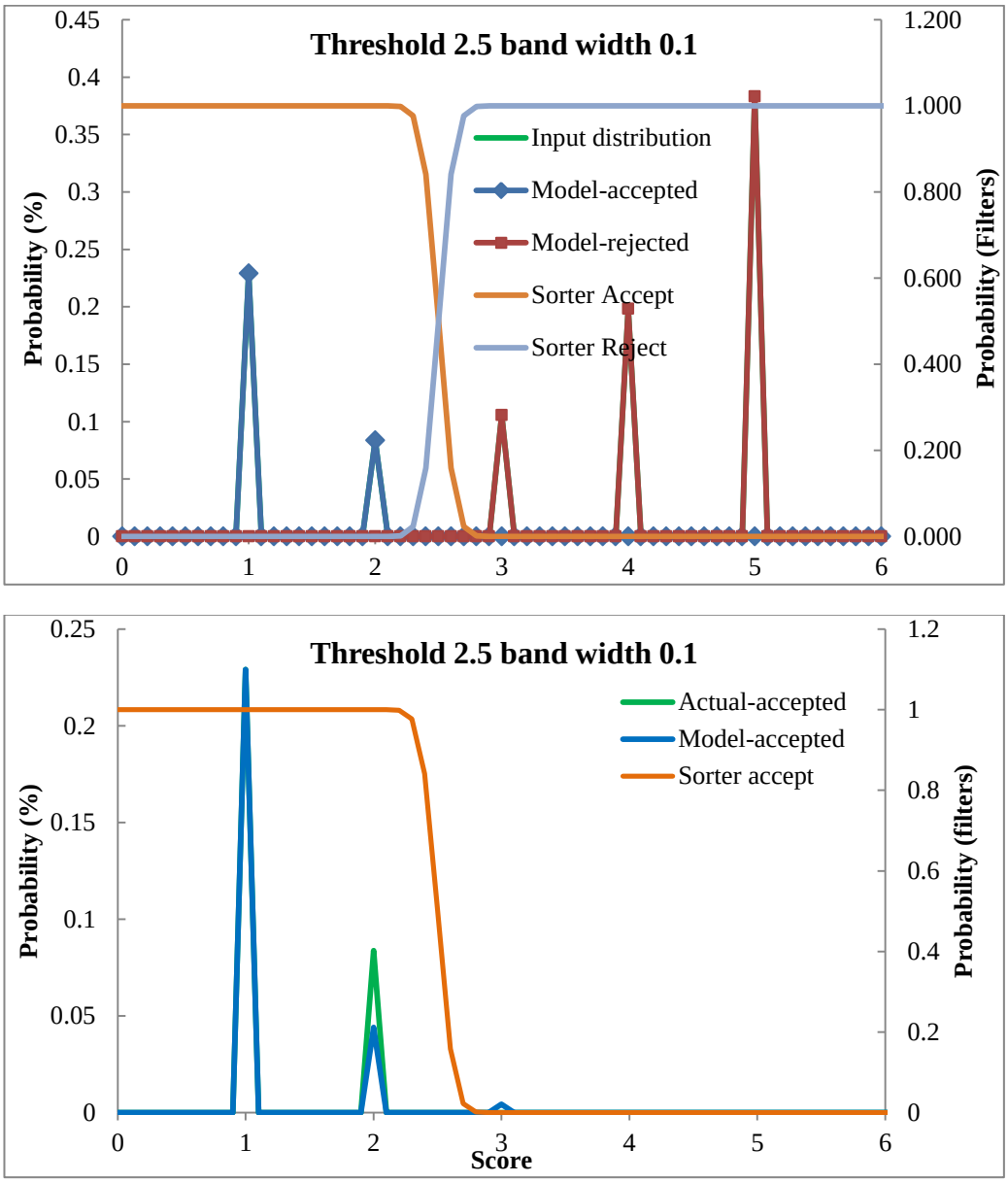



**Figure 5.8.** Probability distribution of filters and of classes. Sorter accept and reject curves represent the transfer function, a normal distribution with a standard deviation equal to the band width setting. The probability distribution for the five score bands is shown for (a) the initial (unsorted) population, (b) the modelled accepted population, (c) the modelled rejected population, (d) the actual accepted population and (e) the actual rejected population. Bottom panel represents the sorter accept curve with model-accepted data with addition of data on the actual-accepted.

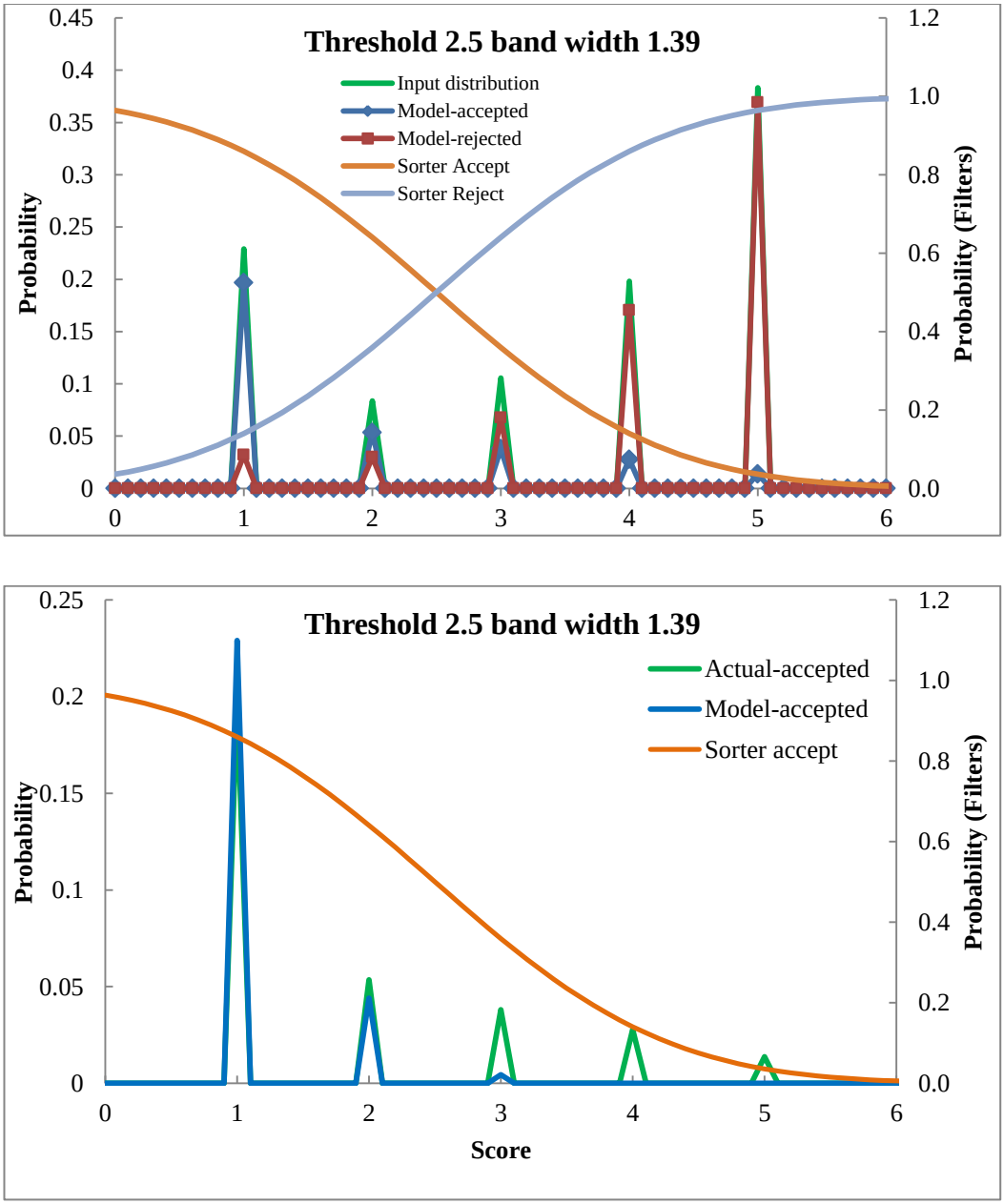



**Figure 5.9.** Probability distribution of filters and of classes. Sorter accept and reject curves represent the transfer function, a normal distribution with a standard deviation equal to the band width setting. The probability distribution for the five score bands is shown for (a) the initial (unsorted) population, (b) the modelled accepted population, (c) the modelled rejected population, (d) the actual accepted population and (e) the actual rejected population. Band width changed to 1.39.

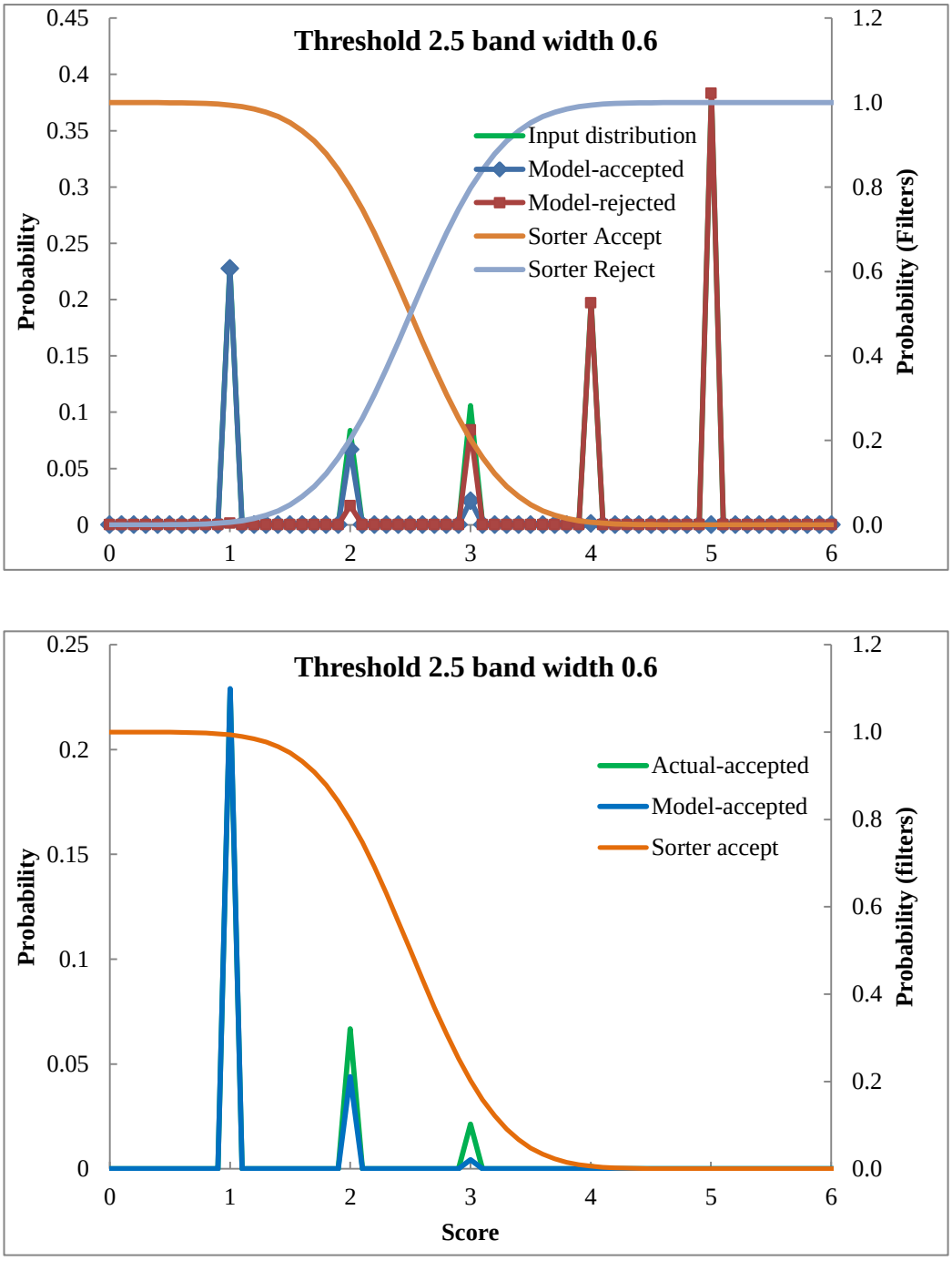



**Figure 5.10.** Probability distribution of filters and of classes. Sorter accept and reject curves represent the transfer function, a normal distribution with a standard deviation equal to the band width setting. The probability distribution for the five score bands is shown for (a) the initial (unsorted) population, (b) the modelled accepted population, (c) the modelled rejected population, (d) the actual accepted population and (e) the actual rejected population. Band width changed to 0.6.

**Table 5.2.** Optimum bandwidth for the given threshold setting for model and actual sorting operation.

| Model Threshold | Actual Threshold | Band width | Sum of squares root |
|---|---|---|---|
| 0.5 | 0.5 | 0.28 | 1.01209E-07 |
| 1 | 1 | 0.29 | 0.001940655 |
| 1.5 | 1.5 | 0.81 | 0.000597161 |
| 2 | 2 | 0.71 | 0.000905361 |
| 2.5 | 2.5 | 0.6 | 0.000807934 |
| 3 | 3 | 0.75 | 0.000443311 |
| 3.5 | 3.5 | 0.68 | 0.00021429 |
| 4 | 4 | 0.96 | 1.93852E-05 |
| 4.5 | 4.5 | 1.75 | 0.004545492 |
| 5 | 5 | 0.5 | 0.005943266 |

### 5.3.7 A Sorting Optimisation Curve with pricing data as a decision support aid

In sorting of fruit for internal browning, two groups of fruit are created, one with lower incidence of defect and so higher value, and one with higher incidence of defect and so lower value. As noted earlier, 'harsher' sorting (lower threshold value) will result in a lower false discovery rate but also a lower yield. A practical or economic decision on sorting pressure (threshold value setting) requires consideration of the three parameters of input distribution, misclassification level and pricing. Ooms et al. (2010) describe the compromise between yield and a quality factor and suggest use of a 'sorting optimisation curve' (SOC) to guide the choice of the threshold value used in the classifier. In this approach, yield refers to (TP+FP)/(P+N), while quality refers to TPR or TPR with an assigned value (e.g. price) to sorted objects based on the proportion of defect items present in the lot.

