The Application of Spectroscopic Methods to Predict Sugarcane Quality Based on Stalk Cross-sectional Scanning

Nazmi Mat Nawi¹,³, Troy Jensen¹,², and Guangnan Chen¹*

¹Faculty of Engineering and Surveying, University of Southern Queensland, Toowoomba, QLD 4350, Australia
²National Centre for Engineering in Agriculture, University of Southern Queensland, Toowoomba, QLD 4350, Australia
³Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author: chengn@usq.edu.au

ABSTRACT

With the increasing adoption of Precision Agriculture (PA) technique in the sugarcane industry, there is a growing need for a reliable method of in-field quality measurement. However, current PA monitoring systems can only monitor cane yield and have no ability to measure the product quality. Thus, the purpose of this study was to evaluate the ability of the spectroscopic techniques as a rapid and non-destructive tool to predict quality properties of sugarcane in the field. Both handheld Vis/NIR (350-1075 nm) and full range (350 - 2500 nm) spectroradiometers were used to determine the quality attributes of sugarcane by scanning the cross-sectional surface of the stalk. NIR calibration models were constructed using a set of 100 stalks, each which were further cut into three sections of top, middle and bottom sections. After preprocessing treatments, Partial Least Squares (PLS) method was used to interpret spectra and to develop calibration model for sugarcane quality. The overall coefficient of determination ($r^2$) for Brix, Pol, CCS and fibre as predicted by the Vis/NIR for all sample sections were 0.68, 0.71, 0.70 and 0.56 respectively. The corresponding $r^2$ for Brix, Pol, CCS and fibre as predicted by the FRs were 0.76, 0.76, 0.81 and 0.68 respectively. It was found that by using Vis/NIR, the top section can achieve $r^2$ of 0.89 for CCS prediction. The results suggested that spectroscopy based on stalk cross-sectional scanning is a feasible method for mapping and predicting sugarcane quality in the field.

Key words: sugarcane, commercial cane sugar (CCS), quality attributes, NIR, spectroradiometer
INTRODUCTION

Sugarcane (*Saccharum* spp.) is an important crop in Australia, valued between AUS $1.5 to 2.5 billion per year (Anon. 2011). Recently, substantial emphases have been given to the use of precision agriculture (PA) technologies in the sugarcane industry to increase its productivity and quality performance. However, most of the current PA technologies can only monitor the yield and have no ability to measure the quality of the product (Davis et al., 2009). This is a serious limitation because sugarcane quality varies significantly across the paddock (Lawes et al., 2003).

Furthermore, measuring sugarcane quality in the field is also important to improve the current payment system. At present, sugarcane quality is only randomly sampled and measured at the mills using sophisticated laboratory equipment. Hence, it is difficult to ensure growers are appropriately paid for the quality of their product. The current practice does not strongly motivate growers to pay more attention to improve the quality of their product.

The common technologies used to measure sugar content in a laboratory are refractometric, polarimetric, chromatographic and Near Infrared Reflectance (NIR) spectroscopic methods (McCarthy, 2003). Unfortunately, most of these methods have great limitations for use in the field because they are often time-consuming, operator-dependent, and require hazardous reagents (Mehrotra and Siesler, 2003). Several laboratory studies have shown that spectroscopic methods can be successfully used to predict quality of sugarcane based on juice samples (Madsen et al., 2003; Valderrama et al., 2007 and Doherty et al., 2007). Unfortunately, obtaining sufficient juice samples in the field is often very difficult. Meyer and Wood (1988) have also performed spectroscopic analysis on shredded cane. However, the small size of the sample cell and the restricted number of wavelengths of the filter instrument limited the accuracy of their work. Berding et al. (1991) found that inadequate sample preparation, mixing, and sample presentations were the major sources of error in their spectroscopic methods.

Recent studies have shown that the latest portable spectroscopic technology with Vis/NIR spectroradiometer (350-1075 nm) is able to provide a rapid, cost-effective and non-destructive measurement of product quality for several crops in the field (Temma et al., 2002; Montes et al., 2006). For sugarcane, it may be argued that the simplest sample form for direct quality measurement in the field is the stalks.

To date, there is no available published study regarding the use of spectroscopic method to measure sugarcane quality in the field based on fresh stalk samples. This is because sugarcane solids are among the most difficult agricultural materials to analyse by spectroscopic methods (Mehrotra and Seisler, 2003). The high moisture content (up to 681 g kg⁻¹) may mask many spectral features of the quality components (Berding et al., 1991).

