University of Southern Queensland

DEVELOPMENT OF NEW MEASUREMENT METHODS TO DETERMINE SUGARCANE QUALITY FROM STALK SAMPLES

A Dissertation submitted by

Nazmi Mat Nawi

B.Eng (UPM, Malaysia)
M.Eng (USQ, Australia)

For the award of

Doctor of Philosophy

2014

Abstract

Recently, there has been a growing interest within the Australian sugarcane industry to measure sugarcane quality in the field to further improve product quality and value. However, conventional technologies for measuring sugarcane quality in a laboratory have limitations for uses in the field because they require sugarcane to be prepared as either juice or fibrated samples. In-field samples processing is very difficult and time-consuming, especially during harvest. Thus, the development of a rapid and efficient measurement technique which can be performed directly on stalk samples is highly desirable.

In this thesis, a new quality measurement method for fresh sugarcane stalk samples was developed using a visible and shortwave near infrared spectroradiometer (VNIRS) with the wavelength ranging from 350 to 1075 nm. A light-proof measurement box was developed and used as an instrument platform to evaluate the capability of the VNIRS to measure quality parameters of sugarcane samples. The box was used to determine quality parameters using two newly proposed scanning methods: the skin scanning method (SSM) and the cross sectional scanning method (CSSM). These methods were applied on both whole stalk and internode samples. No preparation mechanism was required prior to the quality measurement on stalk samples.

The selection of chemometrics methods used to optimise the regression models between spectral data and sugar content were also investigated. Partial least square (PLS) regression analysis with full cross validation (leave-one-out) technique was chosen to establish regression models between the spectral data and quality parameters. To improve the accuracy of the regression models, the spectral data was first pre-processed using the multiplicative scatter correction (MSC) method. Principal component analysis (PCA) was then used to extract useful information from the spectral data, decrease the noise and determine the optimum number of latent variables (LVs). The pre-processing methods, PLS and PCA exercises were run using Unscrambler V 9.6 software. The RPD (ratio of prediction to deviation) value was also used to evaluate the performance of the models.

For whole stalk samples, it was found that the R^2 for SSM and CSSM were 0.82 and 0.68, respectively. The calibration models for the fibrated, juice and whole stalk samples were developed using quality values obtained by standard industry procedures. For internode samples, the R^2 for SSM and CSSM were 0.91 and 0.87, respectively. The calibration models for internode samples were developed using °Brix values obtained from a handheld refractometer. The RPD values of the prediction models for

internode samples by both SSM and CSSM were 2, indicating that these newly proposed methods can be used for coarse quantitative prediction purposes.

The variation of sugar content (°Brix) along the length of the stalks and internode samples were also assessed. The understanding of these variations can provide a foundation toward the design and development of the quality measurement system in the field. In this study, sugar content was found to vary significantly between the first and last internodes, with their average °Brix values being 22.2 and 7.6, respectively. The variation of sugar content between node and internode areas was 7.6% (SSM method) and 8.7% (CSSM method), respectively.

To demonstrate the possible applications of the proposed methods on a harvester, a basic calculation and conceptual design for a proposed in-field quality measurement system was outlined using the VNIRS mounted on top of the elevator conveyor. The proposed system had the potential to sense billet samples based on SSM either by directly scanning the moving billets on the elevator or by scanning the billets supplied by a sampling mechanism using a vacuum system. This theoretical design has shown that it is technically possible to develop a quality measurement system on a sugarcane harvester. However, more work needs to be done before this proposed method can be successfully mounted on a harvester.

Overall, it is concluded that the accuracy of the new measurement methods based on stalk samples using portable and low-cost VNIRS developed in this thesis is adequate. The proposed methods have significant potential uses as a tool for measuring sugarcane quality parameters from stalk samples in the field.

Certification of Dissertation

I certify that the ideas, experimental works, results, analyses, software and conclusions reported in this dissertation are entirely my own effort, except where otherwise acknowledged. I also certify that this work is original and has not been previously submitted for any other award, except where otherwise acknowledged.

Candidate

_____ _____

Nazmi Mat Nawi Date

ENDORSEMENT

Principal supervisor

_____ _____

Dr Guangnan Chen Date

Associate supervisor

_____ _____

Dr Troy Jensen Date

Acknowledgements

First of all, I would like to express my deepest gratitude to the Ministry of Higher Education, Malaysia and Universiti Putra Malaysia (UPM) for providing me the scholarships to carry out this study. The financial supports from these organisations are very much appreciated.

Many thanks are also extended to both of my supervisors, Dr Guangnan Chen and Dr Troy Jensen for their continuous guidance, advices, comments and encouragement. Their efforts ensured that this study was completed in a reasonable time and is highly appreciated.

Special thanks are also due to staff members of National Centre of Engineering in Agriculture (NCEA), Associate Professor Craig Baillie, Associate Professor Nigel Hancock, Dr Cheryl McCarthy and Dr Saman Abdanan Mehdizadeh for their kindness, help and technical support throughout the journey of this thesis. I also appreciate the support of the staff of Bureau of Sugar Experimental Station (BSES) Bundaberg, especially Mr John Panitz for assistance in collecting sugarcane stalk samples in the field.

I also express my gratitude to the Sugar Research and Development Corporation (SRDC) for providing a funding to the project of CSE022 ‘A coordinated approach to Precision Agriculture RDE for the Australian Sugar Industry’. Without this financial assistance, my work would not have been possible.

Finally, my sincere gratitude is due to my wife Suhaila and children (Muaz, Marissa and Munif) for their patience and support. Without their continuous encouragement, this study might not be as it is today. Thanks are also due to all of my friends and those who are directly or indirectly related to this study.

List of Publications

The following articles have been published or submitted for publication from the research contained within this dissertation.

Refereed international journal articles

Nawi, N. M., Jensen, T. & Chen, G. (2012), 'The application of spectroscopic methods to predict sugarcane quality based on stalk cross-sectional scanning', *Journal of American Society of Sugar Cane Technologists* **32**, 16-27.

Nawi, N. M., Chen, G., Jensen, T. & Mehdizadeh, S. A. (2013), 'Prediction and classification of sugar content of sugarcane based on skin scanning using visible and shortwave near infrared', *Biosystems Engineering* **115**, 154-161.

Nawi, N. M., Chen, G. & Jensen, T. (2013), 'Visible and shortwave near infrared spectroscopy for predicting sugar content of sugarcane based on a cross sectional scanning method', *Journal of Near Infrared Spectroscopy* **21**(4), 289-297.

Nawi, N. M., Chen, G. & Jensen, T. (2013), 'Application of spectroscopic method to predict sugar content of sugarcane from internode samples', *Journal of Tropical Agriculture and Food Sciences* **41**(2), ***in press***.

Nawi, N. M., Chen, G. & Jensen, T. (2013), 'Quality measurement technologies and sampling strategies for precision agriculture in sugarcane industry: A review', accepted for publication in the journal of '*Precision Agriculture*' on 15 July 2013, subject to revisions.

Nawi, N. M., Chen, G. & Jensen, T. (2013), 'Determination of sugarcane quality from juice samples using visible and shortwave near infrared spectroscopy: Comparison between transmittance and transflectance measurement modes', under preparation and will be submitted to Journal of Mechanical Engineering and Sciences.

Refereed full-length conference articles

Nawi, N. M., Jensen, T., Chen, G. & Baillie, C. (2011), 'The application of spectroscopic methods to predict sugarcane quality based on stalk cross-sectional scanning', *SEAg 2011: Diverse Challenges, Innovative Solutions*, 29-30 Sep 2011, Gold Coast, Australia.

Nawi, N. M., Chen, G., Jensen, T. & Baillie, C. (2011), 'Transferring technology of measuring sugarcane quality from the laboratory to the field: what is possible?', *SEAg 2011: Diverse Challenges, Innovative Solutions*, 29-30 Sep 2011, Gold Coast, Australia.

Nawi, N. M., Jensen, T., Chen, G. & Baillie, C. (2012), 'Prediction of sugar content along the height of sugarcane internodes using Vis/NIR', *CIGR-AgEng 2012: Agriculture and Engineering for a healthier Life,* 8-12 July 2012, Valencia, Spain.

Nawi, N. M., Chen, G., Jensen, T. & Baillie, C. (2012), 'Potential technologies to measure sugarcane quality in the field', *CIGR-AgEng 2012: Agriculture and Engineering for a healthier Life*, 8-12 July 2012, Valencia, Spain.

Nawi, N. M., Chen, G. & Jensen, T. (2013), 'Application of visible and shortwave near infrared spectrometer to predict sugarcane quality from different sample forms', *Proceedings of SPIE 8881, Sensing Technologies for Biomaterial, Food, and Agriculture 2013*, 23-25 April 2013, Yokohama, Japan.

Nawi, N. M., Chen G. & Jensen, T. (2013), 'Visible and shortwave near infrared spectroscopy to predict sugarcane quality from clear and raw juice samples', *SEAg 2013: Innovative Agricultural Technologies for a Sustainable Future*, 22-25 Sep 2013, Perth, Australia.

Table of contents

Lists of Figures

Lists of Tables

Abbreviations

ACFA	Australian Cane Farmers' Association
ANN	Artificial neural network
AOTF	Acoustic optic tunable filter
ASM	Assisted (Pneumatic) scanning method
ASD	Analytical spectral device
BCH	Burnt cane harvesting
BSES	Bureau of Sugar Experimental Station
CA	Cluster analysis
CAS	Cane analysis system
CCD	Charge-coupled device
CCS	Commercial cane sugar
CJ	Clarified juice
CSSM	Cross sectional scanning method
CV	Coefficient of variation
DPLS	Discriminant partial least squares
DSM	Direct scanning method
EWs	Effective wavelengths
FOV	Field-of-view
FR	Full range
FRS	Full range spectroradiometer
FS	Fibrated sample
FWHM	Full width at half maximum
GCH	Green cane harvesting
GIS	Geographic information system
GLC	Gas-liquid chromatographic
GPS	Global positioning system
HPLC	High performance liquid chromatography
InGaAs	Indium Galium Arsenide
KNN	K-nearest neighbours
LDA	Linear discriminant analysis
LVs	Latent variables
MA3	Moving average with three segments
MA9	Moving average with nine segments
MIR	Mid infrared
MLR	Multiple linear regressions
NIR	Near infrared
MN	Mean normalization
MSC	Multiplicative scatter correction
PA	Precision agriculture
PbS	Lead sulfide
PCs	Principal components
PCA	Principal component analysis
PCR	Principal component regression

PDA	Photodiode array
PLS	Partial least square
QMSS	Quality measurement system for sugarcane
R^2	Coefficient of determination
RJ	Raw juice
RMSEC	Root mean square error of calibration
RMSEP	Root mean square error of prediction
RPD	Ratio of prediction to deviation/residual predictive deviation
SD	Standard deviation
SEC	Standard error of calibration
SEP	Standard error of prediction
SG1	Savitzky-Golay first derivative
SG2	Savitzky-Golay second derivative
Si	Silicon
SIMCA	Soft independent modelling of class analogy
SNV	Standard normal variate correction
SSM	Skin scanning method
SVM	Support vector machine
SWNIR	Shortwave near infrared
SW-NIR	Shortwave near infrared
TGM	Ternary growth model
UV	Ultraviolet
Vis	Visible
Vis-NIR	Visible and near infrared
VNIR	Visible and shortwave near infrared
VNIRS	Visible and shortwave near infrared spectroradiometer

Chapter 1

Introduction

1.1 Background

Sugarcane (*Saccahrum spp*.) is an important crop in Australia, with the economic value ranging between AUD$1.5 and 2.5 billion per year (Canegrowers 2011). The sugarcane industry is also one of Australia's largest rural industries with the annual sugar production of more than 5 Mt (Canegrowers 2011). In this industry both yield and quality components form the basis of the payment system to growers. The quality of sugarcane is determined based on its sugar content, known as commercial cane sugar (CCS). CCS is derived from °Brix (soluble solids content), pol (sucrose content) and fibre content.

Recently, substantial emphasis has been given to the use of precision agriculture (PA) technologies in the sugarcane industry as a means of increasing its productivity and quality performance. PA is a valuable management tool for maximising farm profits through the efficient application of crop inputs by matching them with variations in crop yield and quality in the field (Wendte et al. 2001). However, current PA technologies in the sugarcane industry can only monitor the yield but do not have the ability to measure the quality of the product (Bramley 2009). This is a serious limitation for the implementation of PA techniques in this industry because considerable quality variations exist across the paddocks (Lawes et al. 2003).

The ability to measure sugarcane quality parameters in the field would bring many benefits to the industry, especially to improve the current payment system to growers and eliminate possible consignment errors in current field data collection systems (Stafford 1999). In addition, in-field quality measurement could also help maximise the sugarcane value at harvest through the identification of sugarcane blocks that have the highest in-field CCS. The in-field CCS values could also be used to optimise harvest schedules during the crushing season (Staunton et al. 2011). In research programs, the ability to measure quality parameters in the field would bring benefits for clonal evaluation (Berding et al. 1991a and 1991b; Purcell et al. 2005).

There are several published studies regarding the in-field measurement of CCS (Johnson & Richard 2005; Bramley et al. 2012). However, these studies were all carried out with a manual sampling method and the quality parameters were later measured in a laboratory. Unfortunately, sampling the sugarcane stalks from representative sites, pre-treatment, transport to the laboratory, the analysis itself and communication of results back to the growers can be a costly, labour intensive and lengthy process (Abdel-Rahman et al. 2010).

The development of an in-field quality measurement system for the sugarcane industry is hindered due to the absence of a reliable sensor which can survive the harsh field environment and rugged harvesting conditions. The common technologies used to measure sugarcane quality in laboratories such as refractometric, polarimetric, chromatographic and near infrared (NIR) spectroscopic have great limitations for field use because they are often time-consuming, operator-dependent, and require hazardous reagents (Mehrotra & Siesler 2003). These methods also require that sugarcane samples be prepared as juice or fibrated samples. Unfortunately, preparing the sugarcane samples into these forms in the field is time consuming and technically difficult, especially during harvesting. Therefore, a rapid sensing technique which can be performed on the convenient forms of stalk samples is highly desirable.

NIR spectroscopic methods have long been used in agricultural industries because they are fast, simple, low-cost and non-destructive (Day & Fearn 1982). In recent years, several studies have demonstrated that a portable spectroscopy is able to provide a rapid, cost-effective and non-destructive measurement of product quality in the field (Temma et al. 2002; Montes et al. 2006). These studies found that the application of a visible and shortwave NIR (VNIR) spectrometer with the wavelength range from 400 to 1000 nm appear particularly promising because the bands are ascribed to the third and fourth overtones of O–H and C–H stretching modes and the instrument is low-cost and portable enough for in-field measurements (Walsh et al. 2000).

VNIR spectrometers are currently being applied to non-destructively measure the sugar content of crops including apples (Huang & Lu 2010), pineapples (Chia et al. 2012) and guavas (Hsieh & Lee 2005). Unfortunately, there are no published studies regarding the use of VNIR spectroscopy to measure sugarcane quality parameters from stalk samples. Thus, this thesis was undertaken to investigate the potential application of a portable and low-cost VNIR spectroscopic method to predict sugarcane quality parameters from stalk samples.

1.2 Research objectives

The goal of this research is to investigate the potential of using a portable and low-cost visible shortwave near infrared spectroradiometer (VNIRS) in predicting sugarcane quality parameters from non-juice sample forms. The hypothesis of the research is:

"A portable and low-cost visible-shortwave near infrared spectroscopy can be used to predict sugarcane quality from solid samples or other forms".

This hypothesis was related to the following specific objectives:

1. To investigate the feasibility of using the VNIRS to predict sugarcane quality parameters from conventional sample forms (raw juices, clarified juices and fibrated samples).

2. To investigate the feasibility of using the VNIRS to predict sugarcane quality parameters from stalk samples based on cross sectional scanning methods (CSSM).
3. To investigate the feasibility of using the VNIRS to predict sugarcane quality parameters from stalk samples based on skin scanning methods (SSM).
4. To compare the prediction accuracy between the VNIRS with the full range spectroradiometer (FRS).
5. To quantify the variation of sugar content and prediction accuracy along stalks and individual internodes so that the best scanning point on the stalk samples can be determined.
6. To determine the accuracy of assigning quality classes using an artificial neural network (ANN).
7. To develop appropriate design criteria and specification for sampling mechanisms for the installation of a sensor and measurement system on a harvester for use in the field.

1.3 Organization of the thesis

This thesis is organised into nine chapters. A brief discussion of every chapter is presented below.

Chapter 1 presents the research background driving this work. This chapter also defines the research objectives and hypothesis of this study.

Chapter 2 provides an overview of the significance of the sugarcane industry to the Australian economy. The agronomic characteristics and typical harvesting operation for sugarcane are described. The chapter also discusses the importance of an in-field measurement system for the sugarcane industry. Then, current quality measurement methods in the sugarcane industry and their potential applications and limitations for field use are reviewed. The new emerging technologies which have the potential to be applied in the sugarcane industry are also evaluated.

Chapter 3 discusses the theory of NIR spectroscopy, including available wavelength regions, different types of spectroscopic instrumentation and different sample presentations. A brief review of the application of NIR technologies in the agricultural industry is provided, including the potential applications of NIR spectroscopy to measure sugarcane quality in the field.

Chapter 4 describes the new methodologies developed to investigate the innovative application of the spectroradiometers in measuring sugarcane quality from different sugarcane sample forms. Two applications of the new measurement methods namely SSM and CSSM are discussed. This chapter also describes the construction of the light-proof measurement box, designed to provide a consistent experimental setup for all spectral measurement during this study. The chemometrics methods used to optimise the regression models between spectral data and sugar content are also investigated.

Chapter 5 describes the results of the preliminary studies of the application of the VNIRS to determine sugarcane quality parameters from the conventional sugarcane sample forms: raw juice, clarified juice and fibrated samples. This chapter also discusses the potential application and limitation of each sample form for quality measurements in the field.

In **Chapter 6**, the results of the calibration and prediction models developed from CSSM, for both whole stalk and internode samples are presented. The variation of sugar content and prediction accuracy for both samples are also quantified. Also discusses is the selection of the optimum quality parameter to be used for predicting and representing sugar content from stalk samples. The influence of varieties on prediction accuracy is assessed.

Chapter 7 compares the performance of both spectroradiometers (VNIRS and FRS) in determining sugarcane quality parameters from sugarcane stalks based on the SSM. The SSM is applied on whole stalk samples as well as internode samples. The variation of sugar content and prediction accuracy for both sample types are quantified. The optimum method for quality prediction in the field is identified, including the comparison of prediction performance between CSSM and SSM. The performance of an artificial neural network (ANN) algorithm to classify sugarcane quality into several classes is assessed.

Chapter 8 outlines a basic calculation and conceptual design of the proposed measurement system to determine sugar content using the VNIRS on a sugarcane harvester. This chapter covers the specification for both measurement and sampling systems. The integration between the sampling system, measurement system, harvester and GPS to produce a quality map is explored.

Chapter 9 summarises the main conclusions and novelties of this research. Further studies are also recommended.

Chapter 2

The Australian sugarcane industry and the adoption of precision agriculture (PA)

The purpose of this chapter is to give an overview of the significance of the sugarcane industry to the Australian economy. The agronomic characteristics and typical harvesting operation for this crop are also described. This chapter will also discuss the importance of the in-field measurement system for the industry. Next, the chapter reviews current quality measurement methods in the sugarcane industry, their potential applications and limitations for field use. The new emerging technologies which have the potential to be applied in sugarcane industry are also evaluated. This chapter also discusses the adoption of precision agriculture (PA) methods in the Australian sugarcane industry.

2.1 The economy of the Australian sugarcane industry

The economic value of the sugarcane industry in Australia is ranging between AUD$1.5 and 2.5 billion per annum (Canegrowers 2011). The sugarcane industry is one of Australia's largest rural industries, with annual sugar production reaching more than 5 Mt (Canegrowers 2011). Sugarcane is extensively grown in high rainfall or irrigated areas on coastal plains and river catchments along 2100 km of coastline, from far north Queensland to northern New South Wales. The vast majority of sugarcane is grown on family-owned farms ranging from 40 to 250 ha. Around 4000 sugarcane farming business enterprises supplied more than 35 Mt of sugarcane to 24 sugar mills across Australia (Canegrowers 2011).

About 95% of Australia's sugarcane is grown in Queensland, with 5% in New South Wales. In 2010 season, the Queensland's harvested sugarcane covered the area of 291,805 ha. In the season of 2009, the average sugarcane and sugar production for Queensland was 84 t ha^{-1} and 15 t ha^{-1}, respectively. These figures can vary quite drastically from year to year. Australia is the world's second largest exporter of raw sugar after Brazil. Around 20% of the raw sugar produced in Australia is refined for domestic consumption and the remainder is exported (Canegrowers 2011). Australia exports its sugar mainly to Japan, Korea, Malaysia, New Zealand, Canada and the USA.

2.2 Sugarcane botany

Sugarcane is a perennial grass of the family Gramineae of the genus Saccharum and its botanical name is *Saccharum officinarum*. Sugarcane stalks grow to maturity over a period of nine to twelve months. Sugarcane stalk generally grows to about 2.5 m in height with an average diameter of 30 mm. Sugarcane is grown from a section of the stalk containing a bud. The sugarcane plant produces a bud at the leaf axil of the plant. The bud grows on the band surrounding the stalk. When the stalk or section of the stalk

is planted, roots develop and supply moisture to the growing bud. As the shoot grows, roots appear at the bottom of the plant. Buds on the new plant germinate and produce more shoots forming a clump at the stool.

Sugarcane is a tall crop with its stalk comprises a series of internodes which are separated by joints or nodes. The internode skin is made up of a hard outer rind which provides strength and protection for the softer inner tissue or pith (Kroes 1997). The rind which represents 15% of the cross sectional area of the stalk consists of layers of small, highly lignified, and thick-walled cells (Schembri & Harris 1998). The node is an encircling ring from which the leaf sheath grows (Figure 2.1). As sugarcane grows, the leaves of the older nodes die and detach from the stalk and the node is represented by the remaining leaf scar. Above the leaf scar is the root band and growth ring. Within the root band are rows of root primordia and a single bud occurring on opposite sides of the stalk for alternate nodes. The stalk surface is covered with a moisture resistant layer of wax which is thickest just below the node where it forms a wax band. Typically, two types of cracks, 'corky' and 'growth' cracks may occur within an internode (Kroes 1997).

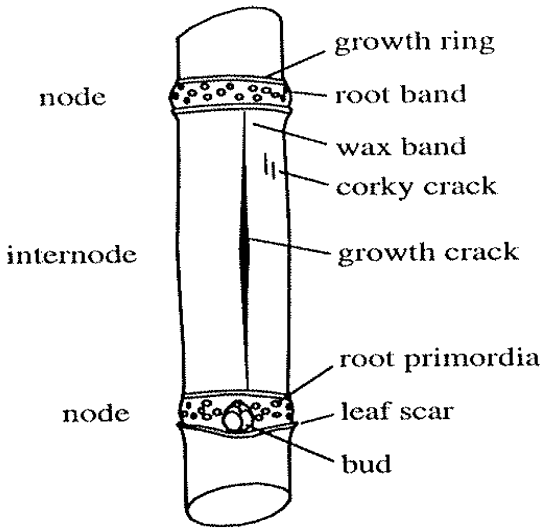


Figure 2.1: Typical sugarcane internode (Kroes 1997)

The cross-section of sugarcane stalks is circular (or oval for some varieties), and ranges in diameter from 12 to 50 mm. The internode is made up of a hard outer rind which provides strength and protection for the softer inner tissue or pith (Figure 2.2). The rind consists of layers of small, highly lignified, thick-walled cells. Conversely the large, juice bearing cells of the pith are thin-walled structure and slightly lignified. Throughout the stalk, the internode fibrovascular bundles run parallel to the stalk axis. At the nodes the fibrovascular bundles bend, branch and knot with those of the following internode (Kroes 1997). The fibrovascular bundles are responsible for the

transportation of the water and food between the photosynthetic parts of the plant (the leaves) and the remainder of the plant.

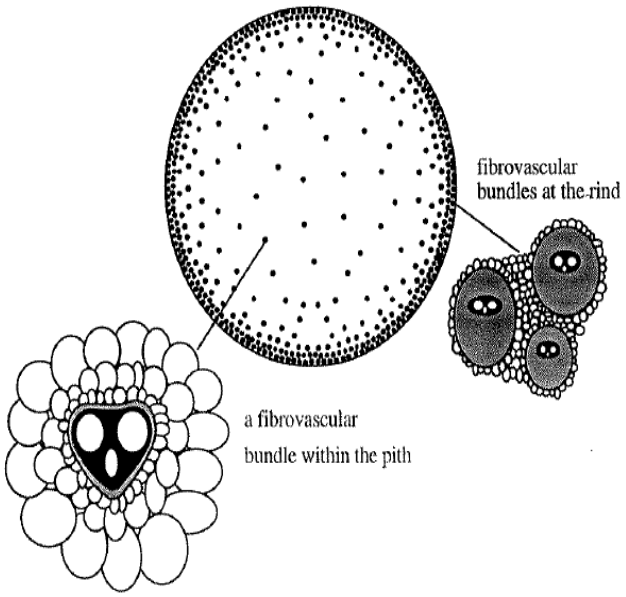


Figure 2.2: Representation of a sugarcane cross-section and the fibrovascular bundles (Kroes 1997)

Sugarcane has an ability to trap the sun's energy and convert that energy into sucrose (sugar), which is stored as sweet juice in the stalk of the plants. Sugar is produced in the leaves of a sugarcane plant through photosynthesis. Sugarcane contains sucrose as prime constituent along with numerous other dissolved substances, cellulose, and fibre. The percentage of sucrose is usually referred to as the polarisation value (pol) and varies from 8 to 10%, depending on the variety of sugarcane, maturity level, soil and climatic conditions (Mehrotra & Siesler 2003). The typical constituents of ripe sugarcane are given in Table 2.1.

Table 2.1: Constituents of sugarcane

Constituents	Percentages (%)
Water	69-75
Sucrose	8-16.4
Reducing sugars	0.5-2
Organic matter other than sugar	0.5-1
Inorganic compound	0.2-0.6
Nitrogenous bodies	0.5-1
Ash	0.3-0.8
Fibre	10-16.5

(Source: Mehrotra & Siesler 2003)

2.3 Sugarcane harvesting

Sugarcane is a seasonal crop with a harvesting period of approximately five months. In Australia, depending on the weather conditions, the harvesting of sugarcane usually commences in late June and is generally completed by late November. Harvesting operations are performed by a conventional chopper harvester (Figure 2.3). Sugarcane is typically harvested at physiological maturity when the total soluble solid content (°Brix) has reached 20-24%, which is usually about 11 to 12 months after planting.

Figure 2.3: Sugarcane being harvested by a conventional sugarcane harvester

Burnt cane harvesting (BCH) and green cane harvesting (GCH) are two common sugarcane harvesting methods in Australia. In the 2008 season, 57% of the industry applied GCH and the remainder used BCH (ACFA 2009). The burning of sugarcane prior to harvesting was popular during the 1940's as a measure for controlling diseases spread by rats and reducing the amount of leaves and weeds from the sugarcane. Lately, however, BCH has become less popular because of the need to retain cover on the ground for moisture and weed control.

GCH allows the leaf and the leafy stalk tops to fall to the ground and provide a trash blanket over the field. In many growing districts, the trash blanket returns nutrients to the soils, helps to prevent weed germination, preserve moisture and nutrition, and reduce soil erosion. GCH involves cutting the crop as it stands without prior preparation and has, in recent times, become the more popular method. In fact, sugar industries throughout the world are increasingly adopting GCH, mainly due to public pressure to reduce standing burns and the potential benefits of increased sugar recovery (Richard et al. 1996).

All commercial sugarcane harvesters work using the same principles. Figure 2.4 shows the series of events of a self-propelled mechanical harvester. During harvest, a harvester will first separate the rows and lift lodged sugarcane with the crop dividers. Directly

above the crop dividers is the topper that removes the leafy material at the top of the stalk (topping) and this material is allowed to fall back onto the field. After topping, a knockdown roller then pushes the sugarcane over to ensure that the butts of the stalks enter the harvester first. Then, the stalk is cut off at ground level with the base cutters, a pair of counter-rotating discs fitted with blades. The inward rotation of the base cutters assists with the transportation of sugarcane towards the feed train. A finned feed roller and a butt-lifter roller then guide the stalks into the feed train to the choppers. The stalks are cut into billets (200 to 300 mm) by the choppers, and are sprayed into the bowl of the harvester elevator across the base of the extraction chamber.

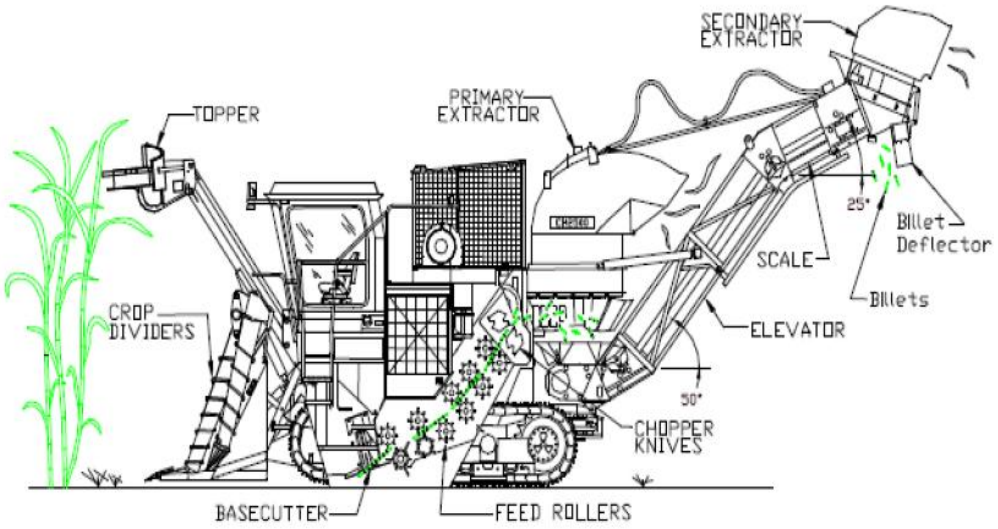


Figure 2.4: Typical sugarcane harvester (Caryn et al. 2002)

Billets are often contaminated by soil and leaf material (Cargnello & Fuelling 1998). A majority of the leaves, tops and dirt are separated from the sugarcane in the extraction chamber by the primary extractor fan. The primary extractor can remove up to 90% of the foreign material from the billets (Mailander et al. 2010). Then, the billets travel up the elevator conveyor to the secondary fan. The secondary extractor fan further removes extraneous matter from the billets. The elevator conveyor is a chain-driven slat system moved at a typical speed of 2.7 m s^{-1}, the required speed which is being controlled by the operator. There is a 3 mm clearance between the slats and the elevator floor. The elevator can be divided into two sections. When raised to its highest position during harvesting, the first section is at an angle of 50° with the ground and the second section is at an angle of 25°. The first section floor is constructed of expanded metal to allow mineral material to fall through it.

After the billets have passed through the second extractor, they are deposited into a waiting haul-out bin that is pulled by a tractor beside the harvester. Sometimes, a haul-out truck is also used to collect billets from a harvester and take them to the railway siding (Figure 2.5). Each harvester will generally have multiple haul-out vehicles to

reduce lengthy stoppage times. The full bins are then taken from the field to a sugarcane railway siding or a road haulage delivery point for transport to a sugarcane mill.

Figure 2.5: Haul-out truck to carry harvested sugarcane from a harvester to rail sidings

2.4 Mill measurement system for the sugarcane industry

When harvested sugarcane is received at a mill, it is known as a consignment of sugarcane. The consignment usually represents a block of sugarcane field. From this consignment, the average sugarcane yield (t ha^{-1}) and CCS are determined and calculated. The sugarcane yield and CCS are the primary factors of economic value to the sugarcane industry. Sugarcane yield is an estimate of the weight of sugarcane stalks on a fresh weight basis obtained at crop harvest (Lawes & Lawn 2005). CCS is an estimate of commercially recoverable sugar content at the mill and includes the effects of impurities (non-sucrose substances dissolved in sugarcane juice) that cause sugar loss as molasses (BSES 2001). CCS is expressed as a percentage of sugarcane fresh weight and is derived from measurements of °Brix in juice, pol in juice and the fibre content. The standard algorithm for calculating CCS as given by BSES (2001) is as follows:

$$\text{°Brix in cane} = \text{°Brix in juice} \times \frac{100 - (\text{fibre\%} + 3^{*})}{100} \tag{2.1}$$

$$\text{Pol in cane} = \text{Pol in juice} \times \frac{100 - (\text{fibre\%} + 5^{*})}{100} \tag{2.2}$$

$$\text{Impurities in cane} = \text{°Brix in cane} - \text{Pol in cane} \tag{2.3}$$

$$\text{CCS} = \text{Pol in cane} - 0.5 \times (\text{impurities}) \tag{2.4}$$

Where number 3* (Eqn 2.1) and 5* (Eqn 2.2) are correction factors used to correct the °Brix and pol measurements in first expressed juice to more accurately represent those of the total juice in sugarcane.

2.4.1 °Brix

°Brix (a measure of sucrose and soluble impurities) of a solution is the concentration (in g solute per 100 g solution) of a solution of pure sucrose in water, having the same density as the solution at the same temperature (BSES 2001). °Brix can be measured based on density or refractive index (% soluble solids in juice). The refractive index is an index relating the amount of light that is bent when it passes through two mediums. Typical equipment to measure °Brix value is Brix spindle, Brix hydrometer, density meter or refractometer. A typical desktop refractometer used in a research station is shown in Figure 2.6.

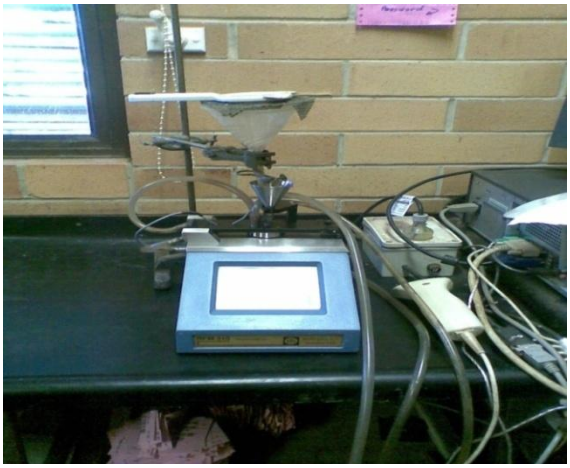


Figure 2.6: A typical desktop refractometer for measuring sugarcane °Brix

2.4.2 Pol

Pol is a measure of the sucrose present and other optically active substances (% sucrose present in juice). Pol of a solution is the concentration (in g solute per 100g solution) of a solution of pure sucrose in water having the same optical rotation as the sample at the same temperature (BSES 2001). Since sucrose is optically active, it can optically rotate linearly polarised light. This optical property allows the concentration of sucrose in a substance to be calculated. Typical laboratory equipment to measure pol value is a desktop polarimeter (Figure 2.7).

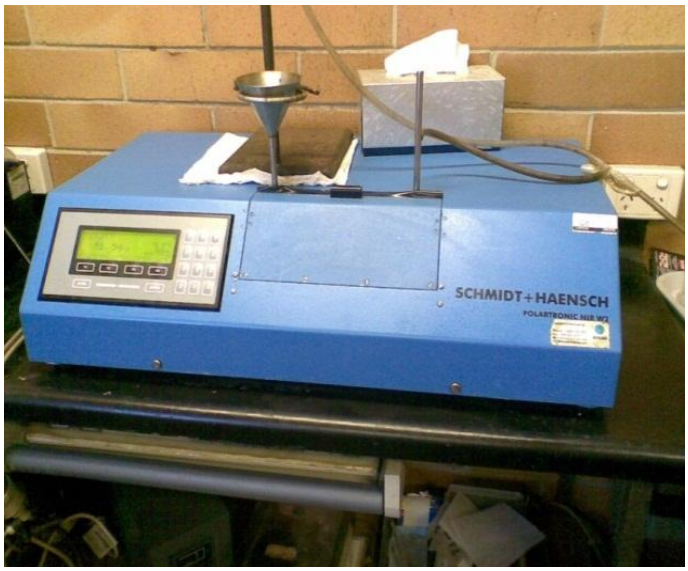


Figure 2.7: A typical desktop polarimeter for measuring sugarcane pol

A polarimeter is widely used as a saccharimeter in sugar analysis. Even though sugarcane juice contains glucose, fructose and sucrose, only sucrose was used by the industry for the calculation of sugar content in the crop. In the determination of sucrose in a substance containing no other optically active substance, 75.2 g dissolved in 100 mL of water, using a 200 mm tube the rotation in degrees is numerically equal to the concentration of the sucrose in percent by weight. However, since sugarcane juice also contains other optically active sugars such as glucose and fructose, a polarimeter for the industry has been standardized that 26 g of sucrose dissolved in 100 mL water gives 100°C rotation. Therefore the rotation observed multiplied by 0.26 directly gives the % sucrose in sugarcane juice sample (Mehrotra & Seisler 2003).

2.4.3 Fibre

Fibre is a dry, water insoluble matter in the sugarcane (BSES 2001). The percentage of fibre is directly determined by breaking open all the fibrous cells containing the juice, washing the juice out and drying and weighing the remaining fibre (BSES 2001). A standard industry method for fibre determination is carried out using either whole stalk or prepared cane method (BSES 2001). More recently, a prepared cane method is increasingly being applied. For prepared cane method, fibre is analysed by a fibrating, washing and drying process, using either special cloth bags to retain the fibre or a can fibre apparatus (Watson et al. 1999).

2.5 Crop production information for growers

Under conventional mill measurement practices, mill managers provide growers with the sugarcane yield and CCS information for their blocks. However, since the crop production data provided by the mills is for each block, it cannot be used to quantify spatial variation within the individual block. As a result, although many growers know that their yield and quality vary across the field, without the exact information or tools

to either quantify or manage this variability, they have tended to manage their farm on the basis of homogeneity, assuming their field as a single large homogenous field (Cook & Bramley 1998). However, if the information regarding block production can be divided into smaller units of characteristic performance (zones), some forms of differential or targeted management could be possibly implemented (Taylor et al. 2007). Growers who have access to field variation data may implement new management strategies in which the inputs to the production system can be closely matched to the desired and/or expected outputs.

For this reason, precision agriculture (PA) has been introduced to increase the likelihood of farmers making a 'correct decision' (McBratney et al. 2005) and thereby gaining beneficial outcomes through the targeting of inputs (Cook & Bramley 1998) or selectively harvesting outputs (Bramley et al. 2005). Bramley (2009) has broadly defined 'inputs' for the sugarcane industry as fertilisers, irrigation water, pesticides, chemical ripeners, herbicides, labour and the timing of harvesting.

2.6 Variation of sugarcane production across the field

The level of crop productivity in agricultural fields can vary considerably, depending on several factors. Among the major factors that influence the crop production in the field are soil properties, topography, rooting depth, nutrition, agronomic management, and the interaction of these factors with climate (Runge & Hons 1999). As a result, the input-output relationships for the agricultural production system can vary, often over distances of only a few meters (McBratney & Pringle 1999). The sugarcane field is also subject to these variations, where both yield and quality components were found to be considerably varied across the field (Lawes et al. 2003; Johnson & Richard 2005).

The first study on variability in the sugarcane production system was carried out by Kingston & Hyde (1995) using the hand sampling method. The study was conducted by measuring CCS of six stalk samples from a 50x50 m intra-field sampling grid. They demonstrated that, within a single 8.8 ha sugarcane block, variation in CCS was up to 6.5 units. In another study, Lawes et al. (2002) found that in an individual block, the variation of CCS was up to 8 or 9 units. This was greater than the reported within-season variation in district average CCS of approximately 2.5 units (Kingston 2002) and between seasonal variations of approximately 4 units (Leslie & Wilson 1996).

Another study conducted by Bramley et al. (2012) reported that CCS varied from 13.29 to 16.84 with a mean of 15.73 and median of 15.80, and the coefficient of variation (CV) was 3.5% (Figure 2.8). In their study, the sugarcane stalk samples were manually collected from around 200 geo-referenced locations in a 6.8 ha block of sugarcane in the Bundaberg district. The authors concluded that the variation in CCS was spatially structured - that is, it was not random.

In terms of yield variations, Bramley & Quabba (2001) demonstrated that the range of variation in sugarcane yield, as measured by the coefficient of variation (CV), was of the order of 30 - 45%, which was similar to a range of other crops for which yield

monitoring equipment was available (Pringle et al. 2003). Jhoty et al. (2003) reported that the range of yield variation in Mauritius was between 60 and 150 t ha^{-1}, which was similar to that seen in Australia (Bramley 2009). In the US, Johnson & Richard (2005) reported that the yield variation ranged from 36 to 134 t ha^{-1}.

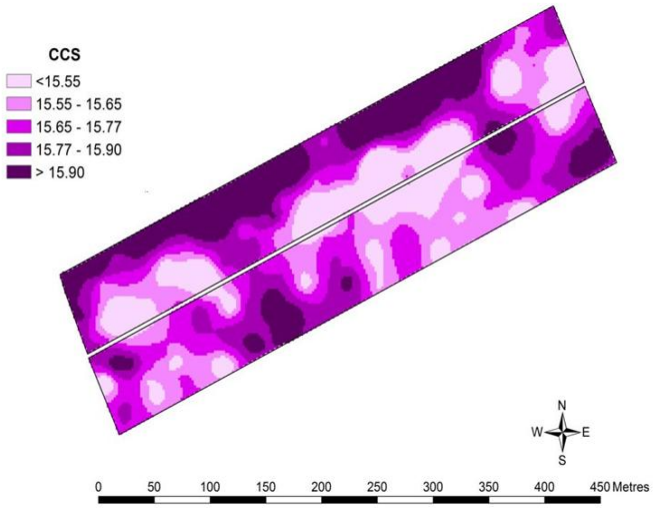


Figure 2.8: CCS variation in a 6.8 ha sugarcane blocks in the Bundaberg region (Bramley et al. 2012)

In Australia, Lawes et al. (2002) observed that the variation of sugarcane yield within an individual block ranged from 10 to 160 t ha^{-1}. In Brazil, Magalhães & Cerri (2007) mapped 43 ha sugarcane field using a yield monitoring system called a Simprocana (Figure 2.9). The authors reported that the spatial variability found in their study ranged from 6 to 150 t ha^{-1}.

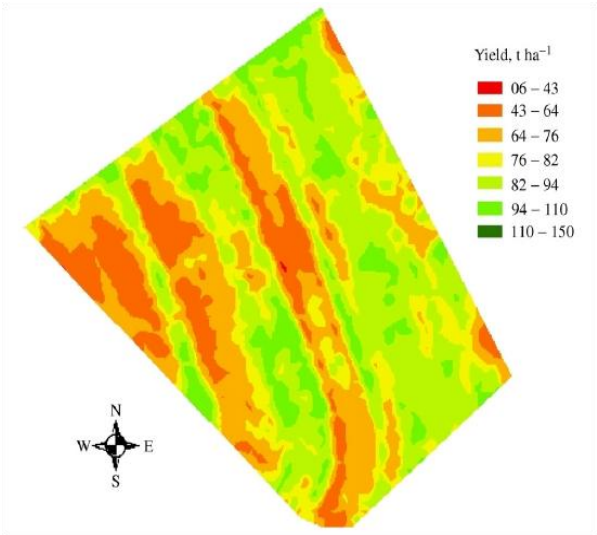


Figure 2.9: Sugarcane yield map in an area of 43 ha in Brazil (Magalhães & Cerri 2007)

In conclusion, all studies reviewed in this section showed considerable magnitude of variation for both yield and quality within a single sugarcane field. The findings of these studies highlight the need for growers to use PA in their farming practices. Bramley et al. (2008) suggested that both zone-based and continuous variable rate management may have potential applications in sugarcane production. Thus, the management of these variations with the targeted management method, as opposed to the blanket application method, will be more efficient and cost-effective for growers.

2.7 PA for the Australian sugarcane industry

PA is an all-encompassing term given to a suite of technologies which promote improved management of agricultural production through recognition that the potential productivity of agricultural land can vary considerably, even over very short distances (Bramley 2009). Recently, substantial emphasises have been given to the use of PA technologies in the sugarcane industry due to the on-going need to improve production and demonstrate the use of environmentally sustainable best practice (Wrigley & Moore 2006). PA is a valuable management tool to maximise farm profits through efficient application of crop inputs by matching them with potential variations of crop yield and quality in the field (Wendte et al. 2001). PA is also defined as an innovative, integrated and standardized approach to increase the efficiency of resource use and to reduce the uncertainty of decisions required to control variation on farms (Schellberg et al. 2008).

In general, the adoption of PA could improve the production of sugarcane and better manage the environment in which sugarcane is grown and harvested. PA has the potential to increase profits and decrease environmental impacts through variable rate application of agricultural inputs such as fertilisers, seeds, lime and herbicides (Adamchuk et al. 2004). Several researchers have demonstrated that variable rate application methods can reduce the total amount of a nutrient applied to a given field and also reduce variability of that nutrient within the field (Mallarino & Wittry 2004). PA could also be used by growers and advisers to improve decision-making in relation to the crop, including the timing of fertilisation and harvesting, crop agronomy and the efficiency of farming practices. However, the adoption of PA will depend on the economic benefits offered, which will, in turn, depend on the degree of variability present in the field, and the opportunity to manage that variability (Bramley 2009).

2.7.1 Technologies and methods in PA

PA is normally applied to manage the crop production at spatial scales smaller than the usual field scale. PA is a cyclical process (Figure 2.10), in which the observation of crop performance is a critical first step, followed by the evaluation and interpolation and targeted management plan (Bramley 2009; Cook & Bramley 1998). During the observation stage, the PA requires yield or quality mapping and the acquisition of complementary information (e.g. a soil map) to be collected. Then, this information is interpreted and evaluated in order to investigate the potential and limitation for crop production management. After the evaluation stage, the collected information is used in the implementation of the targeted management. This is followed by further observation

in order to assess the efficiency of the proposed targeted strategy. The cyclical process continues until the targeted productivity level has been achieved.

The success of PA depends on the availability of the key enabling technologies, including the Global Positioning System (GPS), Geographic Information System (GIS), miniaturized computer components, automatic control, in-field and remote sensing, advanced information processing, and telecommunications (Gibbons 2000). The application of these technologies can be used to collect and process the data relating to site-specific characteristics of the field. Then, site-specific input data can be analysed in a real-time or off-line to generate a prescription map. Yield mapping is usually the first step in developing PA techniques (Jhoty & Autre 2003). A yield map allows growers to determine yield variations within a field due to different management practices being applied on those yields. Thus, the yield monitoring and mapping systems are fundamental components of PA (Heacox 1998). A typical approach for mapping yield variations in the field consists of a combine harvester or a side trailer equipped with a yield monitor sensor to measure the flow rate of crops.

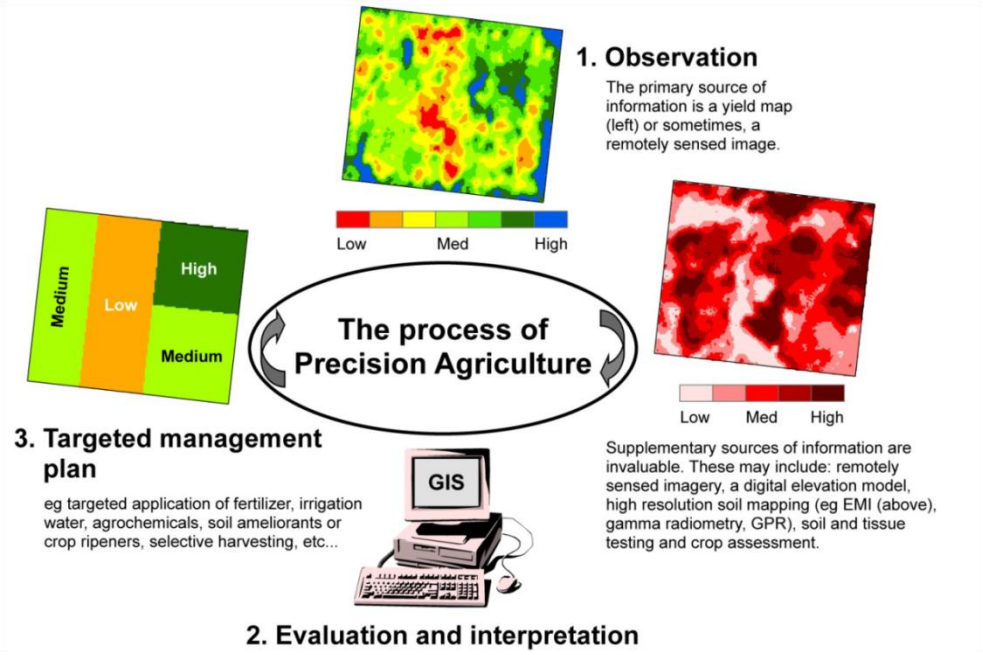


Figure 2.10: The cyclical process of PA (Bramley 2009)

GPS is used for referencing the points where measurements have been made. The GPS receivers have become the most common sensor in PA technology to locate and navigate a combine harvester a field. The GPS receivers can also be used to determine geographic coordinates (latitude and longitude) and measure the altitude (elevation). The resulting data of the measurement can be used to calculate slope, aspect and other parameters relevant to the landscape. When a GPS receiver and a data logger are used to record the position of each grid sample or measurement, a map can be generated and processed along with other layers of spatially variable information. This method is frequently called a ‘map-based’ approach.

The majority of the spatial variability data has some form of geographic information which can be referenced to geographical locations such as points, lines or areas. These systems are known as geographic information systems (GIS), where the database is the fundamental to the GIS system. The database normally comprises of two elements: a spatial database that records the location and an attribute database that describes the characteristics or qualities at each of the locations. Each measured variable from a field that is geo-referenced becomes a unique data layer that can be overlaid and visually and mathematically correlated or interpreted.

2.7.2 Yield monitoring system for sugarcane

Nowadays, PA techniques are being studied in and adopted for a variety of crops. A comprehensive review of PA for different broadacre crops around the world was published by Griffin & Lowenberg-DeBoer (2005). In the review, the authors listed the use of PA by growers of corn (maize), soybeans, potatoes, wheat, sugar beet, barley, sorghum, cotton, oats and rice. In the sugarcane industry, there are also several published studies reporting the application of the PA method in the industry, including Australia (Cox et al. 1997; Jensen et al. 2010, 2012), Mauritius (Jhoty et al. 2003), the US (Johnson & Richard 2002), and Brazil (Magalhães & Cerri 2007). A typical yield monitoring system for a sugarcane harvester called a Simprocana, developed by Magalhães & Cerri (2007), is shown in Figure 2.11. This system consists of several components such as inclinometer, conveyor speed sensor, scale, GPS and monitor.

The first development of a sugarcane yield monitor was conducted at the University of Southern Queensland, Australia (Cox et al. 1996, 1997; Harris & Cox 1997). The system was based on hydraulic pressure determination, hydraulic flow variation and machinery speed. The sensors were placed in the chopper and elevator systems and produced a linear line output with R^2 values equal to 0.96 and 0.95 for the chopper and elevator systems, respectively. Cox et al. (1999) produced yield maps using a direct mass measurement technique with the help of GPS and ArcView. These yield maps displayed a significant amount of yield variability, ranging from 70 to 190 t ha^{-1}. There were also studies that produced the yield map by measuring yield on a trailer, with load cells equipped with a DGPS and a data acquisition system (Earl et al. 1996; Pierossi & Hassuani 1997).

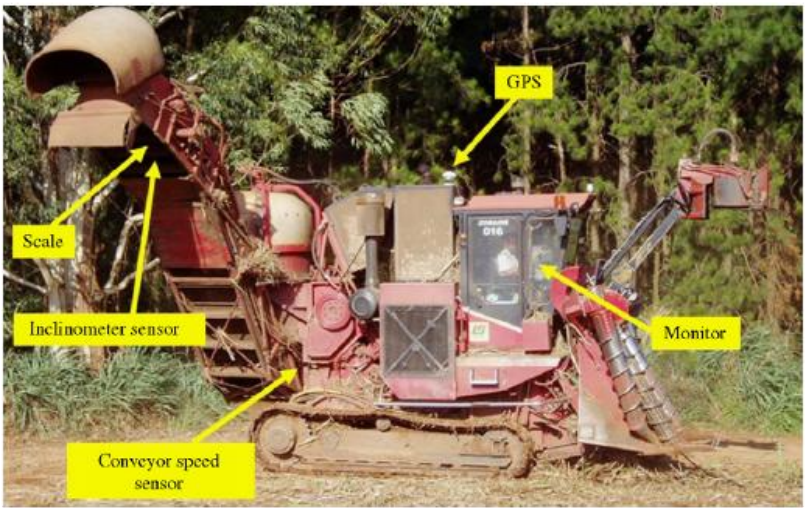


Figure 2.11: A sugarcane harvester equipped with components of yield monitoring system (Magalhães & Cerri 2007)

Molin & Menegatti (2004), Magalhães & Cerri (2007), Cox et al. (1999, 2003), Pagnano & Magalhaes (2001), Benjamin et al. (2001) and Benjamin (2002) tested weight scale systems placed in the elevator floor of the harvester (Figure 2.12). Pagnano & Magalhães (2001) designed, constructed, and tested a mass flow sensor based on load cells transducers, and included detectors mounted on the side conveyor to measure its velocity so that the mass flow could be determined in real time. The recorded data, added to the information provided from DGPS, allowed digital maps to be produced. Wendte et al. (2001) developed a monitor equipped with a torsion deflection plate at the outlet of the chopper harvester's elevator that measured the force and impact of billets spilling from the elevator outlet. In the most recent studies, Jensen et al. (2010, 2012) investigated issues which could have an impact of yield monitoring accuracy, including sensor type, harvester speed and the presentation of the crop to the harvester (e.g lodged versus erect).

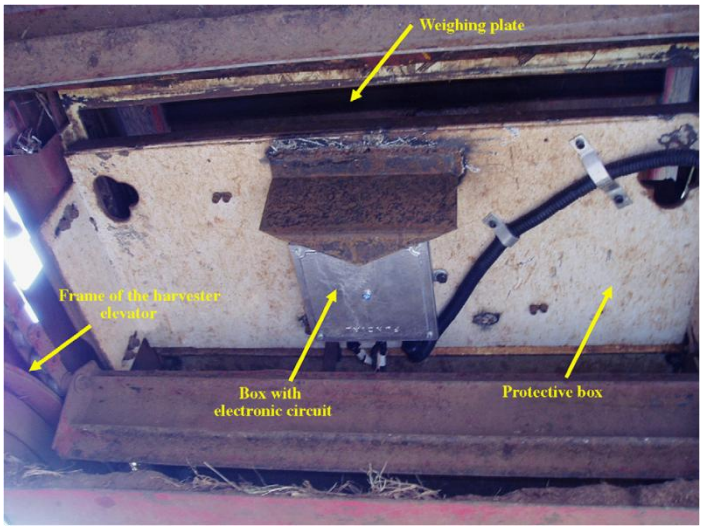


Figure 2.12: Position of a scale set-up on the elevator of the sugarcane harvester from bottom view (Magalhães & Cerri 2007)

Most of the published studies on sugarcane yield monitoring are based on weight scale plate systems. However, a number of potential issues have been highlighted regarding this system. Over time, soil, sediment and debris can accumulate in the gap between the weight scale and the elevator floor, causing the weight plate to lock to the elevator floor (Price et al. 2011). Thus, Price at al. (2011) developed a fibre optic yield monitoring system for a sugarcane harvester that utilises fibre optic sensors mounted in the elevator floor to estimate sugarcane yield.