For example, if price achieved on fruit with <2% incidence level of defect is $35/(9 kg) box with the alternative being sale at $150/tonne ($13.50/9 kg) for juicing (a scenario presented by one grower), then the sorting operator will 'intuitively' know to adjust the sorting threshold so as to achieve specification on false discovery rate, regardless of yield, in order to optimise economic outcome. However, when the

price differential is not as clear, other sorting outcomes may lead to the optimal economic result. For example, in practice growers report that the major retailers display some tolerance in false discovery rate above their published specification levels for apple browning. Additionally fruit at moderate levels of defect may be sold on the general market.

A decision support tool was created to assist the sorting operator in setting of a threshold value. This tool codifies optimum behaviour in terms of price achieved. The tool requires:

(i) knowledge of the distribution of defect in the population. The output of the sorting unit could be accepted.
(ii) knowledge of accuracy of the sorting operation. A representative number of fruit estimated as per 1.3.2 should be destructively sampled and assessed relative to the sorting operation result.
(iii) knowledge of price achieved for fruit with given levels of defect,

The modelling of sorter output accuracy for given threshold settings has been described in the previous section. A function was added to this model to couple user-supplied pricing data to the sorter output distribution data to calculate total value achieved for the consignment. To demonstrate this calculator, consider the following three scenarios:

**Case 1.** Pricing of $ 35 / 9 kg box if defect incidence level is below 2%, $0/box for higher levels of defect incidence.

**Case 2.** Pricing of $ 35 / 9 kg box if defect incidence level is below 5%, $0/box for higher levels of defect incidence.

**Case 3.** Pricing of $ 35 / 9 kg box if defect incidence level is <5% defect incidence, $ 20/box for fruit with up to 40% incidence, $0 for other fruit.

Maximum value is achieved using a threshold of 2, 2.5 and 3.5 in Cases 1, 2 and 3, respectively (Fig 5.11).

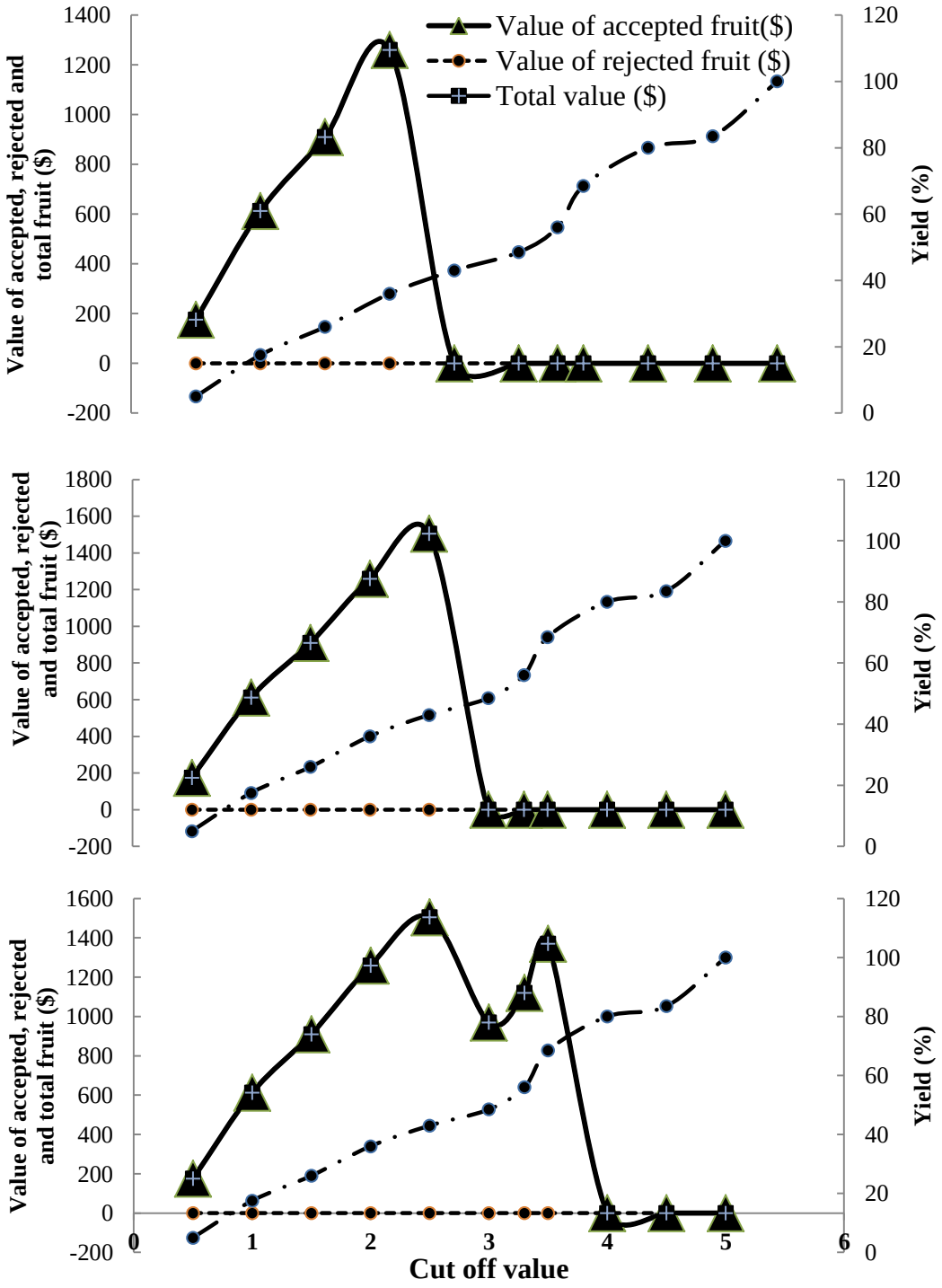



**Figure 5.11.** Value ($) of product achieved with change in sorting threshold level for three arbitrary market pricing scenarios: (a) $35/box for fruit with <2% defect incidence, $0 for other fruit; (b) $35/box for fruit with <5% defect incidence, $0 for other fruit; (c) $35/box for fruit with <5% defect incidence, $ 20 for fruit with up to 40% incidence, $ 0 for other fruit.

A similar optimisation applies for change in the sorting function. For example, should the sorting accuracy decrease, resulting in more misclassification events, the threshold sorting value to achieve maximum value will change.

## 5.4 Conclusion

The commercial front of the application of NIR spectroscopy in quality sorting involves the classification of the entire fruit into groups of desired quality. This process includes transcribing the spectroscopic information retained in the spectra into the number and applies on to the sorting operation. Sorting fruit for internal defects involves an exercise of probabilities, of incidence and of detection. Optimising the sorting process also requires knowledge of price available for fruit with different defect incidence levels and tolerance for those defects in the market based on given consignment criteria agreed between parties. Calculators have been produced to guide the choice of a statistically relevant number of fruit samples to judge defect incidence level, to predict the sorting outcomes given change in threshold sorting values and to estimate the value achieved for the sorted product. This field is ‘fertile’ and further effort is recommended to maximise the practical outcomes possible from sorting technology.

Further work with a range of populations varying in defect level and spread is recommended to develop decision guides around the setting of a threshold sorting level given knowledge of population statistics, to maintain a desired false defect rate, and to allow for interpretation of the value of the bandwidth setting that minimises model error, relative to sorting error as measured by RMSEP. It is recommended that the operation of the sorting process be guided by a destructive sampling of sample lots to gauge errors in class assignment.

# Chapter 6.
# Conclusion and future directions

The need to focus on fruit quality is a mantra within the fruit industry. Such a focus must be applied throughout a developed horticulture supply chain. The use of near infrared spectroscopy (NIRS) in the sorting of internal defects is but one element in a quality 'tool-kit', in which the ultimate aim must be to avoid the development of disorders, and thus avoid the need for sorting processes. However, in practice there is need for efficient sorting processes to at least reduce defect incidence in consignments to a level that is acceptable to the market place.

Seated deeply in the fruit tissue, internal defects pose a major problem for detection and automated sorting. Many research reports on sorting of fruit for internal defect have been based on limited sample sets (one population divided into a calibration and validation set), and so offer optimistic results. Further, while some commercial claims exist for sorting of fruit on internal defect, performance of such instrumentation has yet to be documented.

In approaching the issue of detection of a particular disorder, the following questions should be raised: (i) what is the cause of the disorder (and how can a range of levels of defect be produced/acquired for trial work); (ii) what are the anatomical features of the disorder and how does this inform selection of detection technology and arrangements, e.g. choice of optical geometry; (iii) is a quantitative estimation of defect level required or is a discriminate pass/fail answer sufficient?; (iv) what level of type I and II errors are required?

In this thesis these issues were explored in context of apple browning, mandarin gelling defect and pineapple translucency. Chapters 2 presents on the non invasive detection of apple diffuse browning using spectra acquired from defect and good fruit in interactance, partial transmission and full transmission optical geometries. Defect fruit reveals high absorbance of light at visible-shortwave NIR region compared to normal one with distinctive difference at certain wavelength associated with varied absorbance with aromatic hydrocarbons (brown polymers). This exercise results in improved classification accuracy of good and defect fruit to meet the retailer's standards for defects at the cost of losing less than 10% of the good fruit as

defect using discriminant analysis methods. Similarly, detection of granulation in 'Imperial' mandarin was presented in Chapter 3, based on a transmission optical geometry, although the robustness of the technique to fruit grown under different conditions requires further work. Chemical and anatomical characterisation reveals that the granulation in mandarin is more 'gelling' of the tissues with rapid cell division, formation of multiple epidermal and hypodermal layers. It is likely that the method will be sensitive to changes in fruit optical properties, e.g. resulting from skin 'tightness'. This defect was also characterised as representing a proliferation on the hypodermal cells of the juice sacs. Results for detection of pineapple translucency, presented in Chapter 4, were less encouraging. The complexity of the multiple and aggregate fruit, with presence of bracts and central vascular axis, are likely contributors to the poorer result for this application. Operational issues associated with the instrumentation used in the thesis are presented in Appendix 1.

Further work in the development of instrumentation is anticipated, both in terms of full window spectral devices and devices based on absorbance at only a few wavelengths, given continued advances in both light sources (particularly LEDs) and detectors. Attention to the development of imaging systems is recommended, to deal with the issue of non-uniform distribution of defect within fruit, and the presence of unrelated structures, such as in pineapple fruit. Technologies other than NIRS also hold promise, particularly Xray-CT and MRI, as speed of operation and cost decrease. Aspect Imaging P/L of Israel is one company to watch in this context. The development of chlorophyll fluorescence as an on-line screening tool also has potential. The continued adoption of this technology into controlled atmosphere fruit storage by companies such as Harvest Watch and Bessling is anticipated.

The consumer 'pull' for improved fruit quality, particularly demand for fruit free from internal defect, will extend the use of non-invasive sorting technology to other fruit commodities and other internal defects. Insect infestation and internal rots are likely target applications. For example, Thomas et al. (1995) reported detection of mango seed weevil (*Cryptorynchus mangiferae*) inside mango fruit using a X-ray imaging technique with 100% accuracy, although the sample was relatively small (84). (Xing & Guyer, 2008a) reported use of SWNIR transmittance (550-980 nm) for detection of internal insect infestation in cherry, with higher absorbance at longer

wavelengths in defect fruit. Canonical discriminant analysis yielded a classification accuracy of 80-86% for infested fruit. (Wang et al., 2011) also reported the use of discriminant analysis models for detection on insects within jujube fruit. (Haff et al., 2013) reported detection of mango fruit fly infestation using a hyperspectral imaging technique with 50 mangoes. A 6% misclassification rate was recorded, including 1% defect fruit predicted as good while 11.1% of the good fruit predicted as defect.

The application of a detection technology in context of a sorting operation, to achieve binary classifier through setting of a threshold value, is discussed in Chapter 5. The trade-off between the good fruit rejected as defect (economic loss to the growers) and defect fruit accepted as good (threat to retail specification) are discussed. Various aids to assist the operator in choice of a threshold sorting value are presented, including a Receiver Operator Characteristic (ROC) curve and a Sorting Optimization Curve (SOC). Further work in this area is recommended, to produce tools of practical value to the packing line operator. Further, sorting to remove defect fruit represents economic loss, so with greater knowledge of incidence levels from the sorting operations, increased attention on the control of the development of the defect is likely. A key part of such control is the improvement of models that predict defect incidence, based on genotypic, pre-harvest and postharvest conditions. Future research on internal defects in fruit should focus on;

i. preventing the incidence of internal defects by assessing the interactions among different management and genetic factors using multivariate statistical approach.
ii. quantitative modelling of incidence of internal defects by integrating the biochemical (within fruit) and environmental data (outside fruit).
iii. development of economic, fast, accurate, non-invasive and easy – to-operate methods for measuring and on line sorting of defect fruit for commercial application.