The aim of this paper was to evaluate the possible use of spectroscopy in estimating sugarcane quality under farm conditions. The first objective was to determine the feasibility and accuracy of using spectroscopic equipment to predict sugarcane quality based on non-juice samples of stalks. This information was then further used to determine the best section of the cane stalk for spectra scanning in the field. This information is critical, because a spectrometer
installed on a chopper harvester for the purpose of continuous and real time in-field quality measurement would only be able to scan one point per stalk sample at a time.

**MATERIAL AND METHODS**

The experimental study reported here was conducted on a single, or closed, population. This was to determine the feasibility of using spectroscopic method as a rapid, non-destructive, and accurate technique for analysis of cut cane samples in the field.

In Australia, sugarcane quality known as commercial cane sugar (CCS) is estimated by measuring a sample of sugarcane juice in the mill. In this study, the spectra data and values of quality components (Brix, Pol, CCS and fibre) for fresh stalk from each section was first measured and calibrated. Then, the CCS values for each section and all combined sections were compared. The stalk section with the highest coefficient of determination ($r^2$) for CCS was then selected as a candidate for scanning in the field.

The standard algorithm for calculating CCS in the mill is:

\[
\text{Brix in cane} = \text{Brix in juice} \times (100 - ((\text{fibre} + 3)/100)
\]

\[
\text{Pol in cane} = \text{Pol in juice} \times (100 - ((\text{fibre} + 5)/100)
\]

\[
\text{Impurities in cane} = \text{Brix in cane} - \text{Pol in cane}
\]

\[
\text{CCS} = \text{Pol in cane} - 0.5 \times (\text{impurities})
\]

In the above equations, Pol is a measure of the sucrose present and other optically active substances, while Brix is a measure of sucrose and soluble impurities (BSES, 1984).

**2.1 Sample preparations**

100 fresh sugarcane stalks representing 10 common varieties were taken from the first ratoon crop in Bundaberg, Queensland, Australia. These stalks were manually harvested using machete. The tops of the stalks were removed by cutting each stalk near the growing point and removing all leaf materials. Then, each of them was cut into three sections of top, middle and bottom, each with approximately 600 mm length (Fig. 1). After cutting, using both Visible and Near-Infrared Spectroradiometer (Vis/NIR) and Full Range Spectroradiometer (FRs), the spectra of each section were obtained by scanning the cross sectional surface of cut cane samples (Fig. 1).

One third of these sections were used for fiber determination, while other two third were first used for NIR stalk scanning and then for juice sample extraction. The juice was extracted using a hydraulic press. The extracted juice was used for Brix and Pol determination, with fibre content being estimated using method 4 or 4a (Anon., 2001). These three traits were then used to calculate the final CCS value via the standard procedures outlined above (BSES, 1984). Pol reading was made using Polartronic Universal automatic polarimeter (Schmidt + Haensch,
Berlin, Germany). The measurement for Brix was done using Bellingham and Stanley RFM310 temperature compensated digital refractometer.

2.2 Spectral Acquisitions

A handheld Vis/NIR Spectroradiometer (350-1075 nm) and Full Range (FR) (350 - 2500 nm) Spectroradiometer, both manufactured by FieldSpec® ASD, were applied to scan the cross section of cut cane stalk. The scanning was done using 25° field of-view (FOV) of the spectroradiometer. The probe was located above the sample, and the distance between the sample and probe was 50 mm. The distance was maintained by placing the probe on a tripod and placing the samples in the marked area on the table edge. The reflectance spectra for Vis/NIR from 325 to 1075 nm were measured at 1.5 nm intervals. The spectral range for FRs was 350–2500 nm with a resolution of 1.4 nm in the 350–1000 nm range and 2 nm in the 1000–2500 nm.

Both spectroradiometers were set to record the average reading of 25 scans for each spectrum. For FRs, the ASD used a fibre optic cable for light collection and a notebook computer for data logging. All spectral data were stored in the computer and processed using the RS3 software for Windows (Analytical Spectral Devices, Boulder, USA) designed with a Graphical User Interface. To simulate the field condition, 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 being scanned while taking the spectral measurements. Relative reflectance spectra were calculated by dividing stalk radiance with reference radiance from a spectralon white reference panel for each wavelength.