A yield monitor was developed using optical sensors mounted under the floor of a sugarcane harvester conveyer (Figure 2.13). The system measured weight by estimating the depth of material on the slats using a duty cycle type approach and transforming that information into weight using a calibration line (volume was assumed constant, as a triangular prism shape is formed by the billets on the slats from the step elevator slope during operation) (Price et al. 2011). Then, mass flow rate was determined by dividing the depth value by the total crop area harvested by the sugarcane harvester. This monitoring system compared well with all previously tested methods and was very durable and easy to install. Price et al. (2011) concluded that the optical methods had advantage of being inexpensive and relatively easy to operate.

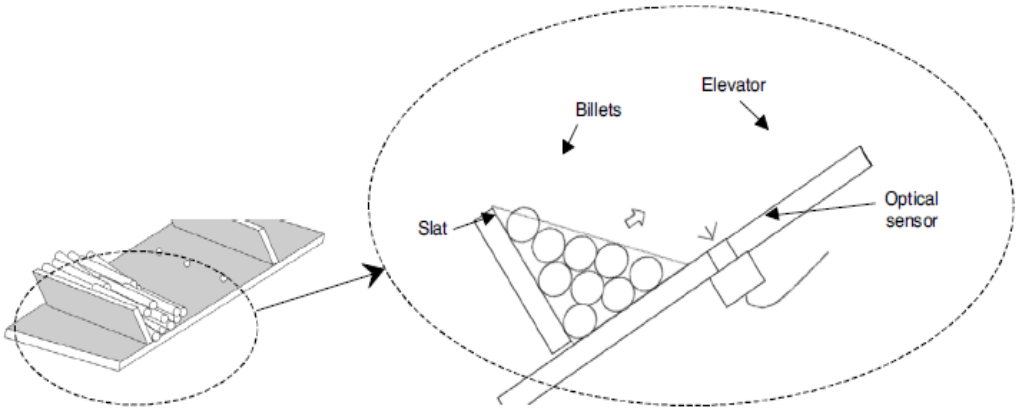


Figure 2.13: Method to detect billets from the elevator floor using optical sensors (Price et al. 2011)

2.8 In-field quality measurement - Missing element in current PA technologies

Bramley (2009) presented an extensive review of the research, development and adoption of PA in the Australian sugarcane industry. It found that the current PA technologies could only monitor the crop yield and did not have the ability to measure the product quality. Even though several studies have been carried out to produce a yield map in sugarcane industries using an on-the-go yield monitoring system (Cox et al. 1996; Wendte et al. 2001; Price et al. 2011), there is no published study regarding the generation of quality map from an on-the-go system. This is a serious barrier to the

full implementation of a PA technique in this industry. Since the quality parameter is as important as yield for the sugarcane industry, it is highly desirable to obtain information of sugarcane quality along with yield information in the field. This will enable management and planning for consistent production quality.

Thus, it can be seen that there is a critical need to produce a separate quality map to complement a yield map in order to manage the spatial variation of both components. Quality maps could also provide essential information for spatial analysis and evaluation of crop production management within a field. This information can be used to make decision for field operations in the next growing season. Furthermore, improvements in CCS offer the industry greater industry economic benefits than for the same proportional improvements in sugarcane yield (Jackson et al. 2000). Jackson et al. (2000) reported that an increase in sucrose yield due to improved CCS is up to 2.5 times more valuable to growers than a sucrose yield increase due to improved sugarcane yield. Therefore, growers using variable rate technology to optimise yield will not want the benefit of doing so to be offset by a possible related CCS penalty. Thus, it would be very useful if the in-field quality measurement and the in-field yield measurement could be done concurrently during harvesting.

Several studies have been undertaken to investigate the relationship between sugarcane yield and CCS (Sukhchain & Saini 1997; Lawes et al. 2003; Johnson & Richard 2005; Rattey et al. 2009). These studies found that there was no significant relationship between CCS and yield. For example, Rattey et al. (2009) reported that on average, CCS and sugarcane yield were not associated. Sukhchain & Saini (1997) also reported that there was a negative association between sugar content and yield. Lawes et al. (2003) studied this issue at the regional scale in the Tully district. They found no spatial relationship between yield and CCS, although spatial variation in both was temporally stable. Thus, understanding the spatial and temporal interactions between yield and quality could be a very important step for implementing a PA method.

Besides targeted management, there are other benefits which can be achieved if the sugarcane quality can be measured in the field. For example, during harvest, the low-cost and rapid in-field sugarcane measurement system can be used to optimize the sugarcane values at harvest through the identification of sugarcane blocks that have the highest in-field CCS. In addition, tracking of in-field CCS for a period leading up to the crushing season would allow for the optimisation of harvest schedules for maximum returns (Staunton et al. 2011).

Furthermore, the measurement of sugar content in the fields would improve a data collection system as the industry is required to record data of tonnage, CCS and the originated farm for the purposes of payment, and for estimating varietal performance, crop class or fallow practice. The collection and processing of these consignment data is complex and errors, referred to as ‘consignment errors’, often occur during the data handling process (Lawes & Lawn 2005). Thus, by measuring and collecting all required data in the field using the same system over the region, consignment errors could be eliminated.

Finally, the development of an in-field quality measurement system would also benefit sugarcane breeding plant programs. Sugarcane plant breeding and selection programs are important as they provide the primary means of developing new cultivars and delivering increased productivity and other quality advances (Cox et al. 2000). Currently, the measurement of quality components for plant breeding with standard laboratory methods is labor intensive and expensive. Thus, rapid quality screening techniques which could be applied in the field, without the need for the complex laboratory procedure, would be more practical for the breeding programs. The development of tools to enable plant breeders to rank plants at an earlier stage in the selection cycle would significantly improve the overall efficiency and cost of the breeding and selection program (Purcell et al. 2005).

2.9 Current in-field measurement methods for research programs

Research programs in the sugarcane industry sometimes require the measurement of quality components from crop samples in the field. For example, in sugarcane breeding and research programs, the determinations of the crop quality components commonly rely on the hand cutting of lightly topped sound whole stalk samples (McRae et al. 1996). Typically, for sampling purposes, a sample of six whole stalks is collected at random from across all rows of each study plot. The green canopy is removed by cutting each stalk near the growing point and removing all leaf material. The quality parameters of the stalk samples are later determined in a laboratory using the standard laboratory procedures. This hand sampling method has been applied for clonal evaluation (Berding et al. 1991a; 1991b) and mapping CCS across the field (Bramley et al. 2012).

Unfortunately, sampling the whole stalk samples from the paddock, performing pre-treatment on the stalks, transport to the laboratory, the analysis itself, and communication of results back to growers can be very costly, labour intensive and a time consuming (Abdel-Rahman et al. 2010). There is also pressure on a laboratory staff to process a huge number of mature sugarcane stalk samples within a limited time, before quality deterioration takes place, especially for clonal evaluation (Purcell et al. 2005). On the other hand, for mapping purposes, a grid sampling generated from the traditional hand sampling method is not dense enough to illustrate detailed variation across the field. Thus, a new method to measure sugarcane quality parameters in the field without the necessity of exhaustive manual measurements would be beneficial to the industry.

Recently, a significant amount of research has been focused on data collection during normal field operations using appropriate sensors mounted on grain combine harvesters. Measuring quality data during the harvesting operation is advantageous over manual measurement method because less labour is required, more data samples may be taken and samples may be easier to correlate with specific positions (Wendte et al. 2001). The availability of sensors for an on-the-go yield monitoring system during harvest (Cox et al. 1999; Bramley & Quabba 2001; Jensen et al. 2010, 2012; Price et al. 2011) has

enabled the measurement of yield in the industry. However, the development of an in-field quality measurement for this industry has been hindered due to the absence of a sensor which could survive the harsh field environment and rugged harvesting conditions. Thus, this research has been carried out to identify the most suitable sensor for field use.

2.10 Common technologies for quality measurement in a laboratory

To identify a suitable instrument which is reliable for field use, an investigation on the existing technologies used in the laboratory is useful. Among the key factors to be considered for selecting a suitable instrument are: the sampling time, sample preparation, sample form, amount of sample required for an individual measurement and cost of equipment. The comparison of these factors among the common laboratory technologies are summarised in Table 2.2. From the table, it can be seen that all laboratory equipment is designed to measure juice samples. Typically, a juice sample is obtained by crushing stalks using a roller mill or hydraulic press. The common technologies used for juice measurement are refractometer, polarimeter, chromatography and NIR spectrometer.

Table 2.2: Comparison of common methods for sugarcane quality determination

Method	Sampling time (min)	Sample form	Amount of samples required	Approximate cost of equipment
Refractrometry	5-20	Juice (raw or clarified)	50 - 100 ml	Hand-held (AUD$700) Laboratory (AUD$5k)
Polarimetry	10-20	Juice (clarified)	100-200 ml	AUD$16k
Chromatography	30	Juice (clarified)	100-200 ml	AUD$13-33k
Biosensor	5	Raw juice must be mixed with a mediator	50 - 200 mmol l^{-1}	AUD$6k
Brix hydrometer	15-20	Juice (raw or clarified)	100-200 ml	AUD$40
Wet chemical methods	20	Clarified juice	Each 0.0047g of sucrose to be mixed with 1ml of Fehling's solution	Only standard laboratory equipment is needed
NIR spectroscopy	0.2 to 1 (after calibration)	Juice (raw or clarified) and macerated sugarcane samples, or possibly billet samples	50 - 100 ml	AUD$10-18k (350 to 1075 nm) AUD$100-140k (350 to 2500 nm)

2.10.1 Refractometer

A refractometer determines sugarcane °Brix by measuring a refraction index of the juice sample which is related to the composition of the material in the samples. This equipment is a relatively cheap (AUD$700-5000) and has been used for the rapid determination of °Brix in juices and syrups. However, this method may require longer preparation time (5-20 minutes) before juice samples can be measured. A desktop refractometer may also require complicated procedures and complex sample preparation prior to a measurement (Mehrotra & Siesler 2003).

2.10.2 Polarimeter

A polarimeter is an optical device used to measure the concentrations of active materials (pol value) in sugarcane juice. This instrument requires clarified juice for measurement. Clarified juice samples are obtained when raw juice is treated with lead acetate and then filtered to remove impurities (Mehrotra & Siesler 2003). Preparing juice samples can be difficult and time-consuming task because this process requires chemical reagents to be used for clarification purposes. Thus, this method seems to be unsuitable for use in the field. The robustness of this method is also affected by the presence of contaminants (McCarthy 2003).

2.10.3 Chromatography

Besides the above two methods, sugar content can also be measured using chromatographic methods such as high performance liquid chromatography (HPLC) (Campbell et al. 1999) or gas-liquid chromatographic (GLC) analysis (Meade & Chen 1985). These methods, however, can be affected by the presence of interfering compounds. Laborious sample pre-treatments are therefore required (Filho et al. 1996). These methods also require highly skilled personnel and relatively expensive equipment. They are often time-consuming, operator-dependent, and involve the use of hazardous reagents which can only be properly used in a laboratory (Mehrotra & Siesler 2003).

2.10.4 Biosensor

Another technology used to measure sugarcane quality is a biosensor. Biosensors are a reliable option for sucrose determination as they are relatively convenient, precise, fast and inexpensive. Several biosensors have been developed to measure sucrose in the sugarcane industry (Gouda et al. 2002; Kumar et al. 2006; Kennedy et al. 2007). However, most of the conventional biosensors are physically designed for laboratory use and are not intended for practical use in the field. The system of the sensor is also complicated with several small and fragile components such as a buffer reservoir, pump and injector (Figure 2.14), making this instrument insufficiently robust for field use. The operation of the device also requires a considerable level of skill because a reagent

is needed to be mixed with samples prior to measurement. For example, active graphite paste mixed with *tetracyanoquinodimethane* is required as a mediator for the sucrose measurement (Kennedy et al. 2007).

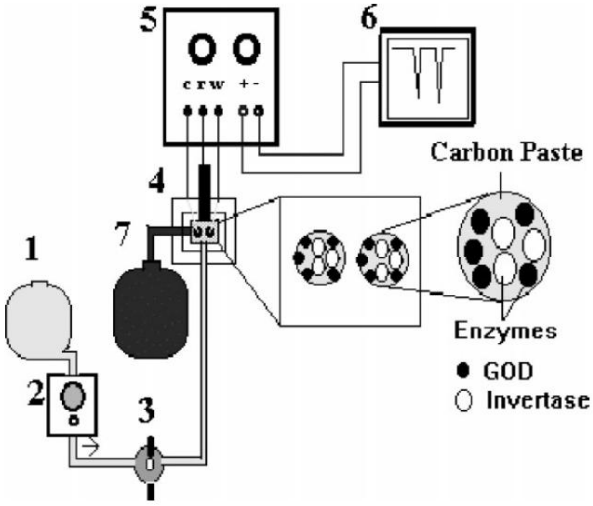


Figure 2.14: Flow injection manifold of biosensor for the determination of sucrose: (1) buffer reservoir, (2) peristaltic pump, (3) injector, (4) electrochemical cell, (5) potentiostat, (6) record, (7) waste reservoir (Kennedy et al. 2007)

2.10.5 NIR Spectrometer

Spectroscopy is an established method for determining chemical constituents in various agricultural products (Carlini et al. 2000). This method has been successfully used for both qualitative and quantitative measurements in the sugarcane industry (Madsen et al. 2003; Valderrama et al. 2007; O'Shea et al. 2011). Development of NIR applications in Australia is mainly due to the emergence of two NIR instruments developed by Bureau of Sugar Experimental Station (BSES) Ltd: the CAS (Cane Analysis System) and SpectraCane. The CAS instrument is an online system mainly used at sugar mills while the SpectraCane is an at-line system which measures prepared sugarcane stalks. Bevin et al. (2002) applied CAS instrument to analysis sugarcane quality property (pol) and reported excellent calibration and validation results with R^2 being 0.97 and 0.98 respectively.

From Table 2.2, it can be seen that NIR spectroscopy is capable of measuring sugarcane quality from both juice and non-juice (macerated) sample forms, whereas the other technologies are only capable of measuring juice samples. For examples, Berding et al. (1991b), Madsen et al. (2003) and Valderrama et al. (2007) used NIR spectroscopy to measure sugarcane quality from juice samples. The measurements of quality parameters from non-juice samples (fibrated sample forms) have been reported by Meyer & Wood (1988), Berding et al. (1991a) and Mehrotra & Siesler (2003). This is a special characteristic of NIR spectroscopy due to its ability to measure the variation of the light intensity either reflected or transmitted by both solid and liquid samples as a response to different chemical constituents of the samples.

However, instead of their powerful performances and ability to be employed for different purposes, there is no published study addressing the application of this technology for measuring sugarcane quality in the field, either from juice or non-juice sample forms. This is mainly because these laboratory spectroscopic technologies are not suitable to be used in the field considering the potential damage from the harsh and dusty field environment. The equipment is also heavy, fragile and expensive, thus requiring careful handling.

It can be concluded that the common technologies used to measure sugarcane quality in laboratories, such as refractometry, polarimetry, chromatography and laboratory NIR spectroscopy have great limitations in the field because they are often time-consuming, operator-dependent and require hazardous reagents (Mehrotra & Siesler 2003). Most of these existing technologies also require juice samples for the quality measurement. However, obtaining sufficient juice samples in the field is often very difficult, especially during harvesting. Another technical challenge for juice measurement in the field is to process a raw juice into clarified juice. Thus, a non-juice sample form can be regarded as a better alternative for field measurement.

The only technology that can be used to measure sugarcane quality from a non-juice sample (macerated forms) is the NIR spectroscopic method. However, the preparation of the sugarcane sample into a homogenous macerated form for field measurement is still technically difficult and can be a time-consuming task. This sample form also requires proper equipment to macerate the sugarcane samples prior to the measurement. Thus, the sugarcane sample in a macerated form still could not be regarded as the best sample form for field use.

2.11 Alternative technologies to measure sugarcane quality in the field

The previous section has discussed the potential and limitation of the existing laboratory technologies for measuring sugarcane quality in the field. In this section, alternative measurement technologies which have potential applications for measuring sugarcane quality parameters in the field are reviewed. The appropriateness of the identified technologies for field used is discussed below.

2.11.1 Electronic refractometer

An early attempt to measure sugar content during harvest was reported by McCarthy & Billingsley (2002). The authors developed a low-cost electronic refractometer together with a signal conditioning algorithm to measure sugar content for determining the optimum topping height of a sugarcane stalk during harvesting (Figure 2.15). The system worked well in a laboratory, producing repeatable and accurate results for sucrose concentrations. It was then mounted on the topper of a harvester in the field. Substantial field trials showed limited results because insufficient juice sample was deposited onto the sensor during harvesting. The poor results of this study were

attributed to a large amount of trash and leaf materials that hindered the freshly topped stalks from wiping across the sensor.

Hence, a mechanism which can discriminate trash from samples and squeeze a sufficient amount of juice is needed before this device can be used in the field. The study also indicated that a sufficient juice sample is very difficult to obtain during harvest. Thus, an alternative sample form, potentially sugarcane stalk, which is relatively easier to measure, should be investigated.

2.11.2 Microwave

Another potential technology for field application is the microwave. The applications of a microwave technology in the sugarcane industry to measure moisture level and sugar content have been studied by several researchers (Nelson 1987; Klute 2007; Shah & Joshi 2010). In a microwave measuring instrument, the components of agricultural products are polarised at varying strengths during microwave transmission through the product, resulting in the microwave signal losing speed and energy (Klute 2007). This phenomenon causes a phase shift in the signal, and the weakening of the energy affects an attenuation of the microwave. Then, if polar materials like water molecules are present in the product, this effect is disproportionately increased. The water influence on the signal is around 40-times greater than the influence of other components such as sugar (Klute 2007). Hence, since the microwave measurement is very sensitive to water concentration, the correlation between water content and dry substance could allow a very precise measurement of the sugar content for sugarcane samples.

Figure 2.15: Location of the electronic refractometer installed to measure sugarcane juice samples

The use of the microwave sensor to measure the moisture content of forage has also been reported by Marcotte et al. (1999). Klute (2007) reported that microwave technology can give an accurate and reliable measurement of most agricultural products, particularly sugar beet and sugarcane. The dry substance of sugar syrup can also be measured online using microwave technology. Overall, it appears that although this technology has become increasingly popular for laboratory use, there is no study regarding the use of this technology for in-field moisture or sugar content measurement. Microwave technology has the potential for use in the field, provided a suitable juice extraction mechanism can be developed.

2.11.3 Portable spectrometer

Another potential technology for field uses is a low-cost and portable photodiodes array (PDA) spectrometer. PDA spectrometers have been mounted on combine harvesters to measure the quality parameters of maize (Montes et al. 2006) and forage (Digman & Shinners 2008) during harvesting. PDA spectrometer is equipped with silicon (Si) or Indium Galium Arsenide (InGaAs) detectors. The wavelength range for Si detector is from 350 to 1100 nm while for InGaAs detector is between 400 and 1700 nm. Of these two detectors, the Si detector which is usually installed in a portable VNIR spectrometer is gaining popularity for field use because it is low-cost and portable enough for in-field measurements (Walsh et al. 2000).

The VNIR spectrometer has been used to predict sugar content from solid samples such as apples (Huang & Lu 2010), guavas (Hsieh & Lee 2005) and pineapples (Chia et al. 2012). The ability to measure quality parameters from solid samples can allow rapid and non-destructive measurement to be performed on intact fruit as complex sample preparation prior to the measurement is no longer required. This equipment has also been employed for the online measurement of soil moisture in the field by Kweon & Maxton (2013). Therefore, considering the cost, portability, ability to perform the measurement on intact samples and the ability to be used for online measurement, this review has concluded that a low-cost and portable VNIR spectrometer is a promising technology for field use, thus it is worth investigating as a best candidate for field application.

2.12 Conclusions

The sugarcane industry is very important to the Australian economy. In order to improve its crop production, PA techniques have been introduced into this industry. However, existing sugarcane yield monitors have no ability to monitor the quality of harvested sugarcane during harvesting. Since both yield and quality components are key economic parameters to the industry, the measurement of both components in the field are very important.

This chapter has reviewed the potential and limitation of various measuring technologies for in-field quality measurements. From the review, it has been found that a reliable sensor and suitable sample form are the determining factors for the

development of the in-field quality measurement system. A low-cost and portable VNIR spectrometer has been identified as the most promising technology to measure sugar content in the field. In terms of sample forms, a non-juice sample form (sugarcane stalk) is considered as the best sample form for field uses. Therefore, the following chapters will further investigate the application of the VNIR spectroscopic method to predict sugarcane quality parameters especially from stalk samples.

Chapter 3

Introduction to NIR spectroscopy and its application in the sugarcane industry

Chapter 2 explained why the development of an in-field quality measurement system is very important to the sugarcane industry. It also found that NIR spectroscopy may be the most promising method for field use. In this chapter, the theory of NIR spectroscopy, including available wavelength regions, different types of spectroscopic instrumentation and different sample presentations will be discussed. This chapter will also provide a brief review of the application of NIR technologies in the sugarcane industry. The application of the chemometrics methods is also described.

3.1 Introduction

The quality parameter of crops, along with the quantity is one of the most important criteria used by the agricultural industry to determine the price of the product and the payment to growers. An accurate and reliable measurement method for crop quality determination, especially in the field is critically required by the industry. Recently, optical techniques, particularly NIR spectroscopy have received significant attention and are widely adopted as a means to non-destructively determine the internal quality attributes such as sugar content, moisture content, firmness and the dry matter of agricultural product. NIR spectroscopic methods have long been used in agricultural and other industries because they are a fast, simple, low-cost and non-destructive analytical technique (Day & Fearn 1982; Yan et al. 2005). This method was first used in agricultural applications by Norris (1964) to measure moisture in grain. Since then, it has been widely used to measure quality parameters for a wide variety of fruits and vegetables (Bobelyn et al. 2010).

NIR spectroscopy has been found useful for the agricultural industry because, in most cases, it does not need any sample preparation prior to quality measurement (Huang et al. 2008). NIR spectroscopy is also easy to implement in online and off-line applications. Moreover, NIR spectroscopy has the potential for simultaneously measuring multiple quality attributes of intact fruits, making it suitable for use with solid samples (Lu & Ariana 2002). All of these characteristics of NIR spectroscopy are very useful for the CCS measurement of sugarcane which is made up of several quality parameters including °Brix, pol and fibre content.

3.2 Theory of NIR spectroscopy

The conceptual basis of optical spectroscopy, also known as vibrational spectroscopy, is that at temperatures above absolute (0°K or – 273°C), all atoms in molecules are in continuous vibration with respect to each other. When the frequency of a specific vibration is equal to the frequency of the NIR radiation directed at the molecule, the

molecule absorbs the radiation (Sherman 1997). When these vibrating molecules absorb light of a particular frequency, they are excited to a higher energy level (Ingle & Crouch 1988). These atom to atom bonds within molecules vibrate with frequencies that may be described by the laws of physics and can therefore be quantified.

At room temperature, when most molecules are vibrating at the least energetic state allowed by quantum mechanics, the fundamental frequency of any two atoms connected by a chemical bond can roughly be calculated by assuming that the energies arise from the vibration of a diatomic harmonic oscillator and obey Hooke's law. The vibrational frequency can be calculated as follows (Afara 2012):

$$v = \frac{1}{2\pi}\sqrt{k\frac{(m_1 + m_2)}{m_1.m_2}} \tag{3.1}$$

Where v is the vibrational frequency, k is the classical force constant, m_1 and m_2 are the mass of vibrating/bonding atoms.

Unlike the classical spring model in Eqn. 3.1, the vibration of molecules, E_{vib} can be described using the harmonic oscillator model, by which the energy of the different, equally spaced levels can be calculated from (Blanco & Villaroya 2002):

$$E_{vib} = \left(v + \frac{1}{2}\right)\frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \tag{3.2}$$

Where v is the vibrational quantum number, h is the Planck constant, k is the force constant, and m is the reduced mass of the bonding atoms.

Only those transitions between consecutive energy levels ($\Delta v = \pm 1$) that cause a change in dipole moment are possible. Thus, E_{vib} can be rearranged as follow:

$$\Delta E_{vib} = \Delta E_{rad} = hv \tag{3.3}$$

where v is the fundamental vibrational frequency of the bond that yields an absorption band in the MIR region. However, the harmonic oscillator model cannot explain the behaviour of actual molecules as it does not take into account the Coulombic repulsion between atoms or dissociation of bonds (Blanco & Villarroya 2002). The Coulombic repulsion is generated when the respective electron clouds, as well as the charges on the nuclei, of the two bound atoms limit the approach of their nuclei during the compression step, creating an energy barrier, while dissociation occurs at the extension step (Ciurczak 2001). As a result, the behavior of molecules more closely resembles the model of an anharmonic oscillator, by which energy levels are not equally spaced. Thus, energy difference decreases with increasing vibrational frequency (Eqn. 3.4).

$$\Delta E_{vib} = hv\,[1 - (2v + \Delta v + 1)y] \quad (3.4)$$

where *y* is the anharmonicity factor.

The anharmonicity can result in transitions between vibrational energy states where $\Delta v = \pm 2, \pm 3, \ldots$ As a result of this anharmonicity, additional absorption bands are generated by the appearance of overtones (approximately integral multiples of the fundamental absorption frequencies), combinations of fundamental frequencies, differences of fundamental frequencies, coupling interactions between two fundamental absorption frequencies, and coupling interactions between fundamental vibrations and overtones or combination bands (Fermi resonance) (Sherman 1997). These frequencies are particular to NIR and are much less likely than the fundamental transitions.

These transitions between non-contiguous vibrational states yield absorption bands known as overtones (first and second overtones, respectively) at multiples of the fundamental vibrational frequency (Afara 2012). Also, they are much less likely than the fundamental transitions, so the bands are much weaker (the band for the first overtone is 10–100 times weaker than that for the fundamental frequency, depending on the particular bond) (Blanco & Villarroya 2002). These overtone bands typically occur in the wavelength region between 780 and 2000 nm.

In polyatomic molecules, two or more vibrational modes can interact between them to cause simultaneous energy changes and give rise to absorption bands called combination bands; the frequencies of which are the sums of multiples of each interacting frequency (Blanco & Villaroya 2002). The combination bands of NIR occur in the wavelength region from 1900 to 2500 nm. The intensity of NIR bands depends on the change in dipole moment and the anharmonicity of the bond. Because the hydrogen atom is the lightest, and therefore exhibits the largest vibrations and the greatest deviations from harmonic behavior, the main bands typically observed in the NIR region correspond to bonds containing this and other light atoms (namely C–H, N–H, O–H and S–H).

Interactions between atoms in different molecules can change vibrational energy states, thus shifting existing absorption bands and giving rise to new ones, through differences in crystal structure. This allows crystal forms to be distinguished and physical properties (such as density, viscosity, and particle size in pulverulent solids) to be determined (Blanco and Villarroya 2002). This indicates that the NIR spectrum contains not only the chemical information of the sample, but also physical information that can be employed to determine physical properties of samples. Thus, a specific spectral region that illustrates the energy absorbed by each, or group, of vibrating molecules in sugarcane stalk samples can be obtained using a spectrometer.

In relation to the principles discussed above, it is clear that the sugar content of a sugarcane crop, which is represented by O-H bonding, can be determined using the NIR spectroscopy. This spectroscopic method works based on the principle that, when a light beam hits the stalk, a small fraction is reflected at the surface as a specular reflectance

and the rest will penetrate into fruit tissues. In the tissues, photons are absorbed or migrate in different directions where radiation will be scattered backward to the surface as diffuse reflectance, while the remaining radiation migrates forward into the tissues as absorbance (Qing et al. 2007). Light absorption is related to certain chemical constituents, such as sugar, acid, water, etc. (Williams & Norris 2001). A reflectance spectrometer measures the aggregate amount of light reflected from a sample, from which light absorption may be estimated and then related to certain chemical constituents of the crop (Lu 2004).

3.3 Wavelength selection

The human eye can only respond to electromagnetic radiation within a wavelength range between 400 and 700 nm (visible spectrum). However, the electromagnetic spectrum extends from the extremely short wavelength of gamma radiation to the long wavelength of radio waves (Figure 3.1). Along these wavelength regions, there are several different regions often used for quality measurement including ultraviolet (UV: 100 - 350 nm), visible (Vis: 350 - 780 nm), near infrared (NIR: 780 - 2500 nm), visible and shortwave NIR (VNIR: 350 - 1100 nm), full range (FR: 350 - 2500 nm) and mid infrared (MIR: 2500 - 25000 nm). Out of these, the most prominent regions used for quality measurement are those based on the infrared region of the electromagnetic spectrum, notably NIR and MIR.

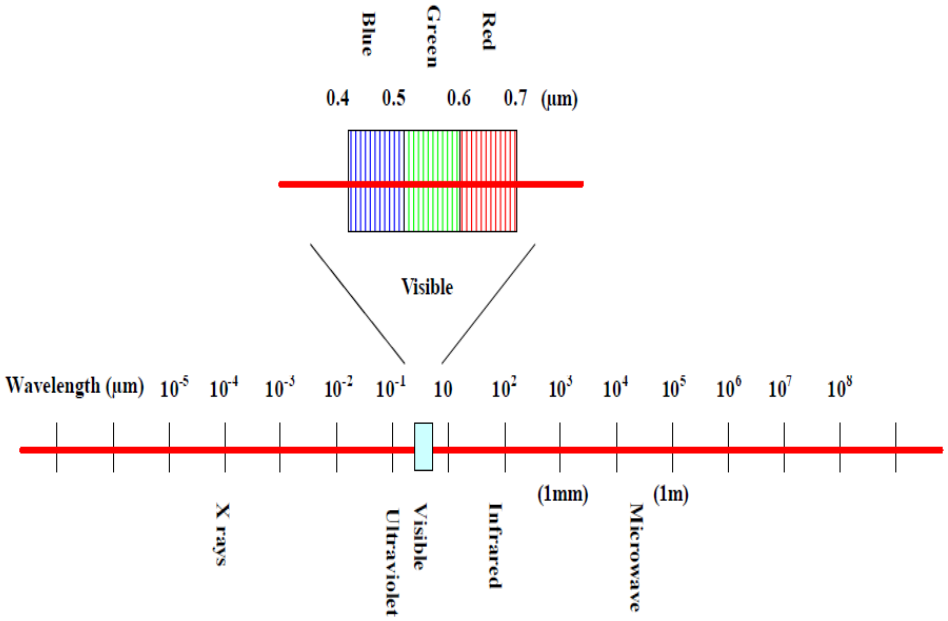


Figure 3.1: The schematic diagram of the electromagnetic spectrum

In the range of visible light (Vis), absorptions are caused by electronic transitions. Absorptions in the NIR region are caused by vibrations and combination overtones of the fundamental O–H, C–H and N–H bond vibrations that occur in the MIR region

(Williams & Norris 2001). In the MIR range, the absorption bands can be interpreted by various vibration modes of analysed molecules (Quilitzsch et al. 2005). The NIR region can be divided into shortwave NIR (SW-NIR) and common NIR at 1300 nm. The SW-NIR region is considered as the absorption band of high overtones, while the latter belongs to the first or second overtones. The absorption intensity will decrease when the overtone increases. Thus, SW-NIR is usually applied in the transmission analysis with long path length, and common NIR is used in diffuse reflection analysis (Cen & He 2007). Walsh et al. (2000) found that the application of the SW-NIR spectrometer (700-1100 nm) was promising for the following reasons:

1. The bands are ascribed to the third and fourth overtones of O–H and C–H stretching modes and are expected to be separated due to anharmonicity,
2. Lower absorbance at these wavelengths allows for transmission optics,
3. The corresponding instrumentation is low-cost and suited to process control, and portable enough for *in situ* field measurements.

Reflectance spectrometers have been extended from the NIR region (780 - 2500 nm) to the Vis region (350 - 780 nm) and these instruments have been available for research and industrial purposes (McCaig 2002). Combining both the Vis and NIR regions in one instrument provides sufficient improvement in efficiency related to instrumental, sampling and analytical cost (McCaig 2002). The Vis-NIR has mainly been used for evaluating internal fruit quality in terms of internal defects and biochemical compounds such as acidity, dry matter, starch, total soluble solids (Butz et al. 2005).

In addition to Vis-NIR spectrometer, recent spectrometers have been developed to function in both Vis and SW-NIR (VNIR) regions for better field application. The VNIR region (350 - 1050 nm) is weakly absorbed by biological material, with this absorption typically associated with second, third, and fourth overtone vibrations. Therefore, this range of region can be applied to thicker samples (path lengths of 1 to 100 mm) and can yield information on the internal attributes of biological material such as the sugar and water content of fruit (Kawano et al. 1998). Furthermore, shorter wavelengths allow better penetration of biological samples (Kawano et al. 1993) which will be very useful in the assessment of sugarcane stalk samples.

Several studies have been carried out using spectroscopic techniques in the VNIR region to measure the sugar content of intact fruits including apples (Ventura et al. 1998, Huang & Lu 2010), kiwifruits (Moghimi et al. 2010), guavas (Hsieh & Lee 2005) and pineapples (Chia et al. 2012). Moons et al. (1997) suggested that improved predictions could be achieved by measuring reflectance in both visible and NIR region from 400 nm to 2500 nm. However, from a technical viewpoint, such a system is difficult and costly to implement for in-field applications because it will require two different detectors to cover the entire spectral region (Lu et al. 2000). Thus, even though a greater spectral region beyond 1100 nm appeared to give improved prediction results in sugar content and other quality attributes (Moons et al. 1997; Lu et al. 2000), the equipment may be significantly more expensive.

3.4 Instrumentation

Regardless of the wavelength regions, all spectrometers are essentially identical to each other. A spectrometer collects spectral data by detecting changes in absorption intensity as a function of frequency. NIR spectroscopy instrumentation has evolved dramatically in response to the need for speed in analyses and flexibility in adapting to different samples and environments. Latest NIR equipment usually includes a variety of devices and accessories such as a light source (usually a tungsten halogen light bulb), sample presentation accessory, monochromator, detector, and optical components, such as lenses, collimators, beam splitters, integrating spheres and optical fibres. The uses of these accessories depend on the characteristics of the samples and the particular analytical conditions and needs, such as speed, sample complexity and environmental conditions, making the technique very flexible (Blanco & Villarroya 2002). The common types of spectrometer used in the agricultural industry as listed by Nicolaï et al. (2007) are discussed in the following section.

3.4.1 Filter instrument

Filter instruments are the simplest NIR instruments which contained a limited number of wavelengths, usually between six and twenty of interference filters. These filters are chosen to represent the absorptions used for the most popular applications, e.g. protein, moisture and oil in agricultural samples (Osborne et al. 1993). Filter instruments are designed for a limited range of routine analyses, either in the laboratory or online.

3.4.2 Scanning monochromator instrument

The scanning monochromator typically comprises a dispersive medium, entrance and exit slits, and imaging components that produce a parallel beam path. To record a spectrum, a detector located behind the exit slit must sequentially record the incident light while the dispersive component or the exit slit is moved (Mouazen et al. 2005). In a scanning monochromator instrument, a grating or a prism is used to separate the individual frequencies of the radiation either entering or leaving the sample. The wavelength separator rotates so that the radiation of the individual wavelengths subsequently reaches a detector (Nicolaï et al. 2007). Scanning monochromators can be used to measure the Vis and NIR spectrum, and may be used in either transmittance or reflectance mode. Usually, monochromator instruments are equipped with three different detectors: a Si detector covering the range of 400-1100 nm, an InGaAs covering the range of 800-1700 nm and a lead sulfide (PbS) covering the range of 1100-2500 nm (Osborne et al. 1993).

3.4.3 Fourier transform spectrophotometer

Fourier transform spectrophotometers use an interferometer to generate modulated light. The time domain signal of the light reflected or transmitted by the sample onto the sample can be converted into a spectrum via a fast Fourier transform. Often a

Michelson interferometer is used, but polarisation interferometers are also used in some spectrophotometers ((Nicolaï et al. 2007).

3.4.4 Photodiode array (PDA) spectrophotometer

PDA spectrophotometers employ an array of IR-emitting diodes. The PDA functions as both the light source and the wavelength selection system. PDA instruments typically cover the range of 400-1700 nm. These instruments are very useful to be used when a high sample throughput or ultra-rapid online measurements are required. PDA instruments use no moving parts for wavelength selection and have the advantage of being able to scan at high speed (Osborne et al. 1993).

3.4.5 Acoustic optic tunable filter (AOTF) instrument

AOTF is another type of dispersive monochromator developed for spectroscopic measurements. An AOTF comprises a *birefringent* crystal of TeO_2 through which a plane travelling acoustic wave is generated at right angles to the incident light beam. This causes the crystal to behave as a longitudinal diffraction grating with a periodicity equal to the wavelength of sound across the material (Osborne et al. 1993). The main advantages of AOTF over grating instruments are their mechanical simplicity (i.e. no moving parts) and their wavelength stability. The absence of moving parts in AOTFs ensures more reliable, reproducible wavelength scans than those provided by other devices. This makes AOTFs suitable for equipment subject to aggressive conditions such as in production plants (Blanco & Villarroya 2002).

3.5 Instrument for field applications

The application of NIR spectrometers in a laboratory has been well established for a variety of crop production. However, the application of such technology in the field still faces some technical difficulties, including size limitation, power requirement and high sensitivity to dust and vibration. As there is an increasing demand for measuring the quality of agricultural product in the field, finding a suitable technology for field use is essential. From Section 3.4, it was found that the PDA spectrometers are gaining considerable attention for field use because of their high acquisition speed (the integration time is typically 50 ms but can be as low as a few milliseconds) and the absence of moving parts enables them to be mounted on online fruit grading lines (Nicolaï et al. 2007). Fernández-Ahumada et al. (2008) suggested the use of PDA instruments in real online industrial environments because of their speed of response and the low price advantage. In addition, the specification requirement for spectroscopic assessment of fruit in an in-line setting includes high signal to noise ratio, relatively high sensitivity (signal per amount of light), good repeatability and tolerance to vibration and dust (Walsh et al. 2000).

Furthermore, PDA spectrometer equipped with a Si detector is approximately 100 times less sensitive to light than charge-coupled device (CCD) Si detectors. The higher saturation level of the photodiode supports a 10-fold higher maximum signal to noise

ratio for this detector, relative to CCD detectors (i.e. 10 000 *cf*. 1000) (Walsh et al. 2000). Thus, CCDs are preferred for very low light applications, while PDA spectrometers are the better choice for accurate absorbance measurements when higher light levels are available (Oriel 1997). PDA spectrometer operating in the NIR region are potentially useful for process measurement because they are more rugged and better suited to online applications, even under aggressive conditions (Fernández-Ahumada et al. 2008). Some early studies addressed the use of this technology mounted on harvesters for the analysis of grains and forages (Dardenne & Femenias 1999; Von Rosenberg et al. 2000; Paul & Pfitzner 2004).

In PDA spectrophotometers, a fixed grating focuses the dispersed radiation onto an array of Si or InGaAs detectors. In addition, a lead sulfide (PbS) detector can also be employed for the wavelength region between 1100 and 2500 nm. However, the InGaAs detector has a much faster response time and better sensitivity than the PbS. In terms of cost, the Si detectors are preferred because they are relatively cheap, as the cost of the instrument rises as the spectral range increases (Shenk 2004). Recent decreases in cost and improvement in performance of Si detectors has offered a range of potential applications of relevance to plant physiologists, associated with spectral analysis in the VNIR region (Walsh et al. 2000).

As well as the cost of a detector, the size and portability of the equipment are also important factors to be considered for field use. During the mid- to late 1990s, several low-cost (<A$10 000), miniature (spectrometer size <500 cm^3) PDA spectrometers in the VNIR region became commercially available (Walsh et al. 2000). These types of instruments have been applied in a range of instrumentation of interest to plant physiologists (e.g. portable spectroradiometers). The availability of low-cost and portable spectrophotometers has opened up the possibility of using such technologies in the orchard for monitoring the maturity of fruit (Nicolaï et al. 2007). The applications of portable spectrometers for field uses have been described by Temma et al. (2002), Saranwong et al. (2003), Miller & Zude-Sasse (2004), and Zude et al. (2006). Portable spectrometers include hand-held instruments and equipment that can be carried in a backpack or mounted on a tractor or combine harvester (von Rosenberg et al. 2000). Detailed applications of the PDA spectrometer for quality measurement in the field are discussed in Chapter 8.2.

Most of the portable spectrometers use Si detectors as they are much cheaper than InGaAs detectors. In most cases, spectrometers with Si detectors yield poorer prediction accuracy than InGaAs detectors. This is due to issues of robustness, such as temperature fluctuations and the limited wavelength range for Si detector (Walsh et al. 2000). However, despite the lower accuracy obtained by a spectrometer with Si detector, due to its low-cost, robustness for field use and portability, the application of this instrument warrants further investigations for application in the sugarcane industry. Improvements in sample presentation, spectral measurement and statistical methods could improve the prediction accuracy of this equipment, thus making the proposed method more attractive to the industry for a field test.

3.6 Common quality measurement systems in the field

One of the main challenges in developing a quality monitor for the sugarcane industry is to find a robust sensor which can survive harsh harvesting operation over long periods of time. Zhang et al. (2002) highlighted that a robust, low-cost, and, preferably, real-time sensing systems are needed for the quality measurement system in the field. The literature review in Chapter 2 indicated that the most suitable technologies for measuring crop quality parameters in the field would be a spectrometer. The specific requirements for a spectrometer for field use include; rapid spectral acquisition, high signal to noise ratio, relatively high sensitivity, and tolerance to vibration and dust (Walsh et al. 2000).

To meet some of the above requirements, PDA spectrometers with a stationary dispersive element and a fixed detector array can be used as they are very robust in terms of wavelength reproducibility, and very rapid in terms of spectral acquisition (Walsh et al. 2000). PDA spectrometers have already been widely used for in-field quality measurement (Table 3.1). The high acquisition speed (integration time well below 100 ms) of modern PDA spectrometers have finally made in-line quality measurement possible (Welle et al. 2003).

Table 3.1: Commercially available spectrometers used for quality measurement in the field

Authors	Wavelength (nm)	Detector type (Brand)	Measurement time	Sample	Quality parameters
Montes et al. (2006)	960 - 1690	PDA (Zeiss Corona 45 NIR)	1 scan/sec	Maize	Dry matter Crude protein Starch content
Wright et al. (2002)	400 - 1700	PDA (Not indicated)	Not indicated	Grain	Moisture content Starch content Protein content
Digman & Shinners (2008)	950 - 1680	PDA (Zeiss Corona 45 NIR)	5 scan/sec	Forage	Crop moisture
Welle et al. (2003)	960 – 1690	PDA (Zeiss Corona 45 NIR)	1 scan/sec	Maize	Dry matter Starch content Soluble sugar
Mouazen et al. (2005)	307 - 1711	PDA (Zeiss Corona 45 NIR)	0.2 scan/sec	Soil	moisture content
Kweon & Maxton (2013)	660 - 940	PDA (Not indicated)	1 scan/sec	Soil	Organic matter

From Table 3.1, it can be seen that PDA spectrometers have been successfully applied to measure quality parameters of different samples in the field. Even though a majority of the studies used InGaAs detectors, a study reported by Kweon & Maxton (2013) applied a Si detector for detecting soil organic matter. The authors did not use the InGaAs detector because it was more expensive than the Si detector and requires additional hardware to control the temperature (Kwoen & Maxton 2013). For the same

reasons of cost and durability, the PDA spectroradiometer with a Si detector is proposed for further investigation in this study.

3.7 Sample presentation

The application technique of spectroscopic methods is varied depending on the purposes of use. Typically, the measurement of optical properties of a fruit is based on reflectance, transmittance, absorbance, or scatter of light (light scattering) by the product (Abbott 1997). When a fruit is exposed to light, about 4% of the incident light is reflected at the outer surface, causing specular reflectance or gloss, and the remaining 96% of incident energy is transmitted through the surface into the cellular structure of the product where it is scattered by the small interfaces within the tissue or absorbed by cellular constituents (Figure 3.2) (Birth 1976).

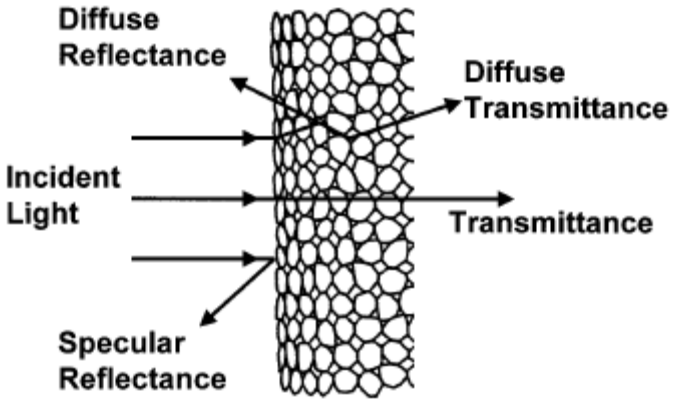


Figure 3.2: Incident light on a fruit results in specular reflectance (gloss), diffuse reflectance, diffuse transmittance, or absorbance (Abbott 1999)

The complex physical structure of tissues creates an optically dense zone that is difficult to penetrate and alters the path length travelled by the light so that the amount of tissue interrogated is not known with certainty. Most light energy penetrates only a very short distance and exits near the point of entry, while some light penetrates deeper (usually a few mm, depending on optical density) into the tissues and is altered by differential absorbance of various wavelengths before being exited (Abbott 1997). Such light may be called diffuse reflectance, diffuse transmittance, or interactance and contains useful chemometrics information (Figure 3.2) The typical experimental setup for reflectance, transmittance and interactance is shown in Figure 3.3.

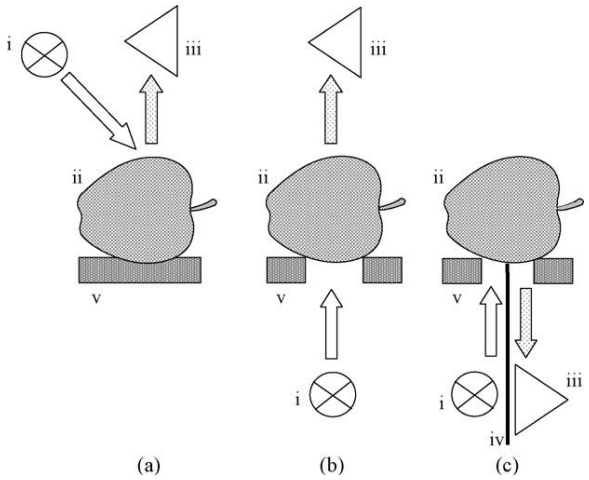


Figure 3.3: Setup for the acquisition of (a) reflectance, (b) transmittance, (c) interactance spectra; (i) the light source, (ii) fruit, (iii) monochromator/detector, (iv) light barrier, (v) support (Nicolai et al. 2007)

In reflectance mode (Figure 3.3a), a light source and a detector are mounted at a specific angle, e.g., 45°, to avoid specular reflection. In this mode, the field of view of the detector includes parts of the fruit surface directly illuminated by the light source (Mowat & Poole 1997). The spectral information is collected from within, and in the vicinity of, the incident area. Reflectance mode measurements are the easiest to perform because they require no contact with the fruit and light levels are relatively high (Schaare & Fraser 2000). This sensing mode has been used in several studies for measuring the quality attributes of intact apples (Lammertyn et al. 1998; Moons et al. 1997; Lu et al. 2000).

In transmittance mode (Figure 3.3b), a light source and a detector are positioned opposite each other and the incident light has to pass through the whole sample being measured. Therefore, because of long photon path lengths as well as strong scattering and absorption in the fruit, transmittance measurements require a higher intensity light source to penetrate the sample, which depends on the absorption coefficient of the sample and the sample thickness. Transmittance mode is commonly applied for liquids, solutions and suspensions. Transmittance measurement on solid samples has also been investigated. For instance, Kawano et al. (1993) used NIR transmittance mode to determine the sugar content of mandarins.

Typically, reflectance or transmittance spectrum of fruit contains information on both absorption as well as scattering. While absorption is related to the presence of chemical components, scattering is related to the microstructure of the tissue. The main scattering elements in fruit and vegetables are the cell wall interfaces since they induce abrupt changes in the refractive index (McGlone et al. 1997). The scattering is also dependent on the size, shape and microstructure of the particles. Scattering may also appear due to heterogeneities such as pores, openings, capillaries that are randomly distributed through the sample. Multiple scattering events largely determine the intensity of the scattered light that is emitted (McGlone et al. 1997).

In interactance mode (Figure 3.3c), the light source and detector are positioned parallel to each other in such a way that light, due to specular reflection, cannot directly enter the detector. This can be achieved by means of a bifurcated cable in which fibres leading to the source and detector are parallel to each other and in contact with the product, or by means of a special optical arrangement (McGlone et al. 2002). Interactance mode provides a compromise between reflection and transmission modes in each of these characteristics, but obtaining a light seal may be problematic at the high conveyor speeds used in modern fruit grading systems. This type of arrangement is particularly useful for large samples such as intact fruit.

In addition, transflectance mode is also applied for liquids measurement (Kawano 2002). Transflectance is actually a combination of transmittance and reflectance modes. In transflectance mode, the incident light is transmitted through the sample, reflected from the reflector (normally made of ceramic or aluminium) which is placed beneath the samples, then transmitted back through the sample before finally reaching the detector (Osborne et al. 1993).

Since this thesis will mainly focus on the measurement of sugarcane quality from stalk samples, the reflectance method which requires no contact with fruit samples has the potential application in this study. While both transmittance and transflectance measurement modes have the potential to be used for juice measurement. Therefore, this study will explore the potential application of those three methods for measuring sugarcane quality.

3.8 Spectroscopic application in the sugarcane industry

In the sugarcane industry, pioneering work carried out by Meyer & Wood (1988) established the fact that NIR spectroscopy was a rapid method for sugarcane juice analysis. Analysis of expressed juice from fibrated sugarcane for °Brix, pol, fibre, and moisture content was also conducted by Berding et al. (1991b). Since then, many studies have been undertaken using NIR spectroscopy for both qualitative and quantitative measurements (Madsen et al. 2003; Valderrama et al. 2007; O'Shea et al. 2011).

Laboratory NIR systems have been developed for applications in various factory streams including molasses, mixed juice and massecuites (Schäffler & De Gaye 1997; Simpson & Naidoo 2010). The development of online NIR systems for the analysis of sugarcane, bagasse and sugar in factories has also been reported (Staunton & Wardrop 2006). Furthermore, data from cane analysis systems (CAS) have been used to develop online analysis systems for fibre analysis (Staunton et al. 1999), sugarcane payment systems (Pollock et al. 2007), sugarcane quality schemes (Pope et al. 2004) and process control purposes using fibre rate control (Jones et al. 2002).

Key published studies regarding the application of spectroscopic methods in the sugarcane industry are summarised in Table 3.1. Different studies have been conducted

using different wavelength regions, measurement modes and sample forms. A suitable sample form and reliable equipment with the right wavelength regions are two important factors to be considered for field measurement. The wavelength region of the spectrometer would normally determine the price, detector type, prediction accuracy, size and portability of the equipment. The price and size of the equipment would certainly influence the industry's decision toward the adoption of any proposed in-field measurement method. Expensive equipment will put financial pressure on the industry, particularly on growers. The complexity of the instrument would also limit the application of the instrument in the field or mounted on a harvester. The most suitable spectrometer to measure sugarcane quality in the field should first be identified. From Table 3.2, the selection of a suitable wavelength regions and a sample form are given special attention.

Table 3.2: Typical application of spectroscopic methods in the sugarcane industry

Authors	Wavelength, nm	Measurement mode	Sample form	Prediction accuracy
Berding et al. (1991a)	NIR (1445 - 2348)	reflectance	Fibrated	°Brix (R^2 = 0.91) CCS (R^2 = 0.91) Fibre content (R^2 = 0.89) Pol (R^2 = 0.96)
Berding et al. (1991b)	NIR (1445 - 2348)	transflectance	Clarified juice	°Brix (R^2 = 0.97) CCS (R^2 = 0.97) Pol (R^2 = 0.98)
Cadet & Offman (1997)	MIR (8000 - 12500)	Reflectance	Raw juice	Pol (R^2 = 0.98)
Mehrotra & Siesler (2003)	NIR (1111 - 2222)	Transmittance	Raw juice	Pol (R^2 = 0.96)
	NIR (1111 - 2500)	Reflectance	Fibrated	Pol (R^2 for the calibration model = 0.93) * Accuracy for prediction model was not reported
Madsen et al. (2003)	NIR (1100-2500)	Reflectance	Fibrated	Pol (R^2 = 0.96) °Brix (R^2 = 0.97) Fibre content (R^2 = 0.90)
Valderrama et al. (2007)	NIR (1100 - 2500)	Transmittance	Clarified juice	°Brix (R^2 = 0.99) Pol (R^2 = 0.99)
Taira et al. (2010)	NIR (1100 - 2498)	Reflectance	Fibrated	Pol (SEP = 0.21%)
Taira et al. (2013)	VNIR (600 – 1100)	Transmittance	Stalk	Pol (RMSEP = 1.1% Pol)

3.8.1 Selection of the wavelength region

Except for work reported by Taira et al. (2013), all of the studies reported in Table 3.2 were carried out using laboratory NIR spectrometers with different wavelength regions. Unfortunately, traditional and modern NIR laboratory instruments have a number of constraints (e.g., lack of robustness leading to poor adaptability to harsh environments, low scanning speed, high cost, non-portability) that hinder their use in the field (Fernández-Ahumada et al. 2008). However, recently developed PDA spectrometers

offer a range of advantages, including ability to record a full spectrum at high speed, lack of moving parts, wavelength repeatability and compatibility with fibre optics for flexible process interfacing, and lower price (Lindström et al. 2004).

Latest PDA spectrometer is normally equipped with a Si detector, working in the VNIR region (350-1100 nm). The VNIR PDA spectrometer features relevant to the sugar (C-H) groups include a third overtone band at 910 nm (Golic et al. 2003). The bands at 954 and 976 nm are assigned to the second overtones of sucrose (O-H) stretching vibrations (low and high H bonded states, respectively), while the 906 nm band is ascribed to a third overtone of the CH stretching vibrations (Golic et al. 2003). Due to its capability, low-cost and portability, this type of spectrometer is potentially more economical and practical for use by farmers in the field. However, there is no published study reporting the use of this low-cost and portable VNIR PDA spectrometer for predicting sugarcane quality parameters in the field especially using reflectance measurement mode. Reflectance mode has the potential to be adopted for scanning the moving billets on the elevator.