# References


Abbott, JA 1999, 'Quality measurement of fruits and vegetables', *Postharvest Biology and Technology*, vol. 15, pp. 207-225.

Abdullah, H, Wills, R, M., AR, Othman, Z, Fatt, LP & Smith, M 2010, 'Blackheart disorder in fresh pineapple ', *Fresh Produce*, vol. 4 pp. 29-35.

Abe, S 2010, *Support vector machines for pattern classification*, Advances in Pattern Recognition, Springer Verlag London Limited.

ABS 2015, *Agricultural commodities, Australia, 2013-2014*, Australian Bureau of Statistics viewed 16 August 2015, http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/7121.0Main%20Features32013-14?opendocument&tabname=Summary&prodno=7121.0&issue=2013-14&num=&view=

Acharya, UK, Walsh, KB & Subedi, PP 2014, 'Robustness of partial least-squares models to change in sample temperature: II. Application to fruit attributes ', *Journal of Near Infrared Spectroscopy*, vol. 22, pp. 287-295.

Al-Najjar, NI & Pai, MM 2014, 'Coarse decision making and overfitting', *Journal of Economic Theory*, vol. 150, pp. 467-486.

Albrigo, LG, Kawada, PW, Hale, PW, Smoot, JJ & Hatton, TTJ 1980, 'Effect of harvest date and preharvest and postharvest treatments on Florida grapefruit condition in export to Japan', *Proceedings of the Florida State Horticultural Society*, vol. 102, pp. 236-239.

Andreassen, CN 2015, 'A simulated SNP experiment indicates a high risk of over-fitting and false positive results when a predictive multiple SNP model is established and tested within the same dataset', *Radiotherapy and Oncology*, vol. 114, pp. 310-313.

Ariana, DP & Lu, R 2010, 'Evaluation of internal defect and surface color of whole pickles using hyperspectral imaging', *Journal of Food Engineering*, vol. 96, pp. 583-590.

Arpaia, ML 2007 *Best management practices, food safety and postharvest* pp. 1-7, California Avocado Commission California, USA

Baranowski, P, Mazurek, W, Witkowska-Walczak, B & Sławiński, C 2009, 'Detection of early apple bruises using pulsed-phase thermography', *Postharvest Biology and Technology*, vol. 53, pp. 91-100.

Baranowski, P, Mazurek, W, Wozniak, J & Majewska, U 2012, 'Detection of early bruises in apple using hyperspectral data and thermal imaging', *Journal of Food Engineering*, vol. 110, pp. 345-355.

Batt, PJ & Sadler, C 1999, 'Lables on apples : winners and losers', in J Cadeaux & M Uncles (eds), *Marketing in the Third Millennium*, University of New South Wales, Sydney, Australia

Beaudry, R 2014, 'Watercore in apples: causes, concerns, detection and sorting ', Postharvest Laboratory, Michigan State University Michigan, USA.

Bechar, A, Mizrach, A, Barreiro, P & Landahl, S 2005, 'Determination of mealiness in apples using ultrasonic measurements', *Biosystems Engineering*, vol. 91, pp. 329-334.

Bee, SC & Honeywood, MJ 2004, 'Optical sorting systems', in M Edwards (ed.), *Detecting Foreign Bodies in Food*, Woodhead Publishing, Cambridge, United Kingdom.

Benjakul, S, Eadkhong, T, Limmun, W & Danworaphong, S 2013, 'Probability of finding translucent flesh in mangosteen based on its electrical resistance and capacitance ', *Food Science Biotechnology* vol. 22 pp. 413-416.

Benkeblia, N, Tennant, DPF, Jawandha, SK & Gill, PS 2011, 'Preharvest and harvest factors influencing the postharvest quality of tropical and subtropical fruits', in EM Yahia (ed.), *Postharvest Biology and Technology of Tropical and Subtropical Fruits*, Woodhead Publishing, Cambridge, United Kingdom.

Bergman, HE, Crouch, I, Crouch, MJ & Majoni, J 2012, 'Update on the possible causes and management strategies of flesh browning disorders in 'Cripps Pink' apples', *South African Fruit Journal*, pp. 56-59.

Bielza, C, Barreiro, P, Rodríguez-Galiano, MI & Martín, J 2003, 'Logistic regression for simulating damage occurrence on a fruit grading line', *Computers and Electronics in Agriculture*, vol. 39, pp. 95-113.

Boeker, P 2014, 'On Electronic Nose methodology', *Sensors and Actuators : Chemical*, vol. 204, pp. 2-17.

Bollen, AF & Prussia, SE 2009, 'Sorting for defects and visual quality attributes', in W Florkowski, J., RL Shewfelt , B Brueckner & SE Prussia (eds), *Postharvest Handling (2nd edn.)*, Academic Press, San Diego, USA.

Bollen, AF & Prussia, SE 2014, 'Sorting for defects', in W Florkowski, J., RL Shewfelt , B Brueckner & SE Prussia (eds), *Postharvest Handling (3rd edn.)*, Academic Press, San Diego, USA.

Bugaud, C, Ocrisse, G, Salmon, F & Rinaldo, D 2014, 'Bruise susceptibility of banana peel in relation to genotype and post-climacteric storage conditions', *Postharvest Biology and Technology*, vol. 87, pp. 113-119.

Burdon, JN, Moore, KG & Wainwright, H 1991, 'Mineral distribution in mango fruit susceptible to the physiological disorder soft-nose', *Scientia Horticulturae*, vol. 48, pp. 329-336.

Burns, JK & Achor, DS 1989 'Cell wall changes in juice vesicles associated with 'section drying' in stored late-harvested grapefruit', *Journal of the American Society for Horticultural Science*, vol. 114, pp. 283-287.

Calders, T & Jaroszewicz, S 2007, 'Efficient AUC optimization for classification ', in J Kok, N., J NKoronacki, RLd Mantaras, S Matwin, D Maladenic & A Skowron (eds), *Knowledge discovery in Databases* Springer, Berlin, Germany

Castro, E, Barrett, DM, Jobling, J & Mitcham, EJ 2008, 'Biochemical factors associated with a $CO_2$-induced flesh browning disorder of Pink Lady™ apples', *Postharvest Biology and Technology*, vol. 48, pp. 182-191.

Castro, ED, B. Biasi, E. Mitcham, S. Tustin, D. Tanner & Jobling J 2007, 'Carbon dioxide induced flesh browning in Pink Lady™ apples', *Journal of Amercian Society of Horticulture Science*, vol. 132 pp. 713-719.

Chakrawar, VR & Singh, R 1978 'Studies on citrus granulation: IV effect of some growth regulators on granulation and quality of citrus fruits ', *Haryana Journal of Horticulture Science*, vol. 7, pp. 130-135.

Chayaprasert, W & Stroshine, R 2005, 'Rapid sensing of internal browning in whole apples using a low-cost, low-field proton magnetic resonance sensor', *Postharvest Biology and Technology*, vol. 36, pp. 291-301.

Chen, CC & Paull, RE 2000, 'Sugar metabolism and pineapple flesh translucency ', *Journal of Amercian Society of Horticulture Science*, vol. 125 pp. 558-562.

Chen, K, Xu, C, Li, F & Zhang, S 2005, 'Postharvest granulation of 'Huyou' (*Citrus changshanensis*) fruit in response to calcium', *Israel Journal of Plan Sciences*, vol. 53, pp. 35-40.

Chen, L & Opara, UL 2013, 'Texture measurement approaches in fresh and processed foods: a review', *Food Research International*, vol. 51, pp. 823-835.

Chen, Q, Zhang, C, Zhao, J & Ouyang, Q 2013, 'Recent advances in emerging imaging techniques for non-destructive detection of food quality and safety', *TrAC Trends in Analytical Chemistry*, vol. 52, pp. 261-274.

Chia, KS, Abdul, RH & Abdul, RR 2012, 'Prediction of soluble solids content of pineapple via non-invasive low cost visible and shortwave near infrared spectroscopy and artificial neural network', *Biosystems Engineering*, vol. 113, pp. 158-165.

Chia, KS, Abdul, RH & Abdul, RR 2013, 'Evaluation of common pre-processing approaches for visible (VIS) and shortwave near infrared (SWNIR) spectroscopy in soluble solids content (SSC) assessment', *Biosystems Engineering*, vol. 115, pp. 82-88.

Choi, JH, Yim, SH, Cho, KS, Kim, MS, Park, YS, Jung, SK & Choi, HS 2015, 'Fruit quality and core breakdown of ‘Wonhwang’ pears in relation to harvest date and pre-storage cooling', *Scientia Horticulturae*, vol. 188, pp. 1-5.

Chongchatuporn, U, Ketsa, S & Van Doorn, WG 2013, 'Chilling injury in mango (*Mangifera indica*) fruit peel: relationship with ascorbic acid concentrations and antioxidant enzyme activities', *Postharvest Biology and Technology*, vol. 86, pp. 409-417.

Clark, CJ, MacFall, JS & Bieleski, RL 1998, 'Loss of watercore from Fuji apple observed by magnetic resonance imaging', *Scientia Horticulturae*, vol. 73, pp. 213-227.

Clark, CJ, McGlone, VA & Jordan, RB 2003, 'Detection of brownheart in 'Braeburn' apple by transmission NIR spectroscopy', *Postharvest Biology and Technology*, vol. 28, pp. 87-96.

Clark, CJ, Hockings, PD, Joyce, DC & Mazucco, RA 1997, 'Application of magnetic resonance imaging to pre and postharvest studies of fruits and vegetables ', *Postharvest Biology and Technology*, vol. 11, pp. 1-21.

Clark, CJ, White, A, Jordan, RB & Woolf, AB 2007, 'Challenges associated with segregation of avocados of differing maturity using density sorting at harvest', *Postharvest Biology and Technology*, vol. 46, pp. 119-127.

Collins, JL 1968, *The Pineapple.*, Leonard Hill London

Conzen, JP 2003 *Multivariate calibration* Bruker Optics Germany.

Cunnigham, DC & Walsh, KB 2002, 'Galactomannan content and composition in *Cassia brewsteri* seed', *Australian Journal of Experimental Agriculture*, vol. 42, pp. 1081-1086.

DAF 2015, *Pineapples* Department of Agriculture, Fisheries, Queensland Government, Australia, viewed 13 August 2015, https://www.daf.qld.gov.au/plants/fruit-and-vegetables/fruit-and-nuts/pineapples

DAFF 2013, *Queensland AgTrends 2012-2013, Forecasts and trends in Queensland agricultural, fisheries and forestry production*, Department of Agriculture, Fisheries and Forestry, Brisbane, Australia

Dart, JA & Newman, SM 2005 'Watercore of apples ', *Primefacts*, no. 49,

de Freitas, ST, do Amarante, CVT & Mitcham, EJ 2015, 'Mechanisms regulating apple cultivar susceptibility to bitter pit', *Scientia Horticulturae*, vol. 186, pp. 54-60.

de Freitas, ST, Amarante, CVTd, Labavitch, JM & Mitcham, EJ 2010, 'Cellular approach to understand bitter pit development in apple fruit', *Postharvest Biology and Technology*, vol. 57, pp. 6-13.

de Freitas, ST, Amarante, CVT, Dandekar, AM & Mitcham, EJ 2013, 'Shading affects flesh calcium uptake and concentration, bitter pit incidence and other fruit traits in "Greensleeves" apple', *Scientia Horticulturae*, vol. 161, pp. 266-272.

Dhanoa, MS, Lister, S, Sanderson, R & Barnes, RJ 1994, 'The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra ', *Journal of Near Infrared Spectroscopy*, vol. 2, pp. 43-47.

Donis-González, IR, Guyer, DE, Pease, A & Barthel, F 2014, 'Internal characterisation of fresh agricultural products using traditional and ultrafast electron beam X-ray computed tomography imaging ', *Biosystems Engineering*, vol. 117, pp. 104-113.

Egidio, VD, Sinelli, N, Limbo, S, Torri, L, Franzetti, L & Casiraghi, E 2009, 'Evaluation of shelf-life of fresh-cut pineapple using FT-NIR and FT-IR spectroscopy', *Postharvest Biology and Technology*, vol. 54, pp. 87-92.

Eksteen, GJ & Truter, AB 1987, 'Controlled atmosphere storage of deciduous fruit in South Africa', *International Journal of Refrigeration*, vol. 10, pp. 14-17.

Fawcett, T 2006, 'An introduction to ROC analysis', *Pattern Recognition Letters*, vol. 27, pp. 861-874.

Felicetti, E, Mattheis, JP, Zhu, Y & Fellman, JK 2011, 'Dynamics of ascorbic acid in ‘Braeburn’ and ‘Gala’ apples during on-tree development and storage in atmospheres conducive to internal browning development', *Postharvest Biology and Technology*, vol. 61, pp. 95-102.

Ferguson, I, Volz, R & Woolf, A 1999a, 'Preharvest factors affecting physiological disorders of fruit', *Postharvest Biology and Technology*, vol. 15, pp. 255-262.

Ferguson, I, Volz, R & Woolf, A 1999b, 'Postharvest factors affecting physiological disorders of fruit', *Postharvest Biology and Technology*, vol. 13, pp. 255-262.

Florissen, P, Ekman, JS, Blumenthal, C, McGlasson, WB, Conroy, J & Holford, P 1996, 'The effects of short heat-treatments on the induction of chilling injury in avocado fruit (*Persea americana* Mill)', *Postharvest Biology and Technology*, vol. 8, pp. 129-141.