2.3 Spectral Data Preprocessing

For Vis/NIR spectroradiometer, in order to reduce low signal-to-noise ratio, the first and last 75 nm data points were removed from the spectral data produced. Therefore, only the data in the regions of wavelengths (400–1000 nm) were used for the calculations. For full range spectra, in addition to the first 75 nm data points, the spectral regions between 1355–1450, 1800–1950 and 2420–2500 nm which have known water absorption features (Kumar et al., 2003; Mutanga et al., 2004) were also excluded from the analysis and removed from the spectra for causing excessive noise (Abdel-Rahman et al., 2010).

Before the calibration stage, the spectral was pre-processed for an optimal performance. This was because, due to light scattering in the samples, some persistent baseline shifts and bias appeared in the spectra. The original light from the light source traveled different distances in the samples before the spectral measurement. 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 the accuracy of analysis results. Several common chemometrics pretreatment methods, such as moving average smoothing (MAS), Savitzky-Golay second derivative (2nd-D), multiplicative scatter correction (MSC), and standard normal variate transformation (SNV) could be applied to reduce the noise and normalize the data.
In this study, the raw NIR spectra data was converted to the multiplicative scatter correction (MSC). MSC technique was chosen because it was the most popular normalisation technique offered by most chemometrics software packages (Næs et al., 2004). MSC was used to compensate for additive (baseline shift) and multiplicative (tilt) effects in the spectral data, which were 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. The method attempted to remove the effects of scattering by linearizing each spectrum to some ‘ideal’ spectrum of the sample, which, in practice, corresponded to the average spectrum.

The most frequently used multivariate-regression methods in NIR spectroscopy are principal component regression (PCR) and PLS regression (Martens & Næs, 1989). 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 (Jong, 1993). A proper model should have a low root-mean-square error of prediction (RMSEP) and a high coefficient of determination between the predicted and measured value of each property. The maximum model that is acceptable is usually 10 PCs (Moghimi et al., 2010). A low number of PCs were desirable in order to avoid inclusion of signal noise in the modeling (Xiaobo et al., 2007). All of the pretreatments were implemented by “The Unscrambler V 9.2” (CAMO PROCESS AS, Oslo, Norway).

2.4 Calibration and Validation Sets

The sample set comprised of data for 300 cut sections. After elimination of extreme outliers, 290 billet samples were available for both Vis/NIR and FRs. The method of outlier rejection included visual inspection of light reflection spectra where any spectra that appeared abnormal when compared to the whole dataset were rejected (Digman and Shinners, 2008). It was found that less than 5% of the spectra were abnormal or contained no reflectance information. The same number of billet samples was used for both methods in order to afford a better comparison between these two methods. This sample set was further divided into two sets. One set (75%) was used to develop a prediction equation (calibration set) while another part (25%) 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 spectra.

RESULTS AND DISCUSSION

3.1 Determination of Sugarcane Quality Components

Table 1 shows the predicted quality components as obtained from FRs for all combined samples of top, middle and bottom sections. It can be seen that among all considered quality components for all samples, CCS had the highest $r^2$ value with the lowest RMSEC/RMSEP values in both calibration and validation models. In contrast, fibre had the lowest square of the coefficient of determination $r^2$ value in both models. In the calibration model, values of $r^2$ ranged from 0.72 for fibre and 0.77 for CCS. In the prediction model, values of $r^2$ ranged from 0.65 for fibre and 0.81 for CCS. This showed that CCS value had the highest correlation with spectra
data. This seems reasonable because CCS is the main measurement factor to measure sugar value in sugarcane stalk as determined by the industry. The graphs of calibration and prediction models for CCS are shown in Fig. 2 and 3. The values for fibre in both models were low due to the fact that fibre value was not directly measured from the sample, but estimated using an equation (Anon., 2001).

For Vis/NIR, the trend of its model performance for all components was similar to the FRs (Table 2). For prediction models, however, the highest $r^2$ value of FRs belonged to CCS while for Vis/NIR this belonged to Pol. The table also shows that FRs was better than Vis/NIR in estimating CCS in the field. This was because Vis/NIR had only 600 data points while FRs had 2500 data points which can be manipulated in order to produce a good prediction model. Overall, for FRs, the $r^2$ values of Brix, Pol, CCS and fibre were 0.76, 0.76, 0.81 and 0.68 respectively. For Vis/NIR, the $r^2$ values of Brix, Pol, CCS and fibre were 0.68, 0.71, 0.70 and 0.56 respectively.