3.8.2 Selection of the suitable sample forms

Table 3.1 shows that the spectroscopy methods have been applied to measure quality parameters using different sample forms including fibrated samples (Berding et al. 1991a; Mehrotra & Siesler 2003; Madsen et al. 2003, Taira et al. 2010), raw juices (Cadet & Offman 1997; Mehrotra & Siesler 2003), stalk samples (Taira et al. 2013) and clarified juices (Berding et al. 1991b; Valderrama et al. 2007). Since different sample forms required different sample preparations and gave different prediction accuracies, the selection of the optimum sample form for field use is very important. Nevertheless, the optimum sample form could not simply be selected from the previous studies because those studies were carried out using different wavelength regions, measurement modes and the calibration models were developed against different quality parameters. Thus, the selection of the optimum sample form relative to the type of selected spectrometer should be the first step in this investigation. The identification of the optimum sample form will be a basis for the design and development of a sampling and measurement system at the appropriate location on a harvester.

The optimum sample form should be assessed based on several criteria including ease of measurement (requires no or minimal sample preparation), speed of measurement (can be performed rapidly on moving samples, e.g. billet samples on the elevator) and easy to be collected on a harvester. From these criteria, a non-destructive and rapid technique which could be performed on solid samples (non-juice) would be highly desirable. It may be argued that the simplest method for quality measurement in the field is by directly scanning sugarcane stalks or billets. During harvest, stalk samples in the form of billets (solid samples) can potentially be scanned on its skin or on cross sectional surfaces with relative ease as no preparation mechanism is needed prior to measurement. To date, there is no published study regarding the use of spectroscopic methods to determine quality parameters from fresh stalk samples based on reflectance measurement.

3.9 Introduction to chemometrics

NIR instruments provide vast amounts of spectral data which require efficient, rapid and robust data processing procedures in order to yield useful analytical information. The analytical information contained in the typically broad and extensively overlapped band of NIR spectra is hardly selective and is often influenced by a number of physical, chemical and structural variables (Blanco & Villaroya 2002). In addition, differences between samples may cause spectral variations that are difficult to distinguish with the naked eye (Nicolaï et al. 2007). The NIR spectrum may be further complicated by wavelength dependent scattering effects, tissue heterogeneities, instrumental noise, ambient effects and other sources of variability (Nicolaï et al. 2007). Therefore, the application of chemometrics in NIR analysis is very important in order to extract the hidden information from spectral data and to reduce irrelevant information.

Multivariate data analysis (also called chemometrics) is defined as the development and application of mathematical and statistical methods to extract useful chemical information from chemical measurements (Barton & Kays 2001). The whole idea of chemometrics is to reduce the data that contain redundant information. The reduced data are easier to understand and are more stable (Geladi 2003). The new latent variables (LVs) in the reduced data set provide an overview of the main variation in the raw data. The outliers can be identified and the noise can be separated from the original data (Wold et al. 2001). The chemometrics method has been widely applied in NIR analysis to perform spectral data pre-processing, reduce irrelevant variables (variable reduction) and build calibration models for quantitative and qualitative analysis (Cen & He 2007).

3.9.1 Pre-processing of the spectra

The spectral data of solid samples are influenced by the physical properties (skin roughness, etc) of the solid samples. This poses some problems in evaluating the internal quality attributes of the samples. Furthermore, the spectral data normally contains background information such as light scattering, path length variations and random noise as well as sample information. In order to obtain reliable, accurate and stable calibration models, it is essential to pre-process spectral data before modeling (Cen & He 2007). Spectral pre-processing techniques are required to remove any irrelevant information including noise, uncertainties, variability, interactions and unrecognized features. Several pre-processing techniques have been developed recently. Nicolaï et al. (2007) divided the pre-processing methods into categories described below:

3.9.1.1 Smoothing

Smoothing techniques have been proposed to remove random noise from NIR spectra. Smoothing is also necessary to optimise the signal-to-noise ratio (Cen & He 2007). The

most common smoothing techniques are moving average filters and the Savitzky–Golay algorithm (Næs et al. 2004).

3.9.1.2 Standardisation

A standardisation technique is used to divide the spectrum at every wavelength by the standard deviation of the spectrum at this wavelength. Typically, variances of all wavelengths are standardized to 1, which results in an equal influence of the variables in the model (Næs et al. 2004).

3.9.1.3 Normalisation

A normalisation technique is applied to compensate for additive (baseline shift) and multiplicative (tilt) effects in the spectral data, which are induced by physical effects such as the non-uniform scattering throughout the spectrum as the degree of scattering is dependent on the wavelength of the radiation, the particle size and the refractive index. Multiple scatter correction (MSC) and standard normal variate correction (SNV) are the most popular normalisation techniques (Næs et al. 2004). MSC could remove the effects of scattering by linearising each spectrum to some 'ideal' spectrum of the sample, which, in practice, corresponds to the average spectrum. In SNV, each individual spectrum is normalized to zero mean and unit variance.

Light scattering (mainly caused my particle size) and nonlinearity of instrument response are very important sources of inaccuracy in the determination of chemical and structural information by diffuse reflectance spectroscopy (Boysworth & Booksh 2001). In a sugarcane solid sample, the scattering effect is more obvious due to their glossy skin. Therefore, the MSC method is used to correct for the significant light scattering problems.

Scattering theory is based on the proposition that scattering should have a multiplicative effect on the reflectance signal. That is, the observed spectra will contain a broad, changing background from differential scattering at each wavelength. The MSC algorithm was developed to reduce this light effect on diffuse reflectance NIR spectra. This scattering was determined by regressing each spectrum onto the mean spectrum, where the scattering at the *j*th wavelength of the sample can be modeled by (Afara 2012):

$$x_j = a + b\bar{x}_j + \epsilon_j \quad (3.5)$$

Where a and b are constants for all j wavelengths in the sample. The scatter-corrected spectra are determined by the scaled deviations about the regression, as shown below:

$$x_j, MSC = \frac{(x_j, raw - a)}{b} \quad (3.6)$$

SNV removes the multiplicative interferences of scatter, particle size, and the change of light distance as well as MSC (Barnes et al. 1989). SNV corrects both multiplicative

and additive scatter effects. To remove slope variations on an individual spectrum basis, each object is transformed independently using the following equation (Barnes et al 1989):

$$x_{i,\,snv} = \frac{x_i - \bar{x}}{\sqrt{\frac{(\sum(x_i - \bar{x})^2)}{(n-1)}}} \tag{3.7}$$

Where $x_{i,snv}$ is the transformed spectrum, x_i is the original spectral, $\bar{x}$ is the average value of the variable, and n is the number of the variables in the spectrum.

3.9.1.4 Differentiation

Differentiation methods including the first and second derivative are employed to remove background and increase spectral resolution (Cen & He 2007). They are usually calculated according to the Savitzky–Golay algorithm (Næs et al. 2004). First derivatives are normally applied to remove additive baseline effects and second derivatives are used to remove slope and additive baselines (Swierenga et al. 1999). Derivative spectra of order two are most popular as they can correct for both additive and multiplicative effects, similar to the function of MSC.

3.9.2 Variable-reduction: Principal component analysis

Since the NIR spectrometers provide vast amount of spectral data, the substantial number of samples required to construct calibration models, and the high correlation among the spectral data (multi-collinearity), most multivariate analytical methods rely heavily on variable-reduction techniques (Afara 2012). Variable-reduction technique reduces the dimensions of the original data to a few uncorrelated variables containing relevant information from the samples. Variable-reduction techniques have become a pertinent part of NIR spectral analytical protocol, with the best known and most widely used being principal component analysis (PCA) (Wold et al. 1987).

PCA involves a mathematical procedure that transforms a number of correlated variables into a smaller number of uncorrelated variables in searching for directions of maximum variability in sample groupings and uses them as new axes called Principal Components (PCs) (Blanco & Villarroya 2002). By this means, the relevant information from the original spectral data is transformed into new orthogonal variables (referred to as PCs) can be used as new variables, instead of the original data, in subsequent calculations. The first principal component accounts for the greatest possible statistical variability (or entropy) in the data, and each succeeding component accounts for as much of the remaining variability as possible (Afara 2012). Usually, PCA is used to reduce dimensionality of the data set by identifying new meaningful underlying variables, i.e, patterns in data, and re-expressing the data in such a way as to highlight their similarities and differences.

3.9.3 Multivariate analysis methods

The primary purpose of multivariate analysis methods is to construct models capable of accurately identifying, with the aim of classifying or predicting, the characteristics and properties of unknown samples (Blanco & villarroya 2002). Multivariate analysis allows samples with similar characteristics to be grouped in order to establish classification methods for unknown samples (qualitative analysis), or to develop models for predicting some property of unknown samples (quantitative analysis). The process involves the steps described in Table 3.3.

Table 3.3: Steps in multivariate analysis techniques

Step	Purpose
Choosing the calibration samples	To select a set of samples representative of the whole population
Determining the target parameter by using the reference method	To determine the value of the measured property in an accurate, precise manner. The quality of the value dictates that of the calibration model
Recording the NIR spectra	To obtain physical–chemical information in a reproducible manner
Subjecting spectra to appropriate treatments	To reduce unwanted contributions (such as shifts and scatter) to the spectra
Constructing the model	To establish the spectrum–property relationship using multivariate methods
Validating the model	To ensure that the model accurately predicts the property of interest in samples not subjected to the calibration process
Predicting unknown samples	To predict rapidly the property of interest in new, unknown samples

(Source: Blanco & Villaroya 2002)

A number of common multivariate-analysis methods can be classified according to their purpose and the algorithms or computational procedures used (Figure 3.4). The method of choice will depend on the purpose of the analysis, the characteristics of the samples and the complexity of the system concerned (for example its non-linearity). Once models are constructed, their predictive capacity must be checked on samples subjected to the same treatment (spectrum recording conditions and spectral pre-processing) as those used for calibration but not employed to construct the model (Afara 2012). The calibration model is a critical component of the spectroscopic analysis system. Multivariate analysis techniques have been used to develop the calibration model for NIR analysis. Calibration models are called robust when the prediction accuracy is relatively insensitive towards unknown changes of external factors.

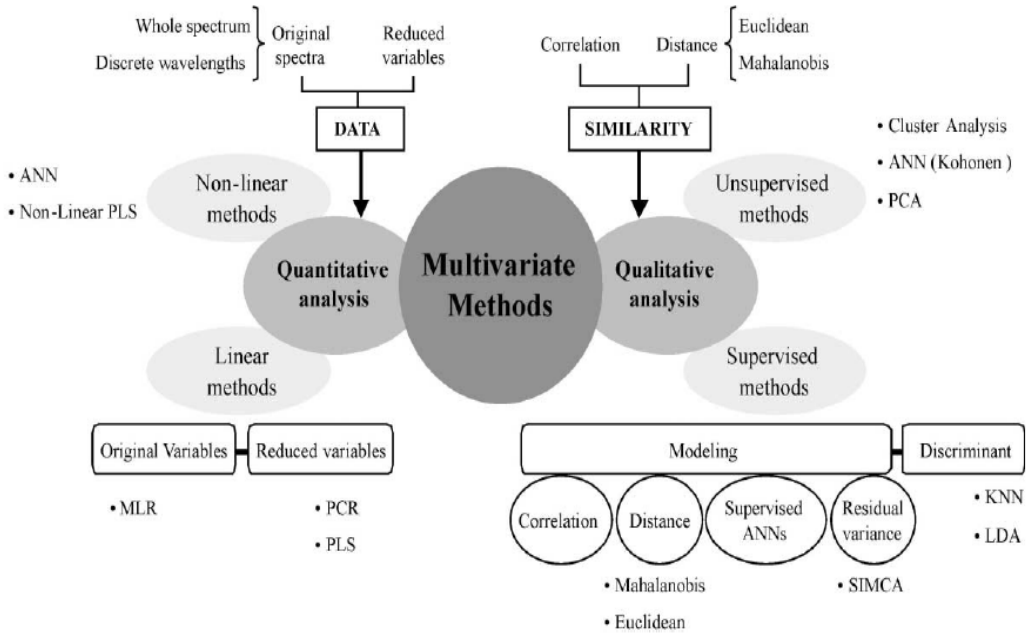


Figure 3.4: Classification of the major qualitative and quantitative multivariate analysis techniques used in NIR spectroscopy (Blanco & Villarroya 2002)

Multivariate analytical methods can be divided into two categories, namely quantitative and qualitative analysis techniques. Qualitative analysis techniques are also known collectively as 'pattern-recognition methods' (Massart et al. 1988), which are labeled 'supervised' or 'unsupervised', depending on whether or not the class to which the samples belong is known. These methods establish mathematical criteria which allow similarity between two samples, or a sample and a class, to be quantified quantitatively. Usually, similarity is expressed as the coefficient of correlation between samples or as a distance (Mahalanobis or Euclidean) measurement, and both type of parameters can be calculated using spectra or PCA results. The different types of classification methods establish boundaries between the different classes, or they model the space occupied by a class and determine whether a sample belongs to it on the basis of distance measurements or the residual variance (Blanco & Villarroya 2002). Among the popular classification methods are linear discriminant analysis (LDA), K-nearest neighbours (KNN), cluster analysis (CA) discriminant partial least squares (DPLS) soft independent modelling of class analogy (SIMCA) artificial neural network (ANN) and support vector machine (SVM) (Cen & He 2007).

Most of the quantitative analysis techniques are developed based on linear regression techniques to model a set of independent and possibly correlated data (NIR spectra) to a set of dependant variables of crop quality parameters such as sugar content, moisture content, and firmness. The most frequently used multivariate-regression methods in NIR spectroscopy are principal component regression (PCR) and partial least-squares (PLS) regression (Martens & Naes 1991).

PCR uses the PCs provided by PCA to perform regression on the sample property to be predicted, while PLS finds the directions of greatest variability by considering both

spectral and target-property information, with the new axes called 'latent variables (LVs) or 'PLS factors'. The simplest quantitative multivariate-analysis method is multiple linear regressions (MLR) (Hsieh & Lee 2005), which usually uses fewer than five spectral wavelengths. MLR assumes concentration to be a function of absorbance, which entails the knowledge of the concentrations of not only the target analytes, but also all other components contributing to the overall signal.

Multivariate analytical methods are mainly applied to establish a relationship between the $n\times1$ vector of observed response values **y** ('*Y*-variables'; quality attributes of interest, such as soluble solids content and firmness) and the $n\times N$ spectral matrix **X** ('*X*-variables'), with n the number of spectra and N the number of wavelengths (Nicolaï et al. 2007).

3.9.3.1 Multiple linear regressions (MLR)

In MLR method, a value of **y** is approximated by a linear combination of the spectral data at every single wavelength. The N regression coefficients are estimated by minimising the error between predicted and observed response values in a least squares sense. A number of variables for the MLR equation can be identified by a step-wise multiple linear regression (SMLR) technique (Kweon & Maxton 2013). SMLR first selects the variable with the largest correlation with the *Y*-data, then the second variable that results in the best improvement of the accuracy of the prediction, and then the third variable and so on until no further significant improvement can be obtained (Nicolai et al. 2007). MLR models typically do not perform well because of the often high co-linearity of the spectra, and easily lead to over-fitting and a loss of robustness of the calibration models (Næs et al. 2004).

3.9.3.2 Principal component regression (PCR)

PCR is a two-step procedure, which first decomposes the **X** by a PCA analysis and then fits a MLR model, using a small number of PCs instead of the original variables as predictors (Nicolaï et al. 2007). The advantage of the PCR compared to MLR is that the *X*-variables (PCs) are uncorrelated, the noise can be filtered and require only a small number of PCs. A drawback of the PCR is that the PCs are ordered according to decreasing explained variance of the spectral matrix, and that the first principal components which are used for the regression model are not necessarily the most informative with respect to the response variable (Wold et al. 2001).

3.9.3.3 Partial least squares regression (PLS)

To overcome the disadvantage of PCR method, PLS was introduced (Wold et al. 2001). In PLS regression, an orthogonal basis of LVs is constructed one by one in such a way that they are oriented along directions of maximal covariance between the spectral matrix **X** and the response vector **y** (Nicolaï et al. 2007). This procedure ensures that the LVs are ordered according to their relevance for predicting the *Y*-variable. In the regression model, the interpretation of the relationship between *X*-data and *Y*-data is

then simplified as this relationship is built based on the smallest possible number of LVs. PLS models are slightly better than the PCR because they do not include latent variables that are less important to describe the variance of the quality parameter, thus requiring fewer LVs compared to PCR for producing a similar model performance (Jong, 1993). PLS method is commonly applied for spectral data of intact biological material when the various *X*-variables express common information, i.e., when there is a large amount of correlation, or even co-linearity (Nicolaï et al. 2007).

3.10 Conclusions

This chapter has reviewed the application of NIR technologies in the sugarcane industry, including the possible best sample forms, wavelength regions and chemometrics methods. It has been demonstrated that the spectroscopic methods were successfully applied to measure sugarcane quality parameters. When considering the cost and portability of the existing spectrometers, a portable and low-cost VNIR PDA spectroradiometer with a Si detector was found to be the best candidate for field use. Stalk samples were also found to be the best sample form used for predicting sugarcane quality as they require the least sample preparation compared to fibrated and juice samples. The PLS method could be used for both calibration and prediction models throughout this thesis. The reflectance measurement mode could be used in dealing with stalk, internode and fibrated samples, while the transmittance and transflectance measurement modes could be applied to predict sugarcane quality parameters from sugarcane juice samples.

Chapter 4

Materials and methodologies

Previous chapters have established the rationale for the application of spectroscopic methods to determine sugarcane quality from stalk samples. This chapter will describe the new methodologies developed to investigate the innovative application of the proposed technologies in measuring sugarcane quality from stalk samples. Two new measurement methods, the skin scanning method (SSM) and the cross sectional scanning method (CSSM) are applied on whole stalk samples and internode samples to assess their sugar content. This chapter also describes the development of the light proof measurement chamber, designed to provide a consistent experimental setup for all spectral measurements throughout this research. The chemometrics methods used to optimise the regression models between spectral data and sugar content are also discussed.

4.1 Instrumentation

In this study, two spectroradiometers equipped with two different detectors were used for spectral measurements from different sugarcane samples. The performance of these spectroradiometers in predicting sugar content from different sugarcane samples was compared and evaluated. Detailed explanations about the spectroradiometers are provided below.

4.1.1 Visible and shortwave near infrared spectroradiometer (VNIRS)

The primary equipment used in this study was a handheld visible and shortwave near infrared spectroradiometer (FieldSpec® Pro HandHeld spectroradiometer, Analytical Spectral Devices (ASD), Inc., Boulder, CO, USA). This visible and shortwave near infrared spectroradiometer (VNIRS) works in the wavelength region ranging from 325 to 1075 nm in 1.5 nm intervals. The equipment has a spectral resolution (full width at half maximum (FWHM) of a single emission line) of approximately 3 nm at around 700 nm. It uses a 512 channel PDA with Si detector overlaid with an order separation filter. Each channel of an individual detector is geometrically positioned to receive light within a narrow (1.6 nm) nominal bandwidth. The VNIRS was used to collect the spectral data from both reflectance and transmittance measurement modes. Reflectance measurement was applied on whole stalks, internodes, and fibrated samples. Transmittance and transflectance measurement modes were applied on both raw and clarified juice samples.

4.1.2 Full range spectroradiometer (FRS)

The secondary equipment used was a full range spectroradiometer (FRS) manufactured by FieldSpec®, Analytical Spectral Devices (ASD), Inc., Boulder, CO, USA). The

spectral range for FRS is from 350 to 2500 nm with a resolution of 1.4 nm in the 350–1000 nm range and 2 nm in the 1000–2500 nm range. The FRS was built using Si and InGaAs detectors. The Si detector was intended to work within the spectral range from 350 to 1000 nm while the InGaAs was intended to work for the spectral range between 1000 and 2500 nm. The FRS consists of a concave holographic grating and a single thermoelectrically cooled InGaAs detector. The gratings are mounted about a common shaft which oscillates within a period of about 200 ms (100 ms/scan). The FRS was equipped with a fibre optic cable for light collection. However, access to this equipment was limited due to scheduling commitments. The FRS was used to compare the prediction performance between Si detector (VNIRS) and InGaAs detector (FRS) in predicting sugar content of sugarcane from stalk samples.

4.1.3 Sensitivity and repeatability of the spectroradiometers

Sensitivity and repeatability are two important factors in determining the right spectrometer for field uses. The FRS can record a complete 350 - 2500 nm spectrum in 0.1 seconds while the VNIRS can record a complete 350 - 1050 nm spectrum in as little as 17 milliseconds (ASD, 1999). This high speed allows the convenience of collecting more data in shorter time, as well as minimizing errors associated with clouds and wind under solar illumination. Another thing to consider is the number of scans which can be collected in a specified time period. The combination of speedy measurement with extremely low Noise-equivalent-Radiance (NeDL) will optimise the performance of the spectroradiometers. Figure 4.1 and 4.2 shows the sensitivity of the both VNIRS and FRS as illustrated by the typical NeDL and Signal-to-Noise plots, respectively.

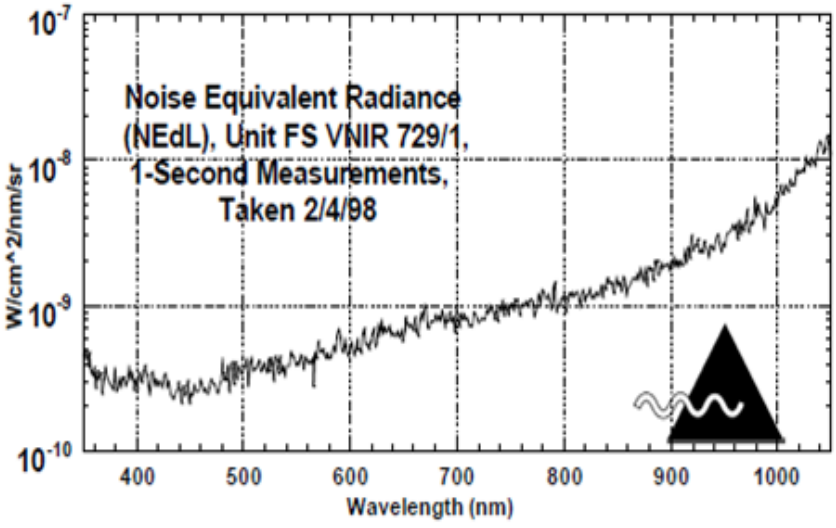


Figure 4.1: Sensitivity figure of VNIRS as indicated by NeDL (ASD 1999)

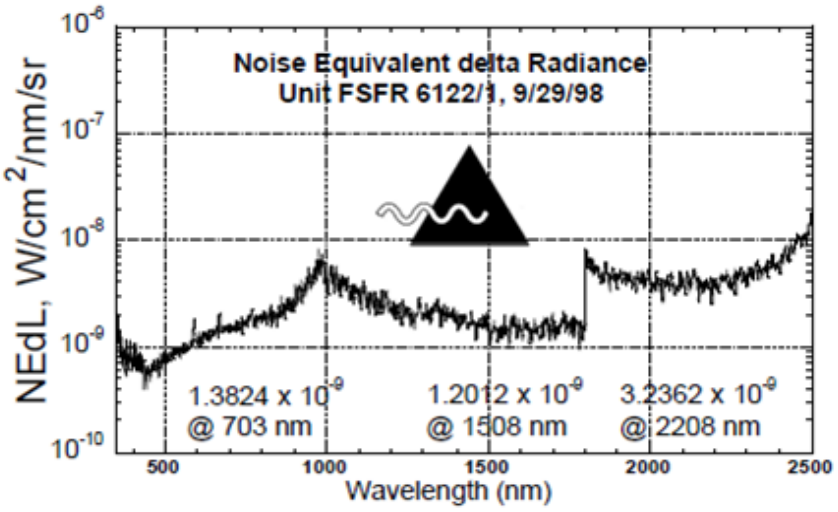


Figure 4.2: Sensitivity figure of FRS as indicated by NeDL (ASD 1999)

In terms of repeatability, high signal-to-noise ratio and superior repeatability of the equipment will provide accurate and precise results of any measurements. The FRS has an excellent repeatability as shown in Figure 4.3. Since VNIRS is manufactured by the same company, it is assumed that it has the same high quality repeatability performance.

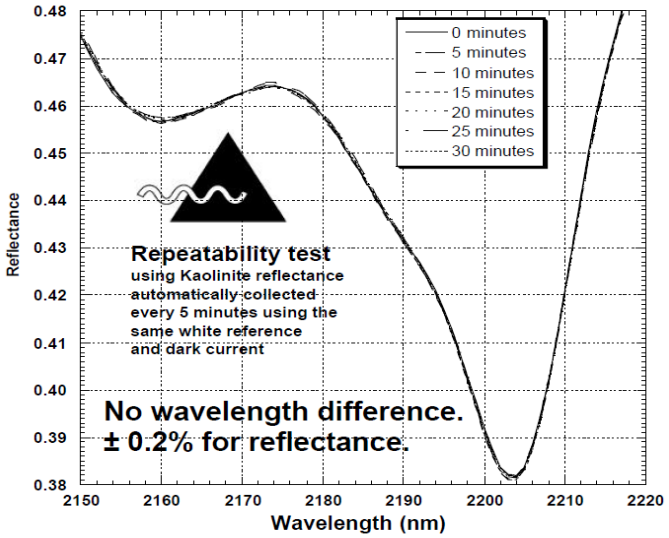


Figure 4.3: Repeatability figure for FRS unit (ASD 1999)

4.2 Light-proof measurement box

Several spectroscopic studies have developed various measurement boxes as the instrument platform for collecting spectral data. These boxes were built for many purposes, including the enclosure of the light sources, probes and samples from ambient light (Guthrie & Walsh 1997; Lu et al. 2000), to keep a constant distance and

angle between the probe and samples (Ventura et al. 1998), or to keep the measurement system under a constant temperature (Yu et al. 2009). For field use, the box was also needed to protect the measurement system from being damaged and contaminated by crop samples or trash (Taylor et al. 2005; Welle et al. 2003; Abdel-Rahman et al. 2010).

4.2.1 Box construction

In this study, a light-proof measurement box was built (900 mm x 600 mm x 400 mm) (Figure 4.4) to serve as an instrument platform. The box was developed firstly to enclose the light sources, sensors and samples from ambient light. Secondly, it was important to keep a constant measurement distance and angle between the sensors and samples. All surfaces inside the box were painted black to minimise the influence of background surfaces on spectral data (Wu et al. 2008). The box was designed to permit reflectance, transmittance and transflectance measurement modes to be performed on different sample forms under the same measurement conditions with consistent experimental setups. With an on-the-go measurement system on a harvester in mind, this box was regarded as the first step towards the development of a real measurement chamber for predicting sugar content on a harvester. A small hole was cut on the both side of the walls to allow intact stalk scanning. These holes were covered by black cloth to block the ambient light from entering into the box.

Figure 4.4: Light-proof measurement box in use

4.2.2 Illumination system

For reflectance measurement, two halogen lamps (Lowell Pro-Lam 14.5 V (50W) tungsten bulb, Ushio Lighting, Inc., Japan) were used to provide illumination inside the box. The lamps were powered by AC/DC power adapter plugged into a wall receptacle. The halogen lamps offered a constant light source and covered both the Vis and NIR regions (Wu et al. 2008). The lamps were placed at a distance of 800 mm above the sample at an angle of 45° (Figure 4.5). The lamp was mounted 300 mm away from each other, with the light beam of both lamps aligned to focus on the approximate centre of the sample. For transmittance measurement, only one halogen lamp was used and

placed at 300 mm from the sample (Figure 4.6). The light sources were switched on 60 minutes before the spectral measurements to warm up the system and stabilize the spectral data as recommended by the manufacturer.

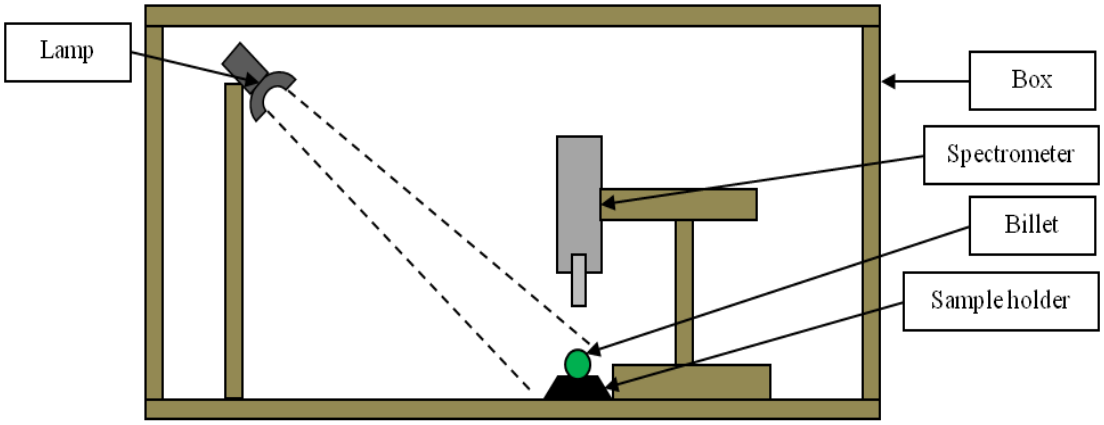


Figure 4.5: Simplified diagram of the reflectance measurement inside the measurement box

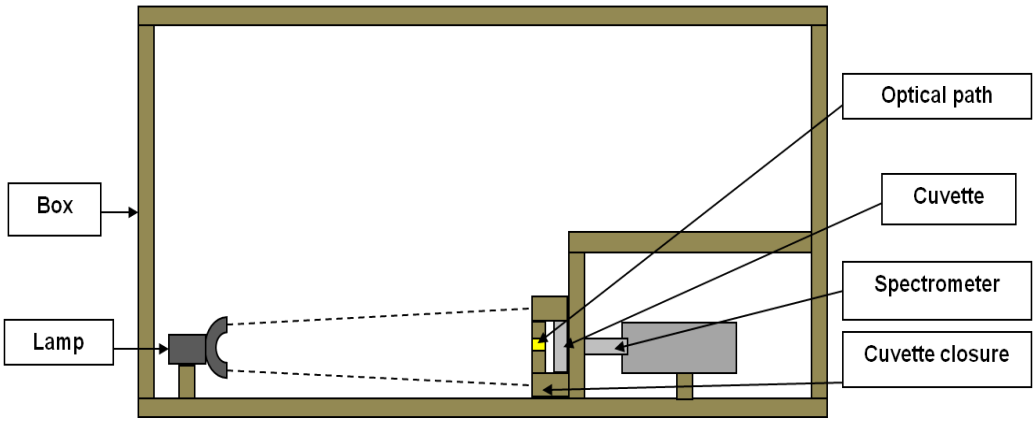


Figure 4.6: Simplified diagram of the transmittance measurement inside the measurement box

4.2.3 Sensor to sample distance

The measurement box was designed to collect spectral data from different sugarcane sample forms. To minimise the spectral variation, the distance between the sensor and sample was maintained by fixing the sensor to a fixed stand while samples were held by a fixed sample holder. The box incorporated two different sample holders, one for the

SSM and another for the CSSM. The sample holders were built to firmly hold the samples against the sensor with the light being positioned at a constant distance and angle. The distance between the probe and sample was chosen, depending on the diameter of the spot size, *Y*, and the field-of-view (FOV) of the spectroradiometers used. The distance between probe and sensor was calculated using the formula provided by ASD (2005) as shown below:

$$Y = D + 2 \times X \times \tan\left(\frac{A}{2}\right) \quad (4.1)$$

Where *Y* is the diameter of the spot size (mm), *A* is the fore optic's angular FOV, *X* is the distance to viewed surface (mm), *D* is the an effective diameter of fore optic lens (mm). The location of each parameter is illustrated in Figure 4.7. The values of *D* at 10° and 25° for FRS are 6.8 and 1.6 while for VNIRS are 2.0 and 0.7 mm, respectively (ASD 2005).

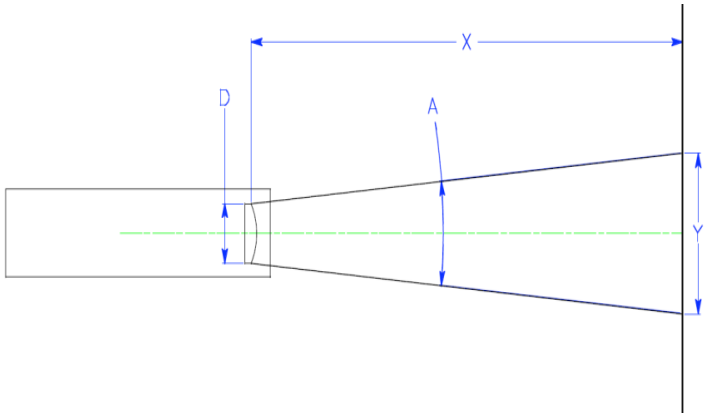


Figure 4.7: Assisted diagram for calculating field-of-view (ASD 2005)

4.2.4 White panel

Throughout this study, reflectance was computed using measurements from both the unknown material (sugarcane samples) and a reference material with known reflectance properties (a spectralon white reference panel). The white reference panel with approximately 100% reflectance was used to reduce the influence of the changing light intensity. White reference measurements were taken after every three sample measurements. Relative reflectance spectra were calculated by dividing sample radiance with reference radiance from a spectralon white reference panel for each wavelength.

4.3 Plant materials and measurement methods

As this research mainly designed to investigate the application of the spectroradiometers (VNIRS and FRS) in predicting sugarcane quality from stalk samples, two measurement methods have been proposed namely the SSM and CSSM. For comparison purposes, the spectral data from the conventional samples forms (juice

and fibrated samples) were also investigated. This comparison study can provide an important opportunity to identify the optimal sugarcane sample form for quality prediction using the spectroradiometers. To give a valid comparison, all sample forms were measured under the same experimental setup (inside the box) and subject to the same data treatment methods.

In this work, all sugarcane samples were supplied by BSES, Bundaberg, Queensland, Australia (25°S, 152°E). All of the quality measurements using standard industry procedures were carried out using laboratory facilities at research station belonged to BSES, Bundaberg. BSES is one of the principal providers of research, development and extension (RD&E) to the Australian sugarcane industry and is funded by a voluntary levy paid by Australian sugarcane farmers and sugar millers.

Samples were collected during the harvest seasons between 2010 and 2012. All stalk samples were manually cut from the field. The green canopy was removed by cutting each stalk at the growing point and removing all leaf materials. Four sample forms were used in this study: whole stalk samples, internode samples, fibrated samples and juice samples (raw and clarified juices). The specific sample preparation for each sample form is discussed below.

4.3.1 Whole stalk samples

The whole stalk samples was collected based on the industry standard practice for quality determination developed by the breeding programs which was consisted of six stalk samples being collected from one sampling point. The collected samples were transported to the research station for a standard quality determination and spectral measurement. Two methods were investigated on the whole stalk samples, namely SSM and CCSM. SSM method was applied by scanning a stalk sample directly on their skin while CSSM was applied by scanning a stalk sample on its cross sectional surface. Both methods were employed under the same experimental setup.

4.3.1.1 Cross sectional scanning method (CSSM)

A total of 100 fresh sugarcane stalk samples representing 10 common varieties (10 stalks for each variety) (Q135, Q138, Q151, Q155, Q190, Q200, Q208, Q232, Q240 and KQ228) were taken from the first ratoon crop in Bundaberg, in 2010. The samples were a commercial varieties cut from a propagation block used to source material for plant breeding trials. They were a 1st ratoon that was harvested in October the previous year (11 months old). The crop was grown under commercial conditions with the fertilisation based on a soil test and the six easy steps nutrition guidelines (Schroeder et al. 2009).

Each stalk was cut into three sections: top, middle and bottom. The length of each cut section was approximately 600 mm (Figure 4.8). After cutting, the cross sectional surface of each section was scanned using both VNIRS and FRS. In order to explore the use of in situ spectroscopic data for predicting the sugar content of sugarcane stalks, the

experiment was carried out in the clear area of the field on a clear sunny day between 10:00 am and 3:00 pm, local time. Care was taken not to cast a shadow over the stalk samples while taking the spectral measurements.

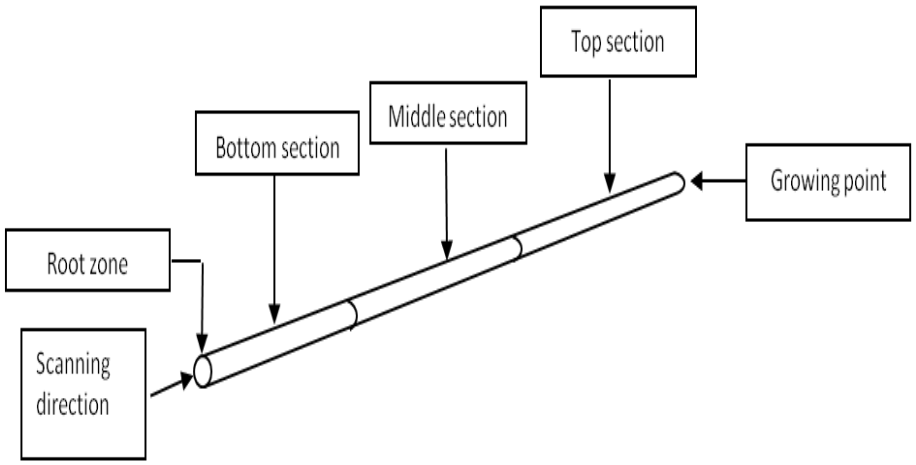


Figure 4.8: The division of fresh sugarcane stalks into three main sections

The scanning was performed using a 25° FOV of the spectroradiometers. The probes were located 50 mm above the sample at a 45° angle. The distance was maintained by placing the probes on a tripod and placing the samples in the marked area on the table edge. The reflectance measurement method was chosen in this part of the study because this method is the easiest to perform as it requires no contact with the samples and the light levels are relatively high (Schaare & Fraser 2000). The equipment was set to record an average of 20 scans for each spectrum in order to increase the signal-to-noise ratio.

4.3.1.2 Skin scanning method (SSM)

A 36 set of millable whole stalk samples from Q238 variety were collected from the BSES research station, Bundaberg in September 2012. Each set contained six stalk samples (216 whole stalks samples used altogether). The experiment was conducted to explore the feasibility of the spectroscopic method to predict sugarcane quality parameters from fresh stalk samples using SSM. Each stalk sample was divided into three labeled sections: bottom, middle and top. Then, each stalk sample was pushed into the measurement box through the holes on the box's walls (Figure 4.9). The spectral data was collected from the mid-internode area of each section, with care was taken to make sure the stalk sample was properly placed on the sample holder located in the measurement box before the spectral measurement was made. For every section, three reflectance readings were taken from each mid-internode sample by rotating the stalk at approximately 120°. The light reflected from the skin surfaces were collected using both VNIRS and FRS.

Figure 4.9: The whole stalk sample being scanned using spectroscopic method inside the measurement box

4.3.1.3 Quality measurement for whole stalk samples

Standard laboratory procedures were applied to determine the quality values of the whole stalk samples. According to this standard, two thirds of the stalk sections were crushed using a small laboratory roller mill to extract the juice samples for °Brix and pol determination. The remaining one third of the stalk sections were used for fibre determination (BSES 2001). The measurement of °Brix (Method 3, BSES 2001) was performed on a raw juice (after the suspended solids settled out and air bubbles escaped from the juice) using Bellingham and Stanley RFM310 temperature compensated digital refractometer with the accuracy of ± 0.1° Brix. A sample of the juice for pol measurement was first clarified by adding lead acetate powder (which coagulates colloidal impurities and removes some colorant) followed by filtration. The pol reading (Method 2, BSES 2001) on clarified juice samples was made using Polartronic Universal automatic polarimeter with reading precision of ± 0.03°Z (Schmidt + Haensch, Berlin, Germany). Fibre content was determined using Method 4 (BSES 2001) by washing the fibrated samples (in a cotton cambric bag) for three cycles in an automatic washing machine, and drying at 100-105°C for approximately one hour until constant weight was achieved. Fibre (%) was calculated directly from fresh and dry sample weights. These three parameters were then used to calculate CCS value using the Eqns. 2.1-2.4, as discussed in Section 2.4.

4.3.2 Internode samples

The ability to measure sugarcane quality from internode samples has the potential application in the field, especially for an on-the-go measurement system on a sugarcane harvester because internode samples are similar to billets. Thus, this proposed method can open up a new possibility of installing a spectrometer on a harvester to predict sugar content of billets in the field during harvesting. Quality measurement from internode samples would also be able to provide the variation of sugar content along the length of the stalk and internode samples. This information is an important parameter of consideration for the design and development of the quality measurement system on a

harvester. This variation could not be quantified from the whole stalk sample analysis, as discussed in Section 4.3.1.

For this purposes, a total of 292 internodes were extracted from 22 sugarcane stalk samples. The stalk samples were collected from the research station of BSES, Bundaberg, in May 2012. The stalk samples were from a plant crop that was planted in September 2011 (8 months old). The crop was grown under commercial conditions with the fertilisation based on a soil test and the six easy steps nutrition guidelines (Schroeder et al. 2009). The stalks were from a commercial variety trial represented three different maturity stages, namely early-maturing (Q155), mid-maturing (Q208) and late-maturing (Q190) crops. The selection of three varieties was designed to ensure that the models developed in this study cover the range of °Brix (7.6 to 22.2) which is commonly experienced during commercial harvesting. The stalk samples were stored at - 18°C in a freezer and were equilibrated to room temperature before being cut into internode samples prior to spectral measurements.

4.3.2.1 Reflectance measurement for internode samples

Reflectance measurements were performed on internode samples using both SSM and CSSM. The VNIRS was used for all 292 internode samples while the FRS, due to its limited availability, was only applied on 128 internode samples. Firstly, each internode was nominally labelled as S1, S2, S3 and S4 following the sequence from S1 to S4 (bottom to the top). Then, each internode sample was scanned on its skin surface using the sensors at four different points (S1 to S4) from bottom to the top (Figure 4.10).

Figure 4.10: Scanning positions along individual internodes

Then, the individual internode sample was cut perpendicularly into four sections (C1, C2, C3 and C4), representing the node areas and the middle of the internode (Figure 4.11). Each cut section was then scanned on cut cross-sectional surface using both sensors. The scanning for both SSM and CSSM were undertaken using the 25° FOV of the spectroradiometers. The distance between the sensor and samples was set at 70 mm, resulting in a measured spot size of 31 mm diameter as calculated using Eqn. 4.1. The

distance between the probe and sample were maintained by fixing the probe to a tripod and holding the samples with a fix sample holder. The equipment was set to record the average reading of 20 scans for each spectrum.

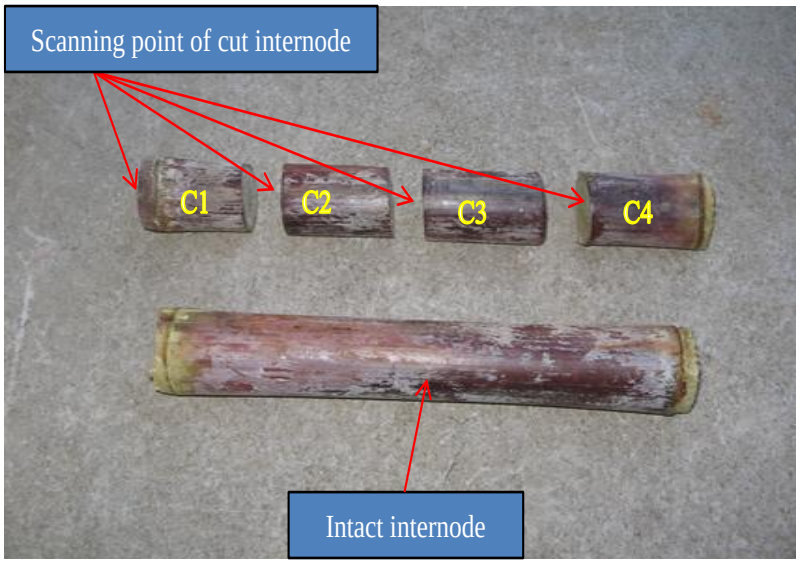


Figure 4.11: Intact internode versus cut internode with scanning positions

4.3.2.2 °Brix measurement for individual internode samples

After the spectral acquisition, all of the cut sections from the same individual internode were squeezed using a clamp to extract a representative juice sample for °Brix measurement. Before squeezing, the internode samples were wrapped in a small plastic bag to avoid cross contamination. The plastic bag was also used to channel the squeezed juice samples into a container. The plastic bag was disposed after each use. For internode samples, the standard laboratory procedures (described in Section 4.3.1.3) could not be carried out to measure quality parameters due to the limited juice amount contained in individual internode. On average, only 20 ml of juice was extracted from individual internodes. However, the measurement of pol requires about 200 ml of juice samples. Therefore, for internode samples, only °Brix values were used for calibration purposes.

°Brix value was the easiest, least expensive quality parameter to be measured and requires little preparation (Staunton et al. 2011). The strong correlation between °Brix and CCS has been previously reported by Staunton et al. (2011). Furthermore, in many applications, the °Brix in juices measured using a refractometer are widely used as an index of total sugar content due to the fact that sugars are the major components of the soluble solids in fruit juices (Chang et al. 1998).

°Brix measurement is a relatively simple process because a refractometer only requires a few drops of juice samples. The extracted juice of each cut section from the same internode was collected and mixed in a container, shaken and poured onto the

refractometer to measure the °Brix. The °Brix measurement was made using a handheld °Brix refractometer (Model: RHB-32ATC, from Huake Instrument Co., Ltd, Baoan, Shenzhen, China, the °Brix range is 0-32% with automatic temperature compensation). The spectral scanning and °Brix measurement were carried out in the same day.

4.3.2.3 Moisture measurement for internode samples

A standard industry method for moisture determination of prepared sugarcane is achieved by using Spencer type drying ovens with air at 100-105°C for an approximately one hour until constant weight was obtained (BSES 2001). However, since this standard method was designed to process high volume of samples per time (e.g. 1000 g samples for SRI type oven), it is not practical to be applied for processing an individual internode sample. Instead, the moisture content of each internode sample was obtained by drying the samples at 60°C for 24 hr (Purcell et al. 2009). The weight of wet internode sample and metal container, W_I was measured before it was squeezed for the °Brix measurement. After the sample was squeezed, it was put into a metal container and placed into a drying oven to determine the weight of dried internode, D_I (Figure 4.12). The drying oven used was Qualtex, Solidstat, model OM18SZ2, manufactured in Australia by Watson Victor Ltd, which operated in the temperature range of 0 to 270°C. The moisture content of each internode was determined based on a wet basis method (MC %, w. b.) using the following formula:

$$MC, \%\ (w.b.) = \frac{W_I - D_I}{W_I} \tag{4.2}$$

Where W_I is weight of the wet internode with metal container and D_I is weight of the dried internode with metal container.

Figure 4.12: Cut section of internode samples which have been oven dried

4.3.2.4 Fibre measurement for internode samples

A standard industry method for fibre determination of sugarcane samples was described in Section 4.3.1.3. However, due to a small volume of individual internode sample, this standard method is not practical to be adopted. In this study, the fibre content for each internode sample was determined using the mathematical method proposed by Watson et al. (1999) as detailed below:

$$°Brix\ in\ sugarcane = Brix\ in\ juice\ x\ \frac{\left(100 - (fibre\ \% + 3^{*})\right)}{100} \tag{4.3}$$

$$Fibre\ \% = 100 - moisture\ \% - °Brix\ in\ sugarcane \tag{4.4}$$

Combining the above equations yields:

$$Fibre\ \% = \frac{(100^{2} - 100\ \times moisture\ \% - 97\ \times °Brix\ of\ juice)}{(100 + °Brix\ of\ juice)} \tag{4.5}$$

Where number 3* (Eqn 4.3) is a correction factor used to correct the °Brix measurement in first expressed juice to more accurately represent those of the total juice in sugarcane.

4.3.3 Juices and fibrated sugarcane samples

A total of 100 sample sets were collected from different locations across Bundaberg, representing different commercial varieties in the 2012 harvest season. Each sample set consisted of a group of six whole stalk samples. These sample sets were collected in conjunction with the BSES agronomy research project. The selection of the sample sets from different harvest locations helped to establish sufficiently robust calibration models that are applicable to all varieties. Each sample set was passed through the Spectracane system at the BSES research station in Bundaberg. Each sample set was first fibrated using a sugarcane disintegrator (Dedini, Model D-2500-II) with a 10HP, 400V motor, and operating at 3340 rpm (Figure 4.13). Two subsamples of an average of 120 g each were taken from each fibrated sample for standard measurement of the fibre content (refer to Section 4.3.1.3).

Figure 4.13: Sugarcane disintegrator used to fibrate stalk samples

The remaining fibrated samples were pressed at 25 MPa for 1 minute using a hydraulic press to extract the first express juice samples. Then, the collected juices of each sample set was divided into two sets; one set to determine °Brix and another set to determine pol. The °Brix measurement on raw juice and pol measurement on clarified juice were carried out using the method as described in Section 4.3.1.3.

4.3.3.1 Reflectance and transflectance measurement modes

To investigate the potential application of VNIRS to predict sugarcane quality from juice [raw juice (RJ) and clarified juice (CJ)] and fibrated samples (FS), a subsample of fibrated and juice samples was taken from each sample set from the above-mentioned standard procedures. For the FS, 100 g of fibrated sugarcane was collected immediately after they passed through the Spectracane system. For juice samples, 50 ml of raw and clarified juice samples from each sample set were also collected. All collected samples (RJ, CJ and FS) were stored at –18°C in a freezer and were equilibrated to room temperature before spectral measurement. RJ was filtered using a cotton ball prior to the spectral measurement to eliminate suspended particles (Valderrama et al. 2007). These sample forms were scanned using the VNIRS inside the measurement box. For spectral measurement, the FS (Figure 4.14) and juice samples (Figure 4.15) were placed in plastic containers with a diameter of 60 mm and 30 mm thick.

Figure 4.14: Fibrated sugarcane samples for reflectance measurement

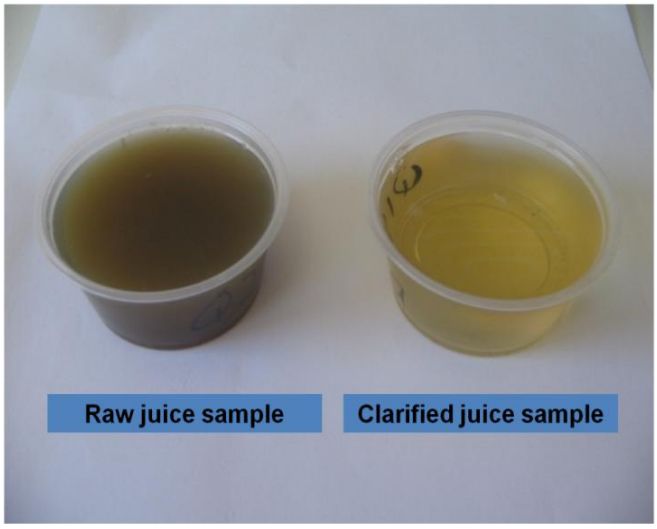


Figure 4.15: Colour difference between the raw and clarified sugarcane juice

The FS was scanned using the reflectance measurement mode. The FS was packed and levelled before scanning to produce a smooth sample surface and uniform presentation. Care was taken not to overfill or excessively compress the samples in the container. Considering the diameter of the container, the distance between the sensor and container was set at 112 mm, resulting in a diameter of the measured spot of 50 mm.

For juice samples, 50 ml of both RJ and CJ were also placed into the container for the spectral measurement using transflectance mode. The spectral measurements were undertaken against the black background of the measurement box. Three reflectance/transflectance readings were taken from each sample by rotating the container at 120°. A standard reference (a spectralon white reference panel) measurement was made every time prior to each sample measurement.

4.3.3.2 Transmittance measurement mode

For juice samples, transmittance measurement mode is commonly employed using fibre optic probes or cuvettes. For example, Liu et al. (2008) applied VNIRS in transmittance mode to determine soluble solids content (SSC) and pH values of rice vinegars. In this study, the transmittance measurement was applied to both RJ and CJ. Both juice samples were poured into a rectangular plastic cuvette with 10 mm of optical path length (path length of the cuvette in which sample is contained). The lamp was placed at a distance of 30 mm from the cuvette. A cuvette holder was built to hold the cuvette throughout the experiments. The probe was placed in opposite direction to the lamp. Considering FOV of 10°, the distance between VNIRS and cuvette was set at 45 mm, giving the scanning diameter of 10 mm.

The transmission spectra from 325 to 1075 nm were measured at 1.5 nm intervals with an average reading of 20 scans for each spectrum. Three spectra were collected for each sample and the average spectrum of these three measurements was used for further analysis. Before spectral measurement, reference spectrum was measured from the cuvette filled with distilled water (Fernández-Novales et al. 2009).

4.4 Data analysis methodology

4.4.1 Spectral data transformation

All collected spectral data in this thesis were stored in the computer and processed using the RS3 software for Windows (Analytical Spectral Devices (ASD), Boulder, CO, USA) designed with a graphical user interface. The transmittance, *T* and the reflectance, *R* spectra were transformed into ASCII format by using the ASD ViewSpec Pro software (ASD, Boulder, CO, USA). The spectral data of transmittance *T*, was calculated using the following formula by comparing NIR energy transmitted through the sample with that through the reference spectrum:

$$T(\lambda) = Is(\lambda)/\, Ir(\lambda) \tag{4.6}$$

Where *Is* (λ) is the intensity of light transmitted through samples at λ (nm), and *Ir* (λ) is the intensity of light transmitted through standard reference at λ (nm).

While for reflectance;

$$R(\lambda) = Ls(\lambda)/Lr(\lambda) \tag{4.7}$$

Where *Ls* (λ) is the light reflected from samples at λ (nm), and *Lr* (λ) is the light reflected from the spectralon white reference panel at λ (nm). Then, both *R* and *T* were transformed into absorbance data, *A*. According to Beer's law, the amount of radiation absorbed by a sample is directly proportional to the concentration of the compound and the path length of the radiation through the sample which can be calculated as:

$$A\ (\lambda) = \log\ (\frac{1}{T(\lambda)}) = \log(\frac{1}{R(\lambda)}) \quad (4.8)$$

Where A (λ) is the absorbance at λ (nm), whereas T (λ) and R (λ) is the transmittance and reflectance at λ (nm), respectively.

4.4.2 Noise removal

Due to potential system imperfection, obvious scattering noises affecting the accuracy of the measurements might be observed at the beginning and the end of the spectral data (Li et al. 2007). In order to avoid a low signal-to-noise ratio, the first and last 75 nm data points were removed from the original reflectance spectral data (Liu et al. 2008). Only the wavelength regions between 400 and 1000 nm were used for the calculations. For transmittance measurement, in addition to spectral noise at the edge, the absorbance in Vis region (400 to 550 nm) was found to be very weak, therefore only the wavelength region between 600 and 1000 nm were used for the analysis (Cozzolino et al. 2007). For FRS, only the first 75 nm data points were removed from the original spectral data (Abdel-Rahman et al. 2008).