Fornasini, P 2008, *The uncertainty in physical measurements: an introduction to data analysis*, Springer Science +Business Media, New York, USA.

Francis, FJ, Bramlage, WJ & Lord, WJ 1965, 'Detection of watercore and internal breakdown in delicious apple by light transmittance', *Journal of Amercian Society of Horticulture Science*, vol. 87 pp. 78-84.

Franck, C, Lammertyn, J, Ho, QT, Verboven, P, Verlinden, B & Nicolaï, BM 2007, 'Browning disorders in pear fruit', *Postharvest Biology and Technology*, vol. 43, pp. 1-13.

Franck, C, Baetens, M, Lammertyn, J, Scheerlinck, N, Davey, MW, Nicolaı, amp, x & M., B 2003, 'Ascorbic acid mapping to study core breakdown development in 'Conference' pears', *Postharvest Biology and Technology*, vol. 30, pp. 133-142.

Fu, X, Ying, Y, Lu, H & Xu, H 2007, 'Comparison of diffuse reflectance and transmission mode of visible-near infrared spectroscopy for detecting brown heart of pear', *Journal of Food Engineering*, vol. 83, pp. 317-323.

Galvis-Sánchez, AC, Fonseca, SC, Morais, AMMB & Malcata, FX 2004, 'Effects of preharvest, harvest and postharvest factors on the quality of pear (cv. Rocha) stored under controlled atmosphere conditions', *Journal of Food Engineering*, vol. 64, pp. 161-172.

Gamble, J, Harker, FR, Jaeger, SR, White, A, Bava, C, Beresford, M, Stubbings, B, Wohlers, M, Hofman, PJ, Marques, R & Woolf, A 2010, 'The impact of dry matter, ripeness and internal defects on consumer perceptions of avocado quality and intentions to purchase', *Postharvest Biology and Technology*, vol. 57, pp. 35-43.

Golic, M, Walsh, KB & P., L 2003 'Short-wavelength near infrared spectra of sucrose, glucose and fructose with respect to sugar concentration and temperature', *Applied Spectroscopy*, vol. 53 pp. 139-145.

Gómez Sanchis, J, Blasco, J, Soria-Olivas, E, Lorente, D, Escandell Montero, P, Martínez Martínez, JM, Martínez Sober, M & Aleixos, N 2013, 'Hyperspectral LCTF-based system for classification of decay in mandarins caused by *Penicillium digitatum* and *Penicillium italicum* using the most relevant bands and non-linear classifiers', *Postharvest Biology and Technology*, vol. 82, pp. 76-86.

Gongal, A, Amatya, S, Karkee, M, Zhang, Q & Lewis, K 2015, 'Sensors and systems for fruit detection and localization: A review', *Computers and Electronics in Agriculture*, vol. 116, pp. 8-19.

Gonzalez, JJ, Valle, RC, Bobroff, S, Biasi, WV, Mitcham, EJ & McCarthy, MJ 2001, 'Detection and monitoring of internal browning development in Fuji apples using MRI', *Postharvest Biology and Technology*, vol. 22, pp. 179-188.

Greensill, CV & Walsh, KB 2000, 'A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit', *Measurement of Science and Technology* vol. 10 pp. 27-35.

Gudenschwager, O, García-Rojas, M, Defilippi, BG & González-Agüero, M 2013, 'Identification and characterization of two putative genes encoding acetyl-coenzyme A carboxylase subunits that are possibly associated with internal browning during cold storage of 'Hass' avocados (Persea americana Mill.)', *Postharvest Biology and Technology*, vol. 84, pp. 74-80.

Guthrie, JA, Wedding, BB & Walsh, KB 1998, 'Roubstness of NIR calibrations for soluble solids in intact melon and pineapple', *Journal of Near Infrared Spectroscopy*, vol. 6,

Haff, RP, Slaughter, DC, Sarig, Y & Kader, A 2006, 'X-ray assessment of translucency in pineapple', *Journal of Food Processing and Preservation*, vol. 30, pp. 527-533.

Haff, RP, Saranwong, S, Thanapase, W, Janhiran, A, Kasemsumran, S & Kawano, S 2013, 'Automatic image analysis and spot classification for detection of fruit fly infestation in hyperspectral images of mangoes', *Postharvest Biology and Technology*, vol. 86, pp. 23-28.

Han, D, Tu, R, Lu, C, Liu, X & Wen, Z 2006, 'Nondestructive detection of brown core in the Chinese pear 'Yali' by transmission visible–NIR spectroscopy', *Food Control*, vol. 17, pp. 604-608.

Hanley, JA & McNeil, BJ 1982, 'The meaning and use of the area under a receiver operating characterstic (ROC) curve', *Diagnostic Radiology* vol. 143, pp. 29-36.

Hannah, JJ 2007, Understanding the flesh browning disorder of 'Cripps Pink' apple, Doctor of Philosophy thesis, Sydney University, Sydney, Australia.

Harker, FR 2001, 'Consumer response to apples', in W-TPI Network (ed.), *Washington tree fruit postharvest conference* WSU-TFREC Postharvest Information Network, Wenatchee, WA.

Harker, FR, Gunson, FA & Jaeger, SR 2003, 'The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples', *Postharvest Biology and Technology*, vol. 28, pp. 333-347.

Harker, FR, Kupferman, EM, Marin, AB, Gunson, FA & Triggs, CM 2008, 'Eating quality standards for apples based on consumer preferences', *Postharvest Biology and Technology*, vol. 50, pp. 70-78.

Hatoum, D, Annaratone, C, Hertog, MLATM, Geeraerd, AH & Nicolai, BM 2014a, 'Targeted metabolomics study of 'Braeburn' apples during long-term storage', *Postharvest Biology and Technology*, vol. 96, pp. 33-41.

Hatoum, D, Buts, K, Hertog, MLATM, Geeraerd, AH, Schenk, A, Vercammen, J & Nicolai, BM 2014b, 'Effects of pre and postharvest factors on browning in Breaeburn ', *HortScience*, vol. 41, pp. 19-26.

Hernández-Sánchez, N, Hills, BP, Barreiro, P & Marigheto, N 2007, 'An NMR study on internal browning in pears', *Postharvest Biology and Technology*, vol. 44, pp. 260-270.

Herremans, E, Verboven, P, Bongaers, E, Estrade, P, Verlinden, BE, Wevers, M, Hertog, MLATM & Nicolai, BM 2013, 'Characterisation of 'Braeburn' browning disorder by means of X-ray micro-CT', *Postharvest Biology and Technology*, vol. 75, pp. 114-124.

Herremans, E, Verboven, P, Defraeye, T, Rogge, S, Ho, QT, Hertog, MLATM, Verlinden, BE, Bongaers, E, Wevers, M & Nicolai, BM 2014a, 'X-ray CT for quantitative food microstructure engineering: The apple case', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 324, pp. 88-94.

Herremans, E, Melado-Herreros, A, Defraeye, T, Verlinden, B, Hertog, M, Verboven, P, Val, J, Fernández-Valle, ME, Bongaers, E, Estrade, P, Wevers, M, Barreiro, P & Nicolaï, BM 2014b, 'Comparison of X-ray CT and MRI of watercore disorder of different apple cultivars', *Postharvest Biology and Technology*, vol. 87, pp. 42-50.

Herve, A 2003 *Partial least square (PLS) regression* Thousand Oaks CA.

Hills, BP & Clark, CJ 2003, 'Quality assessment of horticultural products by NMR', *Annual Reports on NMR Spectroscopy*, vol. 50, pp. 75-120.

Ho, QT, Verboven, P, Verlinden, BE, Schenk, A & Nicolaï, BM 2013, 'Controlled atmosphere storage may lead to local ATP deficiency in apple', *Postharvest Biology and Technology*, vol. 78, pp. 103-112.

Hofman, H 2011, *Management of internal dryness of Imperial mandarin* Horticulture Australia Ltd Sydney, Australia.

Holderbaum, DF, Kon, T & Guerra, MP 2014, 'Dynamics of total phenolic content in different apple tissues and genotypes: impacts and relevance for breeding programs', *Scientia Horticulturae*, vol. 168, pp. 58-63.

Holderbaum, DF, Kon, T, Kudo, T & Guerra, MP 2010, 'Enzymatic browning, polyphenol oxidase activity and polyphenols in four apple cultivars: dynamics during fruit development ', *HortScience* vol. 45 pp. 1150-1154.

Hong, K, Xu, H, Wang, J, Zhang, L, Hu, H, Jia, Z, Gu, H, He, Q & Gong, D 2013, 'Quality changes and internal browning developments of summer pineapple fruit during storage at different temperatures', *Scientia Horticulturae*, vol. 151, pp. 68-74.

Huang, H, Zhu, Q, Zhang, Z, Yang, B, Duan, X & Jiang, Y 2013, 'Effect of oxalic acid on antibrowning of banana (*Musa* spp., AAA group, cv. 'Brazil') fruit during storage', *Scientia Horticulturae*, vol. 160, pp. 208-212.

Huang, J, Romero-Torres, S & Moshgbar, M 2010, 'Practical considerations in data pre-treatment for NIR and rama spectroscopy ', viewed 28 August 2015,

Huque, R, Wills, RBH, Pristijono, P & Golding, JB 2013, 'Effect of nitric oxide (NO) and associated control treatments on the metabolism of fresh-cut apple slices in relation to development of surface browning', *Postharvest Biology and Technology*, vol. 78, pp. 16-23.

Jamal, N 2012, 'Application of acoustic properties in non destructive quality evaluation of agricultural products ', *International Journal of Engineering and Technology* vol. 2 pp. 668-675.

James, HJ & Jobling, JJ 2009, 'Contrasting the structure and morphology of the radial and diffuse flesh browning disorders and $CO_2$ injury of 'Cripps Pink' apples', *Postharvest Biology and Technology*, vol. 53, pp. 36-42.

Jha, SN & Matsuoka, T 2000, 'Non destructive techniques for quality evaluation of intact fruits and vegetables', *Food Science Technology Research*, vol. 6, pp. 248-251.

Joachims, T 2005, 'A support vector method for multivariate performance measures', paper presented at the 22nd International Conference on Machine Learning, Bonn, Germany.

Jobling, J 2005, *Understanding the flesh browning disorder of Cripps Pink apples* Horticulture Australia Ltd, Sydney, Australia.

Jobling, JJ & James, HJ 2008, *Managing the flesh browning disorder of Cripps Pink apples*, Apple and Pear Australia Limited, Australia.

Kauer, JS & White, J 2009, 'Electronic Nose', in LR Squire (ed.), *Encyclopedia of Neuroscience*, Academic Press, Oxford, United Kingdom.

Kaur, H, Chanana, YR & Kapur, SP 1991, 'Effect of growth regulators on granualtion and fruit quality of sweet orange cv. mosambi', *Indian Journal of Horticulture*, vol. 48, pp. 224-227.

Kawano, S, Fjuiwara, T & MIwamoto, MJ 1993, 'Nondestructive determination of sugar content in Satsuma mandarin using near infrared (NIR) transmittance', *Journal of Japananese Society of Horticultural Science* vol. 62 pp. 465-470.

Kendra, PE, Roda, AL, Montgomery, WS, Schnell, EQ, Niogret, J, Epsky, ND & Heath, RR 2011, 'Gas chromatography for detection of citrus infestation by fruit fly larvae (Diptera: Tephritidae)', *Postharvest Biology and Technology*, vol. 59, pp. 143-149.

Khanmohammadi, M, Karami, F, Mir-Marqués, A, Bagheri Garmarudi, A, Garrigues, S & de la Guardia, M 2014, 'Classification of persimmon fruit origin by near infrared spectrometry and least squares-support vector machines', *Journal of Food Engineering*, vol. 142, pp. 17-22.

Kleinbaum, DG & Mitchel, K 2002, 'Introduction to Logistic Regression', in *Logistic Regression*, Springer, New York, USA.

Ko, HL, Campbell, PR, Jobin-Decor, MP, Eccleston, KL, Graham, MW & Smith, MK 2006 'The introduction of transgenes to control blackheart in pineapple (Annanas comosus L.) cv. Smooth Cayenne by microprojectile bombardment ', *Euphytica*, vol. 150 pp. 387-395.

Ladaniya, MS 2008a, 'Fruit quality control, evaluation, and analysis', in MS Ladaniya (ed.), *Citrus Fruit*, Academic Press, San Diego, USA.

Ladaniya, MS 2008b, 'Preharvest factors affecting fruit quality and postharvest life', in MS Ladaniya (ed.), *Citrus Fruit*, Academic Press, San Diego, USA.

Ladaniya, MS 2008c, 'Physiological disorders and their management', in MS Ladaniya (ed.), *Citrus Fruit*, Academic Press, San Diego, USA.

Lammertyn, J, Aerts, M, Verlinden, BE, Schotsmans, W & Nicolaï, BM 2000, 'Logistic regression analysis of factors influencing core breakdown in 'Conference' pears', *Postharvest Biology and Technology*, vol. 20, pp. 25-37.

Lammertyn, J, Scheerlinck, N, Jancsók, P, Verlinden, BE & Nicolaï, BM 2003, 'A respiration–diffusion model for 'Conference' pears: II. Simulations and relation to core breakdown', *Postharvest Biology and Technology*, vol. 30, pp. 43-55.

Laofor, C & Peansupap, V 2012, 'Defect detection and quantification system to support subjective visual quality inspection via a digital image processing: A tiling work case study', *Automation in Construction*, vol. 24, pp. 160-174.

Lattaznzio, V 2003, 'Bioactive polyphenols: Their role in quality and storability of fruit and vegetables ', *Journal of Applied Botany*, vol. 77, pp. 128-146.