In other studies on solid cane samples, Berding et al. (1991) reported that for the calibration model, $r^2$ values of Brix, Pol, CCS and fibre for fibrated cane were 0.94, 0.97, 0.94, and 0.92 respectively. For the prediction model, the value for Brix, Pol, CCS and fibre were 0.91, 0.96, 0.91 and 0.89, respectively (Berding et al., 1991). These data trends are consistent with this study where the value of Pol was the highest among other components. On shredded cane, Meyer and Wood (1988) reported $r^2$ values of 0.88, 0.92, 0.8, 0.86 and 0.89 for Brix, dry matter, fibre, Pol and purity. These studies had higher $r^2$ values than the current study because they used mechanically treated samples (fibrated and shredded) which were homogenous solid samples. Valderrama et al (2007) found that $r^2$ for both Brix and Pol were 0.99 and reducing sugar (RS) was 0.76 for juice. However, when the same author repeated the similar experiment on juice, the values for $r^2$ of prediction model for Brix, Pol and CCS were 0.97, 0.97 and 0.98, respectively.

### 3.2 Variations in different cane sections

Initially, each cane section (bottom, middle and top) had 100 scanned data. The influence of sugar content of sugarcane for each section on the reflectance spectra for FRs and Vis/NIR is shown in Fig. 4 and 5 respectively. Overall, it seemed that each section had its own unique characteristic curve. For the spectra generated by FRs, the curves for the three sections were nearly overlapping with each other. For Vis/NIR, the curve for the top section appeared to have the lowest reflectance values compared to other two curves. This difference may be due to different equipment sensitivity and sensor materials being used for Vis/NIR and FRs. Fig. 5 shows that Vis/NIR has also detected two sensitive wavelengths at 760 and 930 nm. Band around 760 nm represents OH stretching at 3rd overtone of sugar (Golic et al., 2003) and band around 930 nm represents CH2 stretching at 3rd overtone of sugar (Osborne et al., 1993).

In order to evaluate Vis/NIR and FRs as tools to predict the sugarcane quality CCS for different sections, calibration models had been established for both these systems (Tables 3 and 4). It can be seen from these tables that both calibration and prediction models varied for each section. Table 3 shows the models performance for FRs. It can be seen that the differences in
values of $r^2$ for both models for each section are small (+/- 0.05), thus negligible. This indicated that if the FRs is going to be used in the field, any section could be scanned and used as they would have similar prediction values. By using Vis/NIR (Table 4), the top section had better $r^2$ values for both models. This indicated that the top section would best represent the sugar content amount for the whole stalk.

**CONCLUSIONS**

Sugarcane is an important agricultural crop. Sugarcane quality can vary significantly across the paddock. Although there were a number of previous laboratory studies to measure sugarcane quality based on prepared juice or macerated samples, having these samples in the field in the proper amount for quality determinations is often extremely difficult. This paper has explored the use of spectroradiometers as a potentially rapid, non-destructive method to scan sugarcane cross sectional samples to predict the values of Brix, Pol, CCS and fibre. It has been found that the overall coefficient of determination ($r^2$) for Brix, Pol, CCS and fibre as predicted by the Vis/NIR for all combined samples of top, middle and bottom sections were 0.68, 0.71, 0.70 and 0.56 respectively. The corresponding $r^2$ for Brix, Pol, CCS and fibre as predicted by the FRs were 0.76, 0.76, 0.81 and 0.68 respectively. The relatively low prediction values for these models were due to variation between subsamples, reflecting the heterogeneous nature of the raw cane stalks. The high moisture content of the material and the influence of natural daylight in the field environment may also affect the accuracy of the results.

It has also been found that overall, CCS appeared to be able to achieve the best correlation with the spectra. Although this study showed that FRs had an overall better accuracy than Vis/NIR, it was noted that the existing FRs machine was not portable and quite costly so it cannot be realistically used in the sugarcane chopper harvesters to predict sugarcane quality in the field. It has also been shown that the best section for scanning was dependent on wavelength and equipment used. Using a low-cost Vis/NIR, it has been found that the top section can achieved a good $r^2$ of 0.89 for CCS prediction. The future research should focus on developing and evaluating a suitable automatic system for extracting and scanning billet samples in the sugarcane chopper harvesters for direct in-field quality measurement and initial field mapping (Wright et al., 2002; Taylor et al., 2005).