4.4.3 Pre-processing of the spectral data

The spectra of solid samples are influenced by their physical properties with scattering phenomena (which is wavelength-dependent and non-linear) is the most common factor for causing error in absorbance values. Thus, spectral pre-processing method should be used to minimise the influences of irrelevant information into spectra in order to be able to develop more simple and robust models (Blanco & Villarroya 2002). The pre-processing of spectral data is a key part of spectral analysis used to improve the quality and accuracy of the regression models (Wu et al. 2008). The commonly used pre-processing methods have been reviewed in Section 3.9.1.

In this thesis, several pre-processing methods were investigated including mean normalization (Griffiths 1995), moving average (Wu et al. 2008), multiplicative scatter correction (MSC) (Geladi et al. 1985), first and second derivatives according to Savitzky-Golay differentiation (Swierenga et al. 1999), standard normal variate (SNV) and a combination of them. The pre-processing procedures were implemented using the Unscrambler, V 9.6 software (Camo Process AS, Oslo, Norway).

4.4.4 Application of principal component analysis (PCA)

Section 3.9.2 has reviewed the application of PCA method as data reduction method which transforms original data into new orthogonal variables referred to as principal components, or PCs (Purcell et al. 2005). In this thesis, PCA method was used to extract useful information from spectral data, decrease the noise, identify spectral outlier and determine the optimum number of PCs (Wu et al. 2008). The identified outliers found in the data set were removed before the development of the regression models. Outliers were detected by PCA from the influence plot which displays the sample residual x-variances against leverages. Usually, only a few PCs are required to describe most of

the data variance with the first PC accounting for the greatest amount of variance. A low number of PCs are normally desirable to avoid inclusion of signal noise in the modeling (Xiaobo et al. 2007). Throughout this research, the maximum number of PCs was set at ten. PCA was also used for the development of classification algorithms using ANN. PCA method was exercised using the Unscrambler, V 9.6 software (Camo Process AS, Oslo, Norway).

4.4.5 Application of Partial least square (PLS)

The multivariate-regression methods most frequently used in NIR spectroscopy are PCR and PLS regression (Martens & Næs 1991). PLS models are slightly better than the PCR because they do not include latent variables that are less important for describing the variance of the quality parameter (Jong 1993). PLS found the directions of greatest variability by considering both spectral and measured property information, with the new axes, called LVs (Blanco & Villarroya 2002). The LVs were considered as new eigenvectors of the original spectra to reduce the dimensionality and compress the original spectral data (Wu et al. 2008).

PLS is a well-known factor analysis multivariate method which is commonly used in the NIR spectroscopy analysis for predictive purposes (Purcell et al. 2005). It requires a calibration step in which a model is constructed from a number of significant factors, which are selected, for example, by the well-known cross-validation leave-one-out method. PLS analysis was performed to establish a regression model to predict quality parameters of sugarcane (matrix Y) from spectral data (matrix X). This method was used in this study to interpret the spectra and develop both calibration and prediction models for sugarcane quality parameters. The maximum LVs number for an acceptable PLS model is usually ten (Moghimi et al. 2010). Thus, the maximum numbers of LVs used in this study was ten. In this thesis, PLS analyses were performed using the Unscrambler, V 9.6 software (Camo Process AS, Oslo, Norway).

In this thesis, the purpose of PLS regression model was to build calibration method (with respect to NIR spectral data analysis) to predict a sample's properties (e.g. sugar content, moisture content etc) y_i in objects $i = 1,2,\ldots, I$ from a set of spectral absorption x_{ik} at wavelength channels k = 1, 2…, K through a linear predictor equation as shown in Eqn. 4.9 (Afara 2012):.

$$\hat{y} = b0 + \sum_{K=1}^{K} xijbk \tag{4.9}$$

Based on data $\boldsymbol{y} = (y_i, i = 1, 2\ldots.I)^T$ and $\mathbf{X} = (x_{ik}, I = 1, 2\ldots..I; k = 1, 2,\ldots., K)$ from a set of I calibration samples, the parameters in the predictor equation are estimated statistically. Once obtained, this predictor equation can be used for converting **X** data into y estimates in a new sample.

PLS is a bilinear regression method that extracts a small number of "factor", $\mathbf{t}_a$, a = 1, 2, …, A that are linear combinations of the KX variables, and uses these factors as

regressors for $\boldsymbol{y}$. In addition, the X variables themselves are also modeled by these regression factors. Therefore, outliers with abnormal spectra $\boldsymbol{x}_i$ in the calibration set or in future data sets can be detected.

A special feature of PLS compared to PCR is that the y variable is used actively in determining how the regression factors $\mathbf{t}_a$, a = 1, 2, …, A are computed from the spectra X. Each PLSR factor $\mathbf{t}_a$ is defined so that it describes as much as possible of the covariance between **X** and **y** remaining after the previous a-1 factors have been estimated and subtracted (Afara 2012). This robust feature of PLS allows it to construct an efficient model using as few LVs when compared with PCR, thus increasing the speed of analysis and requiring fewer computational resources.

4.5 Calibration and validation methods

Calibration is a mathematical process required to relate spectral data to the desired constituents of the measured crop. Multivariate calibration models play a key role in the analytical measurement. The essence of any calibration procedure is to ensure that the range of spectral variation found in the whole population is represented in the samples selected for analysis for calibration development (Foley et al. 1998). In the development of the PLS model, full cross validation (leave-one-out) was used to evaluate the quality and prevent over fitting of the calibration model (Arana et al. 2005). In leave-one-out cross validation, one sample was removed from the dataset, and a calibration model was constructed for the remaining subset. The removed samples were then used to calculate the prediction residual. The process was repeated with other subsets until every sample has been left out once, and in the end the variance of all prediction residuals was estimated.

The performances of the established calibration equations were further validated using the samples in the validation set. The external validation method was used to evaluate the predictive ability of the PLS calibration model. The external validation procedure determined the predictive ability of an equation, based on a sample set which had not been used in the calibration development. Before the calibration, samples were divided into two sets. One part (75% of samples) was used to develop a prediction equation (calibration set) and another part (25% of samples) was used to validate the predictive equation (validation set). Samples for validation were selected by taking one of every four samples from the entire sample set, taking care to ensure that each set included samples that covered the entire range of quality data. Both calibration and validation models were performed using the Unscrambler, V 9.6 software (Camo Process AS, Oslo, Norway).

The PLS model performance was optimized with regard to attribute quality distribution along the length of the stalk and billet samples, wavelength ranges, pro-processing methods on both absorbance and reflectance data. The improvement in performance of calibration/prediction models in predicting sugarcane quality parameters was assessed using different sample forms, spectroradiometers, sample presentations and measurement mode.

4.6 Statistical assessment for model accuracy

Since the PLS regression models were developed based on correlation between sugarcane quality parameters and many highly correlated spectra, their predictive ability requires assessment. The models should not be over-optimistic, which fit too well to the calibration data set, but also are not robust enough to account for variations in future samples (Afara 2012). The common statistical parameters used to evaluate the models are: standard error of calibration (SEC), root mean square error of calibration (RMSEC), coefficient of determination for calibration (R^2), the standard error of prediction (SEP), root mean square error of prediction (RMSEP), and the coefficient of determination for prediction (R^2) and bias (Williams 2001). The mathematical definition for SEP is shown below:

$$SEP = \sqrt{\frac{1}{N-1}\sum (x_i - y_i - b)^2} \tag{4.10}$$

Where $x_i - y_i$ is the difference between the NIR predicted response (x_i) and the reference method (y_i) on sample i, and b is bias. Bias refers to the difference between the mean of actual and predicted values. Bias is defined as:

$$bias = \frac{1}{N}\sum (x_i - y_i) \tag{4.11}$$

Where N is the total number of samples in the test set.

The RMSE is calculated from the difference between the NIR predicted response and the reference method values. RMSE is calculated using Equation 4.12 (Nicolaï et al. 2007). This equation can be extended to calculate the RMSECV and RMSEP. The RMSECV is calculated in the same way as RMSEP, the difference being that the RMSECV is obtained when cross-validation is used, while RMSEP is calculated when internal or external validation is used (Næs et al. 2004).

$$RMSE = \sqrt{\frac{1}{N}\sum (x_i - y_i)^2} \tag{4.12}$$

RMSE combines SEP and bias in a single relationship given by:

$$RMSE^2 = SEP^2 + bias^2 \tag{4.13}$$

Where the bias is insignificant, the RMSEP tends towards the SEP with an increasing number of samples. Generally, the RMSEP give more realistic estimate of the prediction capability of the calibration than SEP. The SEP and RMSEP are calculated on validation set.

Another useful statistic is the R^2 value. R^2 essentially represents the proportion of explained variance of the response variable in the calibration or validation set. The R^2 gives an indication of how well the predicted response fits the measured response. The R^2 is defined by (Afara 2012):

$$R^2 = \frac{\sum(x_i - \tilde{x})(y_i - \tilde{y})}{\sqrt{\sum(x_i - \tilde{x})^2 \sum(y_i - \tilde{y})^2}} \quad (4.14)$$

Where $\tilde{x}$ is a mean of reference for the calibration population, and $\tilde{y}$ is a mean predicted value of samples in prediction/validation population. A proper model should have a low SEC, SEP and RMSEP and a high coefficient of determination for both prediction and calibration models. In this thesis, the optimal calibration equations were chosen based on the highest R^2 values and the lowest SEP values.

Other statistics used to assess the PLS model performance was the RPD. The RPD is defined as the ratio of the standard deviation of the reference data (SD) to SEP (Williams 2008). An RPD between 1.5 and 2 means that the model can discriminate from low to high values of the response variable; a value between 2 and 2.5 indicates that coarse quantitative predictions are possible, and a value between 2.5 and 3 or above corresponds to good and excellent prediction accuracy (Nicolaï et al. 2007).

4.7 Qualitative measurement for spectral data

Spectroscopic methods have been successfully used in the sugarcane industry for both qualitative and quantitative measurements (Madsen et al. 2003; Valderrama et al. 2007; O'Shea et al. 2011). Throughout this research, the applications of quantitative analysis on the spectral data have also been shown. In this section, the application of qualitative analysis on the spectral data to classify sugarcane quality into several quality groups is investigated. Classification is a predictive method where the response is a category variable. The purpose of the classification is to predict which category a new sample belongs to. Classification of sugarcane quality in real-time during harvesting is a very important step in mapping quality variation across sugarcane paddocks. In their work, Bramley et al. (2012) classified sugarcane quality into several quality classes to map sugarcane quality in the paddock.

There are many classification methods available in the literature, including multivariate analysis of variance (MANOVA), cluster analysis (CA), and Bayesian discriminant analysis (BDA), support vector machine (SVM) and the artificial neural networks (ANN) (Lee et al. 2010). ANN often gives higher recognition and prediction probability than other methods based on statistical classification algorithms (Lee et al. 2010). Therefore, in this thesis, the application of the ANN method was explored.

ANN is a well-known non-linear method which could provide a robust classification model (Lee et al. 2010). Inspired by the biological nervous system, the neural networks are composed of a number of elements operating in parallel. By adjusting the weights (connections) between elements, a neural network can perform several functions, such

as prediction, data filtration, data conceptualization, classification, and data association (Haykin 1994). Most applications of ANNs in post-harvest technology have been for classification purposes (Guyer & Yang 2000; Kim et al. 2000; Hahn et al. 2004). The applications of ANN to classify crop quality have been reported by Wu et al. (2008) and Xing et al. (2010). To date, no published study has reported on the application of ANN classification methods to classify sugarcane quality based on spectral data. Thus, the objective of this section was to attempt to classify sugarcane quality based on spectral data as measured by the SSM using ANN algorithm.

4.7.1 Description of the artificial neural networks (ANN)

The ANN algorithm discussed in this section (including Figure 4.16 and 4.17 and Eqns. 4.15-4.20) follows the procedures developed by Heermann & Khazenie (1992). Detailed of the back-propagation learning algorithm including the derivation of equations can also be found in McClelland & Rumelhart (1986). The ANN used in this chapter was a perceptron model, also known as back-propagation perceptron, selected because it is an excellent pattern classifier (Torrecilla et al. 2004). The most widely used ANN is the multilayer perceptron (MLP) which typically consists of three layers namely: the input, hidden, and output layer. The edge layer, where information is presented to the network, is the input layer. The layer on the far side, where the processed information is retrieved, is known as the output layer. All layers in the middle are known as hidden layers (Figure 4.16). All neurons (often called 'nodes') in the network (Figure 4.17) except for the input neurons perform two functions: collecting the activation of neurons in the previous layer and setting an output activation.

Every neuron of the input layer is connected to every neuron of the hidden layer, and every neuron of the hidden layer is connected to every neuron of the output layer. A neuron is a computational device that calculates the weighted sum of its inputs and calculates the output signal from this using a nonlinear function. The spectral value at every wavelength is fed to the input layer, while the output layer delivers the prediction of the attribute. The weights are estimated using an appropriate algorithm based on a calibration set using cross validation (Kim et al. 2000).

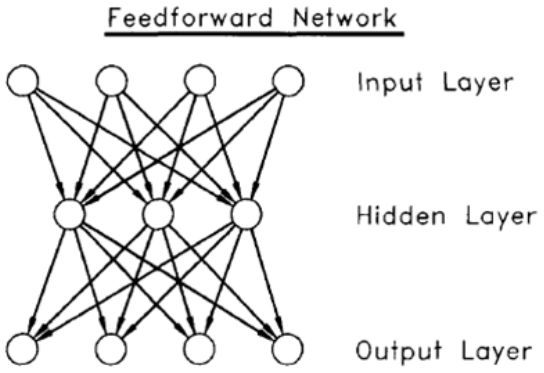


Figure 4.16: Three layers of feed-forward ANN

ANN only allows the information to be passed in one direction through the networks, starting at one side of the network and moving through successive layers. This type of network is known as a feed-forward network and is presented in Figure 4.17. ANN consists of an input layer which receives one piece of spectral data each and distributes it to a hidden layer where the data are transformed before being distributed again to a set of output layers. The neurons/nodes simply provide the mechanism for distribution of the data through the hidden layer which may be considered as a nonlinear function approximation machine.

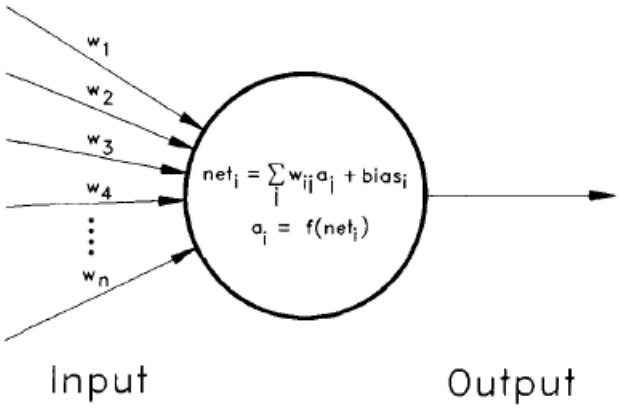


Figure 4.17: Neuron/Node computation

The input nodes activation are determined by the input data. The collection function used in this study is:

$$net_{pi} = \sum_{j} w_{ij}\, a_{pj} + bias_i \tag{4.15}$$

Where variable w_{ij} represents the connection strength for the current node i to a node j in the previous layer, a_{pi} is the activation of node j for pattern p plus a node bias, $bias_i$, which could be considered a connection to a node which is always at full activation. The result of the collection, net_{pi}, is passed to the output section which determined the node's output activation, a_{pi}. The output function was a non-linear function which allows a network to solve non-linear problems. In this study the sigmoid function given in Eqn. 4.16 is used to determine the output state.

$$a_{pi} = \frac{1}{1 + e^{-net_{pi}}} \tag{4.16}$$

A back-propagation network is trained by example. A set of representative input and output patterns was selected. As each input pattern was presented, the connections of the network were adjusted so that the activation of the output nodes more closely matched the desired output pattern. All the patterns were repeatedly presented to the

network until the network 'learnt' the patterns. The foundation of the back-propagation learning algorithm was the non-linear optimization technique of gradient descent on the sum of the squared differences between the activation, O_{pi}, of the nodes in the output layer and desired output t_{pi}. The global error of the net, E, is defined as:

$$E = \sum_{p} \sum_{i} (t_{pi} - O_{pi})^2 \tag{4.17}$$

where p indexes the training patterns and i indexes the output nodes of the network. By adjusting the network connection strength, wij, the above function is minimised and the network 'learns' the patterns. Application of the gradient descent method yielded the following iterative weight update rule:

$$\Delta w_{ij}(n + 1) = \varepsilon (\gamma_{pi} a_{pj}) + \alpha \Delta w_{ij}(n) \tag{4.18}$$

where w_{ij} is the connection strength from node i to node j, γ_{pi} is the node i error for pattern p, a_{pj} is the activation of node j for pattern p, and ε is a parameter known as the learning rate and the parameter α controls the momentum term. The node error, γ_{pi}, for an output node is then given as:

$$\gamma_{pi} = (t_{pi} - a_{pi}) a_{pi} (1 - a_{pi}) \tag{4.19}$$

where the first term is the error between an output node's activation and the target pattern. The other terms (Eqn. 4.19), are the result of the derivative of the activation function. The error at an arbitrary hidden node i is

$$\gamma_{pi} = a_{pi}(1 - a_{pi}) \sum_{k} \gamma_{pk} w_{ki} \tag{4.20}$$

The summation in Eqn. 4.20 collects the errors from the layer below and the other terms the derivative of the activation function.

4.8 Conclusions

This chapter has detailed the methodologies applied to measure sugarcane quality parameters using spectroscopic methods (VNIRS and FRS) from different sample forms: whole stalk samples, internode samples, fibrated sample and juice samples. The measurement setup of two spectroradiometers equipped with different detectors in predicting quality parameters from these sample forms has been described. The SSM and CSSM were developed and applied on both whole stalk and internode samples. The spectral measurement from juice and fibrated sugarcane samples was also described. The light-proof measurement box, serving as an instrument platform, was built to block ambient light from affecting the spectral measurements and to provide a consistent measurement setup for every measurement on different sample forms. The measurement of quality parameters using standard laboratory procedures (reference

values) for each sample form was also discussed. These reference values together with the spectral data were used to build a regression model using the PLS method. The application of chemometrics to pre-process the spectral data and build calibration models was implemented using Unscrambler V 9.6 software. Finally, it has been proposed that the quality of the calibration models developed in this study was assessed using R^2, RMSEP and RPD values.

Chapter 5

Application of VNIRS to predict sugarcane quality from juice and fibrated sugarcane samples

This chapter describes a preliminary study on the application of the VNIRS to determine sugarcane quality parameters from the conventional sugarcane sample forms: raw juice, clarified juice and fibrated sugarcane samples. The spectral measurements for all sample forms were conducted inside the measurement box under the same experimental setups with the same scanning procedures. Based on these results, this chapter also discusses the potential applications and limitations of each sample form for quality measurements in the field.

5.1 Introduction

Chapter 3 discussed the application of the laboratory-type NIR spectrometers to measure sugarcane quality parameters either from juice or fibrated samples. However, all of the laboratory-type spectroscopic technologies reported in that chapter are 'fundamentally' not practical for field uses due to size limitation, power requirement and high sensitivity to dust and vibration. These laboratory-type equipment are also expensive and require careful handling by well-trained personnel. Therefore, the measurement of sugarcane quality in the field using the portable and low-cost VNIRS is more practical and economical than using stationary and expensive laboratory equipment. However, there is no published study regarding the application of the VNIRS as a low-cost tool for predicting quality parameters in the sugarcane industry, especially in the field.

Thus, this chapter aims to explore the capability of the VNIRS to determine sugarcane quality parameters from the conventional sugarcane sample forms: raw juice (RJ), clarified juice (CJ) and fibrated samples (FS). This chapter will quantify the relative accuracy of each sample form as measured under the same experimental setups in order to identify the optimum sample form for field uses. The specific objectives of this chapter are:

1. To investigate the feasibility of using the VNIRS to predict sugarcane quality components from RJ, CJ and FS.
2. To identify the optimum sample form from these three candidates based on the highest prediction accuracy.
3. To compare the prediction accuracy between transmittance and reflectance measurement modes for juice samples.
4. To discuss the potential applications and limitations of each sample form for possible field uses.

5.2 Statistical characteristics of the conventional samples

The statistical characteristics of the quality components for sugarcane samples used in this chapter for both calibration and prediction data sets are presented in Table 5.1. Detailed information on sugarcane samples and the methods used to process the samples were discussed in Section 4.3.3. The table shows that the quality components in the calibration and prediction data sets have similar means, ranges and standard deviations, indicating that selection of the data in each model was appropriate. The ranges of °Brix and pol in this table are comparable to the °Brix (21.5 to 24.7) and pol (82.5 to 99.9) values reported by Berding et al. (1991b). Since the sugarcane stalks were collected from different locations within Bundaberg region throughout the 2012 harvest season, the °Brix and pol values in both data sets could be considered sufficient to represent the variation of typical quality values of sugarcane during harvest.

Table 5.1: Statistical characteristics of sugarcane quality components for juice and fibrated samples

Component	Data sets	Max	Min	Mean	SD
°Brix	Calibration	25.2	19.3	22.7	1.48
	Prediction	24.7	19.3	22.8	1.46
Pol	Calibration	100.6	71.0	88.3	8.07
	Prediction	98.8	71.7	88.8	7.31
Fibre	Calibration	15.2	9.8	12.3	1.25
	Prediction	15.6	10.4	12.7	1.77

5.3 Overview of the spectral curves for the reflectance and transflectance measurements

Typical absorbance spectra for RJ, CJ and FS are shown in Figure 5.1. The absorbance spectra for all sample forms were collected within the same measurement box and constant experimental setups. All of the spectral data were pre-processed with the MSC method. The spectral data for the conventional sample forms were calibrated against the average quality values obtained from the whole stalk samples. From the figure, it can be seen that all curves showed a downward trend as the wavelength increased. A higher absorption level by RJ was due to the presence of impurities, colour pigments and fine fibres (Cadet & Offmann 1997). Except for CJ, all spectral curves displayed absorption peaks at 680 nm. The peak at 680 nm could be related to chlorophyll content (Abbott et al. 1997). For CJ, the chlorophyll content was eliminated through a clarification process. The absorption peaks around 975 to 980 nm could be related to soluble solid content (Shao et al. 2007).

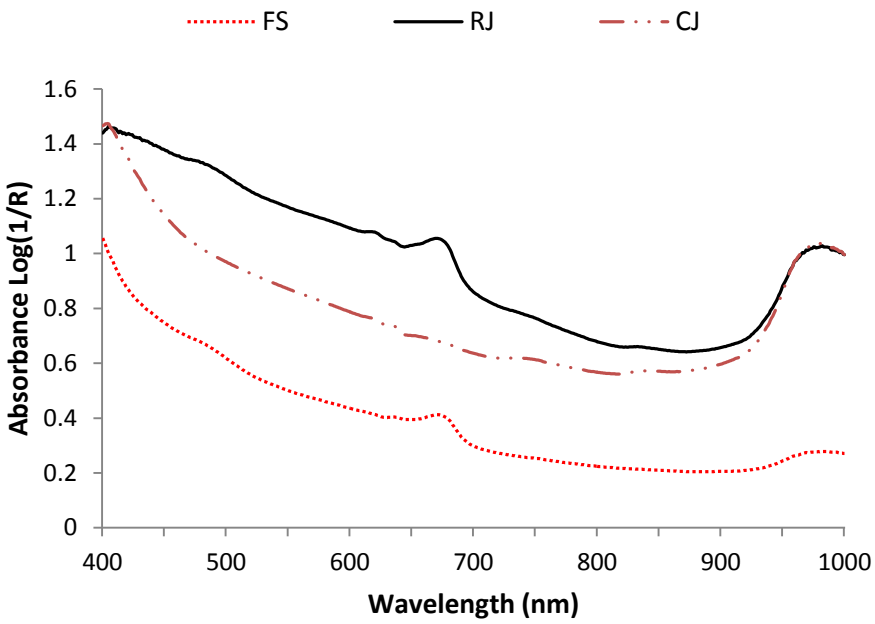


Figure 5.1: Typical absorbance spectra for different sample forms

5.4 Prediction of sugarcane quality parameters from conventional samples

PLS models were developed and applied for both calibration and prediction data sets. The performances of the models were evaluated using R^2 and RMSEP. The performances of the models in predicting quality components from the conventional sample forms are shown in Table 5.2. The table shows that °Brix was the best quality component to be predicted by the VNIRS from all sample forms. It can be seen from the table that FS had the highest prediction accuracy for °Brix prediction (R^2 = 0.86), followed by CJ (R^2 = 0.84) and RJ (R^2 = 0.80).

For pol prediction, the FS and CJ shared the same prediction accuracy with R^2 being 0.81. The prediction of pol from RJ yielded a relatively lower accuracy with R^2 value of 0.76 and RJ had the lowest quality prediction because it was a completely opaque solution containing impurities and fibres. These substances would absorb energy from light source thus influencing the prediction of sucrose content in the juices. Hence, filtration was necessary before RJ was analysed by spectroscopic method. Although CJ could give higher prediction accuracy than RJ, the processing of RJ into CJ required complex sample preparation and was a time consuming task. RJ, which normally contains fibres and impurities would need to be processed with lead acetate which precipitates impurities, and then filtered on a cellulose filter so as to obtain CJ for quality analysis. Since 20 minutes is needed to carry out these procedures, it is not practical for rapid measurement in the field.

Table 5.2: PLS models performance for each quality component of different sample forms

Sample form	Component	LVs	Calibration		Prediction	
			R^2	RMSEC	R^2	RMSEP
FS	°Brix	10	0.85	0.81	0.86	0.84
	Pol	10	0.85	4.28	0.81	5.21
	Fibre	10	0.71	0.96	0.63	1.07
CJ	°Brix	6	0.83	0.81	0.84	0.83
	Pol	9	0.83	4.29	0.81	5.32
	Fibre	5	0.38	1.21	0.19	1.51
RJ	°Brix	8	0.86	0.75	0.80	0.87
	Pol	8	0.76	5.21	0.76	5.11
	Fibre	6	0.60	1.07	0.26	1.68

The estimation of fibre content by the VNIRS from FS was not satisfactory, with R^2 and RMSEP values being 0.68 and 1.07% respectively. The prediction accuracies of fibre content from both CJ and RJ were not reliable since the analysis of expressed juice by NIR spectroscopy did not provide an estimate of fibre content (Berding et al. 1991b). However, juice analysis has the advantage of using a more uniform sample. This approach has potential applications in early selection stages, and also offers considerable promise for process control in manufacturing.

Overall, the prediction accuracies obtained in this chapter were lower than similar studies available in the literature. For example, for quality prediction from FS, Berding et al. (1991a) who used spectroscopic method in the wavelength region between 1445 and 2348 nm, reported the R^2 for °Brix, pol and fibre content were 0.91, 0.96 and 0.89, respectively. For CJ, excellent prediction accuracies for °Brix (R^2 = 0.97) and pol (R^2 = 0.98) were reported by Berding et al. (1991b). In their study, Berding et al. (1991b) used a spectrophotometer with the wavelength range from 1445 to 2348 nm. Cadet & Offmann (1997) used Fourier transformed MIR attenuated total reflectance (8000-12500 nm) to predict pol from raw juice samples. In their study, excellent prediction was reported, with R^2 value being 0.98. A relatively lower accuracy reported in this chapter was expected as the VNIRS had a limited useful wavelength range (400-1000 nm). Lu et al. (2000) reported that a greater spectral region beyond 1100 nm could give improved prediction results for sugar content. However, the relatively lower accuracy obtained in this study may be acceptable as the equipment used is low-cost, portable and more suited for field application.

5.5 Overview of the spectral curves for juice samples using transmittance mode

The previous section has shown the prediction accuracy for several quality components from juice samples (CJ and RJ) using the reflectance mode. However, transmittance mode is the most common spectroscopic measurement method for juice samples. This section discusses the prediction accuracy for °Brix and pol from CJ and RJ based on the transmittance measurement mode. Typical absorbance spectra from transmittance measurement for both RJ and CJ using 10 mm optical path length are shown in Figure

5.2. Both curves show a downward trend as the wavelength increased, similar to the curves of the reflectance measurement. A higher absorption level by RJ at 680 nm could be related to chlorophyll pigments (Abbott et al. 1997). It can be seen from the figure that there were spectral noises for both RJ and CJ in the range of 400-500 nm. Hence, to afford a better comparison with a better predictive performance, only the wavelength region from 500 to 1000 nm was used for further analysis in this chapter.

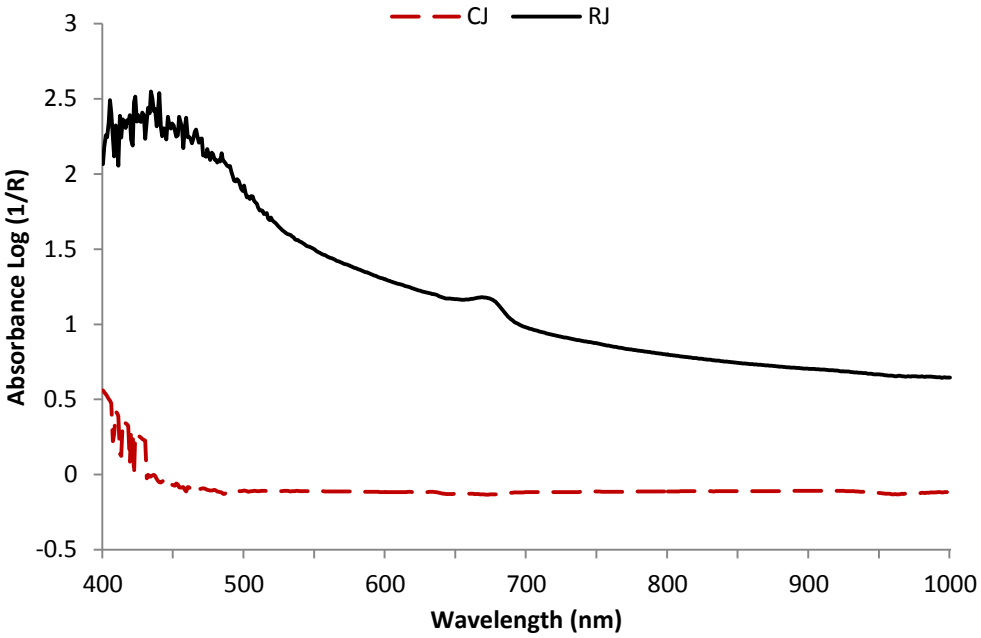


Figure 5.2: Typical absorbance spectra of CJ and RJ as measured by transmittance mode

5.6 Prediction of quality components from juice samples using transmittance measurement mode

PLS models were developed to correlate the spectral data with the sugarcane quality components in both calibration and prediction data sets. Table 5.3 shows that both calibration and prediction models for CJ were good with all R^2 values being above 0.80. The prediction accuracy for °Brix and pol from CJ was good with R^2 values being 0.84 and 0.85, respectively. For RJ, the performance of calibration models for both °Brix and pol were good with R^2 values being 0.84 for both components (Table 5.4). The prediction accuracy for °Brix (R^2 = 0.72) and pol (R^2 = 0.75) from RJ were lower than CJ, but were still acceptable for field screening, as complex sample preparation and processing could be avoided.

Table 5.3: PLS model performances of CJ for transmittance measurement

Parameter	LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP
°Brix	7	0.86	0.80	0.84	1.01
Pol	8	0.88	4.01	0.85	4.49

Table 5.4: PLS model performances of RJ for transmittance measurement

Parameter	LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP
°Brix	10	0.84	0.83	0.74	0.99
Pol	10	0.84	4.62	0.73	5.68

For both juice samples, prediction accuracy for pol was higher than that of °Brix. This finding was consistent with the study reported by Berding et al. (1991b) who also found that the prediction accuracy for pol (R^2 = 0.98) was higher than that of °Brix (R^2 = 0.97). Pol yielded a better prediction performance than °Brix probably because pol is an estimation of sucrose content in juice whereas °Brix is an estimation of soluble solids content in the juices. In comparison to the reflectance measurement mode, the transmittance measurement mode has improved the prediction accuracy (R^2) of pol in CJ from 0.81 (reflectance mode) to 0.85 (transmittance mode). Through the comparison between reflectance and transmittance measurement modes on honey samples, Qiu et al. (1999) also reported that the transmittance measurement mode performed better than the reflectance measurement mode.

5.7 Potential applications of the VNIRS on conventional samples in the field

Section 5.4 has shown that the VNIRS could be used to predict several quality components (°Brix, pol, and fibre content) from the conventional sample forms. Among them, °Brix appeared to be the best quality component to be predicted by the VNIRS. The prediction of °Brix component from FS yielded the highest accuracy (R^2 = 0.86), followed by CJ (R^2 = 0.84) and RJ (R^2 = 0.80). The prediction accuracy for pol from FS and CJ were good with R^2 values of 0.81 for both sample forms. The prediction of fibre content from FS was not satisfactory with R^2 and RMSEP were 0.63 and 1.07%, respectively. The prediction of fibre content from juice samples however, was very poor.

Based on these prediction accuracies, it is suggested that FS could be the optimum sample form for quality prediction using the VNIRS. The application of the VNIRS to predict °Brix and pol from both juice samples were good, where the prediction accuracy of °Brix and pol from CJ was better than that of RJ. Although CJ had better prediction accuracy than that of RJ, the preparation procedure for CJ in the field would be difficult. The requirement for a clarification process to remove both soluble and insoluble impurities using lead acetate would not be practical in the field. Therefore, the measurement of quality parameters from RJ may be more practical even with a lower accuracy. The following sub-sections discuss the potential application and limitation of FS and RJ in the field.

5.7.1 Quality measurement for breeding programs

Recently, there has been a growing desire among the growers and researchers to measure sugarcane quality in the field. Since the VNIRS is a portable instrument, it has great potential to be used in the field to measure sugarcane quality parameters from

either FS or RJ. For breeding evaluation programs, the use of this equipment is possible if a trolley equipped with a sample preparing device powered by battery could be developed. The trolley could be used as a platform to prepare stalk samples into an intended sample form prior to the quality measurement by the VNIRS. A mini shredder installed on the trolley could be used to prepare stalk samples into FS. The shredder is required to macerate the stalk samples into a homogenous state. For RJ, the development of a small scale stalk crusher which can be installed on the trolley will be very useful. The purpose of the crusher is to crush stalk samples and supply a sufficient amount of juice for the measurement.

5.7.2 Quality measurement for an on-the-go system on a harvester

The FS and RJ have the potential to be used for field measurement. Between them, FS is a more promising sample form with higher prediction accuracy than CJ. However, in a real situation on a harvester, preparing both sample forms for the measurement would be difficult and technically challenging. For example, the measurement of sugar content from FS would require a macerator to be installed on a harvester for preparing the billet samples into macerated form. For juice measurement, additional equipment such as a portable or mini stalk crusher or squeezer would be needed to be installed on a harvester. Once the juice has been collected, it needs to be clarified prior to the quality measurement. The clarification process requires the mixing of a chemical, usually lead acetate and brewing for a few seconds. This procedure is time-consuming and technically difficult, especially during harvesting. Therefore, a new sample form which can be measured rapidly in the field without complex sample preparation should be investigated. For that reason, Section 3.8.2 has recommended that quality measurement should be determined directly from stalk samples.

5.8 Conclusions

It has been shown that the VNIRS can be used to measure sugarcane quality components from the conventional sample forms (FS, CJ and RJ). By means of PLS, a relationship was established between the spectral data and the quality components. The prediction performance for °Brix from all of the conventional samples was good. FS was found to be the best sample form for quality prediction by the VNIRS. The potential application and limitation for each sample form in the field have also been discussed.

It has been demonstrated that the portable and low-cost VNIRS could offer the possibility of predicting sugarcane quality in the field without the need for the costly and laborious analysis of conventional methods. However, the development of a preparation device such as a mini crusher or shredder would be needed to prepare stalk samples into an intended sample form. This requirement could limit the application of the VNIRS for quality prediction in the field. Therefore, the quality measurement method which can be performed on stalk samples without any sample preparation is more practical in the field. Thus, the following chapters will investigate the application of the VNIRS to predict sugarcane quality from stalk samples.

Chapter 6

Cross sectional scanning method

This chapter presents the results of the calibration and prediction models developed from CSSM for both whole stalk and internode samples. The variation of sugar content and prediction accuracy for both samples are also quantified. This chapter also describes the selection of the optimum quality component for predicting sugar content from stalk samples and investigates the influence of varieties on prediction accuracy.

6.1 Introduction

There are several published studies addressing the application of spectroscopic methods to predict sugar content from solid samples based on the CSSM, including melons (Sugiyami 1999), kiwifruits (Martinsen & Schaare 1998), sugar beet (Panigrahi & Hofman 2003), carrots (Quilitzsch et al. 2005) and pineapples (Chia et al. 2012). Panigrahi & Hofman (2003) used CSSM to develop an on-the-go measuring system to perform real-time measurement of sugar content for sugar beet during harvest. The system consisted of a spectrometer and sample preparation mechanism with a knife to slice a cross section of the sample for the measurement. Once the sample has been sliced, an illumination chamber radiates the exposed surface, and a sensor receives the reflected radiation. A spectrometer then converts the reflected radiation into a spectral signal.

In theory, CSSM was applied by looking at the interaction between the incident light and the physical and chemical properties of the sliced samples. This measurement method could potentially be adopted in the sugarcane industry by scanning the stool after it is cut by the base cutter during harvesting.

The goal of this chapter was to investigate the ability of the spectroradiometers to predict sugar content from the cross sectional surface of sugarcane stalk samples. The specific objectives of this chapter were:

1. To investigate the potential of spectroradiometers coupled with PLS models to determine sugarcane quality components from stalk samples using CSSM.
2. To quantify the variation of sugar content and prediction accuracy along the stalks and individual internode (between the node and internode).
3. To compare the prediction accuracy between the PLS models developed from the average spectral data and an individual spectrum data.
4. To investigate the effect of different spectral pro-processing methods on models accuracy.
5. To investigate the influence of different sugarcane varieties on PLS model performance.

6.2 Whole stalk samples

6.2.1 Statistical characteristic of the whole stalk samples

Table 6.1 shows the values of the quality components of stalk samples for each variety used in this study. Two sample sets were collected for each variety, and each sample set contained six stalk samples. The measurements were carried out over two consecutive days under the same experimental conditions. In order to determine the variation of sugar distribution along the length of the stalk, the whole stalks were cut into three sections (bottom, middle and top) with each section being scanned on the cut surface.

Table 6.2 shows the statistical characteristics of the whole stalk samples used for CCSM. The statistical values of °Brix, pol, and fibre content were obtained from the standard mill procedures as described in Sections 4.3.1.3, with the CCS being calculated using Eqns. 2.1-2.4.

Table 6.1: Values of quality components for each variety

Variety	°Brix		Pol		Fibre		CCS	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
KQ228	24.4	24.2	98.2	98.0	13.0	12.7	16.7	16.8
Q135	21.7	22.0	87.0	88.4	14.5	14.1	14.5	14.9
Q138	23.0	23.3	91.8	92.3	13.3	12.6	15.5	15.6
Q151	21.9	23.0	87.7	93.3	12.5	13.7	15.0	15.9
Q155	23.1	23.2	93.5	94.1	13.4	13.8	15.9	16.0
Q190	22.1	21.8	88.9	88.2	15.6	15.1	14.7	14.7
Q200	22.8	22.7	91.6	89.7	16.0	14.4	15.0	14.9
Q208	22.5	23.2	90.6	94.0	15.8	15.1	14.9	15.7
Q232	22.1	22.0	89.1	87.7	12.8	13.0	15.3	14.9
Q240	22.5	22.7	90.5	92.0	12.8	12.8	15.5	15.8

Table 6.2: Statistical characteristics of the whole stalk samples for CSSM

	°Brix	Pol	Fibre	CCS
Max	24.40	98.20	15.96	16.83
Mean	22.71	91.31	13.84	15.41
Min	21.70	86.97	12.48	14.54
SD	0.75	3.21	1.15	0.65

6.2.2 Determination of sugarcane quality components

Table 6.3 shows the performance of the calibration and prediction models in predicting sugarcane quality components from the spectral data collected from whole stalk samples using both FRS and VNIRS spectroradiometers. The calibration model was developed using 15 sample sets while the prediction model was developed using five sample sets. Each sample set consisted of six stalks. The spectrum for each stalk was computed by averaging the three spectra collected from bottom, middle and top sections

of the stalk. Then, the spectrum for each sample set was calculated by averaging the spectra collected from six stalk samples.

Table 6.3: Performance of PLS models in predicting sugarcane quality components using spectroradiometers

Component	FRS					VNIRS				
	LVs	Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
°Brix	8	0.94	0.23	0.86	0.97	8	0.96	0.21	0.83	0.85
CCS	7	0.91	0.24	0.73	0.71	8	0.98	0.13	0.73	1.03
Fibre	7	0.93	0.42	0.27	2.14	4	0.71	0.73	0.49	1.23
Pol	7	0.86	1.50	0.90	3.44	9	0.99	0.53	0.89	3.21

Table 6.3 shows that different numbers of latent variables (LVs) were needed for different quality components. These optimal LVs numbers were automatically calculated by the Unscrambler 9.6 software. The calibration models for all quality components for both spectroradiometers showed good accuracy. Most of them had R^2 values above 0.90. For both spectroradiometers, it can be seen that pol had the highest prediction accuracy, followed by °Brix, CCS and fibre content. The R^2 values for pol, °Brix, CCS and fibre as predicted by the FRS were 0.90, 0.86, 0.73, and 0.27, respectively. For the VNIRS, the R^2 values for the corresponding quality components were 0.89, 0.83, 0.73 and 0.49, respectively.

The level of accuracy found in this study was consistent with the study on fibrated sugarcane reported by Berding et al. (1991a). In their study, Berding et al. (1991a) reported that the prediction accuracy as indicated by R^2 for pol, °Brix, CCS and fibre were 0.96, 0.91, 0.91, and 0.89 respectively. Their study had higher R^2 values than this study because they used mechanically treated samples (fibrated sugarcane) which were homogenous and solid. However, this study has achieved better accuracy compared to work reported by Meyer & Wood (1988) on shredded sugarcane. In their study, Meyer & Wood (1988) reported that R^2 values for pol, °Brix, fibre and purity were 0.86, 0.88, 0.80, and 0.89, respectively.

Generally, the FRS yielded better prediction accuracy for all quality components compared to the VNIRS, except for fibre content. Pol had the highest R^2 value as predicted by both instruments. This showed that the pol value had the highest correlation with the spectral data. In contrast, fibre had the lowest R^2 values as predicted by both spectroradiometers. The values of R^2 for fibre prediction using the FRS and the VNIRS were 0.27 and 0.49 respectively. Berding et al. (1991a) also reported that the prediction accuracy for fibre content was the lowest among the quality components. The low R^2 values for the fibre content in both models were probably due to the fact that fibre content was not directly measured from the samples, but estimated using an equation. The table also shows that both VNIRS and FRS had the same capability for predicting CCS values.

6.2.3 Variations of prediction accuracy for different sugarcane sections

Table 6.4 shows the relative comparison of prediction accuracy for each cut section by both spectroradiometers. This table was developed by averaging the spectrum collected from the same section of each stalk for every sample set. Then, the averaged spectrum of each section for all sample sets were divided into three groups namely bottom, middle and top. For each group, 75% of the data was used for the calibration model while the remaining data was used for the prediction model. All groups were calibrated against the °Brix value of each sample set.

Table 6.4: PLS models' performance in predicting °Brix values from different cutting sections

Section	FRS					VNIRS				
	LVs	Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
Bottom	6	0.87	0.34	0.79	1.15	2	0.50	0.58	0.76	0.77
Middle	8	0.95	0.20	0.57	1.45	2	0.43	0.61	0.89	0.76
Top	9	0.99	0.10	0.71	0.97	9	0.98	0.12	0.82	0.51

Table 6.4 shows that the relative accuracy of both calibration and prediction models varied for each section. It can be seen that the variation of relative prediction accuracy for the FRS was 38.6%. For the VNIRS, the variation of relative accuracy between the sections was about 17.1%. For the FRS, the bottom section was found to have the highest relative accuracy (R^2 = 0.79) while the middle section had the lowest (R^2 = 0.57). Whereas for the VNIRS, the middle section was found to have the highest accuracy with R^2 of 0.89 while the top section had the lowest accuracy with R^2 of 0.82. The section with a high R^2 value had less spectral variation, thus should be the goal for measurement purposes in the field.

The mean absorbance spectra for each cut section as measured by the FRS are shown in Figure 6.1. The figure clearly demonstrates that all spectra had similar spectral patterns, though the difference between spectra representing bottom, middle and top section can still be clearly identified. The curves for the three sections are nearly overlapping with each other in the wavelength region below 1000 nm. From 1200 nm onwards, the bottom section had the highest absorption level followed by the middle and top sections. This difference may indicate that the bottom section had relatively higher sugar content than middle and top sections.

There are three water band regions (1355-1450, 1800-1950 and 2420-2500 nm) could be clearly identified from the original spectra (Figure 6.1). These regions are known as water absorption bands (Kumar et al. 2003; Mutanga et al. 2004). These wavelength regions were removed from the spectra (Figure 6.2) and excluded from the analysis to prevent excessive noise (Abdel-Rahman et al. 2010). Five major peaks could be seen from the graph: 640, 980, 1200, 1420, 1900 and 2400 nm.

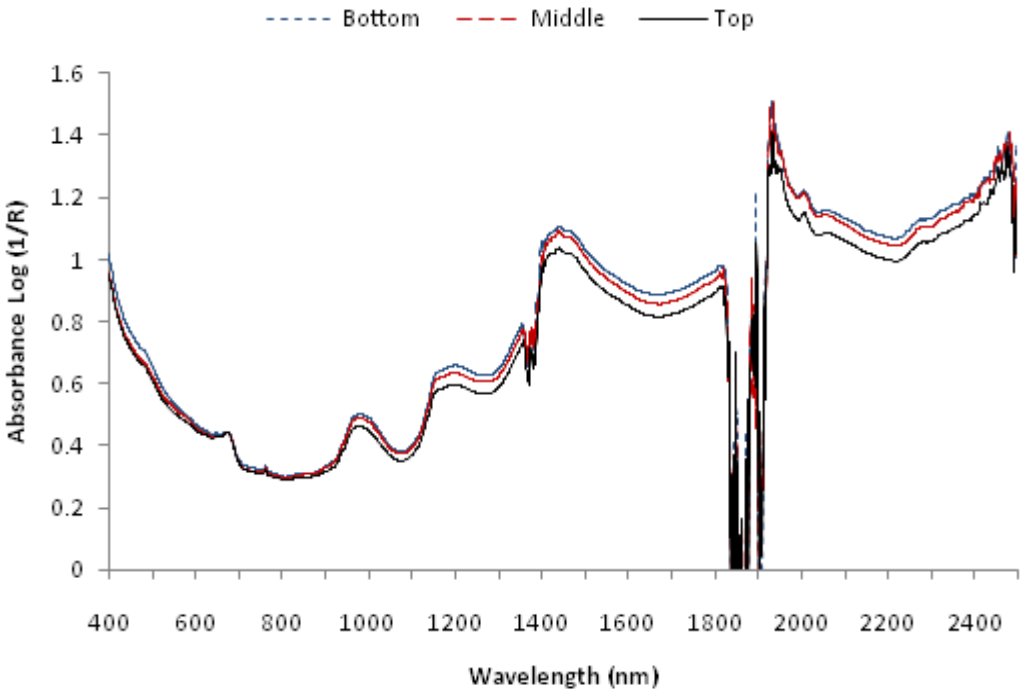


Figure 6.1: Typical absorbance spectra for bottom, middle and top sections as measured by the FRS

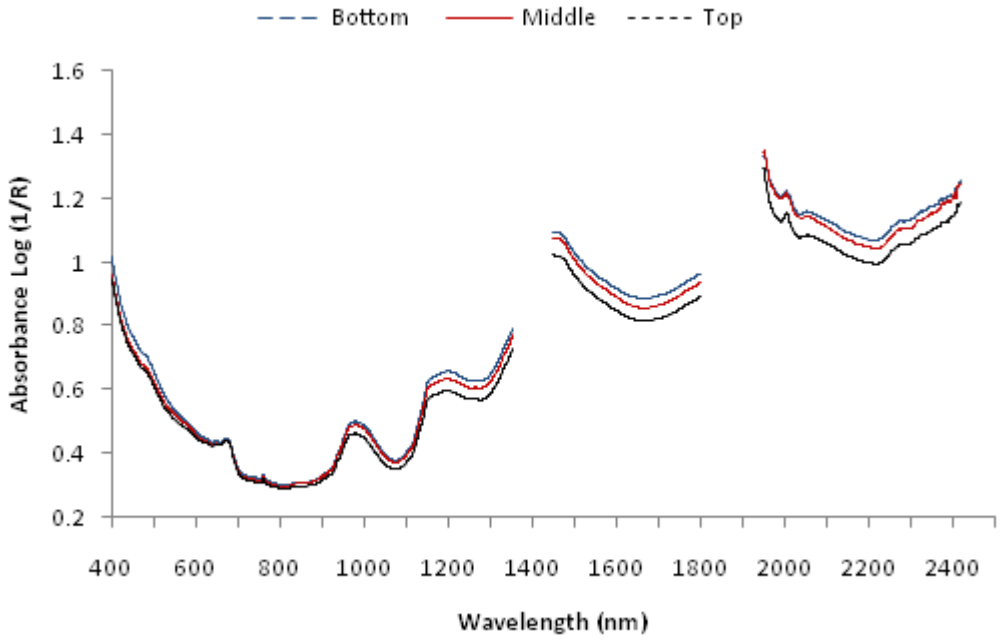


Figure 6.2: Typical absorbance spectra for bottom, middle and top sections as measured by the FRS (without water bands)

Figure 6.3 shows the typical spectral curve for CSSM as measured by the VNIRS. The figure illustrates that the curve for the top section appears to be located similar to the curve of middle section. The spectra for the middle and bottom sections overlap between each other. From this figure, three absorption peaks can be identified at around 680, 760 and 930 nm. The peak at 680 nm could be correlated with chlorophyll content (Abbott et al. 1997), while the peak around 760 nm represents OH stretching at the third overtone of sugar (Golic et al. 2003) and the peak around 930 nm represents CH_2 stretching at the third overtone of sugar (Osborne et al. 1993).

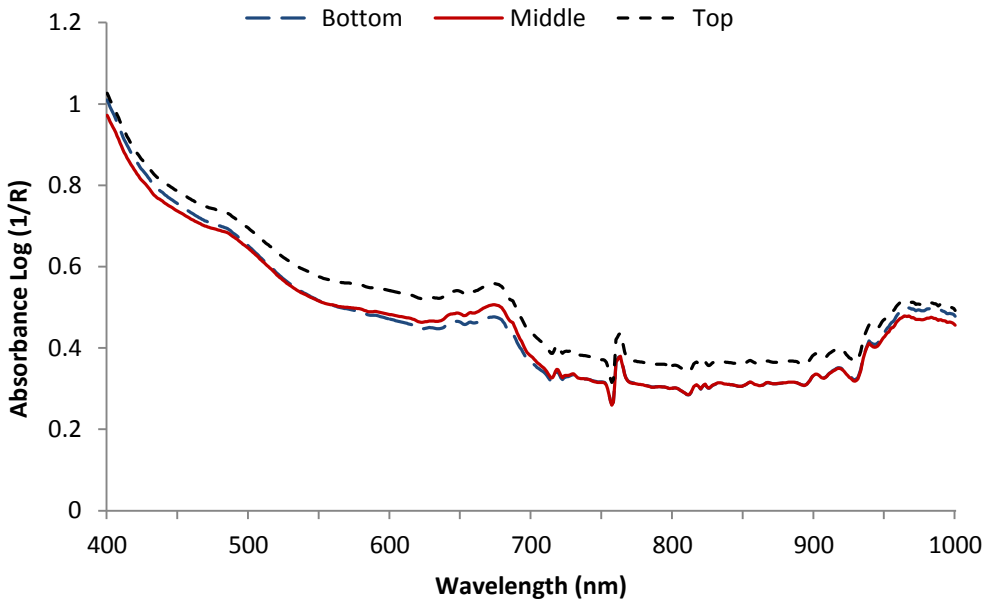


Figure 6.3: Typical absorbance spectra for bottom, middle and top sections as measured by the VNIRS

6.2.4 Relative prediction performance using individual spectrum data

This section investigates the robustness of the PLS models developed from the spectra data of each stalk sample, instead of spectra data from each sample set. It was expected that the model developed from more spectra numbers would be more robust. Instead of using the average spectral data for each sample set (each set consisted of six stalk samples), in this section, the PLS models were developed using the average spectrum of individual stalk samples. Then, each spectrum data from stalk samples was calibrated against the °Brix value of the corresponding sample set. Table 6.5 shows the performance of both the calibration and prediction models developed from the spectrum of each stalk sample. It can be seen that the prediction accuracy was generally lower than that reported in Table 6.3. This was expected as the spectral variation corresponding to °Brix value of each sample set increased.

Table 6.5: Quality prediction from stalk samples

Component	LVs	FRS				VNIRS				
		Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
°Brix	7	0.73	0.48	0.74	0.50	9	0.83	0.40	0.79	0.50
CCS	7	0.75	0.40	0.76	0.43	10	0.89	0.28	0.78	0.43
Fibre	5	0.74	0.75	0.70	0.78	10	0.89	0.50	0.76	0.71
Pol	7	0.73	2.08	0.74	2.16	10	0.88	1.46	0.83	1.97

In summary, this section has demonstrated the ability of spectroradiometers to predict sugarcane quality components from whole stalk samples. It has been found that the prediction accuracy of the study was reasonable but not very high ($R^2 < 0.90$). This is probably due to the variation between subsamples, reflecting the heterogeneous nature of raw sugarcane stalks. Moreover, since this experiment was conducted in the field and relied on the natural sunlight, the variation of weather may have influenced the accuracy of results. However, the relatively low accuracy values reported in this section are still acceptable as the complex sample preparation to produce a clarified juice can be avoided.

6.3 Sugar content prediction using internode samples

The previous section has demonstrated the application of the spectroscopic method to predict sugarcane quality components from the whole stalk samples using CSSM in the field. The study found good correlation between the spectral of the stalk samples and sugarcane quality components by both pieces of equipment. Even though there is a promising potential application of this method for breeding programs, the variation of sugar content along the stalk cannot be shown clearly. This method also did not consider the variation of sugar content along the height of an individual internode. Since sugarcane is a tall crop which grows from a series of nodes and internodes, it is expected that some variation may also occur along the length of an internode, especially between the node and internode areas. Thus, the variation within an individual internode should also be understood in order to achieve better accuracy for the actual field measurement. Obtaining reference values from a six stalks a sample is also time consuming and labour intensive.

For an ‘on-the-go’ measurement system, the measurements would be normally expected to be performed on billet samples on a harvester. Detailed discussion about a sampling mechanism to extract billet samples from the harvester’s elevator is presented in Section 8.5.3. Therefore, in order to have a measurement system which is sensitive to billet samples, the calibration model must be developed using internode samples. The reference value for internode samples is also easier to obtain, and could normally be completed using a handheld refractometer.