Lau, OL 1998, 'Effect of growing eason, harvest maturity, waxing, low $O_2$ and elevated $CO_2$ on flesh browning disorders in Braeburn apples', *Postharvest Biology and Technology*, vol. 14 pp. 131-141.

Lee, WH, Kim, MS, Lee, H, Delwiche, SR, Bae, H, Kim, DY & Cho, BK 2014, 'Hyperspectral near-infrared imaging for the detection of physical damages of pear', *Journal of Food Engineering*, vol. 130, pp. 1-7.

Li, G 2011, 'Nondestructive measurement model of apple internal browning based on FT-NIR spectroscopy', *Advanced Material Research* vol. 304, pp. 316-321.

Likert, R 1933, 'A technique for the measurement of attitudes ', *Archives of Psychology*, vol. 22, pp. 5-55.

Lin, H & Ying, Y 2009, 'Theory and application of near infrared spectroscopy in assessment of fruit quality : a review', *Food Quality*, vol. 3, pp. 130-141.

Lorente, D, Escandell-Montero, P, Cubero, S, Gómez-Sanchis, J & Blasco, J 2015, 'Visible–NIR reflectance spectroscopy and manifold learning methods applied to the detection of fungal infections on citrus fruit', *Journal of Food Engineering*, vol. 163, pp. 17-24.

Loutfi, A, Coradeschi, S, Mani, GK, Shankar, P & Rayappan, JBB 2015, 'Electronic noses for food quality: A review', *Journal of Food Engineering*, vol. 144, pp. 103-111.

Lu, R & Ariana, DP 2013, 'Detection of fruit fly infestation in pickling cucumbers using a hyperspectral reflectance/transmittance imaging system', *Postharvest Biology and Technology*, vol. 81, pp. 44-50.

Lu, X, Sun, D, Li, Y, Shi, W & Sun, G 2011, 'Pre- and post-harvest salicylic acid treatments alleviate internal browning and maintain quality of winter pineapple fruit', *Scientia Horticulturae*, vol. 130, pp. 97-101.

Lumpkin, C, Fellman, JK, Rudell, DR & Mattheis, JP 2015, ''Fuji' apple (*Malus domestica Borkh*.) volatile production during high $pCO_2$ controlled atmosphere storage', *Postharvest Biology and Technology*, vol. 100, pp. 234-243.

Luo, X, Takahashi, T, Kyo, K & Zhang, S 2012, 'Wavelength selection in vis/NIR spectra for detection of bruises on apples by ROC analysis', *Journal of Food Engineering*, vol. 109, pp. 457-466.

Ma, Y, Lu, X, Nock, JF & Watkins, CB 2015, 'Peroxidase and polyphenoloxidase activities in relation to flesh browning of stem-end and calyx-end tissues of 'Empire' apples during controlled atmosphere storage', *Postharvest Biology and Technology*, vol. 108, pp. 1-7.

Macheix, J-J, Fleuriet, A & Billot, J 1990, *Fruit Phenolics* CRC Press Boca Raton, Florida, USA

Marrero, A & Kader, AA 2006, 'Optimal temperature and modified atmosphere for keeping quality of fresh-cut pineapples', *Postharvest Biology and Technology*, vol. 39, pp. 163-168.

Martens, H & Naes, T 1989, *Multivariate calibration* John Wiley & Sons Ltd Chichester, United Kingdom.

Mason, SJ & Graham, NE 2002 'Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: statistical significance and interpretation', *Quarterly Journal of the Royal Meteorological Society* vol. 128, pp. 2145-2166.

Mattheis, J, Felicetti, D & Rudell, DR 2013, 'Pithy brown core in 'd'Anjou' pear (*Pyrus communis* L.) fruit developing during controlled atmosphere storage at $pO_2$ determined by monitoring chlorophyll fluorescence', *Postharvest Biology and Technology*, vol. 86, pp. 259-264.

Mattheis, JP & Rudell, D 2011, 'Responses of 'd'Anjou' pear *(Pyrus communis* L.) fruit to storage at low oxygen setpoints determined by monitoring fruit chlorophyll fluorescence', *Postharvest Biology and Technology*, vol. 60, pp. 125-129.

McGlone, VA, Jordan, RB, Seelye, R & Clark, CJ 2003, 'Dry-matter-a better predictor of the post-storage soluble solids in apples?', *Postharvest Biology and Technology*, vol. 28, pp. 431-435.

McGlone, VA, Martinsen, PJ, Clark, CJ & Jordan, RB 2005, 'On-line detection of Brownheart in Braeburn apples using near infrared transmission measurements', *Postharvest Biology and Technology*, vol. 37, pp. 142-151.

Mehta, VB, Haldankar, PM, Burondkar, MM, Jadhav, BB, Shinde, VV, Khaandekar, RG & Kadam, SG 2013, 'Use of a soft X-ray imaging system for on-line detection of spongy tissue in 'Alphonso' mango fruit', *Acta Horticulturae*, vol. 992 pp. 575-578.

Mendoza, F, Lu, R & Cen, H 2014, 'Grading of apples based on firmness and soluble solids content using Vis/SWNIR spectroscopy and spectral scattering techniques', *Journal of Food Engineering*, vol. 125, pp. 59-68.

Meurant, N, Holmes, R, MacLeod, N, Fullelove, G, I., B & Kernot, I 1999, *Mango Information Kit. Agrilink, your growing guide to better farming guide. Manual*, Department of Primary Industries, Queensland Horticulture Institute, Brisbane, Queensland, Australia.

Miqueloto, A, Amarante, CVTd, Steffens, CA, dos Santos, A & Mitcham, E 2014, 'Relationship between xylem functionality, calcium content and the incidence of bitter pit in apple fruit', *Scientia Horticulturae*, vol. 165, pp. 319-323.

Mizrach, A 2000, 'Determination of avocado and mango fruit properties by ultrasonic technique', *Ultrasonics*, vol. 38, pp. 717-722.

Mizrach, A 2008, 'Ultrasonic technology for quality evaluation of fresh fruit and vegetables in pre- and postharvest processes', *Postharvest Biology and Technology*, vol. 48, pp. 315-330.

Mizrach, A & Flitsanov, U 1999, 'Nondestructive ultrasonic determination of avocado softening process', *Journal of Food Engineering*, vol. 40, pp. 139-144.

Moggia, C, Pereira, M, Yuri, JA, Torres, CA, Hernández, O, Icaza, MG & Lobos, GA 2015, 'Preharvest factors that affect the development of internal browning in apples cv. Cripp's Pink: Six-years compiled data', *Postharvest Biology and Technology*, vol. 101, pp. 49-57.

Moscetti, R, Haff, RP, Saranwong, S, Monarca, D, Cecchini, M & Massantini, R 2014, 'Nondestructive detection of insect infested chestnuts based on NIR spectroscopy', *Postharvest Biology and Technology*, vol. 87, pp. 88-94.

Mowat, A & Collins, R 2000, 'Consumer behaviour and fruit quality: supply chain management in an emerging industry ', *Supply Chain Management : An International Journal* vol. 5, pp. 45-54.

Mucherino, A, Papajorgji, P & Pardalos, P 2009, 'k-Nearest Neighbor Classification', in *Data Mining in Agriculture*, vol. 34, Springer New York,

Munshi, SK, Singh, R, Vij, VK & Jawanda, JS 1978, 'Mineral composition of leaves in relation to degree of granulation in sweet orange', *Scientia Horticulturae*, vol. 9, pp. 357-367.

Muramatsu, N, Sakurai, N, Yamamoto, R, Nevins, DJ, Takahara, T & Ogata, T 1997, 'Comparison of non-destructive acoustic method with an intrusive method for firmness measurement of kiwifruit. ', *Postharvest Biology and Technology*, vol. 12,

Neuwald, DA, McCormick, RJ, Kittemann, D & Streif, J 2007, 'Chlorophyll fluorescence behaviour of 'Braeburn' apples under different $CO_2$ stress treatments at-harvest', paper presented at the Novel approaches for the control of postharvest diseases and disorders, Bologna, Italy.

Nicolaï, BM, Beullens, K, Bobelyn, E, Peirs, A, Saeys, W, Theron, KI & Lammertyn, J 2007, 'Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review', *Postharvest Biology and Technology*, vol. 46, pp. 99-118.

Nicolaï, BM, Bulens, I, Baerdemaeker, JD, Ketelaere, BD, Hertog, MLATM, Verboven, P & Lammertyn, J 2009, 'Non-destructive Evaluation: Detection of External and Internal Attributes Frequently Associated with Quality and Damage', in EbJFLSBE Prussia (ed.), *Postharvest Handling (2nd edn.)*, Academic Press, San Diego.

Nicolas, JJ, Richard-Forget, FC, Goupy, PM, Amiot, M & Aubert, SY 1994 'Enzymatic browning reactions in apple and apple products ', *Critical Reviews in Food Science and Nutrition*, vol. 34, pp. 109-157.

Nock, JF & Watkins, CB 2013, 'Repeated treatment of apple fruit with 1-methylcyclopropene (1-MCP) prior to controlled atmosphere storage', *Postharvest Biology and Technology*, vol. 79, pp. 73-79.

Olsen, K, Schomer, HA & Birth, GS 1962, 'Detection and evaluation of watercore in apples by light transmittance'.

Ooms, D, Palm, R, Leemans, V & Destain, MF 2010, 'A sorting optimization curve with quality and yield requirements', *Pattern Recognition Letters*, vol. 31, pp. 983-990.

Pankasemsuk, T, Garner, JO, Matta, F, B. & Silva, JL 1996, 'Translucent flesh disorder of Mangosteen fruit (*Garcinia mangostana* L.)', *HortScience*, vol. 31, pp. 112-113.

Pasquariello, MS, Rega, P, Migliozzi, T, Capuano, LR, Scortichini, M & Petriccione, M 2013, 'Effect of cold storage and shelf life on physiological and quality traits of early ripening pear cultivars', *Scientia Horticulturae*, vol. 162, pp. 341-350.

Pasquini, C 2003, 'Near infrared spectroscopy: fundamentals, practical aspects and analytical applications ', *Journal of the Brazilian Chemical Society* vol. 14, pp. 198-219.

Pathaveerat, S, Terdwongworakul, A & Phaungsombut, A 2008, 'Multivariate data analysis for classification of pineapple maturity', *Journal of Food Engineering*, vol. 89, pp. 112-118.

Paull, RE & Reyes, MEQ 1996, 'Preharvest weather conditions and pineapple fruit translucency', *Scientia Horticulturae*, vol. 66, pp. 59-67.

Pedreschi, R, Hertog, M, Robben, J, Noben, JP & Nicolaï, BM 2008, 'Physiological implications of controlled atmosphere storage of ‘Conference’ pears (*Pyrus communis* L.): a proteomic approach', *Postharvest Biology and Technology*, vol. 50, pp. 110-116.

Peiris, KHS, Dull, GG, Leffler, RG, Burns, JK, Thai, CN & S.J., K 1998, 'Non destructive detection of section drying, an internal disorder in Tangerine. ', *HortScience* vol. 33, pp. 310-312.

Pesis, E, Ackerman, M, Ben-Arie, R, Feygenberg, O, Feng, X, Apelbaum, A, Goren, R & Prusky, D 2002, 'Ethylene involvement in chilling injury symptoms of avocado during cold storage', *Postharvest Biology and Technology*, vol. 24, pp. 171-181.

Pongprasert, N, Sekozawa, Y, Sugaya, S & Gemma, H 2011, 'A novel postharvest UV-C treatment to reduce chilling injury (membrane damage, browning and chlorophyll degradation) in banana peel', *Scientia Horticulturae*, vol. 130, pp. 73-77.

Prange, RK, DeLong, J, M. , Leyte, J, C. & Harrison, P, A. 2002, 'Oxygen concentration affects chlorophyll fluorescence in chlorophyll-containing fruit ', *Postharvest Biology and Technology*, vol. 24 pp. 201-205.

Pusittigul, I, Kondo, S & Siriphanich, J 2012, 'Internal browning of pineapple (Ananas comosus L.) fruit and endogenous concentrations of abscisic acid and gibberellins during low temperature storage', *Scientia Horticulturae*, vol. 146, pp. 45-51.

Rajkumar, P, Wang, N, Eimasry, G, Raghavan, GSV & Gariepy, Y 2012, 'Studies on banana fruit quality and maturity stages using hyperspectral imaging', *Journal of Food Engineering*, vol. 108, pp. 194-200.

Ramos- Garcia, FJ, Valero, C, Homer, I, Ortiz-Canavate, J & Ruiz-Altisent, M 2005 'Non-destructive fruit firmness sensors: a review ', *Spanish Journal of Agricultural Research*, vol. 3 pp. 61-73

Raymond, L, Schaffer, B, Brecht, JK & Crane, JH 1998, 'Internal breakdown in mango fruit: symptomology and histology of jelly seed, soft nose and stem-end cavity', *Postharvest Biology and Technology*, vol. 13, pp. 59-70.

Rinnan, A, Berg, Fvd & Engelsen, SB 2009, 'Review of the most common pre-processing techniques for near-infrared spectra', *TrAC Trends in Analytical Chemistry*, vol. 28 pp. 1201-1222.

Ruiz-Altisent, M, Ruiz-Garcia, L, Moreda, GP, Lu, R, Hernandez-Sanchez, N, Correa, EC, Diezma, B, Nicolaï, B & García-Ramos, J 2010, 'Sensors for product characterization and quality of specialty crops—A review', *Computers and Electronics in Agriculture*, vol. 74, pp. 176-194.