**ACKNOWLEDGEMENTS**

The authors wish to acknowledge the financial support provided by the National Center for Engineering in Agriculture (NCEA), the University of Southern Queensland, Australia. The authors also acknowledge contributions of Mr. John Panitz, Dr Bernard Schroeder and Mr. Phil Netz from Bureau of Sugar Experiment Stations (BSES Limited), Bundaberg, Queensland, for their technical supports and helps in data collection and laboratory quality analysis.
REFERENCES


Nawi et al.: Stalk Cross-sectional Scanning

Fig 1: The division of fresh sugarcane stalks into three main sections.

Fig. 2: Calibration model for CCS using FRs.

Fig. 3: Prediction model for CCS using FRs.
Fig. 4. Comparison between the spectra of bottom, middle and top scanning for FRs.

Fig 5. Comparison between the spectra of bottom, middle and top scanning for Vis/NIR spectrometer.
### Table 1: Predicted quality components obtained from FRs for all combined samples of top, middle and bottom sections.

<table>
<thead>
<tr>
<th>Components</th>
<th>Calibration</th>
<th></th>
<th></th>
<th></th>
<th>Prediction</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r²</td>
<td>RMSEC</td>
<td></td>
<td></td>
<td>r²</td>
<td>RMSEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brix</td>
<td>0.73</td>
<td>0.53</td>
<td></td>
<td></td>
<td>0.73</td>
<td>0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCS</td>
<td>0.77</td>
<td>0.42</td>
<td></td>
<td></td>
<td>0.81</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibre</td>
<td>0.72</td>
<td>0.83</td>
<td></td>
<td></td>
<td>0.65</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pol</td>
<td>0.74</td>
<td>2.24</td>
<td></td>
<td></td>
<td>0.75</td>
<td>2.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R² = coefficient of determination; RMSEC = Root Mean Square Error of Calibration; RMSEP = Root Mean Square Error of Calibration.

### Table 2: Predicted quality components obtained from Vis/NIR for all combined samples of top, middle and bottom sections.

<table>
<thead>
<tr>
<th>Components</th>
<th>Calibration</th>
<th></th>
<th></th>
<th></th>
<th>Prediction</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r²</td>
<td>RMSEC</td>
<td></td>
<td></td>
<td>r²</td>
<td>RMSEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brix</td>
<td>0.76</td>
<td>0.46</td>
<td></td>
<td></td>
<td>0.68</td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCS</td>
<td>0.81</td>
<td>0.36</td>
<td></td>
<td></td>
<td>0.70</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibre</td>
<td>0.79</td>
<td>0.66</td>
<td></td>
<td></td>
<td>0.56</td>
<td>0.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pol</td>
<td>0.78</td>
<td>1.89</td>
<td></td>
<td></td>
<td>0.71</td>
<td>2.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R² = coefficient of determination; RMSEC = Root Mean Square Error of Calibration; RMSEP = Root Mean Square Error of Calibration.

### Table 3: CCS characteristic of wavelength for FRs.

<table>
<thead>
<tr>
<th>Section</th>
<th>Calibration</th>
<th></th>
<th></th>
<th></th>
<th>Prediction</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>RMSEC</td>
<td></td>
<td></td>
<td>R²</td>
<td>RMSEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top</td>
<td>0.91</td>
<td>0.27</td>
<td></td>
<td></td>
<td>0.73</td>
<td>0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>0.89</td>
<td>0.30</td>
<td></td>
<td></td>
<td>0.70</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bottom</td>
<td>0.88</td>
<td>0.31</td>
<td></td>
<td></td>
<td>0.75</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R² = coefficient of determination; RMSEC = Root Mean Square Error of Calibration; RMSEP = Root Mean Square Error of Calibration.

### Table 4: CCS characteristic for wavelength for Vis/NIR.

<table>
<thead>
<tr>
<th>Section</th>
<th>Calibration</th>
<th></th>
<th></th>
<th></th>
<th>Prediction</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r²</td>
<td>RMSEC</td>
<td></td>
<td></td>
<td>r²</td>
<td>RMSEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top</td>
<td>0.89</td>
<td>0.26</td>
<td></td>
<td></td>
<td>0.89</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>0.88</td>
<td>0.30</td>
<td></td>
<td></td>
<td>0.68</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bottom</td>
<td>0.88</td>
<td>0.29</td>
<td></td>
<td></td>
<td>0.64</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R² = coefficient of determination; RMSEC = Root Mean Square Error of Calibration; RMSEP = Root Mean Square Error of Calibration.