Furthermore, even though the previous section showed that the FRS performed better than that of the VNIRS, the application of the FRS in the field is not economical as it is significantly expensive piece of equipment. Therefore, despite the comparatively low (but acceptable) accuracy obtained by the VNIRS, the application of this low cost and portable instrument warrants further investigations. The improvement of sample

presentation, spectral measurement and statistical methods could increase the prediction accuracy of the equipment, thus making this proposed method attractive to the industry.

6.3.1 Sample properties and their spectral characteristics

Sugarcane is a living plant and hence each stalk is unique, varying in size, shape and sugar content. During the experimentation, variations in other qualities such as colour, stalk height, stalk diameter and internode length were noticed. These variations were observed when three commercial sugarcane varieties were used: Q155, Q208 and Q190. The chemical and physical characteristics of each variety are presented in Table 6.6, which summarises the °Brix values, internode length and internode diameter of each variety.

Table 6.6: Characteristics of chemical and physical properties for internode samples

Variety	°Brix				Internode diameter (mm)				Internode length (mm)			
	Mean	Max	Min	SD	Mean	Max	Min	SD	Mean	Max	Min	SD
Q155	18.5	22.2	7.6	1.2	28.1	33.0	24.3	2.5	141.2	215.0	76.0	21.0
Q208	17.3	21.4	8	2.7	28.6	30.9	25.9	1.4	135.9	190.0	66.0	20.3
Q190	17.3	21	8	2.4	32.9	36.5	29.8	2.0	137.4	208.0	82.0	24.3

The table shows that the early maturing variety (Q155) had the widest °Brix ranges. Q155 was also found to have the widest range of internode length. In terms of internode diameters, the late maturing variety (Q190) had the largest diameter compared to other varieties. The internode diameter is an important parameter in determining the scanning distance between the sensor and samples.

The summary of statistical characteristics for the calibration and prediction data sets of the internode samples is shown in Table 6.7. The calibration and prediction data sets showed similar means, ranges and standard deviations, indicating that the selection of samples for each data set was appropriate. A relatively wide range of °Brix values was obtained due to the inclusion of three different varieties with different stage of maturity. The range of °Brix values for internode samples from the top to the bottom of the Q155, Q208 and Q190 varieties were 7.6 to 22.2, 8 to 21.4 and 8 to 21, respectively.

Table 6.7: Statistical characteristics of °Brix values for internode samples

Model	No. of Samples	Min	Max	Mean	SD
Calibration	220	7.5	22.2	17.86	3.04
Prediction	72	8.2	22	17.83	2.93

The average variation of °Brix values along the stalks for each variety as measured from each internode sample is shown in Figure 6.4. The downward trend of °Brix values along the stalk was consistent for all varieties. The graph also shows that different varieties had different internode numbers. It can be seen that Q155 had higher °Brix values than the other varieties especially, for the internodes at the bottom of the stalks. The °Brix variation along the stalk for Q155, Q208, and Q190 was up to 52.4, 93.1 and 48.6%, respectively.

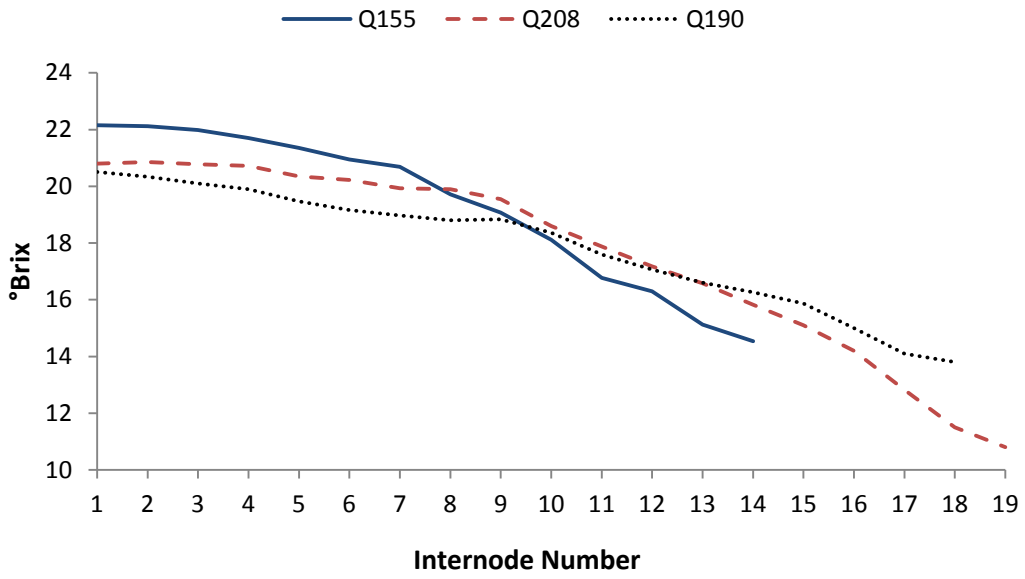


Figure 6.4: Typical average °Brix values of internodes for different sugarcane varieties

6.3.2 Effects of different pre-processing methods on prediction accuracy of CSSM

A large amount of spectral data collected from NIR instruments usually contains much useful analytical and background information as well as noise (Blanco & Illarroya 2002; Osborne et al. 1993). In order to obtain reliable, accurate and stable calibration models, it is necessary to pre-process spectral data before modeling (Cen & He 2007). All pre-processing techniques and calibration methods were carried out using Unscrambler V 9.6. The influence of different spectra pre-processing methods on PLS model performance for both reflectance and absorbance of the CSSM are shown in Table 6.8.

Table 6.8: The effect of different pre-processing methods on the PLS models performance for the CSSM

Pre-processing method	Reflectance				Absorbance			
	Calibration		Prediction		Calibration		Prediction	
	R^2	RMSEC	R^2	RMSEP	R^2	RMSEC	R^2	RMSEP
Raw	0.81	1.79	0.81	1.72	0.83	1.70	0.86	1.52
MSC	0.88	1.44	0.85	1.54	0.87	1.49	0.87	1.45
SNV	0.87	1.48	0.84	1.59	0.85	1.60	0.86	1.49
SG1	0.92	1.22	0.80	1.78	0.91	1.24	0.76	1.90
SG2	0.75	2.20	0.39	2.84	0.70	2.15	0.69	2.09
MN	0.86	1.58	0.82	1.68	0.84	1.66	0.86	1.49
MA (3)	0.80	1.87	0.79	1.77	0.81	1.78	0.86	1.53
MA (9)	0.77	1.94	0.77	1.84	0.79	1.86	0.85	1.60
MSC + SNV	0.87	1.48	0.84	1.59	0.85	1.60	0.86	1.49
MA(3) + SG2	0.91	1.24	0.81	1.75	0.73	2.08	0.49	2.60
SG2 + MSC	0.92	1.20	0.78	1.83	0.49	2.65	0.46	2.58
MSC + SNV + SG2	0.93	1.15	0.79	1.80	0.76	1.99	0.39	2.88

*** n for calibration model=220; n for prediction model=72.**

The pre-processing methods explored in this study included smoothing by moving average with three segments (MA3) and nine segments (MA9), multiplicative scatter correction (MSC), Savitzky-Golay first derivative (SG1), Savitzky-Golay second derivative (SG2), standard normal variate (SNV), mean normalization (MN) and combinations of them. For comparison purposes, the raw spectral data without any pre-processing method was also analysed. The models were compared with one another based on R^2 and RMSEP values. An accurate model should have a low RMSEP and high R^2. Although a low number of LVs are desirable, 10 LVs were used in this study to make the models comparable.

If no pre-processing was applied, the R^2 values for both reflectance and absorbance were 0.81 and 0.86 respectively. Absorbance data has indeed performed better than reflectance data. The highest R^2 value was obtained when the data was treated with MSC, with the R^2 values for reflectance and absorbance being 0.85 and 0.87, respectively. The RMSEP values for reflectance and absorbance were 1.54 and 1.45 °Brix, respectively. Therefore, MSC method was applied throughout this chapter. The MSC technique is the most popular normalization technique offered by most chemometrics software packages (Næs et al. 2004). MSC compensates for additive (baseline shift) and multiplicative (tilt) effects in the spectral data, which are induced by physical effects such as the non-uniform scattering throughout the spectrum because the degree of scattering is dependent on the wavelength of the radiation, the particle size and the refractive index. In contrast, the prediction models show lower accuracy when treated with SG2, or a combination of these methods.

6.3.3 Spectral overview of the internode samples

The typical absorbance and reflectance spectra of three internode samples having high (22), medium (18) and low (14.2) °Brix values, as measured by the VNIRS, are shown in Figures 6.5(a) and 6.6(a), respectively. In both figures, no obvious difference could be seen in the shape of the spectra for different °Brix values. In the raw spectra curves, gaps could be clearly observed among these three spectra in the region of 700 to 1000 nm regions. These spectral differences which were due to samples molecular vibration may assist in constituent analysis in this study.

Due to light scattering in the stalk samples, some consistent baseline shifts and bias appeared in the original spectra. A longer light path produces a lower relative reflectance value. This causes spectral translation and affects the spectral model. Pretreatment of spectral data is a key part of spectral analysis and can improve its accuracy. After applying the MSC pre-processing method, a baseline shift problem that existed in the original absorbance and reflectance spectra was eliminated (Figure 6.5(b) and 6.6(b)).

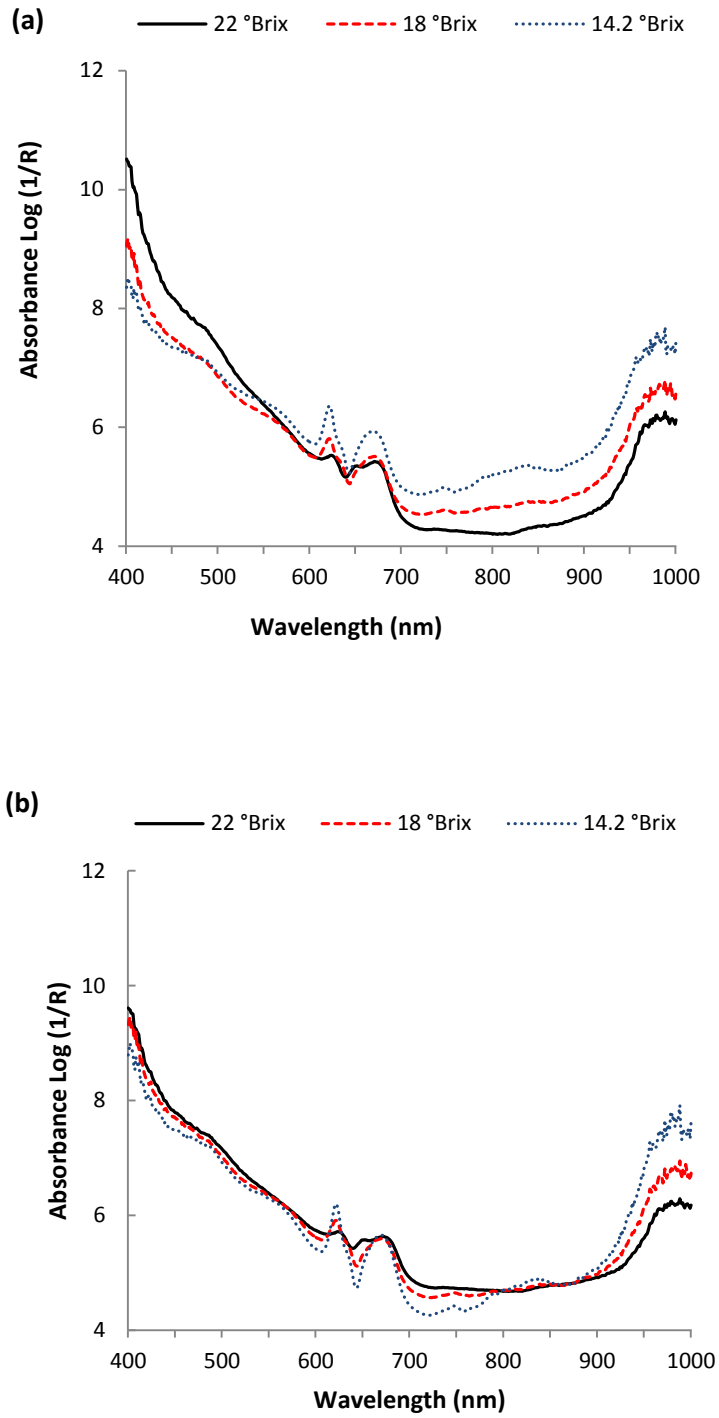


Figure 6.5: Typical absorbance spectra of CSSM at different °Brix values: (a) raw absorbance spectra; (b) absorbance spectra pre-processed with MSC

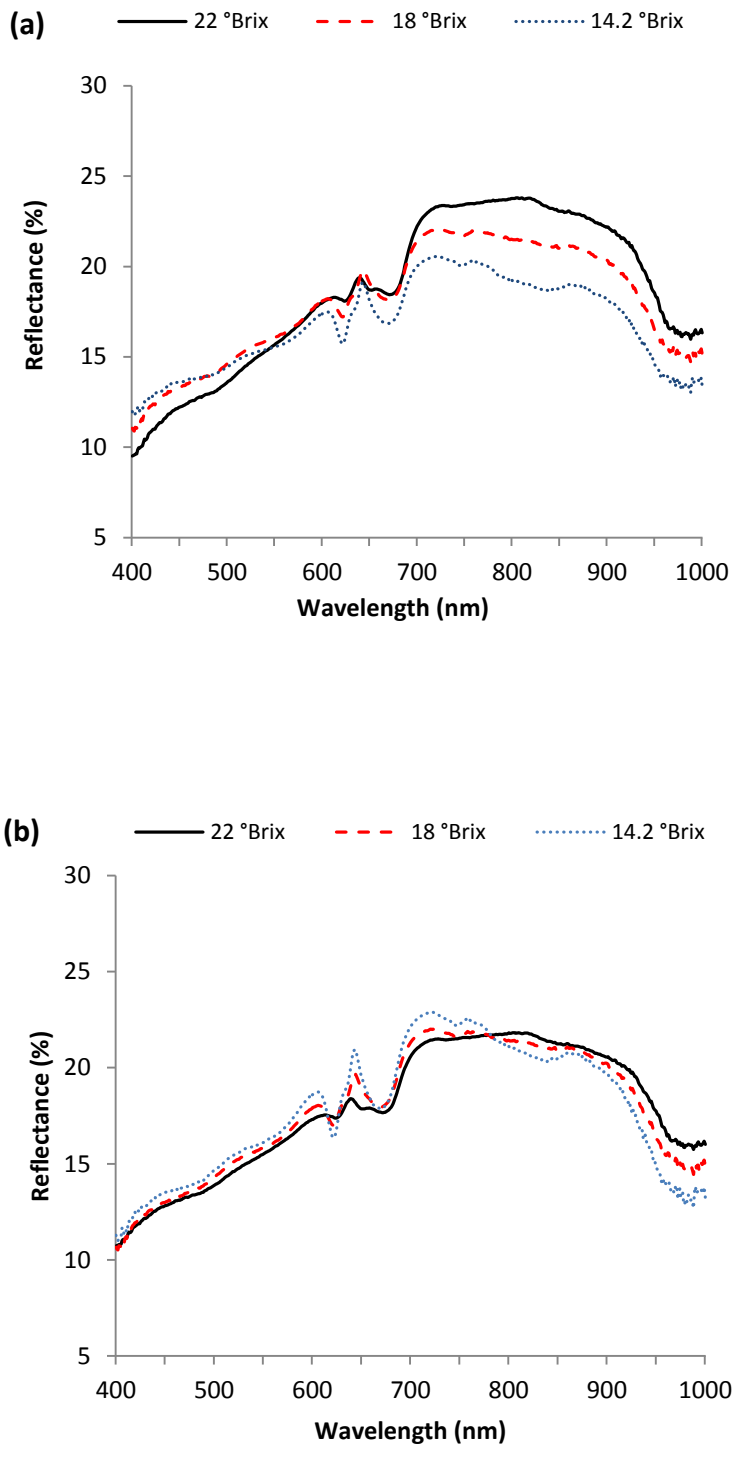


Figure 6.6: Typical reflectance spectra of CSSM at different °Brix values: (a) raw reflectance spectra; (b) reflectance spectra pre-processed with MSC

6.3.4 °Brix prediction by CSSM for the average spectral data

The four absorbance spectra of each internode (C1 to C4) were averaged into one internode spectrum. This internode spectrum data was used for calibration against the °Brix values of the corresponding internode sample. The performances of both calibration and prediction models for absorbance data are presented by the scatter plots in Figure 6.7 (a) and (b), respectively. The R^2 and RMSEC values for calibration model were 0.87 and 1.49 °Brix, respectively. The prediction accuracy of this model was reasonably good with a high R^2 value of 0.87 and a low RMSEP value of 1.45 °Brix, respectively. The RPD value of this model was 2, indicating that this method is suitable for quantitative prediction especially for screening purposes in breeding programs.

For the reflectance data, it was found that the accuracy of the calibration model was also comparable with the absorbance spectra with R^2 and RMSEC values of 0.88 and 1.44 °Brix respectively. However, the prediction model for this type of spectral data was lower than the absorbance spectra with R^2 and RMSEP were 0.85 and 1.54 °Brix, respectively. The performances of both calibration and prediction models for reflectance data are presented by the scatter plots in Figure 6.8 (a) and (b), respectively. The RPD value for the model developed from the reflectance data was 1.9, indicating that this model can only be used to discriminate from low to high values of the response variable (Nicolaï et al. 2007). Since the absorbance data was superior to the reflectance data, only the results for the absorbance data will be discussed in the following sections. Even though Shenk & Westerhaus (1996) suggested that the R^2 value greater than 0.9 indicates excellent quantitative information, the accuracy obtained from this study could be considered as good, noting the heterogeneous nature of the stalk samples. The results indicated that the VNIRS and PLS models could provide a satisfactory method for predicting °Brix from stalk samples using CSSM.

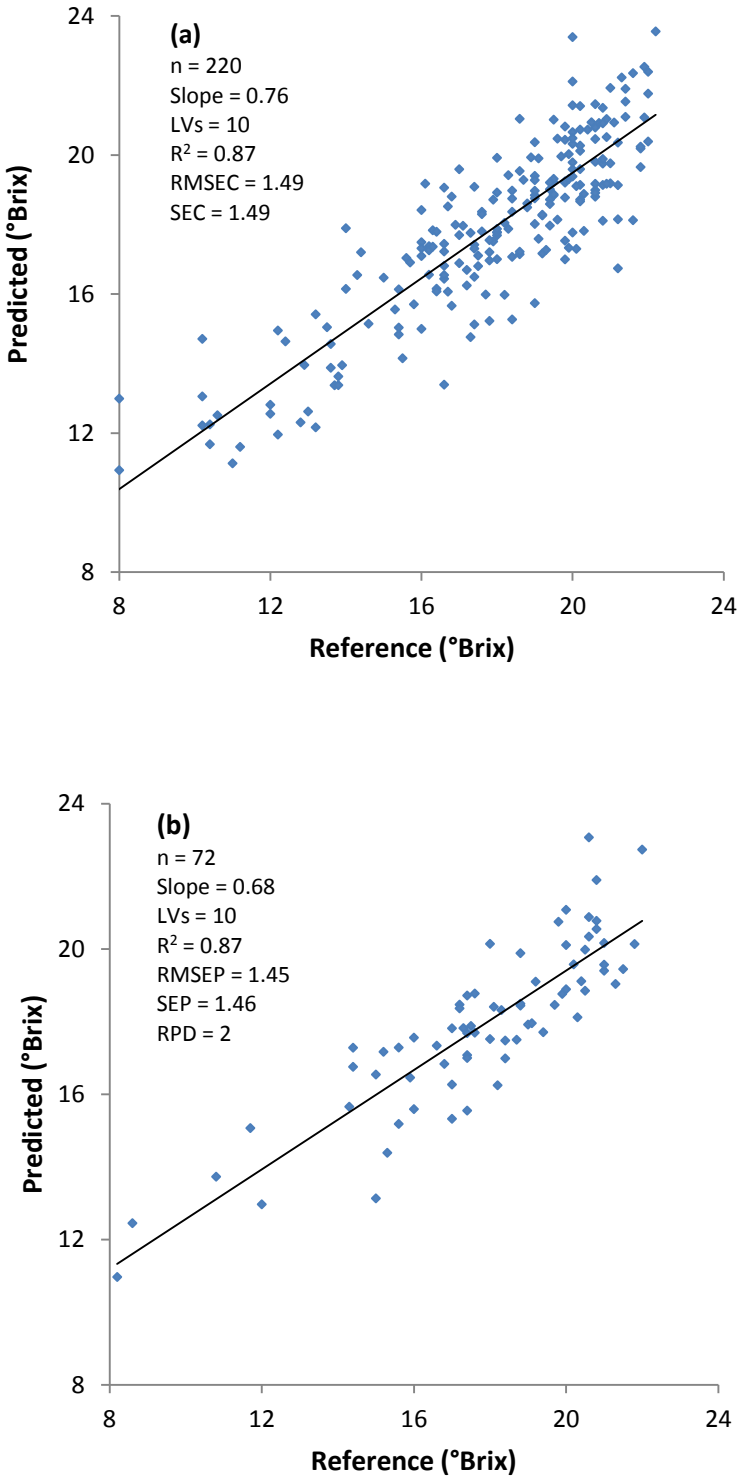


Figure 6.7: Scatter plots of reference versus predicted °Brix for absorbance spectral data by the CSSM: (a) calibration model; (b) prediction model

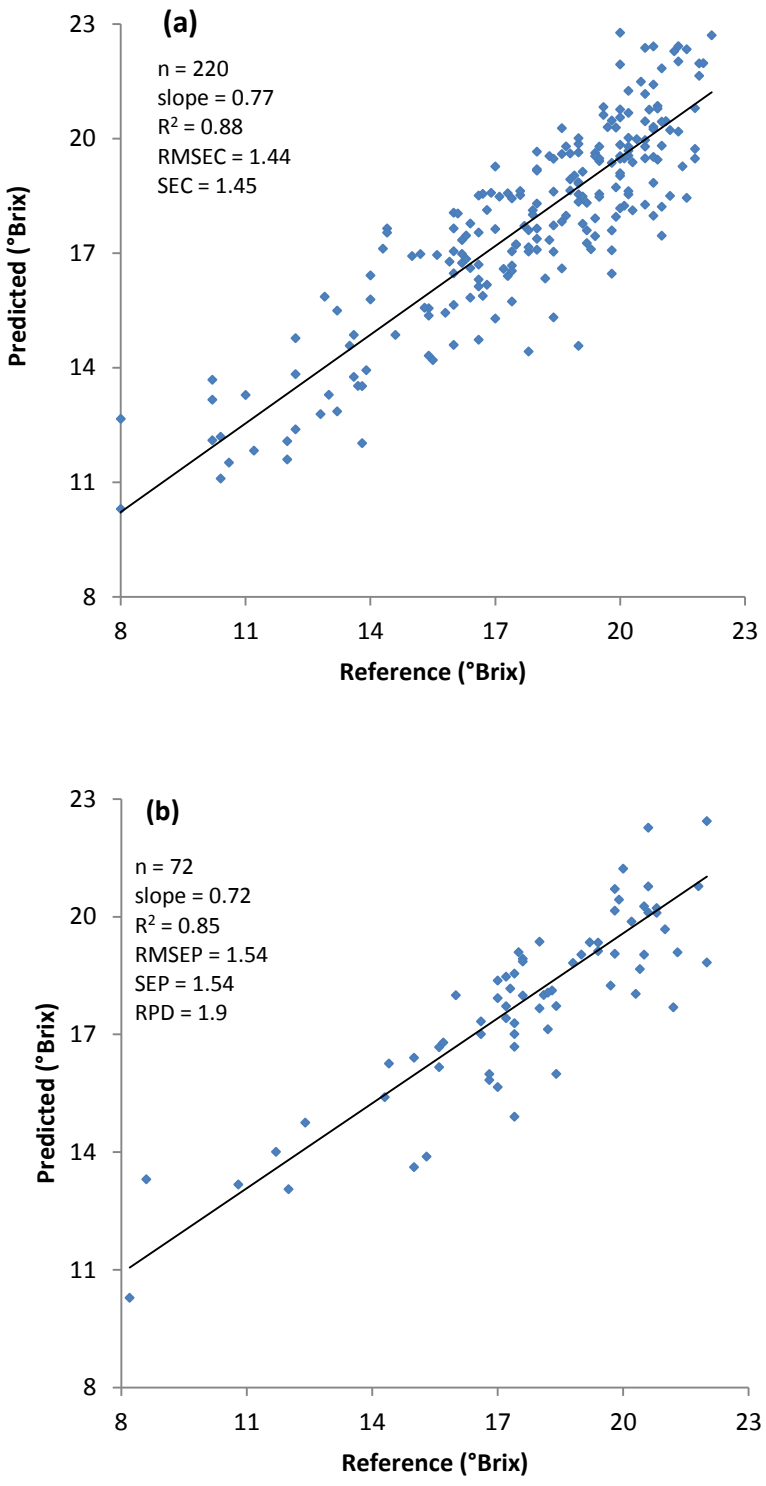


Figure 6.8: Scatter plots of reference versus predicted °Brix for reflectance spectral data by the CSSM: (a) calibration model; (b) prediction model

The R^2 values obtained from the internode samples was found to be better than those obtained for the whole stalk samples. For example, the R^2 for °Brix prediction using the VNIRS from the whole stalk samples was 0.83 compared to 0.87 for the internode samples. These results indicated that the internode approach had successfully improved the prediction capability of the instrument.

In comparison to similar measurement techniques on solid samples of other crops, the result of this study was also better than the study reported by Wu et al. (2008) who used the VNIRS to detect an infection on eggplant leaves. Their model showed prediction ability with 85% accuracy rate. Chia et al. (2012) also reported lower prediction value (R^2 = 0.68) when predicting sugar content from pineapple. However, the accuracy of this study was slightly lower than the results (R^2 = 0.91; SEP = 0.73 °Brix) obtained by Khuriyati & Matsuoka (2004) for non-destructively scanning tomato on the skin to predict sugar content. The accuracy of this study was also lower than the result (R^2 = 0.93; RMSEP = 0.26 °Brix) reported by Moghimi et al. (2010) for predicting the sugar content of kiwifruit and the work reported by Shao et al. (2007) on tomatoes (R^2 = 0.90; RMSEP = 0.38 °Brix). The relatively low accuracy obtained from this study compared to some other crops is mainly due to significant variation of sugar content along a sugarcane stalk from bottom to the top while the sugar content of other crops is relatively uniform within each individual fruit. Sugarcane solids are also regarded as the most difficult agricultural materials to be analysed by spectroscopic methods (Mehrotra & Siesler 2003).

6.3.5 °Brix prediction by the CSSM for an individual internode sample

Manley et al. (2007) suggested that more spectral variation for each reference °Brix resulted in more robust calibration models with better prediction accuracy. Thus, in order to investigate the prediction accuracy from large sample numbers with high spectral variation, every individual spectrum from each cut section was used to develop PLS models in this section. Specifically, the purpose of this section was to compare the prediction accuracy between the PLS model developed from individual spectrum of each cut section and the PLS model developed from average spectra data. The spectrum from each cut section was correlated to the °Brix values of the individual internode. Overall, 1168 spectra data, collected from 292 internode samples with four cut sections (292 x 4 = 1168) were used for both the calibration and validation models. However, 5% of the data in the data set has been identified as outliers by PCA. After the removal of this outlier, only 1109 useful spectral data were used for further analysis.

The performances of the PLS models developed from each spectrum are presented in the scatter plots shown in Figure 6.9 (a) and (b), respectively. For the calibration model, the values of R^2 and RMSEC were 0.73 and 1.99 °Brix, respectively. While the values of R^2 and RMSEP for prediction model were 0.76 and 0.88 °Brix, respectively. The quality of both models obtained from this technique was found to be reasonably good. However, the accuracy of these results was lower than the results obtained from the average spectral data. Thus, for future work, instead of developing PLS models using individual spectrum of each cut section, the average spectra representing four cut

sections of each internode sample should be used. The average spectra value would reduce the level of unnecessary spectral variation within internode samples, thus improving the performance of the prediction model. Alternatively, the prediction accuracy could be improved if each individual spectrum data was correlated to the °Brix values of each cut section.

6.3.6 °Brix prediction by the CSSM along the length of the internode

Figure 6.10 shows typical mean absorbance spectra for each cut section. The curve of the C1 (representing the node portion) showed an obvious difference from the other cut sections (representing the internode sections). Thus, to understand this difference, it is essential to know the °Brix value of each cut section. This knowledge could explain the °Brix variation along the internode, especially between node and internode areas. However, the direct measurement of °Brix from each cut section using the refractometer was not attempted in this study, as this method would be a very time consuming exercise. Therefore, this section describes the prediction of °Brix values of each cut section using the corresponding spectral data from each cut section.

To predict the °Brix value of each cut section, the spectral data belonging to each cut section (C1 to C4) were grouped separately. Each group was assigned to be a new prediction set. Then, the °Brix values of each cut section were predicted from the spectral data of the new prediction set using the calibration model developed from the average spectral data. The predicted °Brix values from each group (cut section) were averaged and are shown in Table 6.9. This table shows that C1 has the highest predicted °Brix (18.6), followed by C4 (17.9), C3 (17.4) and C2 (17.2). The °Brix variation between node (C1) and internode (C2-C4) areas was around 8.7%. The node portion (C1) has high °Brix values possibly because the sugars are concentrated initially at the nodes and later midway between the nodes, where they quickly reaches the highest concentration found in the internode (Fernandes & Benda 1985).

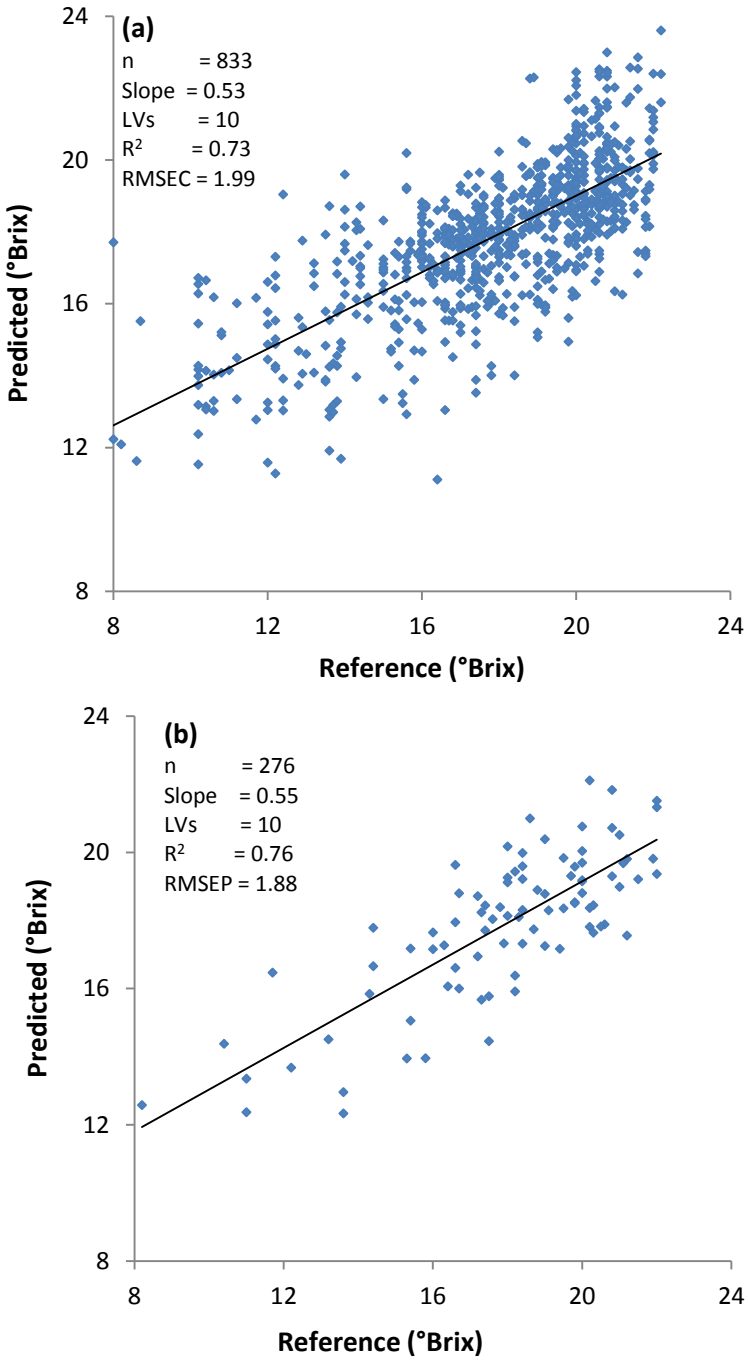


Figure 6.9: Scatter plots of reference versus predicted °Brix for individual spectrum data by the CSSM: (a) calibration model; (b) prediction model

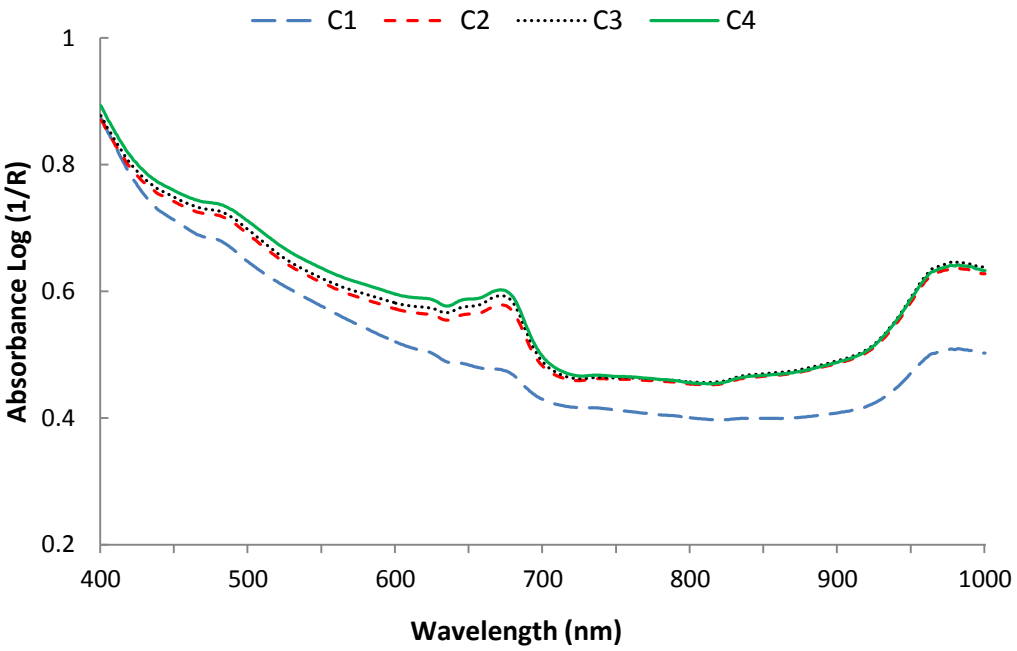


Figure 6.10: Typical absorbance spectral curves of CSSM for each cut section

Table 6.9: Predicted °Brix for each cut section using calibration model from average spectral data (R^2 = 0.87/ RMSEC = 1.49 °Brix)

Cut section	Predicted °Brix	R^2	RMSEP (°Brix)
C1 (node)	18.6	0.58	2.53
C2 (internode)	17.2	0.82	2.00
C3 (internode)	17.4	0.90	2.06
C4 (node)	17.9	0.79	2.07

However, in terms of prediction accuracy, C1 had the lowest prediction accuracy, followed by C4, C2 and C3 with R^2 values of 0.58, 0.79, 0.82 and 0.90, respectively. These results suggested that the node portions had higher spectral variation than internodes. A low absorption level of C1 may suggest that the node portions, which are harder than the internodes, have a low moisture level along with high fibrous materials. Thus, for better prediction accuracy, CSSM should be performed on the internode areas. The application of this measurement method on stalk samples is suitable for screening purposes in breeding programs. However, without a suitable sampling mechanism, this method is difficult to be applied on a harvester.

6.3.7 Spectroscopic performance of different varieties

Peirs et al. (2001) used the spectroscopic methods to predict sugar content on apples with different varieties. They found that the prediction accuracy varied for every apple variety. For sugarcane crops, the sugar content depends on a number of factors, including maturity, variety, soil, and environmental condition. To investigate this issue

in sugarcane crops, the influence of different sugarcane varieties with different maturity stages on prediction performance of PLS models was evaluated. Three main sugarcane varieties used in this study: early-maturing (Q155), middle maturing (Q208) and late maturing (Q190). The selection of these three varieties was made in order to create more robust models for sugar content prediction. To achieve that objective, each internode sample was grouped according to its variety. Then, the PLS regression analysis was carried out to develop both calibration and prediction models for each variety. Roughly, 3/4 of the internode samples were used for the calibration model and 1/4 was used for the prediction model.

The number of samples used to create the model, the calibration and validation correlation, the number of LVs factors, and the RMSEC and RMSEP for each variety are given in Table 6.10. The table shows that the calibration models for all varieties have the same R^2 values of 0.94, however, the prediction accuracy varied for each variety. It was found that the early maturing varieties (Q155) had the highest prediction accuracy with R^2 of 0.90, followed by the middle (Q208) and the late maturing (Q190) varieties with the R^2 values of 0.89 and 0.82, respectively. Thus, the accuracy (R^2=0.76) reported in Section 6.3.5, obtained by combining these varieties, was found to be a little lower, but this may be acceptable as the robustness of the model increased. The downward trend of the accuracy was also consistent with the downward trends of the maturity level of each variety.

Table 6.10: PLS models performance for selected sugarcane varieties

Variety	LVs	Calibration			Prediction		
		n	R^2	RMSEC	n	R^2	RMSEP
Q155	9	61	0.94	1.15	20	0.90	1.18
Q208	10	79	0.94	0.99	26	0.89	1.47
Q190	10	80	0.94	0.88	26	0.82	1.52

The influence of sugarcane varieties on absorbance characteristics is shown in Figure 6.11. The figure shows that the spectral pattern was the same for all varieties with obvious peaks at 680 and 980 nm. However, Q190 had less absorbance than the others. This study has shown that prediction accuracy is dependent on sugarcane variety and the maturity stage. Future studies should test whether the model can be used for new data, using an external validation set of samples which was chosen randomly. For a real application, a robust calibration model should be developed from different varieties, maturity levels, field locations and harvest seasons.

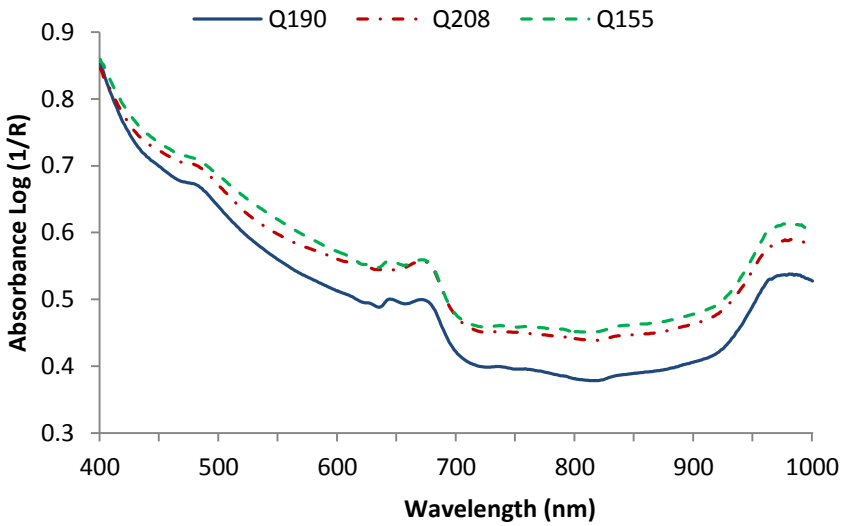


Figure 6.11: Typical absorbance spectra of CSSM for different varieties

6.3.8 The application of the CSSM to predict ternary growth quality components

Recently, Staunton et al. (2011) developed a simple method to estimate the levels of °Brix, pol, moisture content, fibre content and CCS present in sugarcane based on sugarcane biomass ternary relationships (Figure 6.12). This method assumes that the composition of all healthy sugarcane is constrained to follow the same ternary growth curve and that the determination of a sample's position on the growth curve can be used to estimate the levels of all ternary parameters (°Brix, fibre content and moisture content) and non-ternary parameters (pol in juice and CCS). This study has opened up a new dimension for in-field quality monitoring systems. To apply this finding effectively, a study is needed to identify which one of the ternary parameters would be the easiest to measure on a harvester. For calibration purposes, the relative ease of reference value determination would also need to be considered. Once the optimum quality parameter has been identified, the next research step is to focus on the development of the instrument to measure the targeted quality component.

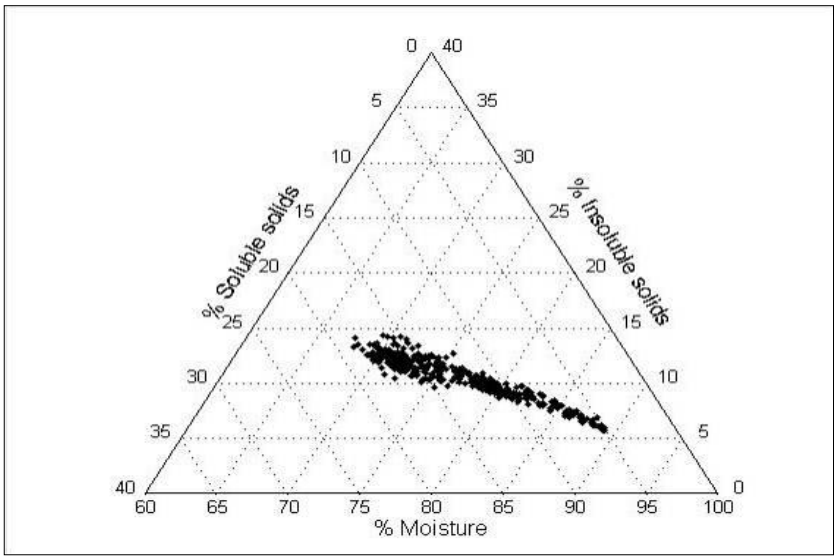


Figure 6.12: Ternary Growth Model (Staunton et al. 2011)

The ability of spectroscopic methods, coupled with a PLS regression analysis to simultaneously determine several quality components by a single scanning has given this equipment capability to measure sugarcane ternary quality components. Therefore, the aim of this section was to explore the potential of using a portable VNIRS to non-destructively predict all ternary quality components from a single scanning of sugarcane internodes. This study used the internode samples for the measurement in order to simulate the billet scanning in a real harvesting environment. The specific objectives of this section were:

1. To investigate the ability of both spectroradiometers to predict all ternary parameters from internode samples using CSSM.
2. To identify the best ternary quality parameter which could be predicted by both spectroradiometers.

6.3.8.1 Statistical characteristics of the internode samples

The summary of statistical characteristics for calibration and prediction data sets of internode samples is shown in Table 6.11. Generally, all ternary parameters in both data sets showed similar statistical characteristics. A relatively wide range of °Brix values (11 to 22.2) was found due to the inclusion of three different varieties with different maturity stages. The range of moisture content (66 to 84.2%) was in the comparable range with the data (62.3 to 89.4%) reported by Staunton et al. (2011). The mean of the fibre content found in this study of 5.4% was lower than the fibre content of 10.6% as reported by Staunton et al. (2011). The difference was due to the fact that this study used the fibre content from individual internode samples while Staunton et al. (2011) used the fibre content measured from prepared sugarcane samples.

Table 6.11: Summary of statistical characteristics of internode samples for predicting ternary growth quality components

Model	Sample no	Component	Max	Mean	Min	SD
Calibration	95	°Brix	22.2	18.4	12.2	2.39
		Fibre (%)	11.5	5.4	0.4	2.51
		Moisture (%)	84.2	75.8	66.0	3.65
Prediction	30	°Brix	21.9	18.2	11	2.60
		Fibre (%)	8.3	5.4	0.2	2.47
		Moisture (%)	83.3	76.0	68.8	3.93

6.3.8.2 Spectral overview of the CSSM for predicting ternary growth quality components

Typical absorbance curves of the spectral data collected from 125 internode samples by CSSM using the VNIRS before and after being treated with the MSC pre-processing method are shown in Figures 6.13 and 6.14, respectively. The MSC method was found to have minimised the baseline shift problem occurring in the original spectra data. In Figure 6.14, the absorptions peak found in the original spectra data around 620, 670 and 960 nm have been maintained. The peak of 680 nm was due to chlorophyll (Abbott et al. 1997). The absorption peak around 958 nm could be related to water in the stalks (Williams & Norris 1987).

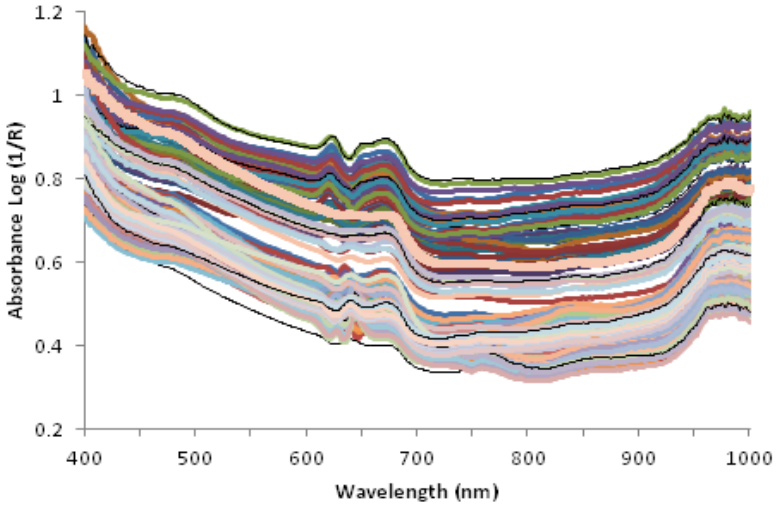


Figure 6.13: Typical raw absorbance spectra for CSSM collected by the VNIRS

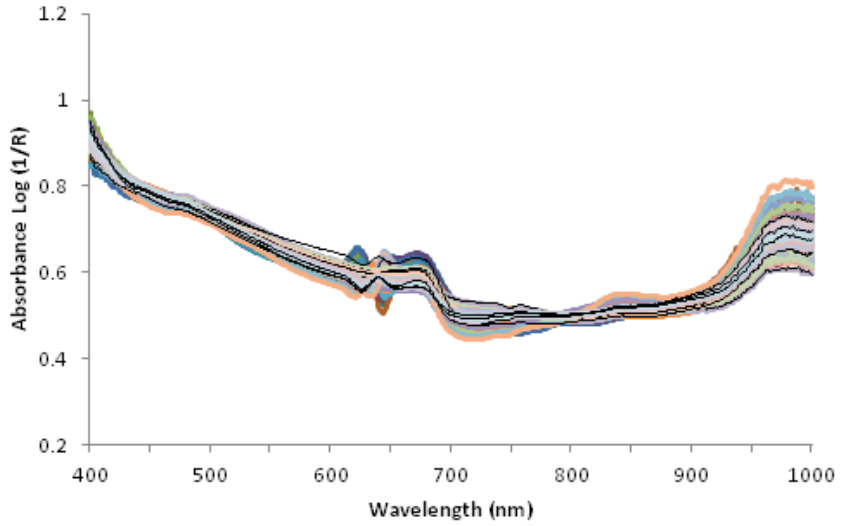


Figure 6.14: Typical absorbance spectra collected by the VNIRS after being treated with MSC method

Typical absorbance spectra of CSSM as measured by the FRS are shown in Figure 6.15. The unwanted baseline shift problems found in the raw spectral data has been substantially reduced by the MSC pre-processing method as shown in Figure 6.17. The spectral curves of both figures exhibit several obvious absorption peaks around 680, 958, 1170, 1400 and 1900 nm. The absorption bands around 680 and 958 nm were also detected by the VNIRS. The absorption peaks around 1400 and 1900 nm are related to water content in the stalks (Williams & Norris 1987). In both figures, it can be seen that the spectral regions between 1800 and 2500 nm exhibit spectral noise. These spectral regions which are known as water absorption features (Kumar et al. 2003; Mutanga et al. 2004) were then excluded from the analysis and removed from the spectra. Therefore, only the wavelength regions from 400 to 1800 nm were used for the analysis in this section.

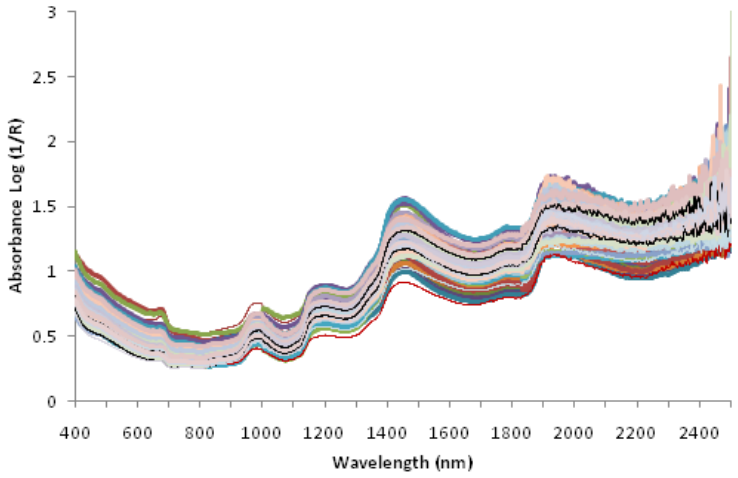


Figure 6.15: Typical raw absorbance spectra collected by the FRS

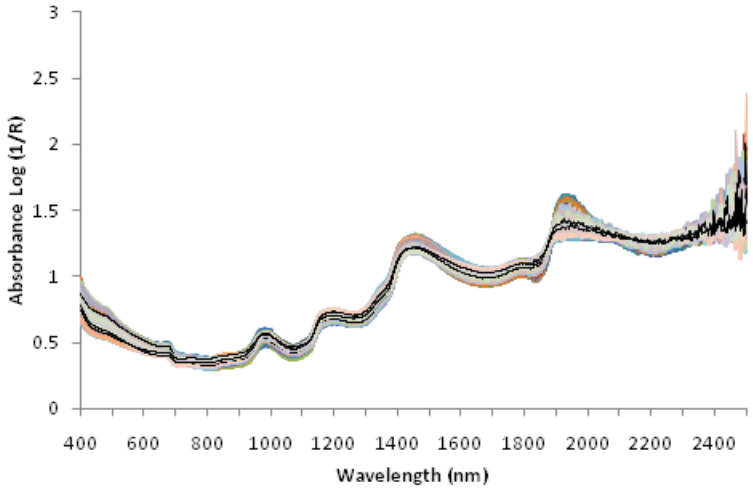


Figure 6.16: Typical absorbance spectra collected by the FRS after being treated with MSC method

6.3.8.3 Selection of the best parameter for spectroscopic prediction

The performance of PLS models in predicting ternary quality components using the VNIRS and FRS based on CSSM are presented in Table 6.12. The calibration models for all quality components were excellent with the R^2 values exceeding 0.90, except for the calibration model of the °Brix component obtained from the VNIRS with R^2 value of 0.87. For prediction performance, °Brix was found to be the best quality component that could be predicted by both spectroradiometers with R^2 values of 0.89. Moisture prediction was the second highest after °Brix with R^2 values for the FRS and the VNIRS were 0.69 and 0.89 respectively. Even though the VNIRS could predict both °Brix and moisture with the same R^2 values, the °Brix should be the first choice as it has a lower RMSEP value (1.14 °Brix) than moisture prediction (1.87%). In terms of the measurement of reference values for calibration purposes, °Brix measurement is also easier to perform than moisture measurement. The prediction accuracies for fibre prediction from both spectroradiometers were lower than °Brix and moisture.

Table 6.12: Performance of PLS models in predicting TGM parameters using the CSSM

Component	FRS					VNIRS				
	LVs	Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
°Brix	10	0.93	0.89	0.89	1.18	10	0.87	1.15	0.89	1.14
Fibre	8	0.90	1.25	0.69	1.99	10	0.93	0.92	0.83	1.63
Moisture	10	0.91	1.77	0.76	2.91	10	0.94	1.26	0.89	1.87

Overall, the prediction performance for all ternary quality components was reasonably good with R^2 values being above 0.80 except for fibre and moisture prediction by the FRS. These results indicated that both spectroradiometers could be successfully used to predict ternary quality components from stalk samples using CSSM. More importantly,

this study has shown that the prediction performances of the VNIRS were comparable with that of the FRS. This finding could justify the application of the portable and low-cost VNIRS in the field as a low-cost alternative to the sophisticated and expensive FRS.

6.4 Discussion and conclusions

This chapter has demonstrated that both spectroradiometers (VNIRS and FRS) coupled with PLS regression analysis could be applied to predict sugarcane quality components from stalk samples based on CSSM. For whole stalk samples, the prediction accuracy obtained by the VNIRS for pol, °Brix, CCS and fibre were 0.89, 0.83, 0.73 and 0.49, respectively. The accuracy obtained by the FRS for the corresponding components were 0.90, 0.86, 0.73 and 0.27, respectively.

For the internode samples, the prediction accuracy of the model developed from average absorbance spectral data of each individual internode sample was good, with the R^2 and RMSEP values being 0.87 and 1.45 °Brix, respectively. The RPD value of this model was 2, indicating that this method is suitable for quantitative prediction especially for screening purposes in breeding programs. The prediction accuracy of the model developed from each individual spectrum of each cut section showed slightly lower results with R^2 and RMSEP values being 0.76 and 0.88 °Brix. The internode sectional analysis has shown that the °Brix variation between the node and internode areas was up to 8.7%. This analysis found that the C1 (node portion) had the highest °Brix value (18.6 °Brix) compared to other sections. However, the prediction accuracy of this section was the lowest with R^2 of 0.58 °Brix. The variation of prediction accuracy within the internode samples could be up to 13%.

This chapter has also shown that the spectroscopic methods could be applied to determine ternary quality components. In this study, it has been found that the prediction performance of the VNIRS was comparable to that of the FRS, where both spectroradiometers have predicted the °Brix values with the same accuracy ($R^2 = 0.89$). This finding suggests that the low-cost and portable VNIRS could be a cheaper alternative to the more sophisticated and expensive FRS for field use. The VNIRS was also applied to predict moisture and fibre contents from the internode samples, resulting in good prediction accuracy with the R^2 values being 0.89 and 0.83, respectively.

Overall, the results presented in this chapter have shown that the VNIRS can be applied for the measurement of sugarcane quality components from stalk and internode samples in the field. The accuracy obtained in this study was acceptable given the heterogeneous nature of the stalk samples and the variation of °Brix values along the stalks. The internode sample technique has the potential to be practically and feasibly adopted for an on-the-go quality measurement system on a harvester, as it is very similar to billet samples available on a harvester.

Chapter 7

Skin scanning method

Chapter 6 discussed the performance of both spectroradiometers (VNIRS and FRS) in determining sugarcane quality parameters from sugarcane stalk samples based on the CSSM. This chapter will discuss the performance of both spectroradiometers to determine sugarcane quality parameters from sugarcane stalks based on the skin scanning method (SSM). The SSM method will be applied on whole stalk samples as well as internode samples. The variation of sugar content and prediction accuracy for both samples will be quantified. The optimum quality parameter for quality prediction in the field will also be identified. Finally, this chapter discusses the comparison of prediction performance between CSSM and SSM.