Santos-Pereira, CM & Pires, AM 2005, 'On optimal reject rules and ROC curves', *Pattern Recognition Letters*, vol. 26, pp. 943-952.

Saquet, AA & Streif, J 2002, 'Chlorophyll fluorescence as a predictive method for detection browning disorders in 'conference' pears and 'jonagold' apples during controlled atmosphere storage ', *Scientia Rural Santa Maria* vol. 32 pp. 571-576.

Schein, AI & Ungar, L, H. 2007 'Active learning for logistic regression: an evaluation ', *Machine Learning* vol. 68 pp. 235-265

Selvarajah, S, Bauchot, AD & John, P 2001, 'Internal browning in cold-stored pineapples is suppressed by a postharvest application of 1-methylcyclopropene', *Postharvest Biology and Technology*, vol. 23, pp. 167-170.

Sharma, RR, Singh, R & Saxena, SK 2006, 'Characteristics of citrus fruits in relation to granulation', *Scientia Horticulturae*, vol. 111, pp. 91-96.

Shivashankar, S 2014, *Physiological disorders in mango fruit* vol. 42, John Wiley and Sons Inc., New Jersey, USA.

Shivashankara, KS & Mathai, CK 1999, 'Relationship of leaf and fruit transpiration rates to the incidence of spongy tissue disorder in two mango (*Mangifera indica* L.) cultivars', *Scientia Horticulturae*, vol. 82, pp. 317-323.

Siedliska, A, Baranowski, P & Mazurek, W 2014, 'Classification models of bruise and cultivar detection on the basis of hyperspectral imaging data', *Computers and Electronics in Agriculture*, vol. 106, pp. 66-74.

Singh, R & Singh, R 1981a, 'Effect of GA3, Planofix (NAA) and Ethrel on granulation and fruit quality in ‘Kaula’ mandarin', *Scientia Horticulturae*, vol. 14, pp. 315-321.

Singh, R & Singh, R 1981b, 'Effect of nutrient sprays on granulation and fruit quality of ‘Dancy tangerine’ mandarin', *Scientia Horticulturae*, vol. 14, pp. 235-244.

Soares, AG, Trugo, LC, Botrel, N & da Silva Souza, LF 2005, 'Reduction of internal browning of pineapple fruit (*Ananas comusus* L.) by preharvest soil application of potassium', *Postharvest Biology and Technology*, vol. 35, pp. 201-207.

Sonego, L, Ben-Arie, R, Raynal, J & J.C., P 1995, 'Biochemical and physical evaluation of textural characteristics of nectarines exhibiting woolly breakdown: NMR imaging, X-ray computer tomography and pectin composition ', *Postharvest Biology and Technology*, vol. 5, pp. 187-198.

Song, Y, Yu- Xin, Y, Zhai, H, Yuan-peng, D.U. , Feng, C & Shu-wei, W 2007, 'Polyphenolic compound and the degree of browning in processing apple varieties ', *Agricultural Sciences in China* vol. 6, pp. 607-612.

Srichamnong, W & Srzednicki, G 2015, 'Internal discoloration of various varieties of Macadamia nuts as influenced by enzymatic browning and Maillard reaction', *Scientia Horticulturae*, vol. 192, pp. 180-186.

Subedi, PP 2007, Non invasive assessment of fruit: atrributes other than sweetness, Research thesis, Doctor of Philosophy thesis, Central Queensland University, Queensland, Australia.

Subedi, PP & Walsh, KB 2009, 'Non-invasive techniques for measurement of fresh fruit firmness', *Postharvest Biology and Technology*, vol. 51, pp. 297-304.

Subedi, PP & Walsh, KB 2011, 'Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy', *Postharvest Biology and Technology*, vol. 62, pp. 238-245.

Subedi, PP, Walsh, KB & Owens, G 2007, 'Prediction of mango eating quality at harvest using short-wave near infrared spectrometry', *Postharvest Biology and Technology*, vol. 43, pp. 326-334.

Subramanian, J & Simon, R 2013, 'Overfitting in prediction models – Is it a problem only in high dimensions?', *Contemporary Clinical Trials*, vol. 36, pp. 636-641.

Sukwanit, S & Teerachaichayut, S 2013, 'Nondestructive prediction of internal browning in pineapple using transmittance short wavelength near infrared spectroscopy', *Acta Horticulturae*, vol. 989, pp. 395-400.

Teerachaichayut, S, Terdwongworakul, A, Thanapase, W & Kiji, K 2011, 'Non-destructive prediction of hardening pericarp disorder in intact mangosteen by near

infrared transmittance spectroscopy', *Journal of Food Engineering*, vol. 106, pp. 206-211.

Teerachaichayut, S, Kil, KY, Terdwongworakul, A, Thanapase, W & Nakanishi, Y 2007, 'Non-destructive prediction of translucent flesh disorder in intact mangosteen by short wavelength near infrared spectroscopy', *Postharvest Biology and Technology*, vol. 43, pp. 202-206.

Terdwongworakul, A, Nakawajana, N, Teerachaichayut, S & Janhiran, A 2012, 'Determination of translucent content in mangosteen by means of near infrared transmittance', *Journal of Food Engineering*, vol. 109, pp. 114-119.

Thomas, P, Kannan, A, Degwekar, VH & Ramamurthy, MS 1995, 'Non-destructive detection of seed weevil-infested mango fruits by X-ray imaging', *Postharvest Biology and Technology*, vol. 5, pp. 161-165.

Torri, L, Sinelli, N & Limbo, S 2010, 'Shelf life evaluation of fresh-cut pineapple by using an electronic nose ', *Postharvest Biology and Technology*, vol. 56, pp. 239-245.

Upchurch, BL, Throop, JA & Aneshansley, DJ 1997, 'Detecting internal breakdown in apples using interactance measurements', *Postharvest Biology and Technology*, vol. 10, pp. 15-19.

Vagin, MY & Winquist, F 2015, 'Electronic noses and tongues in food safety assurance', in AK Bhunia, MS Kim & CR Taitt (eds), *High Throughput Screening for Food Safety Assessment*, Woodhead Publishing, Cambridge, United Kingdom.

Vandendriessche, T, Schäfer, H, Verlinden, BE, Humpfer, E, Hertog, MLATM & Nicolaï, BM 2013, 'High-throughput NMR based metabolic profiling of Braeburn apple in relation to internal browning', *Postharvest Biology and Technology*, vol. 80, pp. 18-24.

Vanoli, M, Rizzolo, A, Grassi, M, Spinelli, L, Verlinden, BE & Torricelli, A 2014, 'Studies on classification models to discriminate ‘Braeburn’ apples affected by internal browning using the optical properties measured by time-resolved reflectance spectroscopy', *Postharvest Biology and Technology*, vol. 91, pp. 112-121.

Varith, J, Hyde, GM, Baritelle, AL, Fellman, JK & Sattabongkot, T 2003, 'Non-contact bruise detection in apples by thermal imaging', *Innovative Food Science & Emerging Technologies*, vol. 4, pp. 211-218.

Verghese, A, Nagaraju, DK, Kamala Jayanthi, PD & Madhura, HS 2005, 'Association of mango stone weevil, *Sternochetus mangiferae* (Fabricius) (Coleoptera: Curculionidae) with fruit drop in mango', *Crop Protection*, vol. 24, pp. 479-481.

Verlinden, BE, de Jager, A, Lammertyn, J, Schotsmans, W & Nicolai, BM 2002, 'PH—Postharvest Technology: Effect of Harvest and delaying Controlled Atmosphere Storage Conditions on Core Breakdown Incidence in 'Conference' Pears', *Biosystems Engineering*, vol. 83, pp. 339-347.

Wall, MM 2013, 'Improving the quality and safety of macadamia nuts', in LJ Harris (ed.), *Improving the Safety and Quality of Nuts*, Woodhead Publishing,

Walsh, KB 2014, 'Postharvest regulation and quality standards on fresh produce', in W Florkowski, J., RL Shewfelt , B Brueckner & SE Prussia (eds), *Postharvest Handling (3rd edn.)*, Academic Press, San Diego, USA.

Wang, J, Nakano, K & Ohashi, S 2011, 'Nondestructive detection of internal insect infestation in jujubes using visible and near-infrared spectroscopy', *Postharvest Biology and Technology*, vol. 59, pp. 272-279.

Wang, J, Nakano, K, Ohashi, S, Takizawa, K & He, JG 2010, 'Comparison of different modes of visible and near-infrared spectroscopy for detecting internal insect infestation in jujubes', *Journal of Food Engineering*, vol. 101, pp. 78-84.

Wang, SY, Wang, PC & Faust, M 1988, 'Non destructive detection of watercore in apple with nuclear magnetic resonance imaging', *Scientia Horticulturae*, vol. 35, pp. 227-234.

Wang, XY, Wang, P, Qi, YP, Zhou, CP, Yang, LT, Liao, XY, Wang, LQ, Zhu, DH & Chen, LS 2014, 'Effects of granulation on organic acid metabolism and its relation to mineral elements in *Citrus grandis* juice sacs', *Food Chemistry*, vol. 145, pp. 984-990.

Wang, Y & Sugar, D 2013, 'Internal browning disorder and fruit quality in modified atmosphere packaged 'Bartlett' pears during storage and transit', *Postharvest Biology and Technology*, vol. 83, pp. 72-82.

Wanitchang, J, Terdwongworakul, A, Wanitchang, P & Noypitak, S 2010, 'Maturity sorting index of dragon fruit: *Hylocereus polyrhizus*', *Journal of Food Engineering*, vol. 100 pp. 409-416.

Wongs-Aree, C & Noichinda, S 2014, 'Postharvest Physiology and Quality Maintenance of Tropical Fruits', in W Florkowski, J., RL Shewfelt , B Brueckner & SE Prussia (eds), *Postharvest Handling (3rd edn.)*, Academic Press, San Diego, USA.

Woolf, AB, Requejo-Tapia, C, Cox, KA, Jackman, RC, Gunson, A, Arpaia, ML & White, A 2005, '1-MCP reduces physiological storage disorders of 'Hass' avocados', *Postharvest Biology and Technology*, vol. 35, pp. 43-60.

Woolworths 2015, *Produce specifications*, Woolworths Australia, viewed 14 March 2015 http://www.wowlink.com.au,

Wu, D & Sun, DW 2013, 'Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review', *Innovative Food Science & Emerging Technologies*, vol. 19, pp. 15-28.

Xing, J & Guyer, D 2008a, 'Detecting internal insect infestation in tart cherry using transmittance spectroscopy', *Postharvest Biology and Technology*, vol. 49, pp. 411-416.

Xing, J & Guyer, D 2008b, 'Comparison of transmittance and reflectance to detect insect infestation in Montmorency tart cherry', *Computers and Electronics in Agriculture*, vol. 64, pp. 194-201.

Yamada, H & Kobayashi, S 1999, 'Relationship between watercore and maturity or sorbitol in apples affected by preharvest fruit temperature', *Scientia Horticulturae*, vol. 80, pp. 189-202.

Yan, S, Li, L, He, L, Liang, L & Li, X 2013, 'Maturity and cooling rate affects browning, polyphenol oxidase activity and gene expression of 'Yali' pears during storage', *Postharvest Biology and Technology*, vol. 85, pp. 39-44.

Yang, C, Lee, WS & Williamson, JG 2012, 'Classification of blueberry fruit and leaves based on spectral signatures', *Biosystems Engineering*, vol. 113, pp. 351-362.

Young, TW 1957, 'Soft nose, a physiological disorder in mango fruits ', *Florida Agricultural Experimental Station Journal series* vol. 662, pp. 280-283.

Zhang, B-Y, Samapundo, S, Pothakos, V, de Baenst, I, Sürengil, G, Noseda, B & Devlieghere, F 2013, 'Effect of atmospheres combining high oxygen and carbon

dioxide levels on microbial spoilage and sensory quality of fresh-cut pineapple', *Postharvest Biology and Technology*, vol. 86, pp. 73-84.

Zhang, B, Huang, W, Li, J, Zhao, C, Fan, S, Wu, J & Liu, C 2014, 'Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review', *Food Research International*, vol. 62, pp. 326-343.

Zhang, L 2012, Measurement of fruit quality using nuclear magnetic resonance, Doctor of Philosophy thesis, University California, Davis

Zhang, Z, Huber, DJ, Qu, H, Yun, Z, Wang, H, Huang, Z, Huang, H & Jiang, Y 2015, 'Enzymatic browning and antioxidant activities in harvested litchi fruit as influenced by apple polyphenols', *Food Chemistry*, vol. 171, pp. 191-199.

Zheng, Y, He, S, Yi, S, Zhou, Z, Mao, S, Zhao, X & Deng, L 2010, 'Predicting oleocellosis sensitivity in citrus using VNIR reflectance spectroscopy', *Scientia Horticulturae*, vol. 125, pp. 401-405.

Zhou, Y, Dahler, JM, Underhill, SJR & Wills, RBH 2003, 'Enzymes associated with blackheart development in pineapple fruit', *Food Chemistry*, vol. 80, pp. 565-572.

## Appendix 1.

## Instrumentation study of machines used in internal defects detection of fruit


## Abstract

This chapter documents aspects of the SWNIRS instrumentation developed and/or used in this thesis. The application of internal defect assessment of intact fruit demands use of high light intensities and/or sensitive detectors. This instrumentation was based on high wattage tungsten halogen lamps and Zeiss MMS1 or Avantes array detectors, or LEDs with a single photodiode detector.



**Keywords:** light emitting diodes, count, light intensity, spectra

## Introduction

The detection of internal defects in fruit requires the measurement of light transmitted through whole fruit. This can be accomplished using high intensities of incident light or through use of sensitive detectors. Long integration times are not an option for the assessment of moving fruit on a packline, for which consideration of timing and position of fruit relative to spectra acquisition is critical. The use of increased illumination levels from a broad output source such as a tungsten halogen lamp is fraught in terms of the incidental heat load. This issue is not trivial – plastic conveyor belts will melt or burn if not moving! Obviously there is also potential for heat damage to fruit if the exposure time increases beyond a few tens of milliseconds. The use of narrow band sources is a solution to this issue, coupled with use of a large area Si photodiode detector.

In this thesis, four instruments have been used: (i) Nirvana, (ii) IDD0; (iii) InSight2; (iv) IDD2. The purpose of this Appendix is to describe these units.

## Nirvana

The Nirvana unit uses the shadow probe optical configuration described by Greensill and Walsh (2000). This is not an appropriate optical geometry (i.e. not $180^{o}$ transmission) for assessment of internal defects; nevertheless the instrument was used for benchmarking purposes as a unit of known specifications (good repeatability etc).


| | |
|---|---|
| **Equipment** | Nirvana |
| **Manufacturer** | Integrated Spectronics, Sydney, Australia |
| **Wavelength Range (nm)** | 308-1150 |
| **Light source** | Single halogen lamp, 32W |
| **Detector** | Zeiss MMS1 NIR enhanced photodiode array |
| **Resolution** | FWHM approximately 10 nm |
| **Sensitivity** | $10^{13}$ Counts/Ws (with 14 bit conversion |
| **Pixel dimensions** | 25x2500 $\mu m^2$ |
| **Optical geometry** | Interactance (‘shadow probe’) |
| **Integration time (ms)** | Automatic variation to achieve >75% detector saturation |
| **Referencing** | Internal gold shutter; referencing on every sample |
| **Repeatability** | <1mA |

## Instrument repeatability

The repeatability of the instrument is measured as the SD of repeated measures in white tile.

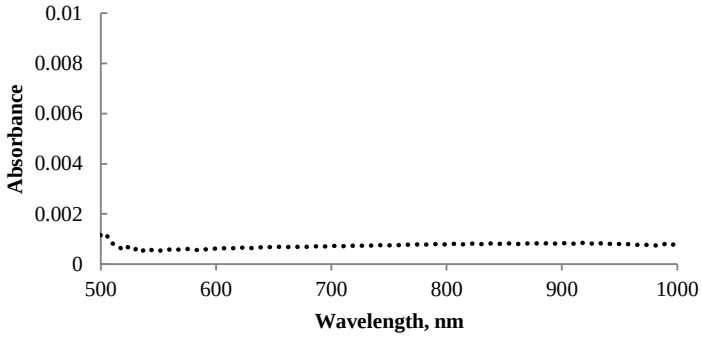



**Figure 6.1** Standard deviation of absorbance for repeated spectra acquired of a white tile, using the internal reference of the Nirvana for calculation of absorbance.

## IDD0

| | |
|---|---|
| **Equipment: IDD0** |  |
| **Manufacturer** | purpose built instrumentation, CQU, Australia (technical assistance of Larry Coleman is acknowledged) |
| **Wavelength Range (nm)** | 300-1100 |
| **Light source** | Single halogen lamp, 300W |
| **Detector** | Zeiss MMS1 NIR enhanced photodiode array |
| **Resolution** | FWHM approximately 10 nm |
| **Sensitivity** | $10^{13}$ counts/Ws for 14 bit conversion |
| **Pixel size** | 256 |
| **Optical geometry** | Partial transmittance $90^{o}$ or varying angle |
| **Integration time (ms)** | 4-2000 ms, varied depending on type of fruit to achieve count >67% of ADC saturation |

The IDD0 instrument was developed *in house* for the purpose of this thesis. It consists of a 300 W halogen lamp as a light source and supplied with MMS 1 spectrometer acquiring the spectra at 300-1100 nm. The instrument allows for work in a full transmission optical geometry with possibility of varying orientation of fruit and detector as desired. A sliding aperture of desired dimensions can be used to control the amount of light incident on the fruit. A light barrier minimizes any light received by the detector.

## Instrument repeatability

The repeatability of the IDD0 unit was assessed as the SD of repeated measures of a white tile.

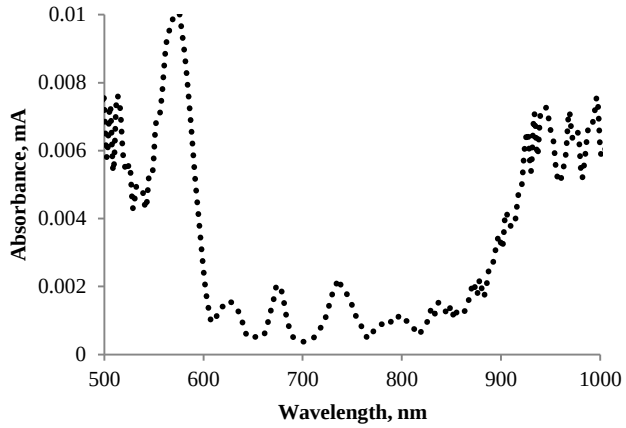



**Figure 2. S**tandard deviation of Absorbance for repeated spectra acquired of a white tile, using spectra of the same white tile for calculation of absorbance.

## Insight2

The MAF RODA Insight2 was designed to assess fruit DM, soluble solids and internal defects using a full transmission optical geometry. It is equipped with the 150 W halogen lamp installed in 180$^{o}$ to the fruit and an Avantes spectrometer.

| | |
|---|---|
| **Equipment** | **InSight2** |
| **Manufacturer** | MAF Agrobotics, France |
| **Wavelength Range** | 600-973 nm |
| **Light source** | Single halogen lamp, 150 W |
| **Detector** | Avantes Avaspec ULS backthinned CCD |
| **Resolution** | approx 10 nm (better resolution posible with narrower slit) |
| **Sensitivity** | 600,000 counts/Ws for 14 bit conversion |
| **Pixel step** | 1.5 nm |
| **Optical geometry** | transmittance 180$^{o}$ |
| **Integration time (ms)** | 4-10 ms, varied depending on type of fruit to achieve count >67% of ADC saturation |
| **Repeatability** | <1mA at 650-975 nm |

## Instrument repeatability

Standard deviation of repeated measurement of white tile reveals the repeatability 0.5 mA for this spectrometer.

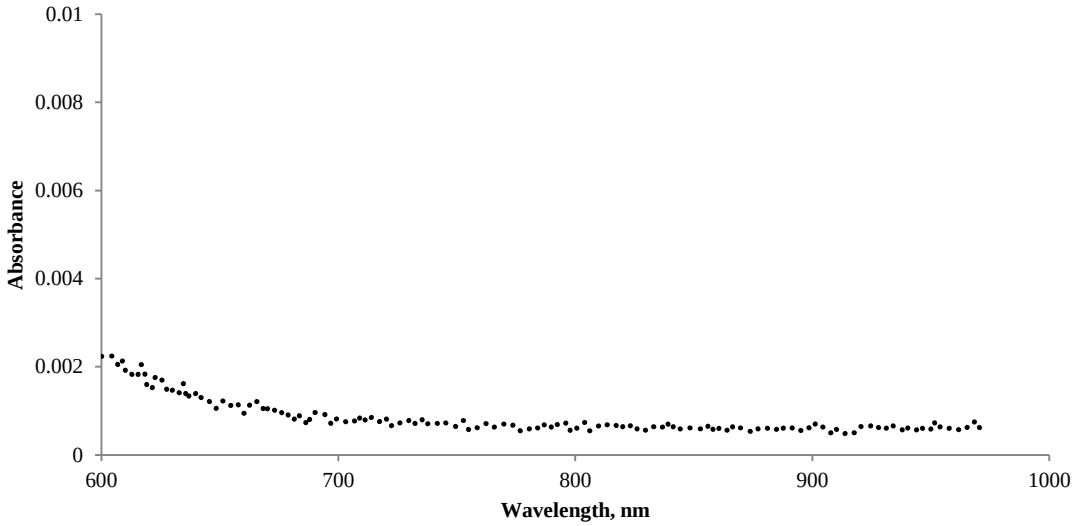



**Figure 2.** Standard deviation of Absorbance for repeated spectra acquired of a white ball, using spectra of the same white ball for calculation of absorbance.

## Lamp stability

Lamp output was stable in terms of spectral quality ad intensity over the period measure (0-76 sec) (Fig. 3).

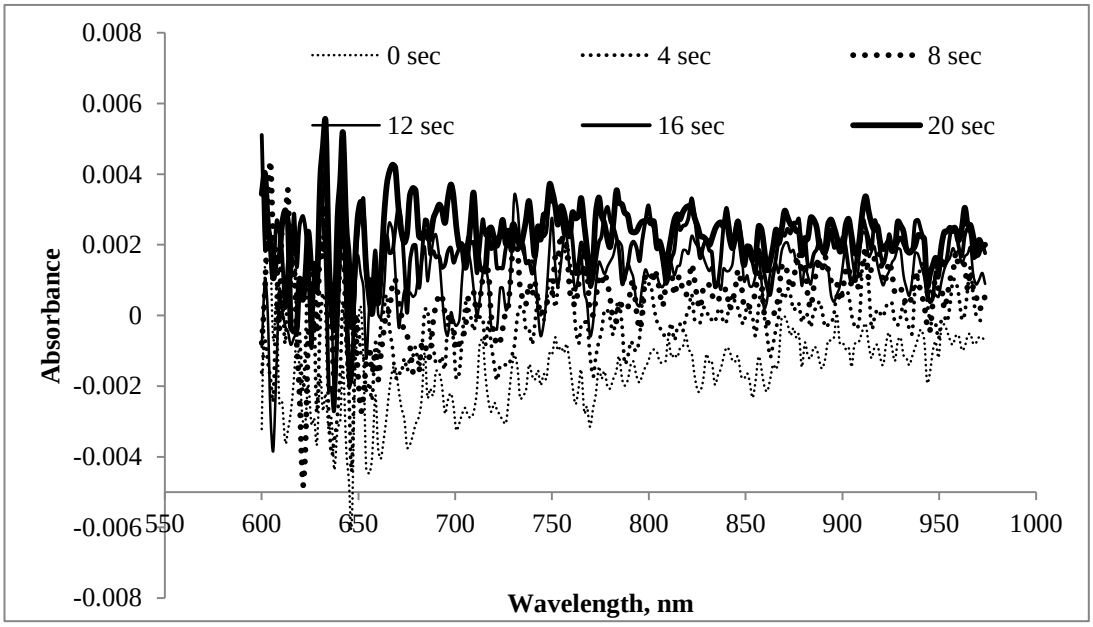



**Figure 3.** Spectra with white tile (dynamic) with PETE reference. Spectra acquired of 8 mm thick PFTE tile from time of lamp power-on.

## Effect of cup on prediction

On a grading machine fruit are conveyed in 'finger cups' or 'hands'. Each cups or fingers are supposed to behave equally in terms of performance. To test the impact of the hand on a spectral prediction of defect level in apple fruit, a population of 50 fruit (apple var. Pink Lady) was repetitively run on each of 6 different cups. The fruit size and weight ranged from 50-80 mm in diameter and 105 - 188 g. The conveyor belt was operated at a speed of 255 cups/min. Fruit were placed on cups with the stem - calyx axis perpendicular to detector optics. The cup did not impact PLSR model performance.

**Table 1.** Partial Least square regression results for defect score (5 point scale) for a population of 50 fruit (4 spectra from each fruit, transmittance 600-973 nm, mean and SD of score = 2.44 ± 1.51) for each of 6 cups.

| | Visual browning score (2.44 ± 1.51) | | |
|---|---|---|---|
| Cup # | $R^2$cv | RMSECV | PCs |
| Cup 1 | 0.83 | 0.63 | 6 |
| Cup 2 | 0.84 | 0.59 | 6 |
| Cup 5 | 0.85 | 0.58 | 6 |
| Cup 7 | 0.85 | 0.58 | 6 |
| Cup 9 | 0.83 | 0.63 | 6 |
| Cup 13 | 0.84 | 0.60 | 6 |

The calibration model for visual score on one cup has similar prediction performance for rest of the cups showing that every cup are behaving equally for the grading of fruit (Table 2).

**Table 2.** Prediction performance of the model developed for visual score by fruit placement in cup 1to other cups.

(Cup 1 score model ($R^2$cv = 0.83;  RMSECV= 0.63) )

| **Cup #** | **$R^2_p$** | **RMSEP** | **Bias** |
|---|---|---|---|
| Cup 1 | 0.84 | 0.6 | $-6.31^{-8}$ |
| Cup 2 | 0.8 | 0.93 | -0.24 |
| Cup 5 | 0.83 | 0.65 | -0.13 |
| Cup 7 | 0.8 | 0.67 | 0.07 |
| Cup 9 | 0.82 | 0.64 | 0.06 |
| Cup 13 | 0.82 | 0.63 | 0.07 |

## IDD2

LEDs offer long life, efficiency and less generation of heat compared to tungsten halogen lamps (e.g. 30000-50000 h compared to 1500-2000 h). LEDs have no warm up requirement and can be powered with short pulses to generate higher intensities of light output.