7.1 Introduction

Recently, many studies have applied NIR spectroscopic methods to non-destructively measure the sugar content by scanning the skin surfaces of the fruit samples. This method, known as a skin scanning method (SSM), has been applied to predict the internal quality of a number of fruits, including kiwifruits (Moghimi et al. 2010), apples (Lammertyn et al. 1998; Peirs et al. 2001; Lu 2004; Park et al. 2003), tomatoes (Khuryati & Matsouka 2004) and guavas (Hsieh & Lee 2005).

The SSM is based on the principle that when a light beam hits the crop surface, a small fraction is reflected at the surface as a specular reflectance and the rest will penetrate into fruit tissues. Photons are absorbed into the tissues or migrate in different directions where radiation will be scattered backward to the surface as diffuse reflectance, while the remaining radiation migrates forward into the tissues as absorbance (Qing et al. 2007). Light absorption is related to certain chemical constituents, such as sugar, acid and moisture content (Williams & Norris 2001). A reflectance spectrometer measures the aggregate amount of light reflected from a sample, from which light absorption may be estimated and then related to certain chemical constituents of the crop (Lu 2004). To the candidate's knowledge, no studies have ever been reported regarding the application of the SSM to determine quality parameters in the sugarcane industry.

As mentioned in Chapter 6, the spectroscopic measurement of the whole stalk sample is very useful for screening the sugarcane quality parameters in the field to suit the need of breeding and clonal evaluation programs. The SSM is very effective as it can be non-destructively applied on a standing stalk in the field. This method will allow researchers and growers to monitor the development/growth of sugarcane without having to cut it down. Non-destructive methods could enable the study of an accumulation of sugar content in the same stalks during the growing season and post-harvest storage, avoiding bias caused by stalk to stalk variation (Peirs et al. 2001).

Therefore, the aim of this chapter is to investigate the ability of the spectroradiometers to predict sugarcane quality parameters from sugarcane stalk samples based on the SSM. The specific objectives of this chapter were:

1. To investigate the potential of spectroradiometers coupled with PLS models to determine sugarcane quality components from stalk samples using the SSM.
2. To quantify the variation of sugar content and prediction accuracy along the stalks and individual internodes (between the node and internode) using the SSM.
3. To compare the prediction accuracy between the PLS models developed from the average spectral data and individual spectrum data.
4. To investigate the effect of different spectral pro-processing methods on model accuracy.
5. To compare the prediction performance between SSM and CSSM for the same samples.
6. To classify sugarcane quality using an artificial neural network (ANN) algorithm based on spectral data of SSM.

7.2 Quality prediction from the whole stalk samples by the SSM

7.2.1 Statistical characteristic of the whole stalk samples for the SSM

This section discusses the application of both VNIRS and FRS to predict sugar content from whole stalk samples using the SSM. A total of 36 sample sets (each sample set consists of six stalk samples) were collected from the BSES research station, Bundaberg in September 2012. From these 36 sample sets (36 x 6 = 216 stalks), 216 individual stalk samples were scanned on their skin surfaces. The detailed measurement procedures for this section were discussed in Section 4.3.1.2. The statistical characteristic of quality components for stalk samples are summarised in Table 7.1, which shows the mean, standard deviation and range for typical sugarcane quality components as measured using the standard mill procedures. Since this preliminary experiment was conducted on a single variety (Q232) in conjunction with a varietal trial, the mean values of °Brix, pol and CCS were found to be lower than that reported by Berding et al. (1991a), with mean values of 23.2, 92.1 and 15.6, respectively. Since the fibre value of sugarcane from the same variety was constant, only one fibre value was used in this section. From the 36 sample set, 27 samples were used for the calibration model and nine for the prediction model.

Table 7.1: Statistical analysis of the whole stalk samples used for the SSM

	°Brix	Pol	Fibre (%)	CCS
Max	23.4	94.74	15.60	15.68
Mean	22.8	91.79	15.60	15.16
Min	22.4	89.04	15.60	14.58
SD	0.2	1.30	0.00	0.25

7.2.2 Determination of sugarcane quality components by the SSM

The performance of the calibration and prediction models in predicting sugarcane quality components from the whole stalk samples using both spectroradiometers based on the SSM is shown in Table 7.2. The calibration models for the FRS were excellent for all quality components with the R^2 values of 0.99 while the RMSEPs ranged from 0.07 to 0.46 °Brix. For the prediction models, °Brix and CCS were found to have the highest R^2 value (0.73) while pol had the lowest (0.69). The R^2 value obtained in this section for °Brix prediction was slightly lower than the R^2 (0.86) obtained from the CSSM. However, the prediction of fibre from the whole stalk samples by the SSM yielded better accuracy ($R^2 = 0.75$) than the CSSM ($R^2 = 0.27$). Overall, the values of R^2 and RMSEP suggested that the FRS has the ability to predict sugarcane quality components from whole stalk samples using the SSM.

Table 7.2: Performance of PLS models in predicting sugarcane quality components based on the SSM

Component	FRS					VNIRS				
	LVs	Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
°Brix	3	0.99	0.07	0.73	0.39	2	0.91	0.27	0.85	0.40
CCS	6	0.99	0.09	0.73	0.43	10	0.89	0.29	0.64	0.62
Fibre	4	0.99	0.07	0.72	0.35	10	0.93	0.17	0.38	0.62
Pol	6	0.99	0.46	0.69	2.19	10	0.89	1.44	0.72	2.73

*** n for calibration model was 27; n for prediction model was 9.**

For the VNIRS, the performance of its calibration models varied with R^2 and RMSEPs values ranging from 0.89 to 0.93 and 0.17 to 1.44 °Brix, respectively. For the prediction models, °Brix was also found to have the highest R^2 value of 0.85, followed by pol, CCS and fibre content with the R^2 values of 0.72, 0.64 and 0.38, respectively. The prediction accuracy for the °Brix found in this chapter ($R^2 = 0.85$) was better, compared to the prediction accuracy obtained by the CSSM ($R^2 = 0.83$). Similar to the CSSM, the performance of the VNIRS in predicting fibre content was not satisfactory. This finding indicates that °Brix is the optimal quality component to be predicted by the spectroradiometers based on the SSM. However, for the CSSM, pol was found to give better prediction accuracy than °Brix. Thus, in the following section, only the °Brix values were used for discussion.

7.2.3 Variations of prediction accuracies for different stalk sections

To apply the SSM in the field, it is important to know which section of the stalk samples should be scanned for the optimum prediction accuracy, as the sugar content varies along the length of the stalk. To identify the optimum section, the stalks were divided into three sections: bottom, middle and top. The relative prediction accuracies of each section as measured by both VNIRS and FRS are shown in Table 7.3. The data in this table was developed by averaging the spectra collected from the same section of each stalk for every sample set. The averaged spectrum of each section for all sample sets was then divided into three groups namely bottom, middle and top. For each group,

75% of the data was used to develop a calibration model while the remaining data was used for a prediction model. The spectral data from all groups were calibrated against the °Brix value of each sample set.

Table 7.3: PLS models' performance in predicting °Brix values from different stalk sections using the SSM

Section	FRS					VNIRS				
	LVs	Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
Bottom	10	0.99	0.10	0.81	0.47	7	0.97	0.16	0.68	0.58
Middle	8	0.98	0.11	0.80	0.39	2	0.93	0.25	0.74	0.44
Top	2	0.98	0.12	0.93	0.28	2	0.91	0.28	0.85	0.51

Table 7.3 shows that the top sections measured by both spectroradiometers had the highest relative prediction accuracies as indicated by R^2 values. It was found that the FRS yielded higher relative prediction accuracy (R^2 = 0.93) than the VNIRS (R^2 = 0.85). The top section had the highest relative prediction accuracy probably because the skin surface on this section, which is far away from the ground level, was cleaner than the bottom section. Thus, the top section had less spectral variation compared to the other sections.

Figure 7.1 shows the mean absorbance spectra for the three sections of the sugarcane stalks measured by the VNIRS based on the SSM. The spectral curves of these three sections showed the same trend with three obvious absorption bands at 640, 670 and 960 nm. The bands around 640 and 670 nm could be related to chlorophyll content while the band at 960 nm could be related to sugar content (Abbott et al. 1997). An absorption peak around 960 nm could also be related to the third overtone of O-H of water (Williams & Norris 2001). With a larger wavelength range as measured by the FRS, three additional major absorption peaks were obvious at 1200, 1420 and 1920 nm (Figure 7.2). The absorption peaks around 1200, 1460 and 1920 nm were related to water content in the stalks (Williams & Norris 1987). From both figures, it can also be seen that the bottom section had the highest absorption level while the top section had the lowest. This absorbance variation may indicate the variation of sugar content along the stalk with the bottom section having relatively more sugar content.

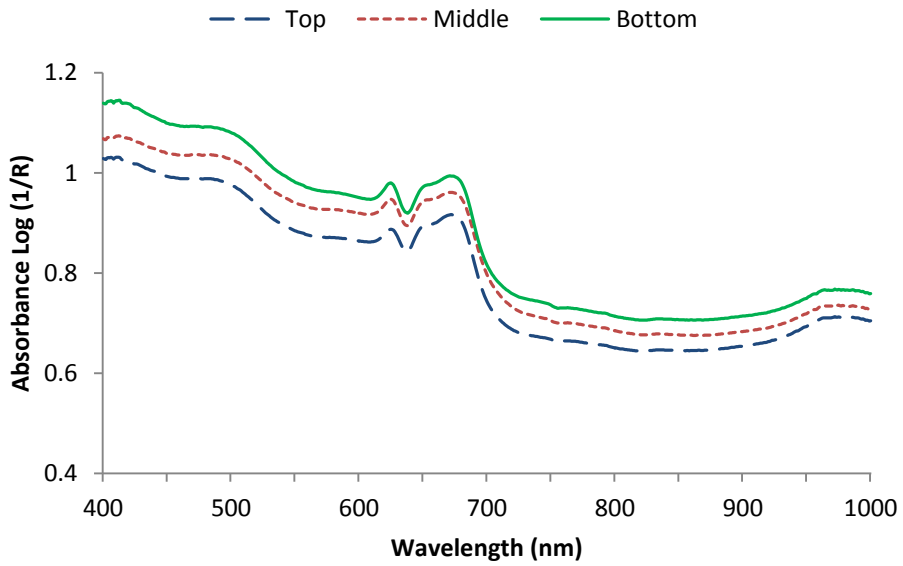


Figure 7.1: Typical absorbance spectra of SSM for different stalk sections as measured by the VNIRS

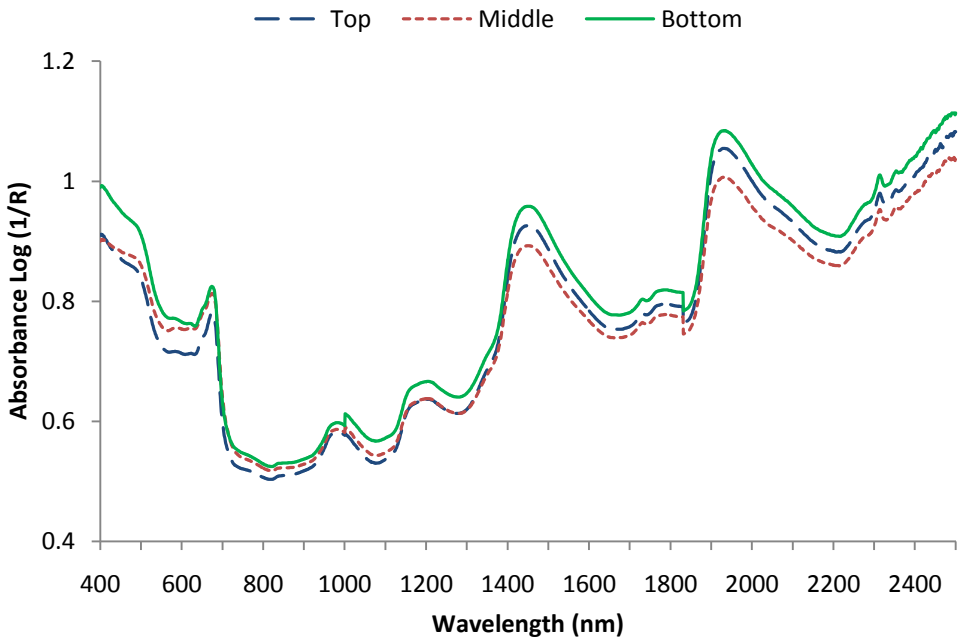


Figure 7.2: Typical absorbance spectra of SSM for different stalk sections as measured by the FRS

7.3 Effect of different pre-processing methods on the SSM

The effect of different pre-processing methods on prediction accuracy for the CSSM was discussed in Section 6.3.2. In this section, the influences of the same pre-processing methods on the prediction accuracy of the SSM are discussed. The results of this

investigation are displayed in Table 7.4. Basic statistical information for this data was presented in Table 6.6.

From the table, it can be seen that if no pre-processing method was applied, the prediction accuracy for the raw reflectance and absorbance spectral data were still good with R^2 values being 0.88 and 0.86, respectively. Consistent with the finding in Section 6.3.2, this section also found that the MSC method and its combination with another method yielded higher R^2 values than the raw spectral data. For the reflectance data, the MSC method improved the prediction accuracy (R^2) of raw spectral data from 0.88 to 0.91. The combination of MSC method with MA3 and MSC with SNV also gave the same improvement. For the absorbance data, the MSC method improved the R^2 of the raw data from 0.86 to 0.87. It is obvious that the reflectance data had higher prediction accuracy that the absorbance data. In contrast, the application of SG2 produced the lowest prediction accuracy for both spectral types. This observation is in good agreement with that reported by Montalvo et al. (1994). Therefore, this MSC method was adopted throughout this chapter.

Table 7.4: The effect of different pre-processing methods on the PLS models performance for the SSM

Pre-processing method	Reflectance				Absorbance			
	Calibration		Prediction		Calibration		Prediction	
	R^2	RMSEC	R^2	RMSEP	R^2	RMSEC	R^2	RMSEP
Raw	0.86	1.43	0.88	1.58	0.83	1.56	0.86	1.69
MSC	0.87	1.35	0.91	1.42	0.83	1.54	0.87	1.68
MA3	0.85	1.46	0.87	1.65	0.82	1.58	0.85	1.71
SNV	0.86	1.43	0.87	1.65	0.84	1.51	0.85	1.73
MN	0.85	1.44	0.87	1.61	0.84	1.53	0.85	1.75
SG(1)	0.91	1.18	0.83	1.85	0.89	1.25	0.84	1.77
SG(2)	0.79	1.69	0.48	2.98	0.74	1.88	0.43	3.06
MA3+MSC	0.86	1.38	0.91	1.43	0.82	1.57	0.86	1.70
SNV+MSC	0.87	1.35	0.91	1.41	0.83	1.54	0.87	1.68
MA3+MSC+SNV	0.86	1.43	0.89	1.53	0.83	1.54	0.85	1.76

*** n for calibration model=220; n for prediction model=72.**

7.4 The SSM for the internode samples

In this study, the same internode samples were used for both CSSM and SSM. For the CSSM, the samples were scanned on the cross sectional surfaces while for the SSM, the samples were scanned on their skin surfaces. Therefore, the statistical characteristics of the samples used for the SSM are the same as presented for the CSSM in Table 6.6. The example of raw reflectance spectra, *R* and absorbance spectra, *A* of three internode samples having high (22 °Brix), medium (18 °Brix) and low (14.2 °Brix) °Brix values as measured by the VNIRS are shown in Figure 7.3 (a) and 7.4 (a), respectively. In both figures, no obvious difference could be seen in the shape of the spectra for different °Brix values. However, due to light scattering in the skin surfaces, some consistent baseline shifts and bias appeared in the raw spectra. These baseline shift problems which existed in the raw reflectance and absorbance spectra were eliminated using MSC method as shown in Figures 7.3 (b) and 7.4 (b), respectively.

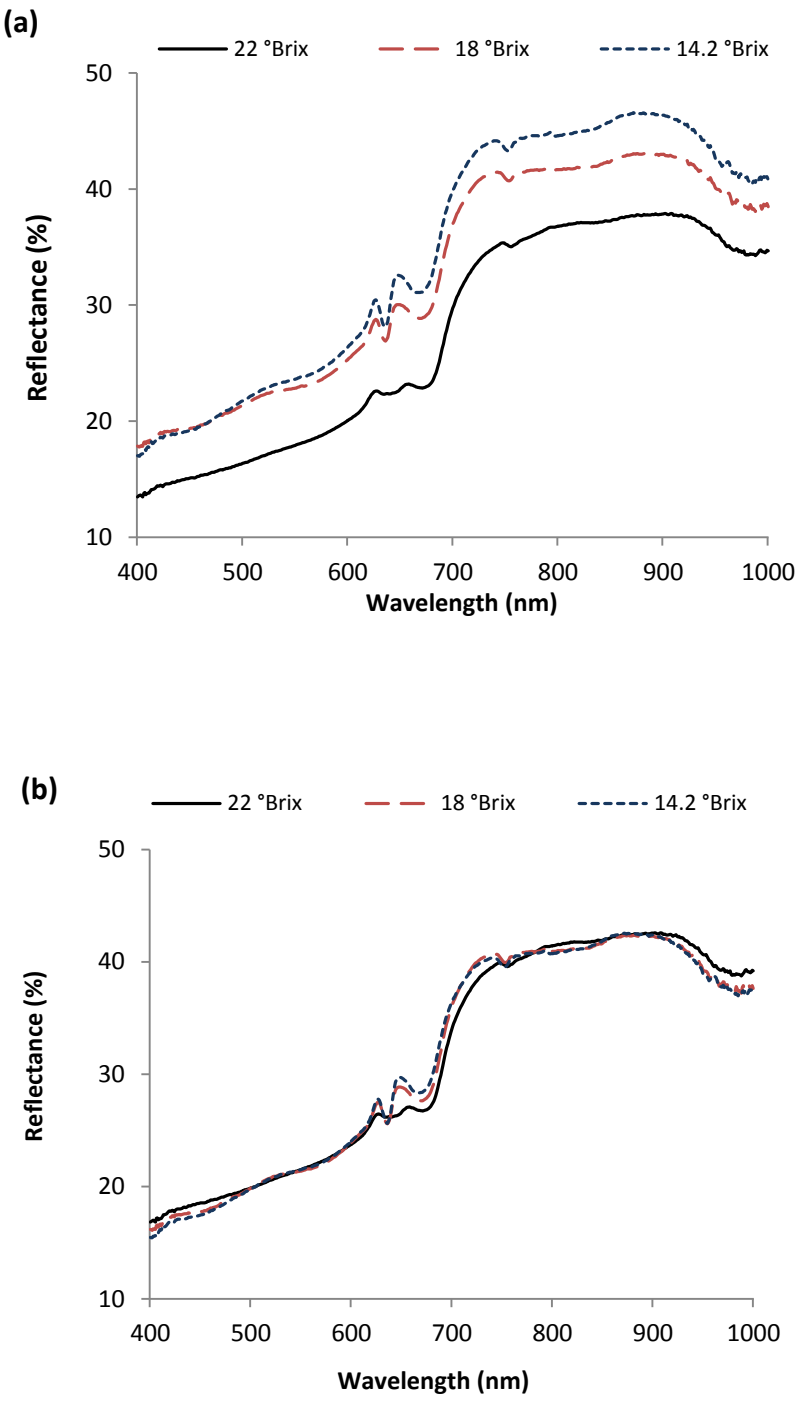


Figure 7.3: Typical reflectance spectra for sugarcane internodes at different °Brix values: (a) raw reflectance spectra; (b) reflectance spectra pre-processed with MSC

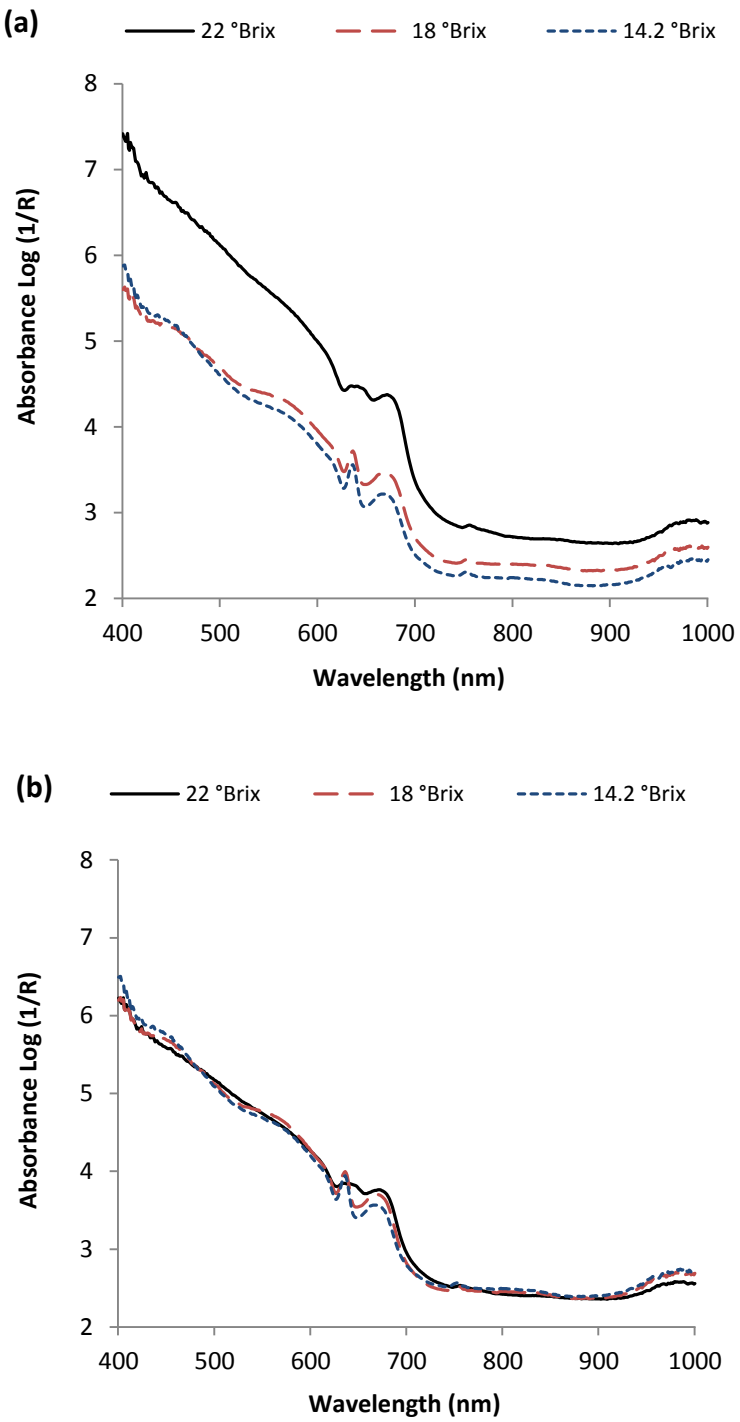


Figure 7.4: Typical absorbance spectra for sugarcane internodes at different °Brix values: (a) raw absorbance spectra; (b) absorbance spectra pre-processed with MSC

7.4.1 °Brix prediction by the SSM from average spectral data

PLS models were developed for both reflectance and absorbance spectral data which were pre-processed using the MSC method. The four spectral data of each internode section (S1 to S4) were averaged into one internode spectrum (as detailed in Section 6.3.4). The performance of the final PLS models developed from the average internode spectrum was evaluated by the RMSEC, RMSEP and R^2 (for both calibration and prediction models) and RPD value. The values of these indices for PLS models developed using the reflectance and absorbance spectra are shown in the scatter plots in Figures 7.5 and 7.6, respectively. In both figures, the ordinate and abscissa represented the predicted and measured values of the °Brix. The R^2 and RMSEP values for the reflectance spectra were 0.91 and 1.42 °Brix, while for the absorbance spectra were 0.87 and 1.86 °Brix, respectively.

The prediction performance of the PLS models for both spectral data showed good agreement between the reference and predicted values. Shenk & Westerhaus (1996) suggested that R^2 value greater than 0.9 indicated excellent quantitative information of the models. The RPD values for the prediction models developed for the reflectance and absorbance spectral data were 2.0 and 1.7, respectively. These RPD values indicates that the SSM method based on the reflectance data has the potential to be applied for coarse quantitative prediction in the field.

From both figures (7.5 and 7.6), the PLS model developed with the reflectance spectra was found to perform better than the model developed with the absorbance spectra as indicated by the R^2 and RMSEP. This was reasonable since the absorbance spectra was related to the presence of chemical components such as sugars, while the reflectance spectra of fruit contained information on both the absorption as well as scattering properties of the tissue (Nicolaï et al. 2008). Since this study was performed on skin surfaces, scattering properties of the tissue would also have contributed to the model performance. Reflectance measurements which were frequently converted to log (1/R), ignoring the fact that light penetration in biological tissue was much more complicated and also involved scattering (Davies & Grant 1987).

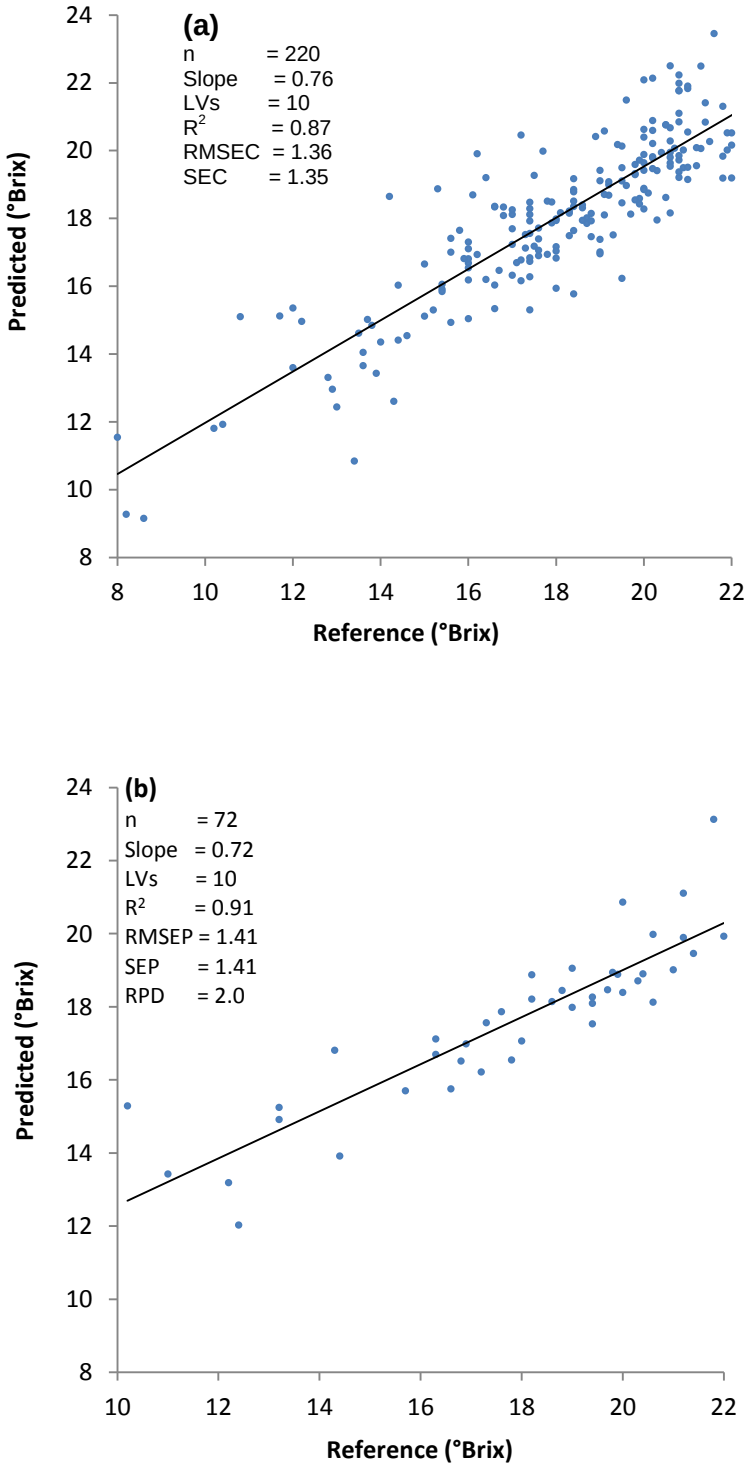
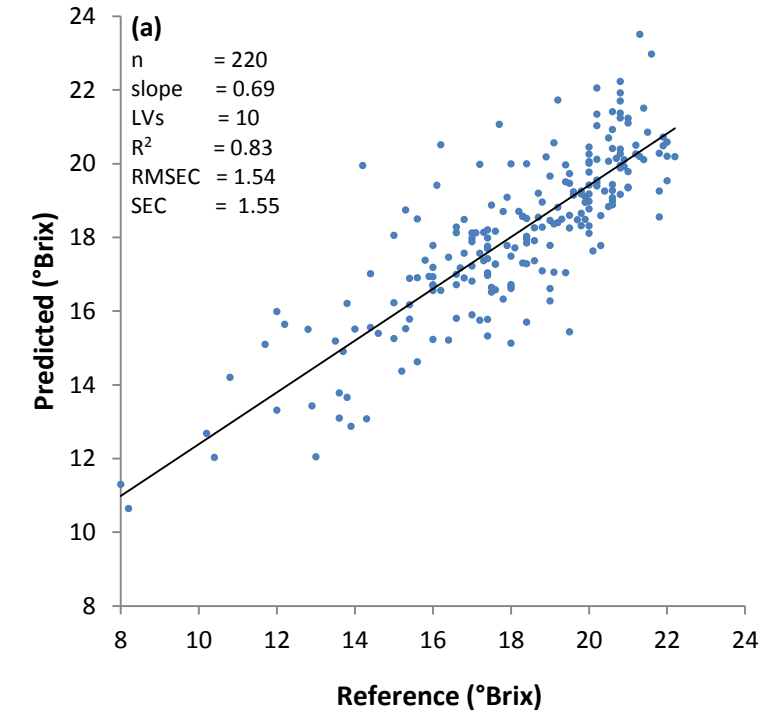


Figure 7.5: Scatter plots of reference versus predicted °Brix for reflectance spectral data by the SSM: (a) calibration model; (b) prediction model



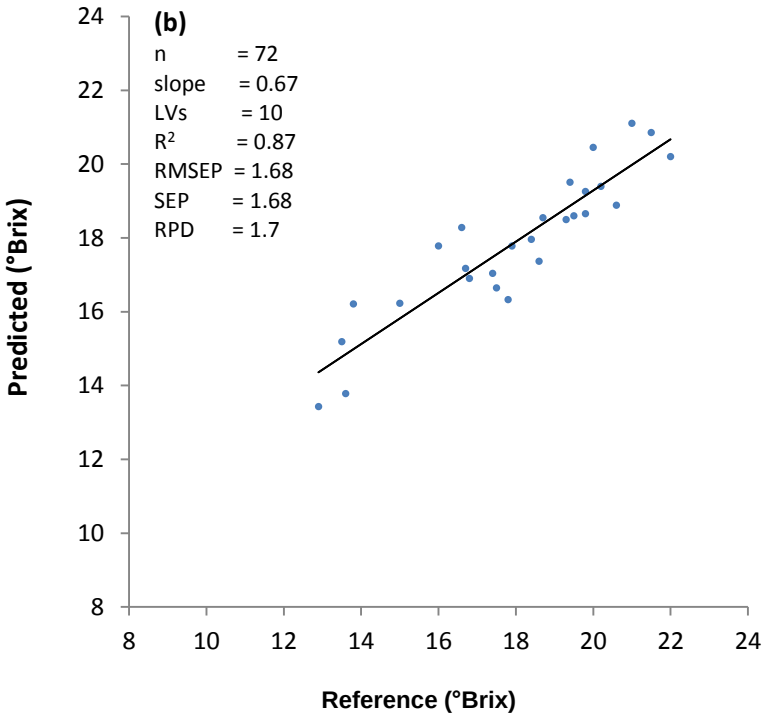


Figure 7.6: Scatter plots of reference versus predicted °Brix for absorbance spectral data by the SSM: (a) calibration model; (b) prediction model

In comparison to the CSSM, the SSM was found to have higher prediction accuracy for both the reflectance and absorbance spectral data. For example, for the reflectance data, the R^2 obtained from the CSSM was 0.85 while for the SSM was 0.91. For the absorbance data, both SSM and CSSM had the same R^2 values of 0.87 while the RMSEP were 1.68 and 1.45 °Brix, respectively. This finding suggests that both SSM and CSSM could give equal accuracy when applied to stalk samples.

In comparison to similar measurement techniques on other fruits, the R^2 value of sugar content (°Brix) prediction reported in this chapter from reflectance spectra was slightly better than those obtained by Peirs et al. (2001) with R^2 values ranged between 0.73 and 0.89 using different apple varieties. The results reported in this chapter are also better than the work reported by Lammertyn et al. (1998) who used apples and obtained R^2 of 0.82 and SEP of 0.6 °Brix. It was also better than that reported by Schaare & Fraser (2000) who obtained R^2 of 0.86 for predicting soluble solids content (SSC) on kiwifruits. This result was the same as obtained by Khuriyati & Matsuoka (2004) for non-destructively scanning tomatoes (R^2 = 0.91) on the skin to predict the sugar content. However, the accuracy level reported in this chapter was lower than the result (R^2 = 0.93) reported by Moghimi et al. (2010) for predicting the sugar content of kiwifruits.

Overall, it has been demonstrated that the VNIRS which operated in the wavelength range of 400 to 1000 nm (Si detector) could be used to predict °Brix values from sugarcane stalk samples based on the SSM. This proposed method has the potential for quality measurement use in the field. With more research, this method has the potential to be applied for predicting sugar content of billet samples on the elevator of a sugarcane harvester during harvest.

7.4.2 °Brix prediction by the SSM for each individual spectrum data

The purpose of this section was to compare the prediction accuracy of the SSM between the PLS model developed from individual absorbance spectrum of each cut section and the PLS model developed from the average spectral data (The corresponding procedures for the CSSM were discussed in Section 6.3.5). The spectrum from each cut section was correlated to the average °Brix obtained for the individual internode. Roughly, 1168 spectral data, collected from 292 internode samples with four cut sections (292 x 4 = 1168) were used for both calibration and validation models. However, around 5% of the data set (59 data) was identified as outliers by PCA. Outliers were detected by PCA from the influence plot which displays the sample residual x-variances against leverages. Thus, after the outliers were removed, only 1114 useful spectral data were used for further analysis in this section.

The performances of the PLS models developed from each spectrum are presented in the scatter plots as shown in Figure 7.7 (a) and (b), respectively. For the calibration model, the values of R^2 and RMSEC were 0.70 and 2.11 °Brix, respectively. On the other hand, the values of R^2 and RMSEP for the prediction model were 0.74 and 1.88 °Brix, respectively. As expected, these results were lower that the results obtained from the average spectral data where R^2 and RMSEP were 0.91 and 1.42 °Brix, respectively

(See section 7.4.1). Comparing this result with the same method from the CSSM, it was found that the accuracy for the prediction model in this study (R^2 = 0.74) is lower than that of CSSM (R^2 = 0.76). This may indicate that the spectra collected from the skin surfaces contained more variation than the spectra collected from cross sectional surfaces.

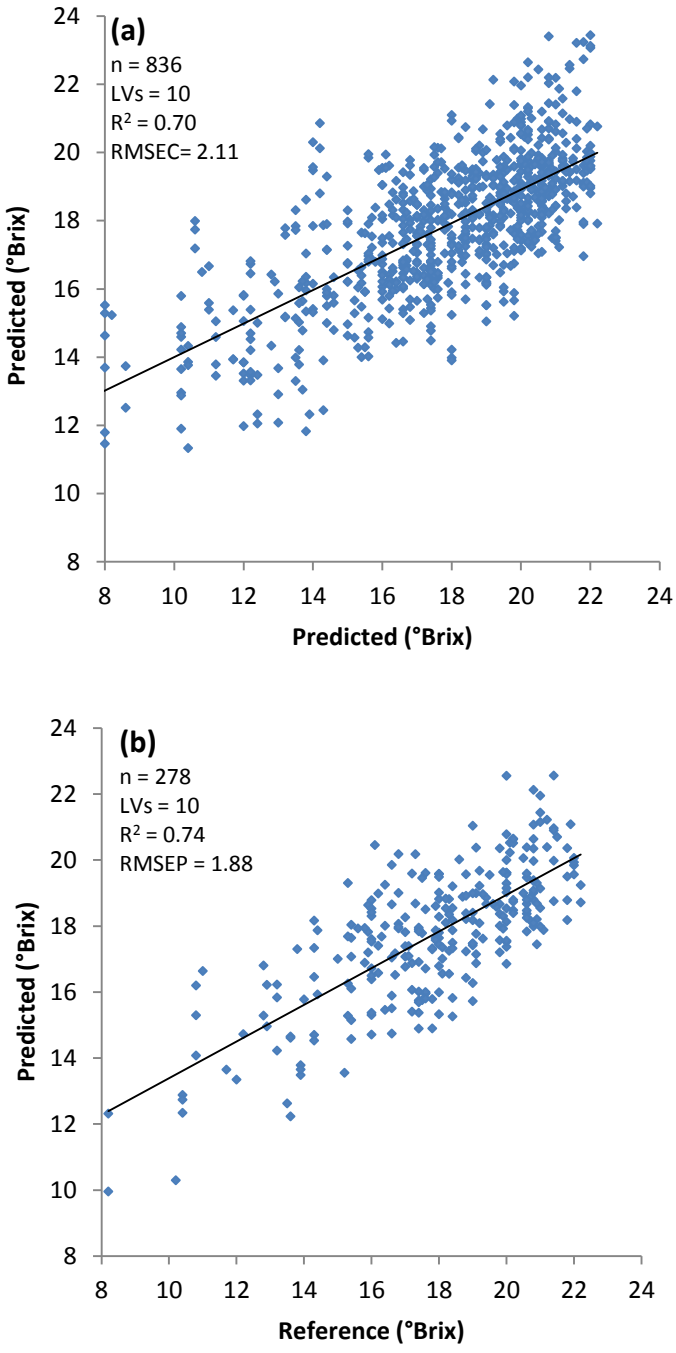


Figure 7.7: Scatter plots of reference versus predicted °Brix for individual spectrum data by the SSM: (a) calibration model; (b) prediction model

7.4.3 °Brix prediction by the SSM along the internode samples

Figure 7.8 shows typical absorbance values for each internode cut section (S1 to S4). The S4 curve which represents the node portion shows obvious difference from other cut sections which represents internode sections. It can be seen that S4 has the highest absorption level than the other cut sections. S1, which also represented the node area at another end, also displayed different absorption levels compared to S2 and S3. S2 and S3 which represents internode sections had similar absorption levels. Different absorption levels as shown by the different cut sections should indicate that the sugar content varied along the individual internode samples. In addition, increasing age of internode will be associated with increasing lignifications of the stalk and different sugar composition and content.

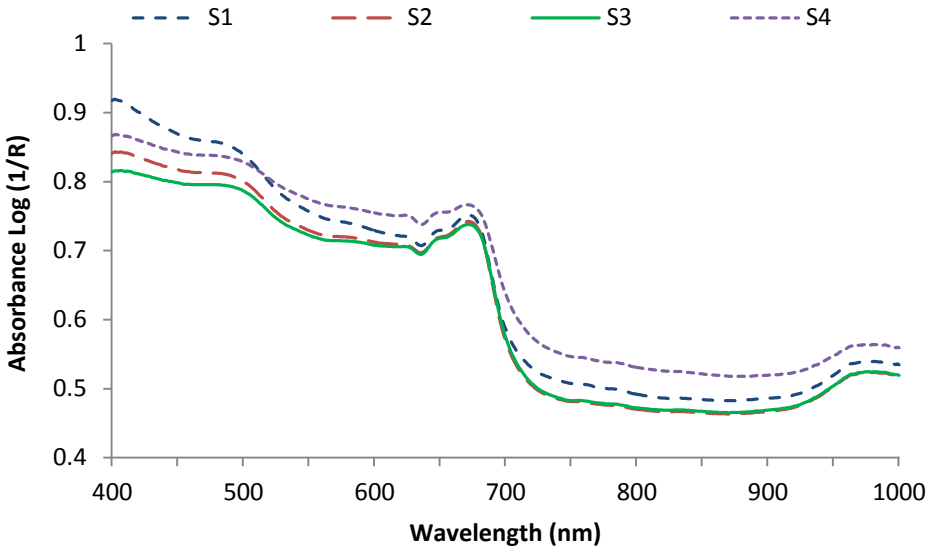


Figure 7.8: Typical absorbance curves of SSM for each internode scanning section

In order to predict the °Brix value of each cut section, the spectral data belonging to each cut section (S1 to S4) were grouped separately. Each group was assigned to be a new prediction set. Then, the °Brix value for each group was predicted from its spectral data using the calibration model developed from the average spectral data. The predicted °Brix values from each group (cut section) were averaged and shown in Table 7.5. The table shows that S4 had the highest predicted °Brix (17.9), followed by S2 (17.4), S3 (17.4) and S2 (17.2). The °Brix variation between node (S1) and internode (S2) areas was around 2.9%. The S4 (node portion) had the highest predicted °Brix values possibly because the sugars were initially concentrated at the nodes and later quickly reached the highest concentration in the internode (Fernandes & Benda 1985).

Table 7.5: Predicted °Brix for each cut section using calibration model from average spectral data (R^2 = 0.83/ RMSEC = 1.54 °Brix)

Cut Section	Predicted °Brix	R^2	RMSEP (°Brix)
S1 (node)	17.2	0.72	2.53
S2 (internode)	17.7	0.75	2.13
S3 (internode)	17.4	0.74	2.00
S4 (node)	17.9	0.67	2.33

However, in terms of prediction accuracy, S4 had the lowest prediction accuracy, followed by S1, S3 and S2 with R^2 values of 0.67, 0.72, 0.74 and 0.75, respectively. These results suggest that node portions have higher spectral variation than internodes. The higher spectral variation in the node portion was probably due to the presence of moisture-resistant wax layer, which was relatively thicker in the node portion compared to internode portion (Kroes 1997).

7.5 The application of the SSM to predict sugarcane ternary quality parameters

Section 6.3.8 discussed the application of both VNIRS and FRS to predict sugarcane ternary quality parameters (°Brix, moisture and fibre content) based on the CSSM. In this section, the potential application of the same equipment to predict sugarcane ternary quality parameters based on the SSM were evaluated. Since the internode samples used in this section were the same as used for the CSSM, a summary of the statistical characteristics of the internode samples can be found in Table 6.11. This section was also designed to identify the most effective equipment between the VNIRS and FRS for measuring sugarcane quality components. Therefore, the direct determination of °Brix, moisture content and fibre content by the SSM using both spectroradiometers were investigated.

7.5.1 Application of the FRS to predict ternary quality parameters by the SSM

Table 7.6 shows the performance of both calibration and prediction models for the FRS in predicting °Brix values based on the SSM from the internode samples. It was shown in Section 5.3.8.2 that the wavelength region of the FRS was subject to spectral noise especially beyond 1800 nm. Therefore, to explore the effects of different wavelength regions on the calibration and prediction models, the wavelength range of the FRS (400 - 2500 nm) was divided into three regions: 400 to 1000 nm, 400 to 1800 nm and 400 to 2500 nm. The comparison shown in Table 7.6 was based on °Brix values.

Table 7.6: PLS model performances for different wavelength ranges for the SSM

Wavelength (nm)	LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP
400 – 1000	10	0.84	1.30	0.79	1.58
400 – 1800	10	0.84	1.31	0.82	1.50
400 – 2500	9	0.86	1.21	0.78	1.63

Table 7.6 shows that the correlations of the calibration models varied from 0.84 to 0.86 while the RMSEC varied from 1.21 to 1.31 °Brix. The prediction results were slightly

lower from those for the calibration models, with the R^2 ranging between 0.78 and 0.82 and RMSEP between 1.50 and 1.63 °Brix. Similar to the CSSM, the wavelength region from 400 to 1800 nm had the highest prediction accuracy with R^2 and RMSEP values of 0.82 and 1.50 °Brix, respectively. The results of the calibration and prediction models for the wavelength region between 400 and 1800 nm, as measured by the SSM are presented in the scatter plots in Figure 7.9 (a) and (b), respectively. In both figures, the ordinate and abscissa represented the predicted and measured values of the °Brix. In contrast, the wavelength region from 400 to 2500 nm had the lowest R^2 value of 0.78 with RMSEP of 1.63 °Brix. Therefore, in the following discussion, only the region between 400 to 1800 nm was used to predict fibre and moisture content from the internode samples.

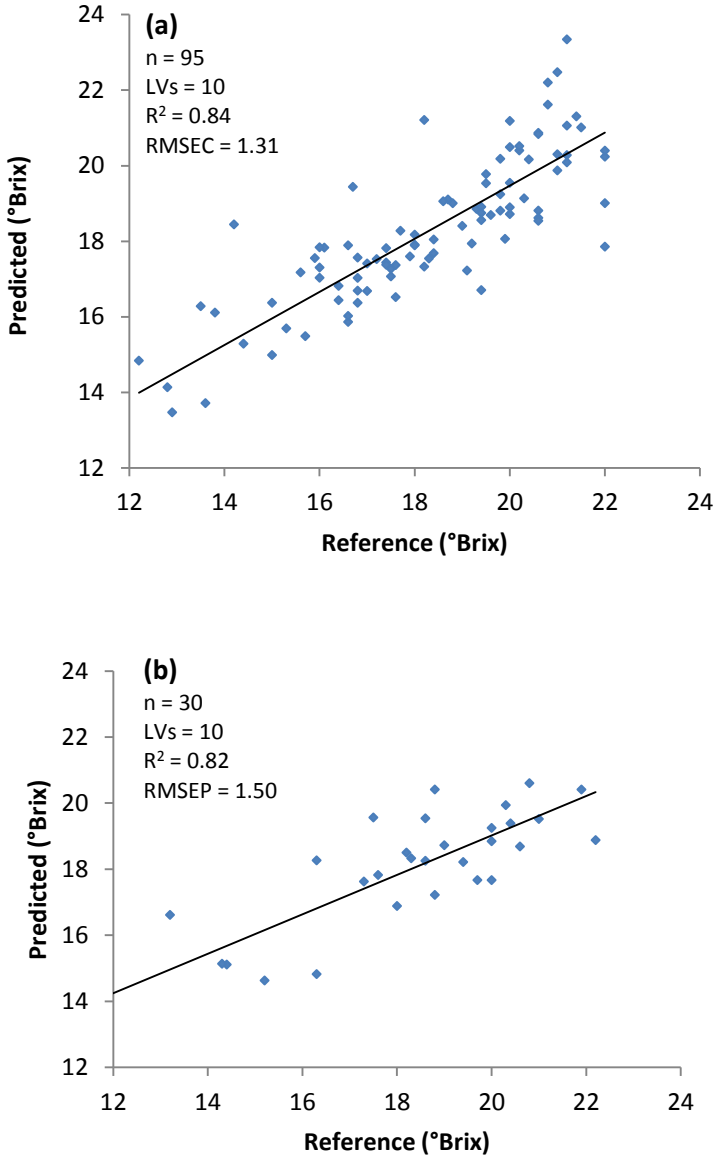


Figure 7.9: Scatter plots of reference versus predicted °Brix by SSM in the wavelength range of 400 to 1800 nm: (a) calibration model; (b) prediction model

7.5.2 Application of the VNIRS to predict ternary quality parameters by the SSM

The performance of the PLS models in predicting ternary quality parameters based on the SSM using both spectroradiometers are presented in Table 7.7. A basic statistical characteristic of the data for this analysis was provided in Table 6.11. The calibration and prediction models for the FRS were developed using a wavelength region between 400 and 1800 nm. The table shows that the calibration models for all components as measured by the VNIRS were very good with R^2 values ranging from 0.89 to 0.90. The RMSECs values for °Brix, fibre and moisture were 1.04, 0.99 and 1.56 °Brix, respectively. In terms of prediction performance, the highest prediction accuracy (R^2 = 0.93) was obtained when the spectral data were used to predict fibre content. The prediction of moisture content and °Brix value gave the R^2 values of 0.90 and 0.88, respectively. This demonstrated that the VNIRS could be used to effectively predict fibre content from internode samples based on SSM. The prediction accuracies for all quality components obtained by the SSM were comparable with the prediction accuracy obtained by CSSM using the same equipment.

Table 7.7: Performance of PLS models in prediction TGM parameters based on the SSM

Component	FRS					VNIRS				
	LVs	Calibration		Prediction		LVs	Calibration		Prediction	
		R^2	RMSEC	R^2	RMSEP		R^2	RMSEC	R^2	RMSEP
°Brix	10	0.84	1.30	0.82	1.50	10	0.90	1.04	0.88	1.19
Fibre	5	0.92	0.99	0.90	1.13	10	0.91	0.99	0.93	0.86
Moisture	8	0.89	1.73	0.91	1.76	10	0.89	1.56	0.90	1.70

Table 7.7 also shows that the prediction performances for all ternary quality parameters by the FRS were reasonably good with all R^2 values being above 0.80. The prediction of fibre and moisture content only required 5 and 8 LVs numbers, respectively. Overall, the prediction accuracy obtained by the FRS in predicting moisture content (R^2 = 0.91) was better than fibre content (R^2 = 0.90) and °Brix (R^2 = 0.82). For moisture content prediction, the use of the FRS gave better prediction accuracy than the VNIRS. While for °Brix and fibre content prediction, the VNIRS performed better than the FRS.

In comparison to the CSSM, the prediction of °Brix using the FRS based on the CSSM gave higher accuracy (R^2 = 0.89) than the prediction based on the SSM (R^2 = 0.82). For fibre prediction, the SSM method yielded higher accuracy (R^2 = 0.90) than the CSSM (R^2 = 0.69). Similar to fibre content, the prediction accuracy for moisture content by the SSM (R^2 = 0.91) was better than the CSSM (R^2 = 0.76). These results suggest that the scanning method had an influence on the prediction accuracy for each quality component. Therefore, the selection of the right scanning method for the right quality parameter is important to achieve better accuracy. For the °Brix measurement however, the VNIRS gives better accuracy than the FRS. This finding is consistent with the results reported for the CSSM. This finding also suggests that the low-cost and portable VNIRS can be used to predict sugarcane quality from the stalk samples based on either CSSM or SSM. The ability to predict sugarcane quality components from stalk samples would allow the application of this technology in the field.

7.5.3 Comparison between CSSM and SSM for the FRS

A comparison of the average absorbance spectra of the internode samples as measured by the FRS based on the SSM and CSSM are shown in Figure 7.10. From the figure, it can be seen that the trend of the curves for both scanning methods look similar with several obvious absorption peaks around 680, 958, 1120, 1400 and 1900 nm. The peak of 680 nm was due to chlorophyll content (Abbott et al. 1997). The absorption peak around 958, 1120, 1400 and 1900 nm were related to water content in the stalks (Williams & Norris 1987).

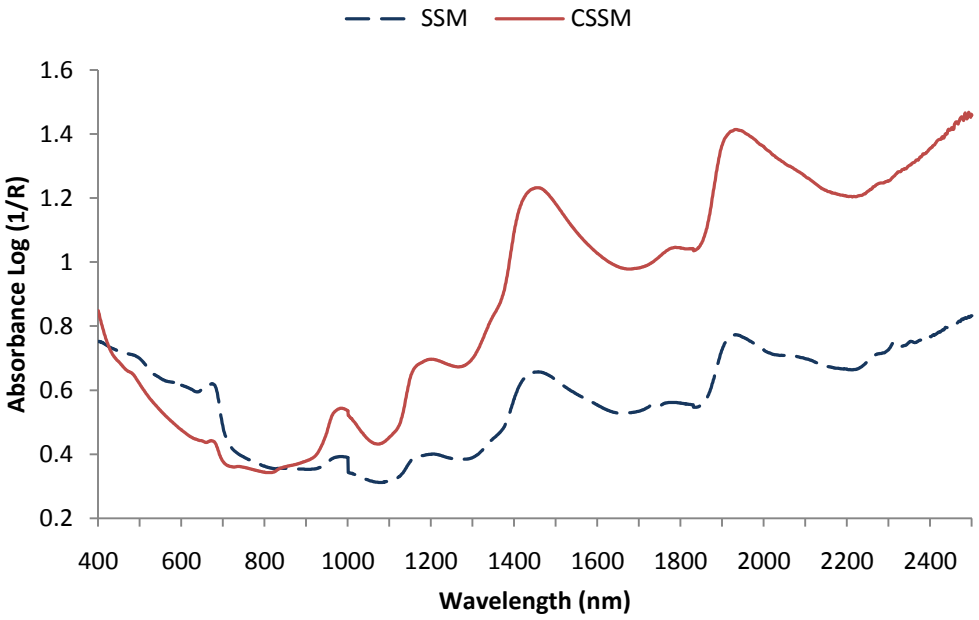


Figure 7.10: Comparison of typical absorbance spectrum between SSM and CSSM

In terms of the absorption level, as the wavelength increased, the absorption level for the CSSM gradually became higher than the SSM curve. The difference indicated that the skin surface had a lower absorption level than the internode flesh because of the presence of the skin. Lower absorption by SSM may indicate that the light source which radiated the skin surface has been scattered away from a detector. The low absorption by the skin is also due to the fact that the whole skin surfaces are covered by a thin protective layer of waxy material which is moisture resistant (Huang et al. 2005). Furthermore, the diffuse reflectance of the skin surface is related to the surface geometry (size and shape) of the individual internode, while the cross sectional surface does not exhibit this problem as measurements were taken from a flat cross sectional surface. This finding is consistent with a study conducted by Lu et al. (2000) on peeled and unpeeled apples.

The difference of the absorption level may have influenced the prediction accuracy of both scanning methods. The higher absorption level for the CSSM caused this scanning

method to have higher prediction accuracy (R^2 = 0.92/RMSEP = 1.03 °Brix) than the SSM (R^2 = 0.82/RMSEP = 1.50). However, both scanning methods could be applied in the field for different measurement activities. For example, the CSSM has the potential to be used to screen the quality of sugarcane crops in the field for breeding and clonal evaluation programs. For an on-the-go quality measurement, the SSM has the potential to be applied to scan the moving billets on the elevator of a sugarcane harvester. However, to realize the potential application of the SSM on a harvester, more research is needed to develop a reliable measurement and sampling system which is suitable for mounting on a sugarcane harvester.

7.6 Classification of sugarcane quality based on spectral data from the SSM

7.6.1 Training and testing data sets for ANN classification

The ANN classification procedures were performed on the spectral data collected from the skin scanning method (SSM). The spectral data from the SSM was selected for classification analysis since the SSM has a better potential to be applied on a harvester during harvesting. The reflectance spectral data were pre-treated with MSC method prior to classification analysis. Then, using the equal interval classification method, the spectral data was first divided into five quality classes (Table 7.8) and then into three quality classes (Table 7.9). These two types of different classes were selected to enable investigation of the effect of class size on classification accuracy. The spectral data of each quality class from both tables were divided into training and testing sets. The training set was used to develop the classifier model while the testing set was used to evaluate the performance of the classifier model. 75% of the data set was selected and used in the training set and the remaining (25%) was used as the independent test set. Samples for the test set were selected by taking one of every four samples from the entire sample set, taking care to ensure that each set included samples that covered the entire range of the spectra.

Table 7.8: Classification table of °Brix values for five classes

Class	°Brix range
High (H)	19.3-22.2
Medium high (MH)	16.5-19.2
Medium (M)	13.5-16.4
Medium low (ML)	10.5-13.4
Low (L)	7.6-10.4

Table 7.9: Classification table of °Brix values for three classes

Class	°Brix range
High (H)	17.5-22.2
Medium (M)	12.6-17.4
Low (L)	7.6-12.5

After removing one outlier from the original training data set (220), the PCA method was then run over the training set consisting of 219 samples while the testing set had 71 samples. For the five-class quality model, the output layer consisted of five neurons

corresponding to each °Brix range as defined in the classification table (Table 8.1). For the three-class quality model, the output layer consisted of only three neurons.

The ANN process was carried out by adjusting the weight of each node, using the transfer function. The transfer function employed is the sigmoid function bounded between 0 and 1 (Eqn. 7.2). The numerical values of the input and output variables used by the ANN were normalised values in the range of 0 to 1. The ANN algorithm including the PCA method for classification purposes was programmed and executed in Matlab (Version 7, The Mathworks Inc. Natick, MA, USA). The source codes for the ANN algorithm is provided in Appendix A.

7.6.2 ANN classification analysis for five quality classes

Five quality classes, with their respective threshold values were predefined from the distribution of °Brix as shown in Table 7.8. Table 7.10 shows the accuracy of classifying sugarcane °Brix using ANN with five PCs. In this model, the accuracy of classification ranged from 50% to 100% with the average accuracy being 83.1%. The accuracy ranges obtained in this study were better than that (45.2 to 93.5%) reported by Park et al. (2003) for classifying soluble solids of Gala apples. However, the result of this study was lower than the result (99.5% accuracy) reported by Mohan et al. (2005) who used ANN to classify grain quality.

Table 7.10: ANN classification results for five quality classes

Class	°Brix range	No. of test samples in each group	No. of correct classification	Accuracy (%)
High (H)	19.3-22.2	30	30	100
Medium High (MH)	16.5-19.2	23	14	60.9
Medium (M)	13.5-16.4	11	10	90.9
Medium Low (ML)	10.5-13.4	4	2	50
Low (L)	7.6-10.4	3	3	100
Total		71	59	83.1

From the table, it can be seen that two dominant classes representing high and low °Brix level (H and L classes) had achieved 100% accuracy. The M class was also good with an accuracy of 90.9%. However, both inter-middle classes (MH and ML) showed lower accuracy. This low accuracy (misclassification) which only affected the boundary samples was probably because these two inter-middle classes had °Brix values which were in a borderline between the main classes. Thus, they had reflectance spectra similar to that of the major classes. However, this misclassification is considered minor because it has just affected the inter-middle group of quality classes. The accuracy of the classification could be improved if only three classes were adopted. Overall, the accuracy of this study can be considered as acceptable and useful especially for mapping the quality regions in the field. The ability to map the quality regions across the paddock could help growers to better manage their fields thus improve crop production in both yield and quality.

7.6.3 ANN classification analysis for three quality classes

The division of the quality classes into three categories was in accordance with the work of Bramley et al. (2012) who suggested that sugarcane quality mapping could be classified into three levels of productivity measure (low, medium and high). In their work, the map created using this classification technique was found to be able to retain the essential spatial structure as compared to the other maps produced using six quality classes. Quality maps with only three levels could be very effective for use in the calculation of variable rate application maps for various crop inputs.

Table 7.11: ANN classification results for three quality classes

Class	°Brix range	No. of samples in test set	No. of correct classification	Success rate (%)
High (H)	17.5-22.2	44	42	95.5
Medium (M)	12.6-17.4	22	19	86.4
Low (L)	7.6-12.5	5	2	40
Total		71	63	88.7

Table 7.11 shows the accuracy of classifying sugarcane °Brix using ANN with three PCs. In this model, the accuracy of classification ranged from 40 to 95.5% with the average accuracy being 88.7%. The average accuracy obtained using three quality groups were better than that of five quality groups. Overall, this study has demonstrated that the Vis/SWNIR spectroscopy technique coupled with ANN has the potential to be used for online quality measurement to fulfill the requirement of PA. The average accuracy increased from 83.1 to 88.7% when a three-group classification was attempted.

7.7 Discussion and conclusions

This chapter has used both spectroradiometers (VNIRS and FRS) to predict sugarcane quality components from either whole stalk samples or internode samples based on the SSM. For the whole stalk samples, the VNIRS could predict the pol and °Brix with reasonable accuracy as indicated by R^2 values of 0.72 and 0.85, respectively. However, the performance of this equipment in predicting CCS was poor (R^2 = 0.64). For fibre prediction, the accuracy obtained was not satisfactory with R^2 value of 0.38. On the other hand, the prediction accuracy obtained by the FRS for the pol, °Brix, CCS and fibre was good with the R^2 being 0.80, 0.85, 0.75 and 0.72, respectively.

For the internode samples, the prediction accuracy of the model developed from the average absorbance spectral data of each individual internode sample was good with the R^2 and RMSEP values being 0.87 and 1.68 °Brix, respectively. For the reflectance spectral data, the R^2 and RMSEP values were 0.91 and 1.41 °Brix, respectively. The RPD values for the prediction models developed for the reflectance and absorbance spectral data were 2.0 and 1.7, respectively. These RPD values indicated that the SSM method based on the reflectance data has the potential to be applied for coarse quantitative prediction in the field.

For the model developed from each individual spectrum of each cut section, its prediction accuracy was lower with R^2 and RMSEP values of 0.74 and 1.88 °Brix. The internode sectional analysis has shown that the °Brix variation between the node and internode areas was up to 2.9%. The analysis also found that the S4 (node portion) had the highest °Brix value (17.9) compared to the other sections. However, the prediction accuracy of this section was the lowest with R^2 being 0.67. The variation of prediction accuracy within the internode samples could be up 11.9%.

Similar to Chapter 6, this chapter has also shown that the spectroscopic methods could be applied to determine ternary quality components based on the SSM. For °Brix prediction, it has been found that the VNIRS gave higher prediction accuracy (R^2 = 0.88) than the FRS (R^2 = 0.80). Similar to °Brix prediction, the VNIRS also performed better than the FRS in predicting the fibre content, with R^2 values being 0.93 and 0.90, respectively. However, for moisture content prediction, the FRS gave slightly better accuracy (R^2 = 0.91) than the VNIRS (R^2 = 0.90). This finding suggests that the low-cost and portable VNIRS has the ability to be used to predict sugarcane quality form the stalk samples. This ability will allow the equipment to be applied in the field.

Using the ANN classification method, the spectral data of SSM have been classified into several quality classes. The division of the quality level into five quality classes has yielded the average classification accuracy of 83.1%. When the same spectral data was divided into three quality classes (low, medium and high), the average classification accuracy was increased to 88.7%. These results suggested that the lower class number could increase classification performance of the model. The classification accuracy found in this study can be considered as acceptable because the main reason for quality mapping in the field is to identify areas with high, medium and low quality levels.

Chapter 8

Basic calculation and conceptual design of in-field quality measurement system

The previous chapters have shown that the spectroscopic methods could be applied to predict sugarcane quality from stalk samples. This chapter will describe how those new methods could be actually applied to develop an in-field quality measurement system on a commercial sugarcane harvester. The description includes the specifications for the development of sampling system. The chapter will also discuss integration procedures for the measurement system, sampling system and GPS mounted on a harvester to perform in-line quality measurement.

8.1 Introduction

The ability of spectroradiometers to measure sugar content from stalk samples has opened up the new possibility of installing the equipment on a sugarcane harvester for an on-the-go quality measurement system. The measurement of quality from stalk samples as opposed to juice or other sample forms could allow a rapid and non-destructive sensing technique to be performed during harvesting as the complex sample preparation is no longer required. The availability of a low-cost, affordable and reliable VNIR Si PDA spectroradiometer makes the proposed system a real possibility for field uses.

The goal of this chapter is to discuss the theoretical development of a quality measurement system for sugarcane (QMSS) and sampling system for a sugarcane harvester. To achieve this goal, existing quality measurement systems and sampling strategies from other crops were critically reviewed. From this review, the specifications for the development of the QMSS, sampling system and in-line data treatment are proposed. The specific objectives of this chapter were:

1. To theoretically develop a quality measurement system mounted on a sugarcane harvester.
2. To theoretically develop a sampling mechanism mounted on a sugarcane harvester.

8.2 Development of conceptual design for the QMSS

The ability to measure sugarcane quality on a harvester during harvest will be very useful for the sugarcane industry. However, the development of such system faces serious technical challenges due to logistical limitation, mechanical vibration, high level of contamination and rugged field environment unsuited to delicate optical and electronic equipment. Since the QMSS for a sugarcane harvester will be working in harsh and dusty harvesting conditions, a number of design factors must be considered,

including ease of mounting on a sugarcane harvester with minimal modifications and has the flexibility to be used in a wide variety of operating conditions (i.e., different harvesting speeds, weather conditions, field conditions and crop yields). The selection of a suitable sensor, sample form, sampling mechanism and sampling location could determine the success of the proposed system.

The selection of suitable and reliable sensor is one of the key considerations for field measurement. The selected sensor should be easy to calibrate, have sufficient precision and accuracy, be low-cost and not present any obstruction to the normal harvesting even when the sensor is damaged (Reyns et al. 2002). To meet some of these requirements, the application of PDA spectrometer has been recommended (Fernández-Ahumada et al. 2008). Section 3.6 has reviewed the application of PDA spectrometers for quality measurement in the field (Table 3.1). The section has suggested that the PDA spectrometers with Si detector had the potential for measuring sugarcane quality on a harvester.

8.2.1 Key components of the QMSS

The QMSS which is proposed in this chapter is composed of an optical module, a controller module and a data logger (Table 8.1). The optical module consists of a Si detector and a halogen illuminator. A sapphire lens could be used to protect optical components the QMSS from wear and tear due to the high speed of billet streams containing abrasive contaminates such as sand, soil, trash etc.

Table 8.1: Basic components of the spectroscopic system for the QMSS

Component	Element	Function
Detector	Si detector	To collect reflected light from billet samples, convert them into a modulated voltage and send them to controller
Illuminator	Halogen	To illuminate billet samples
Filter	Sapphire lens	To provide protection to the system from abrasive soil and trash
Controller module	Signal conditioning circuit	To process the converted modulated voltage from the photodiode which separates each source of reflected light from the photodiode signal and converts the modulated voltage to a direct current (DC) voltage.
	Analog to digital (A/D) converter	To process the DC voltage through a 12 bit A/D converter and the output is then sent serially by a microprocessor to the data logger.
	data logger	To save the digital reflectance values with geo-referenced locations from the GPS.
DGPS	DGPS	To geo-reference quality data using a DGPS receiver. To determine position, latitude and longitude.

8.2.2 Technical specifications of the spectrometer

The experimental works reported in this thesis have been conducted mainly using the VNIR PDA spectroradiometer with a Si detector. However, this spectroradiometer was not designed for in-line measurement. Therefore, a new suitable sensor designed for in-

line application manufactured by Carl Ziess, Jena, Germany, based on a similar PDA spectrometer was proposed for the QMSS. The PDA spectrometer from Carl Zeiss has been a sensor of choice for the QMSS as the sensors have been proved to be reliable for field application by several studies (Table 3.1).

The technical specifications and indicative price of the similar equipment were requested from Carl Zeiss in April 2013. However, the spectrometer that works in the wavelength regions of 400 to 1000 nm as used in this thesis was not available at that time. Instead, the company has given a quotation for the Zeiss Corona Plus 45 NIR (1.7), the alternative unit which was believed to be suitable for scanning the sugarcane stalk in the range of 400 to 1000 nm. The technical specifications of this sensor are given in Table 8.2. Further discussion on the QMSS was based on the information from this table. The price of the Zeiss Corona Plus 45 NIR (1.7) spectrometer which is built using InGaAs detector was AUD $25815.00. As a Si detector is a cheaper option, when this sensor becomes available, its price should be much cheaper than that of the InGaAs.

Table 8.2: Technical specifications of Zeiss Corona Plus 45 NIR (1.7)

Component	Feature
Spectrometer	Double beam diode array
Polychromator	PGS
Measurement range	950 – 1670 nm (InGaAs)
Mean spectral pixel pitch	3 nm
Spectral resolution	≤10 nm
Wavelength accuracy	≤1 nm
Wavelength reproducibility	≤0.1 nm
Light source	Halogen
Dimension (W x H x D)	230 x 110 x 280 mm
Weight	6 kg
Range of operating temperature	5 to 45 °C
Supply voltage	12-24 V

8.2.3 Measurement chamber

A measurement chamber would be required to provide the necessary protection to the sensor and illuminator from damage caused by the weather, billets, trash and dirt. The measurement chamber is also important to protect the system from interference of ambient light and to keep the sensor and samples at a constant angle and measurement distance during field operation. A suitable cradle could be used to mount the measurement chamber in order to minimise the effect of vibration on the system. The proposed installation of the measurement chamber on a harvester is discussed in Section 8.5.2 and 8.5.3.3.

8.3 Possible sample forms and their measurement locations on a harvester

The VNIRS has the potential to measure sugarcane quality from three types of sample forms: juice samples, fibrated samples and stalk samples. Unfortunately, different sample forms require different specific sampling methods and measurement configurations. In author's opinion, in order to develop a proper quality measurement

system on a harvester, at least four essential mechanisms should be designed and developed, namely a sampling, preparation, presentation and measurement mechanism. The function of each mechanism is discussed in Table 8.3. For each mechanism, several possible methods could be adopted. The selection of the best method for each mechanism should be based on criteria related to cost, technical feasibility, minimal mechanical disturbance and ease of operation.

Table 8.3: Primary mechanism needed for in-field quality measurement system*

Mechanism	Purposes	Possible Methods
Sampling	To extract sufficient billet samples from the billet stream on elevator. Sampling mechanism must be designed to suit the machinery parameters (harvester speed, elevator speed) and physical and mechanical parameters of the crop.	Air blower system Vacuum system Deflection flap Mechanical arm
Preparation	To hold samples for solid scanning. To squeeze samples to supply consistent juice sample that could be analysed by the measurement unit.	Mechanical squeezer Mini macerator Hydraulic press Hammer mill
Presentation	To present continuous samples in a uniform, homogeneous and well-defined portion to the sensor.	Sample holder for solid scanning Cuvette for juice scanning
Measurement	To perform real-time measurement on selected samples.	Spectroscopy

* assessed by the author from available information.

However, the sample forms which can be relatively easy to obtain on a harvester are stalk samples and raw juices. The locations where such samples could be measured on a harvester during harvesting are shown in Figure 8.1. The clarified juice samples and fibrated samples will require sample preparation, and are not shown in this figure.

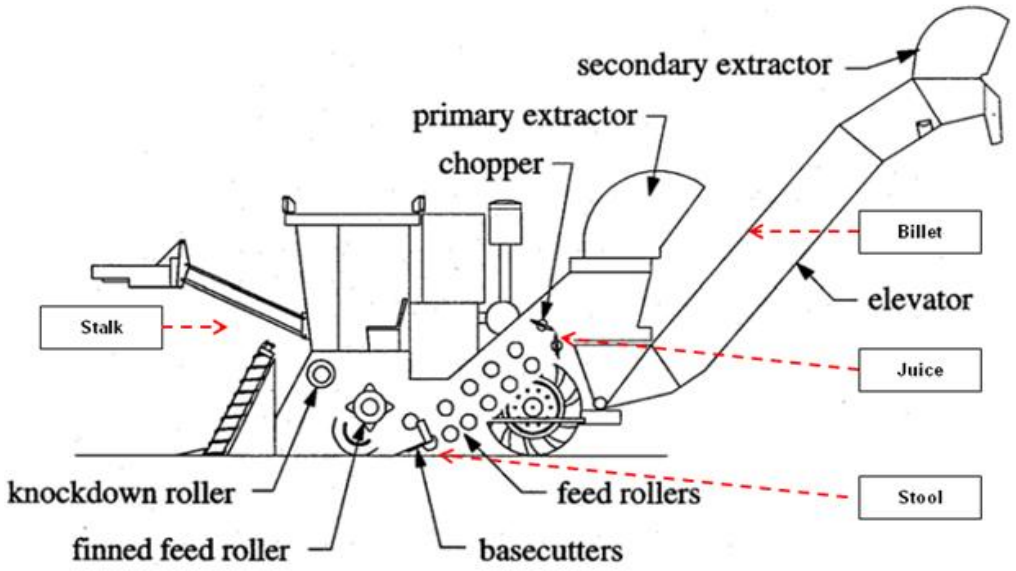


Figure 8.1: Location from which possible samples could be obtained

8.3.1 Juice sample

To measure sugarcane quality from a juice sample without any sample preparation, the chopper drum may be the best place to install the spectrometer. In this area, the cut stalk will be chopped into billets of 200 to 300 mm length. Due to the frequent chopping operation, this section would have the largest amount of juice compared to other sections. The deformation behaviour and resulting damage due to the cutting force (billeting) causes the loss of fibre and juice in the order of 3 to 8% (Gavin 2008). If the lost juice in this section could be collected, it could give enough juice samples for a measurement system. It might, however, be very technically challenging to collect data from this location due to the presence of huge amount of trash and limited space.

8.3.2 Stalk and billet samples

Other than the juice scanning method, quality measurement can also be performed on the stalk samples as a standing stalk, cut stool or billets. For stool scanning, the quality measurement can be done using the CSSM (Described in Chapter 6) by installing a spectrometer on the base cutter. This technique assumes that a spectrometer could immediately scan the stool on the ground after it is cut by the base cutter. The same measurement principle was applied by Panigrahi & Hofman (2003) for harvesting sugar beet. However, since the base cutter cuts the stool very deep into the soil, the moist surfaces of cut sugarcane may attract dirt and soil which could introduce errors into the measurement. This approach was not progressed.

The SSM could also potentially be applied to standing stalks or billet samples. The selection of the best sample types between these two samples should be the result of a number of considerations (Table 8.4). To measure quality from a standing stalk, a spectroradiometer could be installed in front of the crop divider. The sensor would scan the standing stalk while the harvester is moving forwards. This method however might be hindered by the presence of trash, dead leaves and buds. Thus, overall, it may be concluded that measurement of the quality parameter from billets is probably the easiest method to perform on a harvester. The following section will discuss this option in detailed.

8.4 Sampling strategies for field uses

In the Australian sugarcane industry, the first attempt to develop a sampling system during harvest was reported by Robotham (2000). In this study, the author designed, constructed and field-tested an automated sugarcane billet sampler which combined the operations of weighing the total harvested billets, producing representative sub-samples, and unloading of excess billets to a haul-out vehicle. The sampler was mounted on a two wheel drive truck chassis which gave the unit sufficient mobility to test the concept (Figure 8.2). The sampler was designed to receive billets from a harvester. Then, the billet sub-samples were extracted by the sampler elevators before the billets were transferred into the haul-out vehicle.

Table 8.4: Comparison of technical requirements between whole stalk and billet scanning*

Factor	Standing stalk scanning	Billet scanning
Scanning point	Easy to choose the scanning point along the stalk height (assuming the crop height is the same within one plot)	No control over scanning point as the billets will be randomly scanned (multiple detector could average the scanning on individual billets)
Sample preparation	Requires leaf-stripping machine to remove the leaves and sheath from stalks	Not required as the billets can be scanned directly as they are
Trash presence	Standing stalks are relatively clean as they are free from foreign materials (trash, soil, etc)	Since the primary extractor normally removes up to 90% of the trash, there will be some trash left and mixed with billets
Measurement time	Longer measurement time is available due to low forward speed (7 km hr^{-1} = 1.94 m s^{-1})	Shorter measurement time available as the speed of the elevator is high (2.5 m s^{-1})
Potential damage to the system	High density crops would damage the system	No damage as billet scanning applies non-contact measurement method
Dependence on sampling mechanism	Depending on how fast the leaf-stripping mechanism works	Not related
Location	Front area only	Anywhere along the elevator
Mechanical vibration	Relatively high as the measurement system will have to hit the standing stalk to move forwards	Relatively low as the system could be installed external to the harvester

* assessed by the author from available information.

The sampling door was the key element of this type of billet sampler. The sampling door was a moveable floor panel that was mounted on pivoting parallelogram arms. The opening and closing action of the door was provided by a pitman arm system driven by a hydraulic motor as shown in Figure 8.3. Each sampling action took a period of three seconds. Each billet sample was collected in a bag mounted below the sampling door (Robotham 2000). Later, the bags were brought to a laboratory for quality measurement.

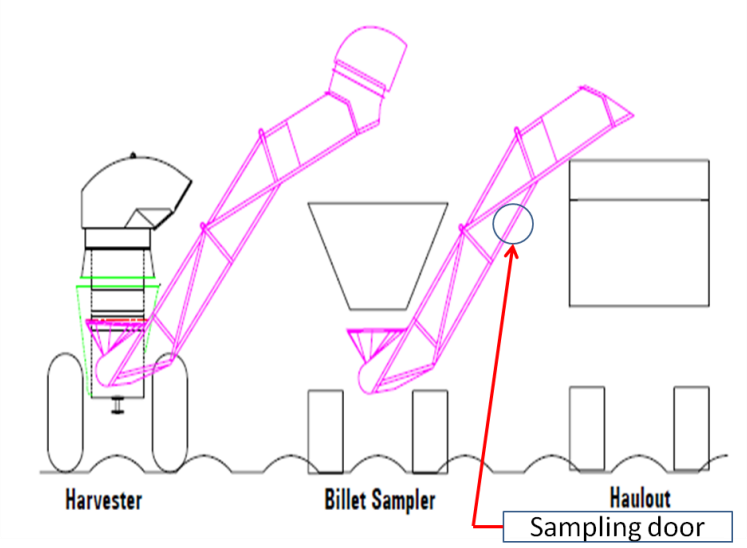


Figure 8.2: Operational configurations for a biller sampler (Robotham 2000)

Figure 8.3: Sampling door and Pitman arm system (Robotham 2000)

The application of this system on a sugarcane harvester was shown to be able to provide adequate samples. However, this sampling system was not designed for an on-the-go quality measurement system as the quality of the collected billets was still determined at a laboratory. Furthermore, it was reported that the machine could not reliably operate in the field with extremely high levels of extraneous matter or excessively long billets. In both cases, uniform feeding of the billets onto the slat elevator could not be maintained and bridging of the sugarcane could occur. Thus, this sampling method was not considered for further investigation in this chapter.

In broadening the sampling system concepts, it is very important to study the existing measurement systems from similar crops, especially forage crops. One of the published

studies regarding the development of a sampling system to be used with a PDA spectrometer to measure forage quality parameters on a plot harvester was reported by Welle et al. (2003). In this study, a subsample (1 kg) was taken from each plot by an auger for reference chemical analysis. After weighing, the samples were deposited onto a moving conveyor belt. The material then passed under overhead rollers in order to produce an even surface. A PDA spectrometer with a wavelength range from 400 to 1700 nm was mounted behind the rollers 120 mm above the sample surface. On a sugarcane harvester, however, the use of an auger to extract billet samples would not be possible due to the size and physical dimensions of the billets. The installation of an auger on a chain-driven slat elevator system of a sugarcane harvester is also technically impractical.

A mobile, PDA NIR spectrometer (950-1680 nm) was integrated by Digman & Shinners (2008) into the spout of a self-propelled forage harvester to measure crop moisture (Figure 8.4). The harvester's spout was fitted with the spectrometer for collection of field spectra during harvest. The sensors location was just above the crop accelerator, where the direction of the crop stream changed from nearly vertical to horizontal. This measurement method could potentially be applied on a sugarcane harvester by installing a spectrometer on top of the elevator. This proposed system would then scan the moving billets on the elevator. This method, however, should be designed to suit the typical speed of a chopper harvester and the elevator at 7 km hr^{-1} and 2.5 m s^{-1} respectively (Cox 2002).

Figure 8.4: Location of a mounted spectrometer to measure crop moisture in forage (Digman & Shinners 2008)

A modular, computer–controlled sampling system was developed by Long & Buckmaster (2003) to directly sample chopped forage material from a forage harvester spout. Pneumatic cylinders powered a deflector flap to redirect crop material from the spout. Then, a hydraulic cylinder compressed deflected forage material into PVC pipes (356 mm long and 102 mm diameter), which served as miniature test silos. This study demonstrated that the pneumatic system could be used to power a deflector flap to

redirect crop materials from the spout. This sampling concept could also potentially be applied to design a sampling mechanism to extract billet samples from the elevator for quality measurement purposes.

A further study that removed the crop samples from the main stream for quality measurement was reported by Sassenrath et al. (2005). In their study, the authors developed a sampling system to remove a portion of the cotton during mechanical harvest to determine fibre properties. The sampler works by diverting cotton from the picker transfer chute to a small sampling station by an air blower system. In contrast to the air blower system, vacuum techniques have been applied in the harvesting of apples (Lehnert, 2010; Schupp et al. 2011).

A vacuum technique (also known as a pneumatic conveying technique) also has the potential to be used to extract the billet samples from the elevator. The pneumatic system has long been used on a sugarcane harvester to extract trash in the primary extractor. The primary extractor, with increased fan speeds, could be used to remove mature billets (Richard et al. 2001). Barber (1997) suggested that a pneumatic conveying system based on the vacuum technique could handle materials ranging from asbestos with a bulk density of 100 kg m^{-3} to crushed stone with a density of 1500 kg m^{-3}. Thus, this system is considered to be capable of handling billet sample with typical density of 353 kg m^{-3} (Mailander et al. 2010).

8.5 Application of the SSM on a harvester

8.5.1 Selection of a sampling location for billet samples

The SSM on billet samples would be the most practical method to be applied on a harvester as they could be obtained along the elevator conveyor. The elevator conveyor is a chain-driven slat system consisting of a slew table, slewing system, bowl and chain conveyor. The typical width, S_w and height, S_h of the elevator slat are 850 and 50 mm, respectively. The typical distance between slats, S_d is 800 mm (Figure 8.5). The elevator receives the billets from a chopper drum and lifts around 1.5 t min^{-1} billets to a 4.5 m unloading height with a typical conveyor speed of 2.5 m s^{-1} (Mailander et al. 2010). Since the elevator is an open section, it seems to be practical to install the QMSS measurement on top of the elevator, at any point along the elevator. Furthermore, the billet samples in this location also have the advantage of being relatively clean as 75 to 90% of the trash has already been blown out by the primary extractor (Mailander et al. 2010).

Figure 8.5: Typical slat arrangement on the elevator of a sugarcane harvester

The QMSS on the elevator can potentially be designed in two ways: by directly scanning the moving billets on the elevator [direct scanning method (DSM)] or by scanning the billet samples which are supplied by a sampling mechanism [assisted scanning method (ASM)]. The mechanism for DSM could be mounted in Location A while the mechanism for the ASM could be mounted in Location B (Figure 8.6). Figure 8.7 shows billets arrangement at the proposed measurement locations. The selection of different locations for each measurement method was due to different floor inclinations. In Location A, the elevator is at an angle of 25° with the ground. At this lower inclination angle, the billets would be distributed more evenly between two slats. This billet arrangement is more suitable for the DSM. In Location B, the elevator is at the angle of 50° with the ground. At this inclination, billets tend to gather on the slat's surface due to the gravitational acceleration. Thus, Location B is suitable for installing a sampling mechanism.

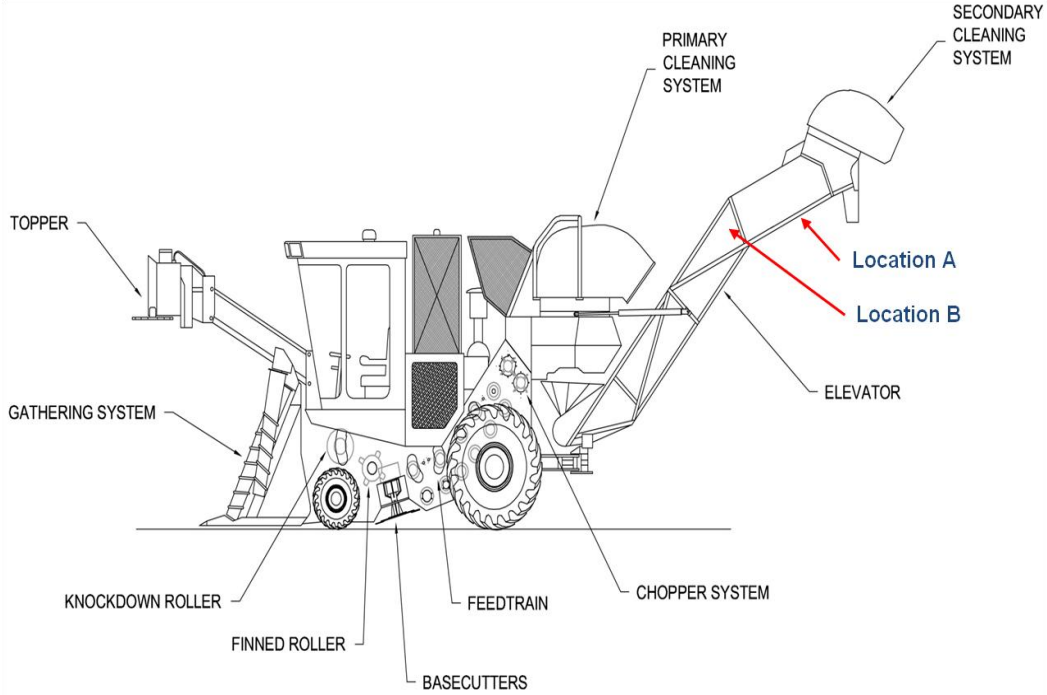


Figure 8.6: Proposed locations of two measurement chambers

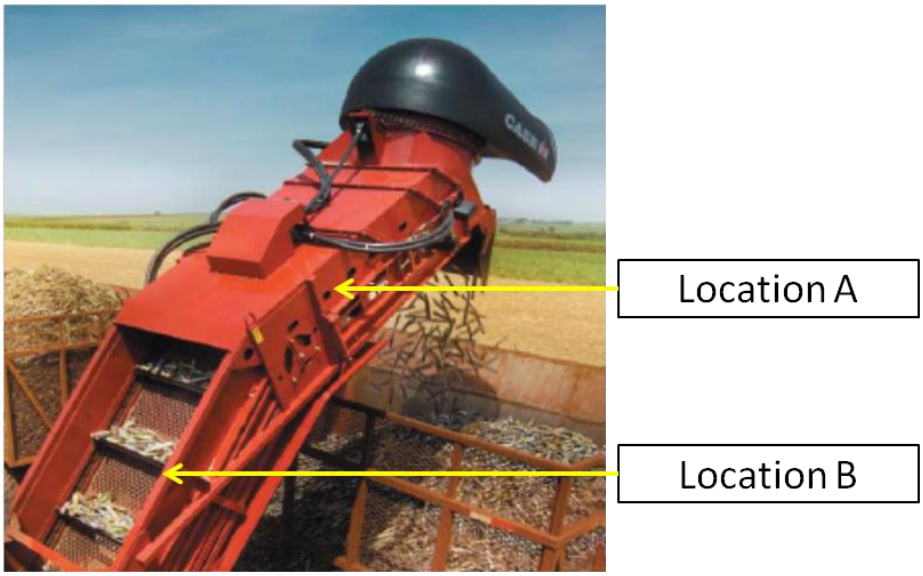


Figure 8.7: Condition of billets on the elevator at the proposed measurement point

There are many crop and machinery parameters which could influence the sensing capability and sampling efficiency of the proposed QMSS. The parameters which should be considered with their fixed values are shown in Table 8.5. The fixed values of these parameters were used for the calculation in the following sections.

8.5.2 Measurement procedure for the direct scanning method (DSM)

The billets moving on the elevator could potentially be scanned on their skin using the proposed measurement setup as shown in Figure 8.8. In this method, the detector should be mounted facing the moving billet samples. The sensor could be mounted in the middle of slat. The illuminator should be installed at an angle of 45° from the detector. The detector is activated when the moving slat touches the trigger button located at the end of the measurement chamber. To minimise the presence of trash with the billets, a hairbrush will be installed from one end of the measurement chamber. The white reference will be kept in the measurement chamber in front of the detector, and comes out electronically once the switch is on.

Table 8.5: Fixed values of parameters used for the QMSS

Parameter	Symbol	Unit	Value
Yield	Y	t ha^{-1}	80
Billets length	B_l	mm	300
Billets Weight	B_w	g	300
Billet diameter	Bd	mm	30
Billet density	B_p	kg m^{-3}	353
Projected area of billet	B_a	mm^2	8670
Terminal velocity of billet	V_t	m s^{-1}	19.23
Air density	A_p	kg m^{-3}	1.223
Row spacing	R_s	m	1.8
Harvester throughput	H_t	t hr^{-1}	60
Harvester ground speed	S_g	km hr^{-1}	7
Elevator speed	E_s	m s^{-1}	2.5
Distance between slats	S_d	mm	800
Slat height	S_h	mm	50
Slat width	S_w	mm	850
Cell size	C_s	m^2	400
Cell length	C_l	m	20

8.5.2.1 Available time for the DSM

This measurement technique requires the spectra to be measured every time a slat with billet samples passes the detector. This technique assumes that the billets collected between the two slats are uniform. It is also assumed that the area between the two slats is full with billets. Thus, the available measurement time, T_m needed to scan the billet samples collected between two slats can be calculated as follows:

$$T_m = \frac{S_d}{E_S} \tag{8.1}$$

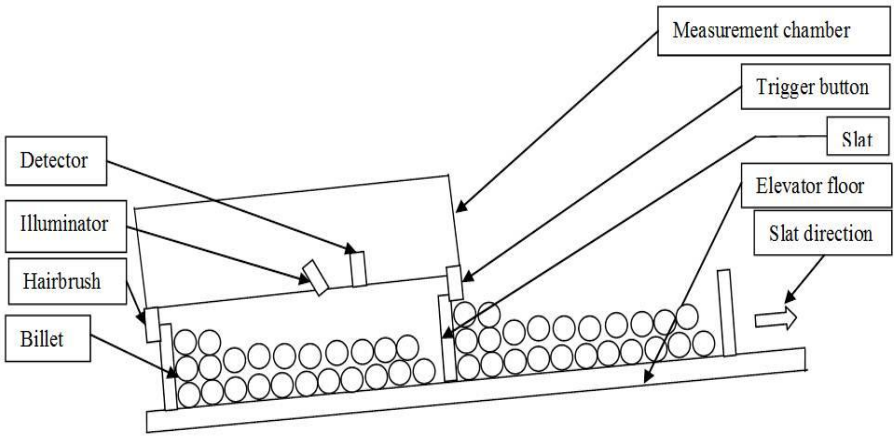


Figure 8.8: Schematic diagram of the QMSS with DSM

Where S_d is a distance between slats (mm), and E_s is an elevator speed (m s^{-1}). In this calculation, the value of T_m is 0.32 seconds. The available measurement time from this proposed system is higher than that reported by Digman & Shinners (2008) at 0.2 second (5 scan in 1 second interval) as shown in Table 8.1. Thus, it is concluded that the QMSS has sufficient time to collect spectral data from moving billets within a reasonable measurement time.

8.5.2.2 Sampling frequency for the DSM

In the sugarcane industry, for mapping purposes, the on-the-go yield monitoring system will divide the harvested area into smaller sections which are usually called a cell. The ideal size of the cell for a harvesting area is approximately 20 x 20 m (400 m^2) (Cox 2002). Thus, this proposed QMSS should be designed to collect spectral data for each cell. Then, the spectral data collected within one cell will be averaged, giving one quality value for each cell. In addition to the quality data, each cell could also contain other information including time, latitude, longitude, harvester speed, and harvested area variables. Using suitable geographical information system (GIS) software, the cells could be interpolated using the Kriging method in order to generate a digital sugarcane quality map.

At a forward speed, S_g of 7 km hr^{-1}, a harvester needs 10.28 seconds to travel over the distance of 20 m (cell length). Within this time, with T_m being 0.32 seconds, approximately 32 spectral data will be collected over the area of 36 m^2 (C_s x R_s). For one cell size (400 m^2), approximately 355 spectra could be collected. These 355 spectra will be averaged in order to have one quality data for one cell.

8.5.3 Assisted (Pneumatic) scanning method (ASM)

The success of the QMSS will depend on how rapidly and precisely a billet sampling mechanism works. This sampling mechanism should be able to provide continuous billet samples to the measurement chamber. The system must not be contaminated by the previous samples. The sample must also be disposed off so that a new reading is obtainable. For these reasons, a non-contact sampling method using a vacuum technique as discussed earlier in Section 8.3 is deemed to be practical. Pneumatic separation relies on the difference in the terminal velocities of two particle types. This technique requires the use of a high-powered fan to obtain an air velocity above the terminal velocity of the billets.

8.5.3.1 Aerodynamic properties of billets

Aerodynamic properties of agricultural materials are needed in the design of handling and processing systems of various agricultural products (Kilickan & Guner 2006). Terminal velocity, V_t is one of the most important aerodynamic properties for separation, pneumatic transportation and cleaning operations (Song & Litchfield 1991). The terminal velocities, V_t of different sugarcane particles were determined by Kroes

(1997) by measuring the velocity of air flow required to just lift and suspend the particles in mid-air. At equilibrium, the drag force is given by:

$$D = m\ g = \frac{1}{2} C_d\ \rho_a A {V_t}^2 \quad (8.2)$$

Where D is the drag force (N), m is the mass of the particle (kg), g is the gravitational acceleration (m s^{-2}), C_d is the coefficient of drag, ρ_a is the density of air (kg m^{-3}), A is the projected area of the particle (m^2), i.e the area perpendicular to the direction of motion or relative air flow, V_t is the terminal velocity (m s^{-1}).

An average value of m, g, ρ_a, A and V_t for typical billet samples as supplied by Kroes (1997) are given in Table 8.6. Based on the results of Joyce & Edwards (1994), Figure 8.8 was derived to show the terminal velocity for whole billets and tops against the sugarcane diameter. From this figure, it can be seen that a vacuum system with air velocity more than 15 m s^{-1} [preferable 19.23 m s^{-1} (Kroes 1997)], would theoretically be able to extract billet samples into the measurement chamber.

8.5.3.2 Power requirement for vacuum system

A measurement box for the QMSS with a vacuum sampling mechanism would consist of a vacuum pump, vacuum nozzle (picking platform) and measurement chamber. The size of this measurement box is expected to be around 0.5 x 0.5 x 0.8 m (0.2 m^3). The development of this measurement box (excluding the NIR sensor) may cost approximately AUD$5000, depending on the materials used. To suit the typical size of billets, the size of the vacuum nozzle should be around 0.4 x 0.05 (0.02 m^2). The vacuum pump and all metering devices could be powered by the hydraulic motor from the sugarcane harvester. A typical sugarcane harvester (model AUSTOFT 8000) comes with 353 hp (260 kW). The primary extractor which is hydraulically driven by a harvester consumes around 30 hp (22.4 kW) for separating billets and extraneous materials with a typical air velocity of 15 m s^{-1} (Figure 8.9).

In earlier research, Joyce & Edwards (1994) reported that the typical requirement of a harvester was about 16 kW of power at an air velocity of 14 m s^{-1} for a cleaning chamber with a cross section of 1.1 m^2, which had an approximate air flow rate, Q of 15.4 m^3 s^{-1}. For the vacuum system proposed in this chapter, as air velocity is increased to 19.23 m s^{-1} and a cross sectional area of the vacuum nozzle is 0.02 m^2, the air flow rate, Q should be about 0.38 m^3 s^{-1}. As the air flow rate decreases, the power required for the vacuum system should also decrease. Thus, since the proposed vacuum method works with the same principle of the primary extractor, the power required for this system which is less than that of the primary extractor, could also be supplied by the harvester.

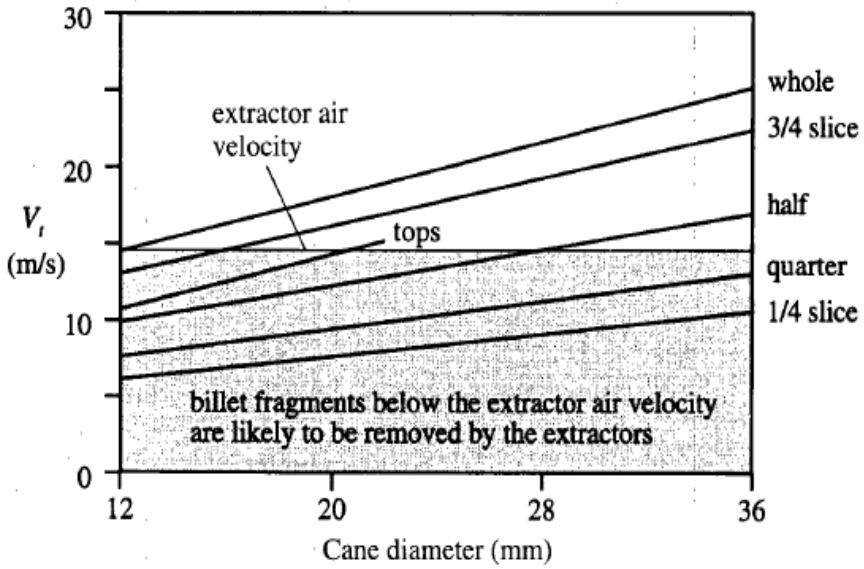


Figure 8.9: Terminal velocities of billets fragments relative to the extractor air velocity (Kroes 1997)

8.5.3.3 Measurement procedure for the ASM

This QMSS with vacuum sampling mechanism could be mounted on the top of the elevator conveyor at the first section of the floor (Figure 8.10). At 50° inclination, the billets tend to gather on the slat wall due to gravitational acceleration. This phenomenon will create an empty space between each slat. The vacuum nozzle will be placed at the top of the slat's tip. The size of the nozzle should be small enough to suck only one billet per time. The billet is assumed to be clean as the hairbrush installed right before the vacuum nozzle will sweep trash from the billet samples. As the slat moves towards the nozzle, it is assumed that only the billet located at the tip of the slat will be sucked by the vacuum nozzle. Once sucked, the billet would go into the measurement chamber. Due to the action of the screw conveyor, the billet will be positioned at the first metering device in a horizontal orientation. Two metering devices will be needed: one before the billet enters the scanning area and another inside the scanning area. The first metering device ensures that only one billet enters the measurement chamber at one time. The second metering device holds the sample for scanning then disposes it back onto elevator. The first and second metering devices should work together, both set to have the same rotation per minute. The white reference will be kept in the measurement chamber opposite the second metering device, and comes out electronically once the switch is on.

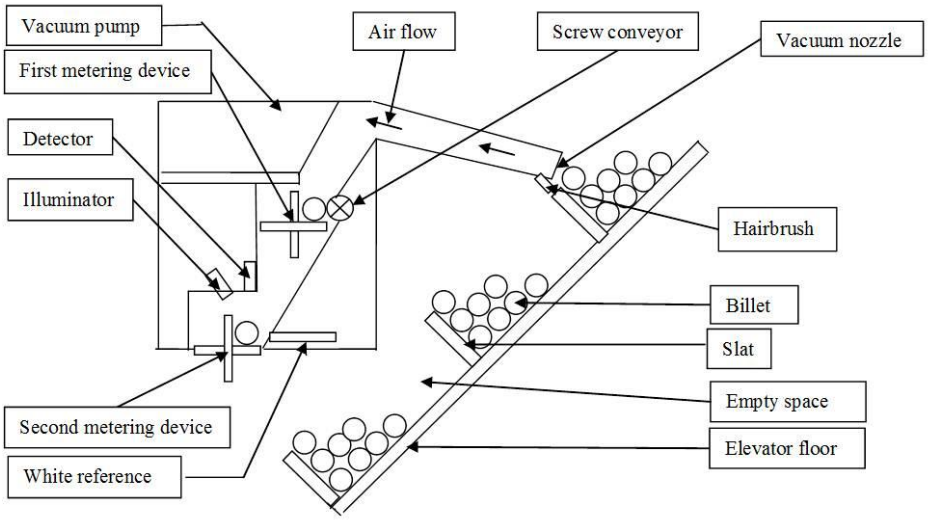


Figure 8.10: Schematic diagram of the QMSS with ASM

8.5.3.4 Sampling frequency with vacuum sampling technique

With typical elevator speed, S_e is 2.5 m^{-1}, it takes about 0.32 seconds for the nozzle to get the billet sample from each slat tip. Thus, the detector can be set to perform the measurement every 0.32 seconds. Both metering devices would also need to rotate based on this time interval. Since the measurement time, T_m and a forward speed, S_f for this measurement method are the same as the measurement method on moving billets, the number of spectra collected over one cell length would be the same (32 spectra). Technically, these 32 spectra represent 32 billet samples. If the weight of one billet sample is 300 g, 32 billets would have the weight of 9.6 kg. To calculate the sampling percentage, the total harvested billets within one cell length, C_l should be calculated. Total harvested billets, Y_d for one cell length, C_l can be calculated using equation below:

$$Y_d = \frac{Y \times C_l \times R_s}{10{,}000} \tag{8.3}$$

$$Y_d = \frac{80 \times 20 \times 1.8}{10{,}000} = 288\ kg$$

Where Y_d is the total harvested billets (kg), Y is the yield (t ha^{-1}), C_l is the cell length (m), and R_s is the row spacing (m). In this calculation, it was found that 9.6 kg of sub-samples would represent about 3.3% of the total harvested billet, Y_d within one cell length, C_l. The sub-samples number could be increased by using multiple detectors, which could also increase the accuracy of the system.

8.6 Quality mapping procedures and data processing

While the quality data is measured and recorded, the position (longitude and latitude) of the harvester is also calculated each second by the DGPS receiver. The GPS receiver

receives a differential GPS correction from the Differential GPS Correction receiver and passes the corrected GPS readings onto the laptop computer through RS-232 communication connections. Both data from spectrometer and GPS are integrated and stored at equal time intervals in the data logger.

The core of the QMSS is a software driver used to provide a communication between the computer, GPS and spectrometer. The software to operate the whole system can be written in MATLAB. Memory cards could be used to save the quality data from a data logger. The data from memory cards can later be transferred into a computer for further processing to produce quality maps. The sensor outputs should be first conditioned and then acquired by the data acquisition card in the laptop computer. By following this procedure, the quality values across the field can be mapped.

8.6.1 Spectral data processing

The collected spectra will be processed according to the procedure described in Chapter 4. In addition to Unscramber software, the algorithm for the preprocessing and multivariate analysis method could also be developed using MATLAB, eigenvector PLS toolbox and EXCEL. The pre-treated and processed spectral data in the form of predicted attribute values will be passed into specialised mapping software. This software is typically a simple Geographic Information System (GIS) specifically designed for PA use (e.g AgLink by Agris, SS Toolbox by SST Development Group, and Farmstar by Fairport Technologies Pty Ltd). This software can display the quality data for every cell as a quality map. Usually, the data requires smoothing to produce a map that is more readily analysed. The smoothing technique involves the placing a 20 m by 20 m grid over the field and calculating the average of all the quality readings in each grid cell (Cox 2002). A time delay of four seconds should be applied to the ground speed reading to account for the difference in time between when the sugarcane is cut from the ground and when it is sensed by the sensors in the elevator.

8.6.2 Outliers removal

A spectra filter could be used to eliminate strange spectra reflected by slat, trash, or rocks which result in extreme sensor readings compared to normal spectra of the billet. These spectra could be recognized by removing all spectra with an absorbance value higher than a threshold set for defined wavelengths. For example, local field outliers would be removed when the value of optical data at each measurement location is greater than twice the standard deviation from the mean at the neighboring 10 sensing points. Global field outliers that are not within three times the standard deviation from the mean of all field data are also removed (Kweon & Maxton 2013). In additions, Hotelling T^2 statistic and Mahalanobis distance methods can also be used to detect outliers.

8.7 Conclusions

The Quality Measurement System for Sugarcane (QMSS) has been theoretically designed and justified for a sugarcane harvester using a low-cost VNIR Si PDA spectroradiometer. It has been shown that this system would be able to predict the quality of the billets based on SSM by either the direct scanning method (DSM) or assisted scanning method (ASM). It has been found that the DSM could be done by scanning the moving billets on the elevator conveyor while the ASM could be performed by extracting the billet samples from the elevator conveyor using a vacuum method. The extracted billet samples are then supplied into the measurement chamber.

This theoretical discussion has also shown that the development of the proposed techniques is technically possible. The available measurement times calculated for both proposed systems (0.32 second) are within the achievable range of a commercially available detector. Thus, the specifications and strategies for the QMSS outlined in this chapter could be used by the industry as a basis for developing a low-cost spectroscopic system to measure sugarcane quality during the harvest. However, it is recommended that this proposed system be first tested in the laboratory using a proper test bed. Further studies will also need to be carried out to investigate the influence of sugarcane variety, crop density, crop moisture content and harvester speed on the performance of the system.

Chapter 9

Conclusions and recommendations

As the quality of sugarcane varies significantly across the field, there is a strong need for the development of a reliable in-field quality measurement system in the sugarcane industry. Most of the existing technologies used to measure sugarcane quality are costly and labour-intensive and can only be used under laboratory conditions, as they require complicated procedures of juice sample preparation and a stable controlled environment. Thus, this thesis aims to deal with and evaluate the most suitable sample form and a reliable sensor which can be used in the field. The primary goal of this study was to develop the new measurement methods to determine sugarcane quality from stalk samples using the spectroradiometers, which can potentially be applied in the field, especially mounted on a harvester. To meet this research goal, seven specific objectives detailed in ***Chapter 1.2*** were addressed in ***Chapters 5*** through ***Chapter 8***. This last chapter presents a summary of the major findings of the research, and offers conclusions and recommendations for future research.

9.1 Summary of findings

This thesis has provided new knowledge and insights on the application of the spectroradiometers, particularly the portable and low-cost VNIR Si PDA spectroradiometer to predict sugarcane quality from stalk samples. These new technologies have the potential to reliably measure sugarcane quality in the field, especially mounted on a harvester. This thesis has yielded fresh information on how this VNIR Si PDA spectroradiometer can be applied to measure sugarcane quality from stalk samples with an acceptable prediction accuracy that had not been possible with conventional laboratory methods.

Chapter 3 discussed the application of NIR spectroscopic technologies to measure sugarcane quality in a laboratory. Various technologies for potential uses of sugarcane in-field quality measurements were critically reviewed and assessed. The selection of chemometrics methods used to optimise the regression models between spectral data and sugar content were also investigated. The VNIRS was identified as a suitable instrument for field used. This chapter also investigated new sample forms which could be measured with relative ease in the field. The major findings of this chapter were as follows:

1. NIR spectroscopic techniques were found to be the best method for field uses since they provide rapid, non-destructive and multiple quality determination from a single measurement.
2. A portable and low-cost VNIR Si PDA spectroradiometer was found to be the best sensor candidate for field use.

3. With mounting the equipment on a harvester in mind, stalk samples were identified as the best sample forms for direct measurement in the field as they did not require complex sample preparation prior to quality measurement.

In ***Chapter 4***, the specific materials and methodologies used to investigate the ability of the spectroradiometers (VNIRS and FRS) to measure sugar content from stalk samples were detailed. This chapter characterized hardware options, suitable wavelength regions, suitable sample form, scanning position and chemometrics approaches (with regards to data pre-treatment, absorbance versus reflectance and regression techniques). The following are the major summaries:

1. The SSM and the CSSM were proposed for applications on both whole stalk and internode samples.
2. The light-proof measurement box as built to block ambient light from affecting the spectral measurements. The box was also designed to provide a consistent measurement setup for every measurement on each sample forms.

Chapter 5 investigated the application of the VNIRS on conventional samples forms: raw juice (RJ), clarified juice (CJ) and fibrated sugarcane (FS). The key findings were:

1. The VNIRS could predict the quality component of all conventional sample forms. The prediction accuracy of °Brix values was good with R^2 for FS, CJ and RJ being 0.86, 0.84 and 0.80, respectively.
2. FS was found to have the highest prediction accuracy while the RJ the lowest.
3. The quality measurement on RJ could reduce the time required for sample preparation because the clarification method could be avoided.

Chapter 6 demonstrated that both spectroradiometers (VNIRS and FRS) could be applied to predict sugarcane quality from stalk samples based on CSSM. The major findings were:

1. For the whole stalk samples, the prediction accuracy obtained by the VNIRS for pol, °Brix and CCS were 0.89, 0.83, and 0.73 respectively. The prediction accuracy obtained by the FRS for the corresponding components were 0.90, 0.86 and 0.73, respectively.
2. For the internode samples, the prediction accuracy for °Brix value was good with the R^2 and RMSEP values being 0.87 and 1.45 °Brix, respectively.
3. The RPD value for the model developed from absorbance spectral data of internode samples using the CSSM was 2, indicating that the model be used for coarse quantitative prediction purposes.
4. The internode sectional analysis showed that the °Brix variation between the node and internode areas was up to 8.7%.
5. The °Brix variation along the stalk, between the bottom and top internodes ranged between 48.6 and 93.1%, depending on variety and maturity level.
6. In comparison with the study carried out on the same samples under the same measurement setup, the prediction performance of the VNIR was found to be

comparable with that of the FRS. Both spectroradiometers predicted the °Brix values with the same accuracy (R^2 = 0.89).

Chapter 7 demonstrated the ability of both spectroradiometers (VNIRS and FRS) to predict sugarcane quality from stalk samples based on the SSM. The notable findings of this chapter were:

1. For the whole stalk samples, the VNIRS could predict the °Brix, pol and CCS with reasonable accuracy as indicated by R^2 values of 0.85, 0.72 and 0.64, respectively. The prediction accuracy obtained by the FRS for the °Brix, pol and CCS was also good with the R^2 being 0.73, 0.69 and 0.73, respectively.
2. For the internode samples, the prediction accuracy for the °Brix value using absorbance spectral data was good with the R^2 and RMSEP values being 0.87 and 1.68 °Brix, respectively. For the reflectance spectral data, the R^2 and RMSEP values were 0.91 and 1.41 °Brix, respectively.
3. The RPD value for the model developed from reflectance spectral data of internode samples using the SSM was 2, indicating that these newly proposed methods can be used for coarse quantitative prediction purposes.
4. The internode sectional analysis showed that the °Brix variation between the node and internode areas was up to 2.9%.
5. In comparison with the study carried out on the same samples under the same measurement setup, for °Brix prediction, the VNIRS gave higher prediction accuracy (R^2 = 0.88) than the FRS (R^2 = 0.80).
6. The spectral data of SSM could be classified into several quality classes using the ANN classification method with the average classification accuracy being 88.7%. This classification procedure is useful for quality mapping purposes.