The IDD-2 system involves a horizontal $180^{o}$ lamp-fruit-detector axis, with lamp and detector mounted above the conveyor cups. The detector is a silicon photodiode equipped with a $1.5 \times 10^{6}$ amplifier. The lDD-2 light source is an array of 32 Epitec (JAPN) LEDs (8 each of four wavelengths, with peak outputs of 780, 880, 700 and 810 nm, designated A, B, C and D). The physical constraints of LED arrangement in a cluster will lead to some non-uniformity of illumination of the sample by the various wavelengths. This may be an issue with change in fruit size or for non-uniform distribution of internal disorder. The LEDs are rated for 2 A maximum current. They are operated using a 27 V power supply with 8 LEDs in series, i.e. operated at 3.3 V. The LEDs are operated in a pulsed (0.5 ms) mode.

The system is designed for an apple fruit of 56-95 mm diameter. Smaller fruit will result in increased illumination of the detector, to the point of detector saturation. Larger fruit will result in lower detector illumination, with lower signal to noise. Conveyor cups are set on 100 mm centres, and the conveyor position is monitored with an encoder giving 64 pulses between each cup (i.e. each pulse represents movement of 1.6 mm). At leading and trailing edges of the fruit, signal will be saturated. The detector signal for the middle three steps are averaged and measured.

## Technical specifications

| | |
|---|---|
| **Equipment** | IDD2 |
| **Manufacturer** | MAF RODA Agrobotics, France |
| **Wavelength Range (nm)** | 700, 810, 780, 880 nm absorption peak wavelengths |
| **Light source** | 4 Light Emitting Diodes (LEDs) |
| **Detector** | A single silicon photodiode array |
| **Optical geometry** | Full transmission, 180 $^{o}$ |
| **White ball repeatability** | SD or 20 repeated measures of B/A = 0.086 |

## Effect of varying current on LED output intensity

The IDD-2 unit offers control of the current to each set of LEDs. To characterise this control, current level to the individual LEDs was varied and photodiode output recorded (Fig. 4). LED intensity was directly related to current for three of the four LEDs for current levels over the range 40 – 240 units (units of IDD control). The exception was the 700 nm LED which increased output for current up to 100 units, then output decreased with increasing current. Presumably the maximum current rating of the 700 nm LED was achieved at a setting of 120.

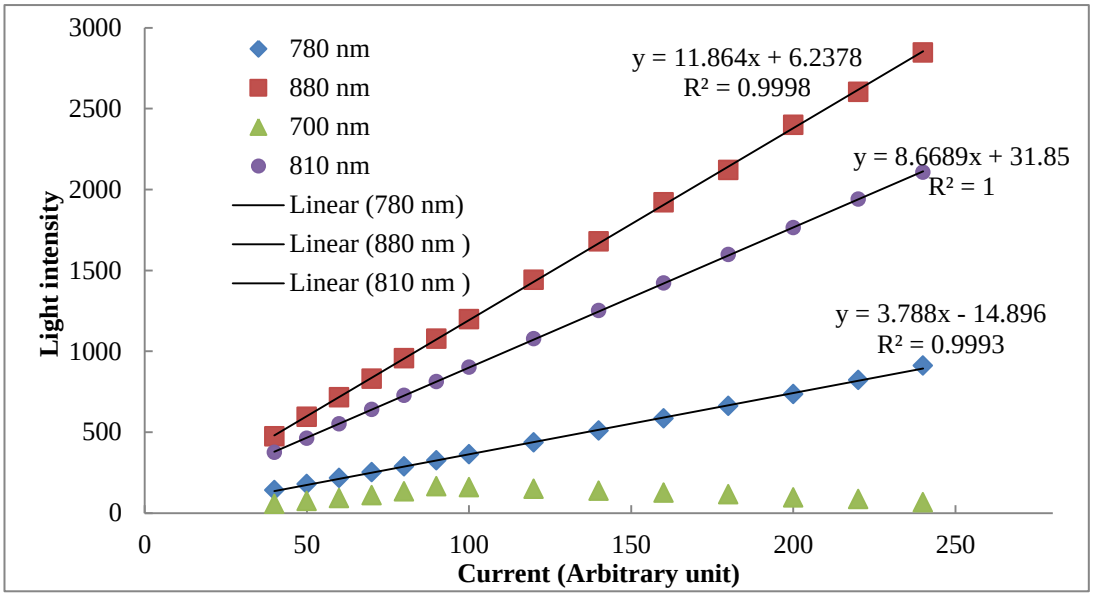



**Figure 4.** Linear relationship of the light intensity received at spectrometer against the increase in current flow in each LEDs for white ball at 700, 780, 810 and 880 nm.

The output intensities of the 4 LEDs were quite different, in the order of 880 > 810 > 780 > 700 nm.

## LED stability

LED output was monitored over time from power on, in terms of signal received from a white reference ball. A drift of approximately 1% was observed after 30 minutes as observed by small change in count readings. Peak wavelength and intensity of output of LEDS is temperature sensitive. The working hypothesis is that observed shifts in intensity result from changes in ambient temperature.

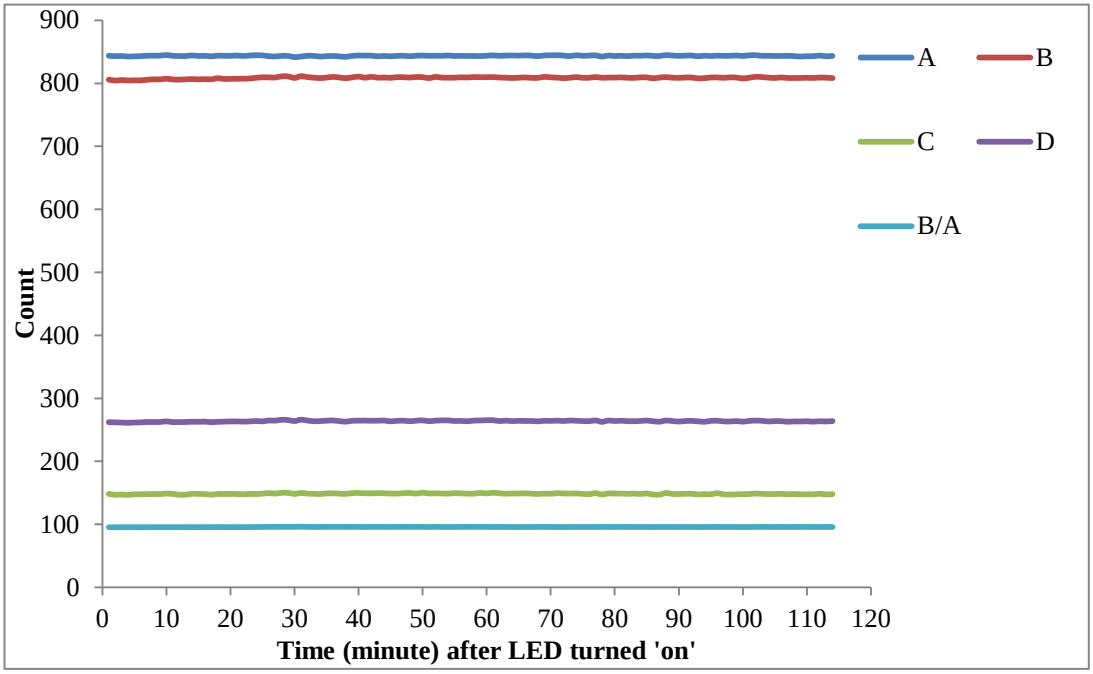



**Figure 5.** Output of four LED channels of IDD-2 unit from time of power-up of LEDs.

## Effect of cabinet temperature

With increase in cabinet temperature, a small decrease in signal was seen for the four channels (Fig. 6), however the ratio of B/A and D/C was stable.

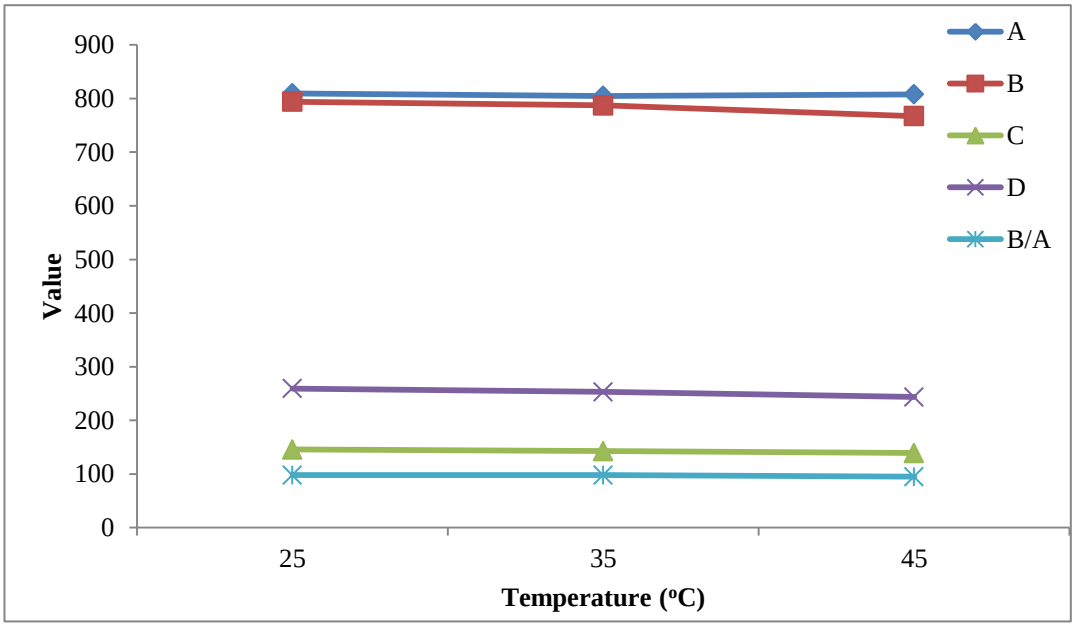



**Figure 6.** Influence of increase in cabinet temperature on luminous intensity with white ball at given peak wavelength.

## Effect of cup on prediction

In a conveyor based sorting system, effectively both the fruit (sample) and the transport cup are assessed by a given sensor. For weight grading, a tare system can be used to remove the weight of each cup from subsequent measurements. For spectroscopic measurements there is the possibility that differences between cups may affect measurement. However, there was no evidence for such an effect, in terms of the $R^2$ of B/A values for fruit run on different cups (Table 3), or the $R^2$ of B/A value and visual score level (Table 4).

**Table 3.** Correlation coefficient of determination ($R^2$) for B/A values of 50 fruit recorded both on cup 1 (following reloading of fruit) and other cups of the conveyor system.

| B/A value | Cup 1 | Cup 2 | Cup 3 | Cup 4 | Cup 5 | Cup 6 |
|---|---|---|---|---|---|---|
| Cup 1 | 0.99 | 0.94 | 0.94 | 0.93 | 0.95 | 0.95 |

**Table 4.** Correlation coefficient of determination ($R^2$) for B/A and visual defect score (50 fruit, average of two readings per fruit).

| B/A vs score | Cup 1 | Cup 2 | Cup 3 | Cup 4 | Cup 5 | Cup 6 |
|---|---|---|---|---|---|---|
| Score | 0.76 | 0.77 | 0.71 | 0.71 | 0.73 | 0.74 |

**Appendix 2**. Repeated measurement error of 1-5 or 1-10 visual score scale for assessment based on cut surface of apple. This involves total of 125 fruit of Cv. Pink Lady.

**1-5 visual score**

| Panel members | Panel Member 1 | Panel member 2 | Panel member 3 | Panel member 4 | panel member 5 | Panel member 6 | average | (PM1-Av)^2 | (PM2-Av)^2 | (PM3-Av)^2 | (PM4-Av)^2 | (PM5-Av)^2 | (PM6-Av)^2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scale 1-5 | | | | | | | | RMSE | RMSE | RMSE | RMSE | RMSE | RMSE |
| Mean | 3.22 | 2.43 | 2.92 | 2.61 | 3.07 | 3.08 | 2.89 | 0.58 | 0.69 | 0.48 | 0.51 | 0.52 | 0.45 |
| SD | 1.39 | 1.52 | 1.20 | 1.53 | 1.61 | 1.39 | 1.37 | | | | | | |

**1-10 scale**

| Panel members | Panel Member 1 | Panel member 2 | Panel member 3 | Panel member 4 | panel member 5 | Panel member 6 | average | (PM1-Av)^2 | (PM2-Av)^2 | (PM3-Av)^2 | (PM4-Av)^2 | (PM5-Av)^2 | (PM6-Av)^2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scale1-10 | | | | | | | | RMSE | RMSE | RMSE | RMSE | RMSE | RMSE |
| Mean | 4.73 | 4.78 | 5.21 | 4.90 | 4.97 | 4.65 | 4.87 | 0.71 | 0.72 | 1.14 | 0.47 | 0.76 | 0.95 |
| SD | 3.10 | 3.01 | 2.54 | 2.95 | 3.40 | 3.24 | 2.94 | | | | | | |

| Statistics of 1-5 scale repeated measurements by a single assessor | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 | Average | (Day 1-Av)^2 | (Day 2 -Av)^2 | (Day 3 - Av)^2 | (PM4-Av)^2 | (PM5-Av)^2 |
| scale 1-5 | mean | mean | mean | mean | mean | mean | RMSE | RMSE | RMSE | RMSE | RMSE |
| Mean | 2.81 | 2.90 | 2.79 | 2.82 | 2.92 | 2.85 | 0.28 | 0.27 | 0.29 | 0.27 | 0.30 |
| SD | 1.67 | 1.62 | 1.64 | 1.68 | 1.65 | 1.63 | | | | | |