Chapter 8 detailed the basic calculations and conceptual design for the new in-field quality measurement system for sugarcane (QMSS) proposed by the candidate. The major findings of this chapter were:

1. The SSM was identified as the best method to scan billet samples on the elevator of a sugarcane harvester.
2. The quality measurement by the QMSS could be done by two methods: The direct scanning method (DSM) or the assisted scanning method (ASM).
3. Development of the QMSS is technically possible as the measurement times calculated for both proposed systems (0.32 sec) within the achievable range of a commercially available detector.

9.2 Conclusions

Based on the outcomes of this research, it has been found that *"A portable and low-cost visible-shortwave near infrared spectroscopy can be used to predict sugarcane quality from solid samples or other forms".*

The application of the relatively inexpensive VNIR Si PDA spectroradiometer to measure sugarcane quality has been confirmed. The proposed measurement technology has been successfully applied to predict sugarcane quality parameters from the conventional sugarcane sample forms (raw juice, clarified juice and fibrated samples) and the newly proposed sample form (stalk samples). The stalk samples might be measured in the forms of intact stalk and internode. The quality measurement from the intact stalk may be useful for breeding programs while the internode (similar to billet) sample would be useful for on-the-go quality measurement on a harvester.

The stalk samples were measured using two new methods: SSM and CSSM. Between these two methods, SSM has been found to be the best method for field use because it does not require sample preparation prior to quality measurement. Thus, a rapid and non-destructive measurement method could potentially be developed and employed to scan billet samples on a harvester. The application of qualitative analysis to classify sugar content into several quality classes based on the SSM method has also been carried out.

Both spectroradiometers (VNIRS and FRS) used in this study have acquired spectral data that could be used to predict sugarcane quality components, including °Brix, pol, CCS, fibre content and moisture content. Out of them, °Brix has been found to be the best quality parameter to be predicted using spectroradiometers from stalk samples. The prediction accuracy obtained from the experiments using both spectroradiometers on stalk samples has been found to be comparable with similar published studies on intact fruits. Furthermore, this study has also quantified the variation of sugar content along the stalk and the internodes.

Finally, based on the findings of the above study, this thesis has also provided the basic calculations and conceptual design of the QMSS on a harvester. It has been demonstrated that the development of this proposed system is technically possible. The proposed new measurement method has the potential for field uses, especially on a harvester.

9.3 Recommendations for future works

This thesis has successfully achieved all of the objectives outlined in ***Chapter 1***. However, since this thesis was designed as a 'proof of concept' study, more work to be done before the proposed methods can be implemented for real field use. Due to issues around seasonality of the sugarcane with limited harvest season, limited availability of the spectroradiometers, limited accessibility of the laboratory facilities, this proof of concept study was not as far progressed as originally hoped. However, through the experience gained from this study, following are the recommendations for future research:

- Most of the results reported in this thesis were based on the experimental works carried out in one harvest season. However, when different crop maturing varieties are sampled, the influence of the diversity of stalk samples would need

to be further studied. It is understood that the calibration models to be used in practice should be based on large datasets, covering different sugarcane fields, climate conditions, seasons and operational conditions. Therefore, for future research, it is recommended that cross calibration using more samples across the field and harvest season (multiple years) should be undertaken.

- This study has quantified the variation of sugar content along the stalk and internode samples. However, it is suggested that a study to investigate the sugar content distribution across the cross sectional surface (variation of sugarcane within the stalk) be carried out. Moreover, since the sugarcane skin has a relatively thick rind, a study on the light penetration depth into the sugarcane skin will also be interesting. Optical sampling volume of the sugarcane skin may also be investigated.

- This study has shown that ANN could be used to classify sugarcane quality into several classes. However, the accuracy of this classification could be improved if more sophisticated classification models such as least squares support vector machine (LS-SVM) is used. Thus, it is recommended for future study to explore different types of classifier algorithms.

- The discussion on the basic calculations and conceptual design of the QMSS and sampling system was a preliminary study. To confirm the proposed concept, an actual pilot study should be carried out in a laboratory using a proper test bed equipped with a suitable spectrometer like the Carl Zeiss Corona which can be used to perform online measurement. Further studies also need to be carried out to investigate the influence of sugarcane variety, crop density, crop moisture content and harvester speed on the performance of this system. Once the laboratory experiments are completed, real investigations of the proposed concept on a harvester could then be undertaken. The sampling and scanning assessment with stalk/leaf etc flowing at 200 t hr^{-1} in a commercial harvester would be challenging. Further development of the specific VNIR spectrometer using a Si detector or other sensors such as microwave or refractometer for the sugarcane harvester is also necessary.

- Currently, it is difficult to identify effective wavelengths which have a strong relationship with sugarcane quality components. If the effective wavelengths could be identified, it could help to improve the speed and accuracy of spectral calculation for predicting sugarcane quality in the field. The effective wavelengths could be used to develop a simple, efficient, and low-cost instrument for field use. Typically, the effective wavelength can be identified by using MLR. Thus, it is recommended that future research identifies the effective wavelengths for the sugarcane quality components.

- The experiments in this thesis were carried out as a ‘proof of concept’ study to evaluate the potential application of the VNIRS to predict sugarcane quality from different sample forms with special focus on stalk samples. Thus, this

study did not attempt to determine detailed statistical relationship and statistical significance of each experiment. However, when the proposed concepts in this thesis are going to be implemented for real applications, it is recommended that the test for statistical significance to determine the probability of the measured relationship and the strength of the relationship be conducted.

REFERENCES

Abbott, J. A., Lu, R., Upchurch, B. L. & Stroshine, R. L. (1997), 'Technologies for non-destructive quality evaluation of fruits and vegetables', *Horticultural Reviews* **20**,1-120.

Abbott, J. A. (1999), 'Quality measurement of fruits and vegetables', *Postharvest Biology and Technology* **15**, 207-225.

Abdel-Rahman, E. M., Ahmed, F. B, & van den Berg M. (2010), 'Estimation of sugarcane leaf nitrogen concentration using in situ spectroscopy', *International Journal of Applied Earth Observation and Geoinformation* **12**, 52-57.

ACFA, (2009), 'About Cane' May 2009, *Australian Cane Farmers' Association (ACFA)*, Australia: **http://www.acfa.com.au.** [last accessed: 25 April 2011].

Adamchuk, V. I., Hummel, J. W., Morgan, M. T. & Upadhyaya, S. K. (2004), 'On-the-go soil sensors for precision agriculture', *Computers and Electronics in Agriculture* **44**, 71-91.

Afara, I. O. (2012), 'Near infrared spectroscopy for non-destructive evaluation of articular cartilage', PhD Thesis, Queensland University of Technology, Australia.

Arana, I., Jaren, C. & Arazuri, S. (2005), 'Maturity, variety and origin determination in white grapes (Vitis vinifera L.) using NIRS', *Journal of Near Infrared Spectroscopy* **13**, 349-357.

ASD, (1999), 'Analytical Spectral Devices: Technical Guide, 3rd Ed: A Boulder, USA: Analytical Spectral Devices, Inc.

ASD, (2005), 'Handheld spectroradiometer: User's guide version 4.05', Suite A Boulder, USA: Analytical Spectral Devices, Inc.

Barber, A. (1997), *Pneumatic handbook*, 8th Eds, Elsevier Science Ltd, Oxford: UK, pp. 387-398.

Barnes, B. J., Dhanoa, M. S. & Lister, S. J. (1989), 'Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra', *Applied Spectroscopy* **43**, 772-777.

Barton, F. E. & Kays, S. E. (2001), *Analytical application to fibrous foods and commodities*, In: Williams P, Norris K (Eds.), Near infrared technology in the agricultural and food industries. American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA. pp. 215-231.

Benjamin, C. E., Mailander, M. P. & Price, R. R. (2001), 'Sugar cane yield monitoring system', ASAE Paper No. 011189. St. Joseph, MI, USA: ASAE.

Benjamin, C. E. (2002), 'Sugar cane yield monitoring system', MSc thesis, Louisiana State University, Baton Rouge, La, USA.

Berding, N., Brotherton, G. A. & Skinner, J. C. (1991a), 'Near infrared reflectance spectroscopy for analysis of sugarcane from clonal evaluation trials: I. Fibrated cane', *Crop Sciences* **31**, 1017-1023.

Berding, N., Brotherton, G. A. & Skinner, J. C. (1991b), 'Near infrared reflectance spectroscopy for analysis of sugarcane from clonal evaluation Trials: II. Expressed juice', *Crop Sciences* **31**, 1024-1028.

Bevin, C., Staunton, S., Stobie, R., Kingston, J. & Lonergan, G. (2002), 'On-line use of near infrared spectroscopy in a sugar analysis system (SAS)', *Proceedings of the Australian Society of Sugar Cane Technologists* **24,** 1-8.

Birth, G. S. (1976), *How light interacts with foods*, In: Gaffney, J.J. Jr. (Ed.), Quality Detection in Foods. American Society for Agricultural Engineering, St. Joseph, MI, USA. Pp. 6-11.

Blanco, M. & Villarroya, I. (2002), 'NIR spectroscopy: a rapid response analytical tool', *Trends in Analytical Chemistry* **21**(4), 240-250.

Bobelyn, E., Serban, A., Nicu, M., Lammertyn, J., Bart, M., Nicolai, B. M. & Saeys, W. (2010), 'Postharvest quality of apple predicted by NIR-spectroscopy: Study of the effect of biological variability on spectra and model performance', *Postharvest Biology and Technology* **55**,133-143.

Boysworth, K. M. & Booksh, K. S (2001), *Aspects of multivariate calibration applied in near infrared spectroscopy*, in Burns, A. D. & Ciurczak, E. W. (eds), Handbook of near-infrared analysis, Marcel Dekker, New York, pp. 209-240.

Bramley, R. G. V. & Quabba, R. P. (2001), 'Opportunities for improving the management of sugarcane production through the adoption of precision agriculture - An Australian perspective', *Proceedings of the 24th Congress of the International Society of Sugar Cane Technologists*'. 17-21 September 2001, Brisbane. Pp. 38-46.

Bramley, R. G. V. (2005), 'Understanding variability in wine-grape production systems. 2. Within vineyard variation in quality over several vintages', *Australian Journal of Grape and Wine Research* **11**, 33–42.

Bramley, R. G. V., Hill, P. A., Thorburn, P. J., Kroon, F. J. & Panten, K. (2008), 'Precision Agriculture for improved environmental outcomes: Some Australian perspectives', *Landbauforschung* **58**, 161-178.

Bramley, R. G. V. (2009), 'Lessons from nearly 20 years of Precision Agriculture research, development, and adoption as a guide to its appropriate application', *Crop and Pasture Science* **60**, 197-217.

Bramley, R. G. V., Panitz, J. H., Jensen, T. & Baillie, C. (2012), 'Within block spatial variation in CCS - another potentially important consideration in the application of precision agriculture to sugarcane production', *Proceedings of the Australian Society of Sugar Cane Technologists*, **34**, 1-8.

BSES (2001), *The laboratory manual for Australian sugar mills, Volume 2. Analytical methods and tables*. BSES Limited, Australia.

Butz, P., Hofmann, C. & Tauscher, B. (2005), 'Recent developments in non-invasive techniques for fresh fruit and vegetable internal quality analysis', *Journal of Food Science* **70**(9), 131-141.

Cadet, F. D. R. & Offmann, B. (1997), 'Direct spectroscopic sucrose determination of raw sugar cane juices', *Journal of Agricultral and Food Chemistry* **45**, 166-171.

Campbell, J. A., Hansen, R. & Wilson, J. R. (1999), 'Cost effective colorimetric microtitre plate enzymatic assays for sucrose, glucose and fructose in sugarcane tissue extracts', *Journal of the Science of Food and Agriculture* **79**, 232-236.

Canegrowers (2011), *Annual Report for 2010/2011*, Brisbane, Australia.

Cargnello, R. & Fuelling, T. G. (1998), 'Cane cleaning systems', *Proceedings of the Australian Society of Sugarcane Technologists*, **20**, Ballina, NSW. pp 28–33.

Carlini, P., Massantini, R. & Mencarelli, F. (2000), 'Measurement of soluble solids in sherry and spricot by PLS regression and wavelength selection', *Journal of Agricultural and Food Chemistry* **48**, 5236-5242.

Caryn, B., Mailander, M. & Price, R. (2002), *Sugar cane yield monitoring system*, Agricultural & Biological Engineering Louisiana State University Baton Rouge, USA.

Cen, H. & He, Y. (2007), 'Theory and application of near infrared reflectance spectroscopy in determination of food quality', *Trends in Food Science and Technology* **18**, 72-83.

Chang, W. H., Chen, S. & Tsai, C. C. (1998), 'Development of a universal algorithm for use of nir in estimation of soluble solids in fruit juices', *Transactions of the ASAE* **41**(6), 1739-1745.

Chia, K. S., Rahim, H. A. & Rahim, R. A. (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* **113**, 158-165.

Ciurczak, E. W. (2001), *Principles of near infrared spectroscopy*, in A. D. Burns and E. W. Ciurczack (eds), Handbook of Near Infrared Analysis, Marcel Dekker, Inc., New York . pp. 7-18.

Cook, S. E. & Bramley, R. G. V. (1998), 'Precision agriculture – Opportunities, benefits and pitfalls', *Australian Journal of Experimental Agriculture* **38**, 753-763.

Cox, G., Harris, H., Pax, R. & Dick, R. (1996), 'Monitoring cane yield by measuring mass flow rate through the harvester', *Proceedings of the Australian Society of Sugar Cane Technologists **18***, 152-157.

Cox, G. J., Harris, H. D. & Pax, R. A. (1997), 'Development and testing of a prototype yield mapping system', *Proceedings of the Australian Society of Sugar Cane Technologists* **19**, 38-43.

Cox, G. J., Harris, H. D., Cox, D. R., Bakker, D. M., Pax, R. A. & Zillman, S. R. (1999), 'Mass flow rate sensor for sugar cane harvesters', Australian Patent No. 744,047.

Cox, G. J. (2002), 'A yield mapping system for sugar cane chopper harvesters', PhD Dissertation, Faculty of Engineering and Surveying, University of Southern Queensland, Australia.

Cox, G. J., Cox, D. R., Zillman, S. R., Simon, R., Pax, R. A., Bakker, D. M., Derk, M. & Harris, H. D. (2003), 'Mass flow rate sensor for sugar cane harvester', U.S. Patent No. 6,508,049.

Cozzolino, D., Kwiatkowski, M. J. & Warers, E. J. (2007), 'A feasibility study on the use of visible and short wavelengths in the near-infrared region for the non-destructive measurement of wine composition', *Analytical and Bioanalytical Chemistry* **387**(6), 2289-2295.

Dardenne, P. & Femenias, N. (1999), 'Diode array NIR instruments to analyse fresh forages', *Near Infrared Spectroscopy: Proceedings of the 9th International Conference*; Davies, A. M. C., Giangiacomo, R. Eds.; NIR Publications: Chichester, West Sussex, UK; pp. 121-124.

Davies, A. M. & Grant, A. (1987), 'Review: near infrared analysis of food', *International Journal of Food Science and Technology* **22**, 191-207.

Day, M. S. & Fearn, F. R. B. (1982), 'Near infrared reflectance as an analytical technique, Part 1. History and development', *Laboratory Practice* **31**, 328.

Digman, M. F. & Shinners, K. J. (2008), 'Real time moisture measurement on a forage harvester using near infrared reflectance spectroscopy', *Transactions of the ASABE* **51**, 1801-1810.

Earl, R., Wheeler, P. N., Blackmore, B. S. & Godwin, R. J. (1996), 'Precision farming - the management of variability', *Journal of Agricultural Engineering Research* **51**(4), 18-23.

Fernandes, A. C. & Benda, G. T. A. (1985), 'Distribution patterns of Brix and fibre in the primary stalk of sugarcane', *Sugar Cane (UK)* **5**, 8-11.

Fernández-Ahumada, E., Garrido-Varo, A. & Guerrero-Ginel, J. E. (2008), 'Feasibility of diode-array instruments to carry near-infrared spectroscopy from laboratory to feed process control', *Journal of Agriculture and Food Chemistry* **56**, 3185-3192.

Fernández-Novales, J., López. M,, Sánchez, M., Morales, J. & González-Caballero, V. (2009), 'Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines', *Food Research International* **42**, 285-291.

Filho, J. L. L., Pandey, P. C. & Weetal, H. H. (1996), 'An amperometric flow injection analysis enzyme sensor for sucrose using a tetracyanoquinodimethane modified graphite paste electrode', *Biosensors and Bioelectronics* **11**, 719-723.

Foley, W. J., McIlwee, A., Lawler, I., Aragones, L., Woolnough, A. P. & Berding, N. (1998), 'Ecological applications of near infrared reflectance spectroscopy - a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance', *Oecologia* **116**,293-305.

Gavin, B. M. (2008), 'Development of a single drum chopper concept for a sugarcane harvester', Master (Research) thesis, James Cook University, Queensland, Australia.

Geladi, P., MacDougall, D. & Martens, H. (1985), 'Linearization and scatter correction for near-infrared reflectance spectra of meat', *Applied Spectroscopy* **39**, 491-500.

Geladi, P. (2003), 'Chemometrics in spectroscopy. Part 1: Classical chemometrics', *Spectrochimica Acta Part B* 58, 767-782.

Gibbons, G. (2000), 'Turning a farm art into science/an overview of precision farming', URL: http://www.precisionfarming.com.

Golic, M., Walsh, K. B. & Lawson, P. (2003), 'Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature', *Applied Spectroscopy* **57**, 139-145.

Gouda, M. D., Kumar, M. A., Thakur, M. S. & Karanth, N. G. (2002), 'Enhancement of operational stability of an enzyme biosensor for glucose and sucrose using protein based stabilizing gents', *Biosensors and Bioelectronics* **7**, 503-507.

Griffin, T. W. & Lowenberg-DeBoer, J. (2005), 'Worldwide adoption and profitability of precision agriculture', *Revista de Politica Agricola* **14**, 20-38.

Griffiths, P. R. (1995), 'Letter: practical consequences of math pre-treatment of near infrared reflectance data: log (1/R) versus F(R)', *Journal of Near Infrared Spectroscopy* **3,** 60-62.

Guthrie, J. & Walsh, K. (1997), 'Non-invasive assessment of pineapple and mango fruit quality using near infrared spectroscopy', *Australian Journal of Experimental Agriculture* **37**, 253-263.

Guyer, D. & Yang, X. (2000), 'Use of genetic artificial neural networks and spectral imaging for defect detection on cherries', *Computers and Electronics in Agriculture* **29**, 179-194.

Hahn, F., Lopez, I. & Hernandez, G. (2004), 'Spectral detection and neural network discrimination of Rhizopus stolonifer spores on red tomatoes', *Biosystems Engineering* **89**, 93-99.

Harris, H. D. & Cox, G. J. (1997), 'Yield mapping - current capabilities in yield mapping for sugarcane. Are these limiting to the introduction of PA to the sugar industry?', *Precision Agriculture – What can it offer the Australian sugar industry?* Proceedings of a Workshop held in Townsville'. 10–12 June, 1997. Pp. 35-41. (CSIRO Land and Water: Townsville)

Haykin, S. (1994), *Neural network: A comprehensive foundation*, Macmillan, New York, USA.

Heacox, L. (1998), 'Precision primer', *American Vegetable Grower* **46**(6), 2-4.

Heermann, P. D. & Khazenie, N. (1992), 'Classification of multispectral remote sensing data using a back-propagation neural network', *IEEE Transactions on Geoscience and Remote Sensing* **30**(1), 81-88.

Huang, L. H., Imou, K., Kaizu, Y. & Yokoyama, S. (2005), 'System for automatic separation of sugar cane tops. *Transactions of the ASAE* **48**(4), 1441-1448.

Huang, H., Yu, H., Xu, H. & Ying, Y. (2008), 'Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review', *Journal of Food Engineering* **87**, 303-313.

Huang, M. & Lu, R. (2010), 'Optimal wavelength selection for hyperspectral scattering prediction of apple firmness and soluble solids content', *Transaction of the ASABE* **53**(4), 1175-1182.

Hsieh, C. & Lee, Y. (2005), 'Applied visible/near-infrared spectroscopy on detecting the sugar content and hardness of *pearl* guava', *Applied Engineering in Agriculture* **21**(6), 1039-1046.

Ingle, Jr, J. D. & Crouch, S. R. (1988), *Spectrochemical analysis*, Prentice-Hall, Englewood Cliff, New Jersey.

Jackson, P., Bonnett, G., Chudleigh, P., Hogarth, M. & Wood, A. (2000), 'The relative importance of cane yield and traits affecting CCS in sugarcane varieties', *Proceedings of the Australian Society of Sugar Cane Technologists* **22**, 23-29.

Jensen, T., Baillie, C., Bramley, G. V., DiBella, L., Whiteing, C. & Davis, R. (2010), 'Assessment of sugarcane yield monitoring technology for precision agriculture', *Proceedings of the Australian Society of Sugar Cane Technologists* **32**, 410-423.

Jensen, T. A., Baillie, C., Bramley, R. G. V., Panitz, J. H. & Schroeder, B. L. (2012), 'An assessment of sugarcane yield monitoring concepts and techniques from commercial yield monitoring systems', *Proceedings of the Australian Society of Sugar Cane Technologists* **34**, 8-15.

Jhoty, I. & Autre, J. C. (2003), 'Precision agriculture: Perspectives for the Mauritian sugar industry', *Proceeding of International Society of Sugar Cane Technologists: Agronomy Workshop* **24, 47-50, Quatre-Bornes, Mauritius.**

Jhoty, I., Ramasamy, S., Autrey, L. J. C. & Blackmore, S. (2003), 'Yield variability in sugar cane', *Proceedings of the 6th International Conference on Precision Agriculture and Other Precision Resources Management*. Minneapolis, Minnesota, 14–17 July 2002. Pp.141-150. (ASA-CSA-SSSA: Madison, WI).

Johnson, R. M. & Richard, Jr, E. P. (2005), 'Sugarcane yield, sugarcane quality, and soil variability in Louisiana', *Agronomy Journal* **97**, 760-771.

Jones, J. N., Schofield, W., McKenzie, N. J. & Olson, B. C. (2002), 'Factory rate control based on more constant fibre rate', *Proceedings of the Australian Society of Sugar Cane Technologists* **24**, 362-366.

Jong, S. (1993), 'PLS fits closer than PCR', *Journal of Chemometrics* **7**, 551-557.

Joyce, K. N. & Edwards, B. P. (1994), 'Modelling the trajectories in pneumatic cane cleaning', *Proceedings of Australian Society of Sugarcane Technologists* **16**, 257-263.

Kawano, S., Fujiwara, T. & Iwamoto, M. (1993), 'Non-destructive determination of sugar content in Satsuma mandarin using near infrared (NIR) transmittance'. *Journal of the Japanese Society of Horticultural Science* **62**, 465–470.

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

Kawano, S. (2002), *Application to agricultural products and foodstuffs*, In: Near Infrared Spectroscopy, Principles, Instruments, Application. (Eds: Siesler H., Ozaki Y., Kawata S. and Heise H.). Wiley-VCH, Chichester, UK, pp. 269-287.

Kennedy, J. F., Pimentel, M. C. B., Melo Eduardo, H. M. & Lima-Filho, J. L. (2007), 'Sucrose biosensor as an alternative tool for sugarcane field samples', *Journal of the Science of Food and Agriculture* **87**, 2266-2271.

Khuriyati, N. & Matsuoka, T. (2004), 'Near infrared transmittance method for nondestructive determination of soluble solid content in growing tomato fruits', *Environmental Control in Biology* **42**, 217-223.

Kilickan, A. & Güner, M. (2006), 'Pneumatic conveying characteristics of cotton seeds', *Biosystems Engineering* **95**, 537-546.

Kim, J., Mowat, A., Poole, P. & Kasabov, N. (2000), 'Linear and non-linear pattern recognition models for classification of fruit from visible-near infrared spectra', *Chemometrics and Intelligent Laboratory Systems* **51**, 201-216.

Kingston, G. & Hyde, R. E. (1995), 'Intra-field variation of commercial cane sugar (CCS) values', *Proceedings of the Australian Society of Sugar Cane Technologists* **17**, 30-38.

Kingston, G. (2002), 'Recognising the impact of climate on CCS of sugarcane across tropical and sub-tropical regions of the Australian Sugar industry', *Proceedings of the Australian Society of Sugar Cane Technologists **24***. (CD-ROM).

Klute, U. (2007), 'Microwave measuring technology for the sugar industry', *International Sugar Journal* **109**(1308), 1-6.

Kroes, S. (1997), 'The cutting of sugarcane', PhD Dissertation, Faculty of Engineering and Surveying, University of Southern Queensland, Australia.

Kumar, L., Schmidt, K., Dury, S. & Skidmore, A. (2003), *Imaging spectrometry and vegetation science*. In: van der Meer, F.D., de Jong, S.M. (Eds.), Image Spectrometry, vol. 3. Kluwer Academic Publishers, London, UK, pp. 111-156.

Kumar, A. J., Gowri, N. M., Raju, R. V., Nirmala, G., Bellubbi, B. S. & Krishna, T. R. (2006), 'Study of fiber optic sugar sensor', *Pramana-Journal of Physics* **67**(2), 383-387.

Kweon, G. & Maxton, C. (2013), 'Soil organic matter sensing with an on-the-go optical sensor', *Biosystems Engineering* **115**, 66-81.

Lammertyn, J., Nicolai, B., Ooms, K., Smedt, V. D. & Baerdemaeker, J. D. (1998), 'Non–destructive measurement of acidity, soluble solids, and firmness of Jonagold apples using NIR-spectroscopy', *Transaction of the ASAE* **41**(4), 1089-1094.

Lawes, R. A., Wegener, M. K., Basford, K. E. & Lawn, R. J. (2002), 'Commercial cane sugar trends in the Tully sugar district', *Australian Journal of Experimental Agriculture* 40, 969-973.

Lawes, R. A., Wegener, M. K., Basford, K. E. & Lawn, R. J. (2003), 'A multivariate evaluation of sugarcane farm performance over time based on yield and commercial cane sugar', *Proceedings of the 11th Australian Agronomy Conference: Solutions for a better environment*, 2-6 Feb. 2003, Geelong, Victoria. (Australian Society of Agronomy: www.regional.org.au/au/asa/2003/c/15/lawes.html)

Lawes, R. A. & Lawn, R. J. (2005), 'Applications of industry information in sugarcane production systems', *Field Crops Research* **92**, 353-363.

Lee, W. S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D. & Li, C. (2010), 'Sensing technologies for precision specialty crop production', *Computers and Electronics in Agriculture* **74**, 2-33.

Lehnert, R. (2010), 'Vacuum harvester passes bruising tests', *Good Fruit Grower*.

Leslie, J. K. & Wilson, G. L. (1996), *Productivity trends in sugarcane in the wet tropics*, Sugar Research and Development Corporation (SRDC) and CRC for Sustainable Sugar Production Technical Report 1/96. SRDC, Brisbane.

Li, X., He, Y. & Fang, H. (2007), 'Non-destructive discrimination of Chinese bayberry varieties using Vis/NIR spectroscopy', *Journal of Food Engineering* **81**, 357-363.

Lindström, H., Malinen, J. & Marbach, R. (2004), *Performance evaluation of standard and extended InGaAs detector array spectrometers*. In Near Infrared Spectroscopy: Stretching the NIR Spectrum to the Limit; Davies, A. M. C., Garrido-Varo, A., Eds.; NIR Publications: Chichester, West Sussex, UK, 99-104.

Liu, F., He, Y. & Wang, L. (2008), 'Comparison of calibrations for the determination of soluble solids content and pH of rice vinegars using visible and short-wave near infrared spectroscopy', *Analytica Chimica Acta* **610**, 196-204.

Long, J. P. & Buckmaster, D. R. (2003), 'Development of an automated system for sampling crop material from a forage harvester', *Applied Engineering in Agriculture* **19**(2), 133-140.

Lu, R., Guyer, D. E. & Beaudry, R. M. (2000), 'Determination of firmness and sugar content of apples using near-infrared diffuse reflectance', *Journal of Texture Studies* **31**, 615-630.

Lu, R. & Ariana, D. (2002), 'A near infrared sensing technique for measuring internal quality of apple fruit', *Applied Engineering in Agriculture* **18**(5), 585-590.

Lu, R. (2004), 'Multispectral imaging for predicting firmness and soluble solids content of apple fruit', *Postharvest Biology and Technology* **31**, 147-157.

Madsen, L. R., White, B. E. & Rein, P. W. (2003), 'Evaluation of a near infrared spectrometer for the direct analysis of sugar cane', *Journal of American Society of Sugar Cane Technologists* **23**, 80-92.

Magalhães, P. S. G. & Cerri. D. G. P. (2007), 'Yield monitoring of sugar cane', *Biosystems Engineering* **96** (1), 1-6.

Mailander, M., Benjamin, C., Price, R. & Hall, S. (2010), 'Sugar cane yield monitoring system', *Applied Engineering in Agriculture* **26**(6), 965-969.

Mallarino, A. P. & Wittry, D. J. (2004), 'Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH and organic matter', *Precision Agriculture* **5**, 131-144.

Manley, M., Joubert, E., Myburgh, L., Lotz, E. & Kidd, M. (2007), 'Prediction of soluble solids content and post-storage internal quality of Bulida apricots using near infrared spectroscopy', *Journal of Near Infrared Spectroscopy* **15**(3), 179-188.

Marcotte, D., Savoie, P., Martel, H. & Theriault, R. (1999), 'Precision agriculture for hay and forage crops: a review of sensors and potential applications', *American Society of Agricultural Engineers*, St. Joseph, MI, USA. ASAE Paper No. 99-1049.

Mark, H. & Workman, J. J. (2004), 'Derivatives in spectroscopy Part IV-Calibrating with derivatives', *Spectroscopy* **1**, 44-51.

Martens, H. & Næs, T. (1991), *Multivariate calibration*, Wiley, New York.

Martinsen, P. & Schaare, P. (1998), 'Measuring soluble solids distribution in kiwifruit using near infrared imaging spectroscopy', *Postharvest Biology and Technology* **14**, 271-281.

Massart, D. L., Vandenginste, B. G. M., Deming, S. N., Michotte, Y. & Kaufmann, L. (1988), *Chemometrics: A Textbook*, Elsevier, Amsterdam.

McBratney, A. B. & Pringle, M. J. (1999), 'Estimating average and proportional variograms of soil properties and their potential use in precision agriculture', *Precision Agriculture* **1**, 125-152.

McBratney, A., Whelan, B., Ancev, T. & Bouma, J. (2005), 'Future directions of Precision Agriculture', *Precision Agriculture* **6**, 7-23.

McCaig, T. N. (2002), 'Extending the use of visible/near infrared reflectance spectrophotometers to measure colour of food and agricultural products', *Food Research International* **35**, 731-736.

McCarthy, S. G. (2003), 'The integration of sensory control for a sugar cane harvester', PhD Thesis, Faculty of Engineering and Surveying, University of Southern Queensland, Australia.

McCarthy, S. G. & Billingsley, J. (2002), 'A sensor for the sugar cane harvester topper', *Sensor Review* 22, 242-246.

McGlone, V. A., Abe, H. & Kawano, S. (1997), 'Kiwifruit firmness by near infrared light scattering', *Journal of Near Infrared Spectroscopy* **5**, 83-89.

McGlone, V. A., Jordan, R. B. & Martinsen, P. J. (2002), 'Vis/NIR estimation at harvest of pre- and post-storage quality indices for 'Royal Gala' apple', *Postharvest Biology and Technology* **25**, 135-144.

McRae, T. A., Bull, J. K., Robotham, B. G. & Sweetnam, R. C. (1996), 'Measuring sugar content in variety trials', *Sugar 2000 symposium, Sugarcane: Research Towards Efficient and Sustainable Production.* CSIRO Division of Tropical Crops and Pastures, Brisbane. Pp. 55-56.

Meade, G. P. & Chen, J. C. P. (1985), *Sugar Cane Handbook*, 11th ed., Wiley: New York.

Mehrotra, R. & Siesler, H. W. (2003), 'Application of mid infrared/near infrared spectroscopy in sugar industry', *Applied Spectroscopy Reviews* **38**, 307-354.

Meyer, J. H. & Wood, R. A. (1988), 'Rapid analysis of cane juice by near infrared reflectance', *Proceeding of the South African Sugar Technologists Association* **62**, 203-207.

Molin, J. P. & Menegatti, L. A. (2004), 'Field-testing of a sugar cane yield monitor in Brazil', ASAE Paper No. 041099. St. Joseph, MI, USA.: ASAE.

Montalvo, Jr. J. G., Buco, S. E. & Ramey, Jr. H. H. (1994), 'Studies to measure cotton fiber length, strength, micronaire, and color by vis/NIR reflectance spectroscopy', Part II: Principal components regression', *Journal of Near Infrared Spectroscopy* **2**(4), 185-198.

Montes, J. M., Utz, H. F., Schipprack, W., Kusterer, B., Muminovic, J., Paul, C. & Melchinger, A. E. (2006), 'Near-infrared spectroscopy on combine harvesters to measure maize grain dry matter content and quality parameters', *Plant Breeding* 125, 591-595.

Moons, E., Dardenne, P., Dubois, A. & Sindic, M. (1997), 'Nondestructive visible and NIR spectroscopy for the determination of internal quality in apple', *Proceedings of Sensors for Nondestructive Testing International Conference*, pp. 122-131. Orlando, Fla., 18-21 Feb. Ithaca, N.Y.: Natural Resource, Agriculture, and Engineering Service.

Miller, W. M. & Zude-Sasse, M. (2004), 'NIR-based sensing to measure soluble solids content of florida citrus', *Applied Engineering in Agriculture* **20**, 321-327.

Moghimi, A., Aghkhani, M. A., Sazgarnia, A. & Sarmad, M. (2010), 'Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit', *Biosystems Engineering* **106**(3), 295-302.

Mohan, L. A., Karunakaran, C., Jayas, D. S. & White, N. D. G. (2005), 'Classification of bulk cereals using visible and NIR reflectance characteristics', *Canadian Biosystems Engineering* **47**(7), 7-14.

Montes, J. M., Utz, H. F., Schipprack, W., Kusterer, B., Muminovic, J., Paul, C. & Melchinger, A. E. (2006), 'Near-infrared spectroscopy on combine harvesters to measure maize grain dry matter content and quality parameters', *Plant Breeding* **125**, 591-595.

Mouazen, A. M., Baerdemaeker, J. D. & Ramon, H. (2005), 'Towards development of online soil moisture content sensor using a fibre-type NIR spectrophotometer', *Soil and Tillage Research* **80**, 171-183.

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

Mutanga, O., Skidmore, A. K. & Prins, H. H. T. (2004), 'Predicting in-situ pasture quality in Kruger National Park, South Africa, using continuum-removed absorption features', *Remote Sensing of Environment* **89**, 393-408.

Nelson, S. O. (1987), 'Potential agricultural applications for RF and microwave energy', *Transactions of the ASAE* **30**(3), 818-831.

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

Nicolaï, B. M., Verlinden, B. E., Desmet, M., Saevels, S., Saeys, W. & Theron, K. (2008), 'Time-resolved and continuous wave NIR reflectance spectroscopy to predict soluble solids content and firmness of pear', *Postharvest Biology and Technology* **47**, 68-74.

Næs, T., Isaksson, T., Fearn, T. & Davies, T. (2004), *A user-friendly guide to multivariate calibration and classification*, Charlton, Chichester, UK: NIR Publications.

Norris, K. H. (1964), 'Design and development of a new moisture meter', *Agricultural Engineering* **45**, 370.

Oriel, (1997), *Oriel Corporation light sources, monochromators and spectrographs, detectors and detection systems, fiber optics (catalogue)*, Oriel Corporation: Stratford, CT, USA.

Osborne, B. G., Fearn, T. & Hindle, P. H. (1993), *Theory of near infrared spectrophotometry*, In 'Practical NIR Spectroscopy with Applications in Food and Beverage Analysis'. Ed. D Browning . Pp. 13-35. Longman Singapore Publishers Ltd.

O'Shea, M. G., Staunton, S. P., Donald, D. & Simpson, J. (2011), 'Developing laboratory near infrared (NIR) instruments for the analysis of sugar factory products', *Proceedings of the Australian Society of Sugar Cane Technologists* **33**, 1-8.

Pagnano, N. B. & Magalhães, P. S. (2001), 'Sugarcane yield measurement', *Proceedings of the 3rd European Conference on Precision Agriculture*, 839-844. S Blackmore and G. Grenier, eds. Montpellier, France: 18-20 June 2001, AgroMontpellier-ENITA de Bordeaux.

Panigrahi, S. & Hofman, V. (2003), 'On-the-go sugar sensor for determining sugar content during harvesting', U.S. Patent No. 6,624,888.

Park, B., Abbott, J. A., Lee, K. J., Choi, C. H. & Choi, K. H. (2003), 'Near infrared diffuse reflectance for quantitative and qualitative measurement of soluble solids and firmness of delicious and gala apples', *Transactions of the ASAE* **46**(6), 1721-1731.

Paul, C. & Pfitzner, C. (2004), *Analytical use of NIR diode array spectrometers on forage harvesters*, In Near Infrared Spectroscopy: Stretching the NIR Spectrum to the Limit; Davies, A. M. C., Garrido-Varo, A., Eds.; NIR Publications: Chichester, West Sussex, UK, pp. 333–338.

Peirs, A., Lammertyn, J., Ooms, K. & Nicolaï, B. M. (2001), 'Prediction of the optimal picking date of different apple cultivars by means of VIS: NIR-spectroscopy', *Postharvest Biology and Technology* **21**, 189-199.

Pierossi, M. A. & Hassuani, S. J. (1997), 'Cac¸amba instrumentada para pesagem de cana picada [Instrumented infield truck for sugar cane billet weighing]', Semina´ rio Copersucar de Tecnologia Agronoˆ mica. **7**, 263-265.

Pollock, J. S., O'Hara, I. M. & Griffin, K. G. (2007), 'Aligning the drivers in the value chain - a new cane payment system for Mackay Sugar', *Proceedings of the Australian Society of Sugar Cane Technologists* **29**, 1-8.

Pope, G., McDowall, R., Massey, W. & Staunton, S. (2004), 'The use of NIR spectroscopy in a cane quality incentive scheme', *Proceedings of the Australian Society of Sugar Cane Technologists* **26**, (CD-ROM).

Price, R. R., Johnson, R. M., Viator, R. P., Larsen, J. & Peters, A. (2011), 'Fiber optic yield monitor for a sugarcane harvester', *Transactions of the ASABE* **54**(1), 31-39.

Pringle, M. J., McBratney, A. B., Whelan, B. M. & Taylor, J. A. (2003), 'A preliminary approach to assessing the opportunity for site-specific crop management in a field, using yield monitor data', *Agricultural Systems* **76**, 273-292.

Purcell, D. E., Leonard, G. J., O'Shea, M. G. & Kokot, S. (2005), 'A chemometrics investigation of sugarcane plant properties based on the molecular composition of epicuticular wax', *Chemometrics and Intelligent Laboratory Systems* **76**, 135-147.

Purcell D. E., O'Shea, M. G. & Kokot, S. (2009), 'Complex biopolymeric systems at stalk/epicuticular wax plant interfaces: A near infrared spectroscopy study of the sugarcane', *Biopolymers* **91**(8), 642-651.

Qing, Z., Ji, B. & Zude, M. (2007), 'Predicting soluble solid content and firmness in apple fruit by means of laser light backscattering image analysis', *Journal of Food Engineering* **82**, 58-67.

Qiu, P. Y., Ding, H. B., Tang, Y. K. & Xu, R. J. (1999), 'Determination of chemical composition of commercial honey by near-infrared spectroscopy', *Journal of Agriculture and Food Chemistry* **47**, 2760-2765.

Quilitzsch, R., Baranska, M., Schulz, H. & Hoberg, E. (2005), 'Fast determination of carrot quality by spectroscopy methods in the UV-VIS, NIR and IR range', *Journal of Applied Botany and Food Quality* **79**, 163-167.

Rattey, A. R., Jackson, P. A., Hogarth, D. M. & McRae, T. A. (2009), 'Selection among genotypes in final stage sugarcane trials: effects of time of year', *Crop and Pasture Science* **60**, 1165-1174.

Reyns, P., Missotten, B. Ramon, H. & De Baerdemaeker, J. (2002), 'A review of combine sensors for precision farming', *Precision Agriculture* **3**, 169-182.

Richard, C., Jackson, W. & Waguespack, Jr. H. (1996), 'The Louisiana whole-stalk, soldier harvesting system and innovation/modifications to improve its efficiency', *Proceedings of the International Society of Sugar Cane Technologists*. **23**(2), 196-200.

Richard, C., Jackson, W. & Waguespack, Jr. H. (2001), 'Harvester trials and extraneous matter in the Louisiana sugar industry', *Proceedings of the International Society of Sugar Cane Technologists* **24**(2), 263-268.

Robotham, B. G. (2000), Production of an automated cane billet sampler for research trials, Sugar Research and Development Corporation (SRDC) - Project BSS156, Australia.

Runge, E. C. A. & Hons, F. M. (1999), 'Precision Agriculture - Development of a heirachy of variables influencing crop yields', *Proceedings of the 4th International Conference on Precision Agriculture*'. St Paul, Minnesota, 19–22 July 1998. Part A. (Eds PC Robert, RH Rust, WE Larson) pp. 143–158. (ASA-CSA-SSSA: Madison, WI)

Saranwong, S., Sornsrivichai, J. & Kawano, S. (2003), 'Performance of a portable near infrared instrument for Brix value determination of intact mango fruit', *Journal of Near Infrared Spectroscopy* **11**, 175-181.

Sassenrath, G. F., Adams, E. R. & Williford, J. R. (2005), 'Rapid sampling system for determination of cotton fiber quality spatial variability', *Applied Engineering in Agriculture* **21**(1), 9-14.

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

Schäffler, K. J. & De Gaye, M. T. D. (1997), 'Rapid near infrared estimation of multi-components in mixed juice and final molasses: the possibility of day-to-day control of raw sugar factories using NIR', *Proceedings of the International Society of Sugar Cane Technologists* **71**, 153-160.

Schellberg, J., Hill, M. J., Gerhards, R., Rothmund, M. & Braun, M. (2008), 'Precision agriculture on grassland: Applications, perspectives and constraints', *European Journal of Agronomy* **29**, 59-71.

Schembri, M. G. & Harris, H. D. (1998), 'Examining the effectiveness of smashing sane by impact loading in the shredder', *Proceedings of the Australian Society of Sugar Cane Technologists* **20**, 340-348.

Schroeder, B. L., Wood, A. W., Park, G., Panitz, J. H. & Stewart, R. L. (2009), 'Validating the 'Six Easy Steps' nutrient management guidelines in the Johnstone Catchment', *Proceedings of the Australian Society of Sugar Cane Technologists* **31**, 177-185.

Schupp, J., Baugher, T., Winzeler, E. & Schupp, M. (2011), 'Preliminary results with a vacuum assisted harvest system for apples', *Fruit Notes* **76**(4), 1-5.

Shah, S. & Joshi, M. (2010), 'Modeling microwave drying kinetics of sugarcane bagasse', *International Journal of Electronics Engineering* **2**(1), 159-163.

Shao, Y., He, Y., Gomez, A. H., Pereir, A. G., Qiu, Z. & Zhag, Y. (2007), 'Visible/near infrared spectrometric technique for nondestructive assessment of tomato "Heatwave" (*Lycopersicum esculentum*) quality characteristics', *Journal of Food Engineering* **79**(3), 1015-1019.

Shenk, J. S. & Westerhaus, M. O. (1996), *Calibration the ISI way*, In A. M. C. Davies, & P. C. Williams (Eds.), Near infrared spectroscopy: The future waves, pp. 198-202. Chichester: NIR Publications.

Shenk, J. S. (2004), 'Are we there yet?', *In Contents of 2nd International Conference on Embedded Near Infrared Spectroscopy*, Gembloux, Belgium, 18-19th November, 2004.

Sherman, H. C. P. (1997), *Infrared spectroscopy*, in F. A. Settle (Ed), Handbook of instrumental techniques for Analytical Chemistry, Prentice-Hall.

Simpson, R. & Naidoo, Y. (2010), 'Progress in improving laboratory efficiencies using near infrared spectroscopy (NIRS)', *Proceedings of the International Society of Sugar Cane Technologists* **27**, 1-8.

Song, H. & Litchfield, J. B. (1991), 'Predicting method of terminal velocity for grains', *Transactions of the ASAE* **34**, 225-231.

Stafford, A. (1999), 'Some sources of error in block productivity data. In: Keating, B.A., Lawes, R.A. (Eds.), Applications of Productivity Data within the Australian Sugar Industry. CRC for Sustainable Sugar Production Occasional Publication. CRC Sugar, James Cook University, Townsville, Pp. 9-11.

Staunton, S. P., Lethbridge, P. J., Grimley, S. C., Streamer, R. W., Rogers, J. & Mackintosh, D. L. (1999), 'Online cane analysis by near infrared spectroscopy', *Proceedings of the Australian Society of Sugar Cane Technologists* **21**, 20–27.

Staunton, S. P. & Wardrop, K. (2006), 'Development of an on-line bagasse analysis system using NIR spectroscopy', *Proceedings of the Australian Society of Sugar Cane Technologists* **28**, 446–453.

Staunton, S., Donald, D. & Pope, G. (2011), 'Estimating sugarcane composition using ternary growth relationships', *Proceedings of the Australian Society of Sugar Cane Technologists* **33**. Mackay, Australia.

Sugiyama, J. (1999), 'Visualization of sugar content in the flesh of a melon by near-infrared imaging', *Journal of Agriculture and Food Chemistry* **47**, 2715-2718.

Swierenga, H., De Weijer, A. P., Van Wijk, R. J. & Buydens, L. M. C. (1999), 'Strategy for constructing robust multivariate calibration models', *Chemometrics and Intelligent Laboratory Systems* **49**, 1-17.

Sukhchain, S. D. & Saini, G. S. (1997), 'Inter-relationships among cane yield and commercial cane sugar and their component traits in autumn plant crop of sugarcane', *Euphytica* **95**, 109-114.

Tai, P. Y. P., He, H., Gan, H. & Miller, J. D. (1992), 'Variation for juice quality and fiber content in crosses between commercial sugarcane and *saccharum spontaneum*', *Journal of American Society of Sugar Cane Technologists* **12**, 47-57.

Taira, E., Ueno, M. & Kawamitsu, Y. (2010), 'Automated quality evaluation system for net and gross sugarcane samples using near infrared spectroscopy', *Journal of Near Infrared Spectroscopy* **18**, 209-215.

Taira, E., Ueno, M., Saengprachatanarug, K. & Kawamitsu, Y. (2013), 'Direct sugar content analysis for whole stalk sugarcane using a portable near infrared instrument', *Journal of Near Infrared Spectroscopy* **21**, 281-287.

Taylor, J. A., Whelan, B. M., Thylén, L., Gilbertsson, M. & Hassall, J. (2005), 'Monitoring wheat protein content on-harvester: Australian experiences', *Proceedings of the 5th European Conference on Precision Agriculture*, Precision Agriculture; Stafford, JV, ed.): Wageningen Academic Publishers, The Netherlands.

Taylor, J. A., McBratney, A. B. & Whelan, B. M. (2007), 'Establishing management classes for broadacre agricultural production', *Agronomy Journal* **99**, 1366-1376.

Temma, T., Hanamatsu, K. & Smnoki, F. (2002), 'Measuring the sugar content of apples and apple juice by near infrared spectroscopy', *Optical Review* **9**(2), 40-44.

Torrecilla, J. S., Otero, L. & Sanz, P. D. (2004), 'A neural network approach for thermal/pressure food processing', *Journal of Food Engineering* **62**, 89-95.

Valderrama, P., Braga, J. W. B. & Poppi, R. J. (2007), 'Validation of multivariate calibration models in the determination of sugar cane quality parameters by near infrared spectroscopy', *Journal of Brazilian Chemical Society* **18**, 259-266.

Ventura, M., De Jager, A., de Putter, H. & Roelofs, F. P. M. M. (1998), 'Nondestructive determination of soluble solids in apple fruit by near infrared spectroscopy (NIRS)', *Postharvest Biology and Technology* **14**, 21-27.

Von Rosenberg, C., Abbate, A., Drake, J. & Mayes, D. (2000), 'A rugged near-infrared spectrometer for real-time measurements of grains during harvest', *Spectroscopy* **15**, 34-38.

Walsh, K. B., Guthrie, J. A. & Burney, J. W. (2000), 'Application of commercially available, low-cost, miniaturised NIR spectrometers to the assessment of the sugar content of intact fruit', *Australian Journal of Plant Physiology* **27**, 1175-1186.

Watson, L. J., Williams, N. D. & Staunton, S. P. (1999), 'Fibre classification using online near infrared spectroscopy', *Proceedings of the Australian Society of Sugar Cane Technologists* **21**, 401-405.

Welle, R., Greten, W., Rietmann, B., Alley, S., Sinnaeve, G. & Dardenne, P. (2003), 'Near-infrared spectroscopy on chopper to measure maize forage quality parameters online', *Crop Science* **43**(4), 407-413.

Wendte, K. W., Skotnikov, A. & Thomas, K. K. (2001), 'Sugar cane yield monitor', U.S. Patent No. 6,272,819.

Wold, S., Esbensen, K. & Geladi, P. (1987), 'Principal component analysis', *Chemometrics and Intelligent Laboratory Systems* **2**, 37-52.

Wold, S., Sjöström, M. & Eriksson, L. (2001), 'PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**, 109-130.

Williams, P. C. & Norris, K. H. (1987), *Qualitative applications of near infrared reflectance spectroscopy*, In: Near-lnfrared Technology in the Agricultural and Food Industries. (Williams, P.C. & Norris, K.H., eds.), St. Paul, MN: Am. Assoc. Cereal Chemists (AACC).

Williams, P. C. & Norris, K. H. (2001), *Near-infrared technology in the agricultural and food industries*, (2nd Ed.). St. Paul, MN: AACC.

Williams, P (2008), *Near-Infrared Technology - Getting the Best out of Light*. A Short Course in the Practical Implementation of Near-Infrared Spectroscopy for the User". Edn. 5.3, PDK Projects, Inc. Nanaimo, Canada.

Wright, S., Brumb, S., Niebur, T. & Welle, R. (2002), 'Near-infrared spectrometry for real-time analysis of substances', U.S. Patent No. 6,483,583.

Wrigley, T. & Moore, S. (2006), 'Public Environment Report 2006', (*Canegrowers*: Brisbane).

Wu, D., Feng, L., Zhang, C. & He, Y. (2008), 'Early detection of Botrytis Cinerea on eggplant leaves based on visible and near infrared spectroscopy', *Transactions of the ASABE* **51**(3), 1133-1139.

Xiaobo, Z., Jiewen, Z., Xingyi, H. & Yanxiao, L. (2007), 'Use of FT- NIR spectrometry in non-invasive measurements of soluble solid contents (SSC) of "Fuji" apple based on different PLS models', *Chemometrics and Intelligent Laboratory Systems* **87**, 43-51.

Xing, J., Symons, S., Shahin, M. & Hatcher, D. (2010), 'Detection of sprout damage in Canada Western Red Spring wheat with multiple wavebands using visible/near-infrared hyperspectral imaging', *Biosystems Engineering* **106**, 188-194.

Yan, Y. L., Zhao, L. L., Han, D. H. & Yang, S. M. (2005), 'Basement and application of NIR analysis', Beijing: Light Industry Company in China.

Yu, H. Y., Niu, X. Y., Lin, H. J., Ying, Y. B., Li, B. B. & Pan, X. X. (2009), 'A feasibility study on online determination of rice wine composition by Vis–NIR spectroscopy and least-squares support vector machines', *Food Chemistry* **113**, 291-296.

Zhang, N., Wang, M. & Wang, N. (2002), 'Precision agriculture - a worldwide overview', *Computers and Electronics in Agriculture* **36**, 113-132.

Zude, M., Herold, B., Roger, J. M., Bellon-Maurel, V. & Landahl, S. (2006), 'Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life', *Journal of Food Engineering* **77**, 254-260.

Appendix A

MATLAB source codes for classification algorithm based on artificial neural networks (ANN)

```
% neural network classifier

function argout = neurc (a,units)

    prtrace(mfilename);

    mapname = 'Automatic Neural Classifier';
    n_attempts = 3;% Attempting three different random initialisations

    if (nargin < 2)
        units = [];
    end
    if (nargin < 1) | (isempty(a))
        w = mapping(mfilename,units);
        argout = setname(w,mapname);
        return
    end

    [m,k] = size(a);
    if isempty(units)
        cs = classsizes(a);
        units = ceil(0.2*min(cs));
    end

    if (~ismapping(units))

    if isnan(units) % optimize complexity parameter: number of neurons
            defs = {[]};
            parmin_max = [1,30];
            w                                                    =
regoptc(a,mfilename,{units},defs,[1],parmin_max,testc([],'soft'),0);
            return
    end

        islabtype(a,'crisp');
        isvaldfile(a,1,2); % at least 1 object per class, 2 classes
        a = testdatasize(a);
        %a = setprior(a,getprior(a));

        % Second parameter is not a mapping: train a network.
        % Reproducability: always use same seeds.
        rand('seed',1); randn('seed',1); opt_err = inf; opt_mapping =
[];

        % Try a number of random initialisations.
        s = sprintf('%i neural network initializations: ',n_attempts);
        prwaitbar(n_attempts,s);
        for attempt = 1:n_attempts
```

```
            prwaitbar(n_attempts,attempt,[s int2str(attempt)]);
            prwarning(4,'training    with    initialisation    %d    of
%d',attempt,n_attempts);
            t  =  gendatk(a,1000,2,1);  %  Create  tuning  set  based  on
training set.
            w = lmnc(a,units,inf,[],t); % Find LMNC mapping.
            e = t*w*testc;        % Calculate classification error.
            if (e < opt_err)
                % If this is the best of the three repetitions, store
it.
                opt_mapping = w; opt_err = e;
            end
        end
        prwaitbar(0);

        % Output is best network found.
        argout = setname(opt_mapping,mapname);

    else

        nodatafile(a);

        % Second parameter is a mapping: execute.
        w = units;

        data = getdata(w);

        if (length(data) > 1)

        % "Old" neural network - network is second parameter: unpack.
      data = getdata(w); weights = data{1};
        pars = data{2}; numlayers = length(pars);

        output = a;   % Output of first layer: dataset.
      for j = 1:numlayers-1
          % Number of inputs (n_in) and outputs (n_out) of neurons in
layer J.
        n_in = pars(j); n_out = pars(j+1);

          % Calculate output of layer J+1. Note that WEIGHTS contains
both
        % weights (multiplied by previous layer's OUTPUT) and biases
          % (multiplied by ONES).

        this_weights                                                    =
reshape(weights(1:(n_in+1)*n_out),n_in+1,n_out);
          output = sigm([output,ones(m,1)]*this_weights);

        % Remove weights of this layer.
          weights(1:(n_in+1)*n_out) = [];
        end
        else
            % "New" neural network: unpack and simulate.
```

```
            net = data{1};
            output = sim(net,+a')';
        end;

        % 2-class case, therefore 1 output: 2nd output is 1-1st
output.
        if (size(output,2) == 1)
            output = [output (1-output)];
        end

        % Output is mapped dataset.
        argout = setdat(a,output,w);

    end

return
```