FT-NIR spectrometry and automated presentation for high-speed, at-line analysis of disintegrated sugarcane

Nils Berdinga, David Marston[b], W. Fred McClure[c], Maarten van Eerten[b], and Brian Prescott[b].

[a]*Bureau of Sugar Experiment Stations (BSES), P.O. Box 122, Gordonvale, 4865, Queensland, Australia.*

[b]*Biolab Technology, P.O. Box 31044, Lower Hutt, New Zealand.*

[c]*North Carolina State University, P.O. Box 7625, Raleigh, NC 27695-7625, U.S.A.*

Introduction

Measurement of quality components of freshly disintegrated sugarcane (*Saccharum* L. spp hybrids) stalks using near infrared spectroscopy has been well developed at BSES Meringa, on the north-east coast of Queensland, Australia, since 1995.[1] The stalks analyzed are samples removed from crop improvement and agronomy R&D field trials. These earlier applications used a laboratory-model scanning monochromator and a custom-made, semi-automated sample presentation device, or large cassette module (LCM).[2, 3] This system involved taking a sub-sample of ≈ 4 kg from a pre-mixed mass of disintegrated stalk tissue from a sample of six to nine random stalks, depending on trial format, and presenting each sample in a cassette 80 (W) x 80 (H) x 1,000 mm (L) for scanning by a remote reflectance probe fibre-optically coupled to the detector of a scanning monochromator.

There were several motivating factors for the development of a new automated at-line system. An inability to have a high-moisture application, using a scanning monochromator, protected via a standardization procedure and a failure to have the instrument returned to an established calibration after repair were paramount.[4] The need to reduce analytical costs, i.e., process more samples per day using fewer personnel, demanded a more automated operation than was afforded by the LCM. An obvious strategy was avoidance of pre-mixing to minimize intra-sample heterogeneity and subsequent sub-sampling. The design target was to scan up to 400 total samples, each weighing 3 – 18 kg, per day. The catalyst to achieving these goals was afforded by a presentation at the 10th ICNIS, at Kyongju, in June 2001.[5] This featured a high-speed, Fourier-Transform, near infrared spectrometer with highly desirable features for at-line analysis of high-moisture materials. This instrument uses a proven interferometer, a Peltier-cooled InGaAs detector, and was designed for non-contact scanning without use of fibre-optics. The instrument has high spectral precision and minimal instrument-to-instrument variation, attributes that improve calibration longevity and ensure ease of transfer. The instrument software (OPUS™) is user configurable for data acquisition as well as process control. In this paper we compare two instruments: (1) the Matrix-E (Bruker Optik, Germany) with (2) the B6500 scanning monochromator (NIRSystems, Silver Spring, MD). In addition, the paper briefly details the design of a high-speed sugarcane sample preparation system that, together with the Matrix-E, is called the "High-Speed Sugarcane Analyzer" (HSSA).

Near Infrared Spectroscopy: Proceedings of the 11th International Conference.

Methods

Instrument evaluation

In October 2001, a FOSS NIRSystems Model B6500 spectrometer with a 600 mm remote reflectance probe was compared with a Bruker Matrix-E. Stalk samples were drawn from two advanced clonal evaluation trials (2 replicates x 96 plots; 2 replicates x 81 plots) and an agronomy assessment trial (3 replicates x 20 plots). Stalks were prepared using a Codistil Dedini (Piracicaba, Brazil) disintegrator and then mixed in a rotating-drum mixer for 90 s. Samples were presented to the B6500 using the LCM. Trays 25 (D) x 100 (W) x 1,000 mm (L) filled immediately prior to scanning were used for the Matrix-E. These were scanned from above by moving the loaded tray through the instrument's focal point, at a distance of 170 mm from the protective window, on a rail system with an electric winch. Spectral data from 10,000 – 4,000 cm^{-1} (1,000 – 2,500 nm) were taken with the Matrix-E, and from 800 – 2,200 nm with the B6500. The instruments operated in controlled environments of 24°C and < 52% R.H. Samples were scanned in batches of four, with the initial presentation of each batch alternating between instruments, thus avoiding any systematic bias arising from an ordered sample presentation.

Juice was expressed from a sub-sample of ≈ 1,000 g of disintegrated stalk tissue using a hydraulic press (Pinette Emidecau, France) operating at 200 Bar. This juice was analyzed for Brix (soluble solids) using a refractometer (Bellingham and Stanley, U.S.A.) and after clarification with lead sub-acetate, for polariscopic reading using a Universal polarimeter (Schmidt and Haensch, Germany). The pressed cane was analyzed for theoretical fibre[6], except that 'wet' pressed plug weights were not captured. "Plug dry matter" was determined. An "adjusted fibre" value was calculated by correcting for residual plug Brix assuming a mean mass balance of 1,005 g kg^{-1}. Moisture was determined by drying a known weight of ≈ 150 g of fibrated cane at 70°C for 7 days. Commercial cane sugar (CCS), the industry payment statistic, and juice purity, were derived from these basic data. Of the 414 samples used, every fourth sample had two sub-samples scanned by each instrument. This gave a total of 107 samples, and allowed assessment of predictive precision by determination of sub-sampling error between predictions on duplicate spectra. Duplicate routine laboratory analyses (RLAs) also were performed on these samples to allow calculation of the standard error associated with the RLA of each component.

Calibrations were developed using the OPUS™ (Bruker Optik, Germany) software package for spectral data collected with the Matrix-E and the WINISI software package (ISI, PA) was used to develop calibrations from spectral data collected with the Model B6500. Routine calibration techniques, including cross-validation and modified partial least-squares regression, were used in both instances. Duplicate spectra were excluded from all calibrations using cross validation.

HSSA concept and design

High throughput and elimination of sample pre-mixing were pivotal criteria for the design of the HSSA. Significant automation was required to achieve high throughput and presentation of the total sample for scanning avoided sample pre-mixing. The challenge was to deliver stalk tissue, discharged intermittently from a disintegrator, in a continuous, linear, ribbon-like presentation for scanning by the FT-NIR spectrometer.

Under routine operation, when prompted, the sample ID bar code was scanned and the sample stalks were fed into the disintegrator. The output was discharged onto a continuously running, "elevating conveyor" and accumulated in a "receival hopper", all under the control of a programmable logic controller (PLC). When the entire sample accumulated in the hopper (an event monitored by an optical sensor) the contents were deposited onto a "distribution conveyor" that

transformed a 3-D sample mass into a linear sample stream. The linear stream was discharged into a "fluffing chamber", located over a "presentation conveyor", that defined the output profile. The height of this output was continuously with an electronic sensor. The sample was presented to the instrument, also under PLC control, and accumulated in a "decision hopper". On completion of scanning, OPUS™ sent a command to the PLC that resulted in a sample being either dumped or automatically rerouted to an "analysis bin". Various elevator and profile sensors monitored material flows to ensure sample and presentation integrity whilst identifying potential errors. A supervisory computer program (CaneCon©) worked in conjunction with OPUS™ to create bar code data files plus tag samples that OPUS™ had defined as outliers for analysis.

Functions such as reference scanning, instrument diagnostics, and functionality checks were automated by OPUS™, and this information was monitored and logged by CaneCon©. This information enabled the user to check all processing parameters at any stage. Furthermore, certain errors caused a display flag/or alarm that required operator intervention.

Efficacy of HSSA

In November 2002, the efficacy of the HSSA was assessed using partial or full sample sets from three advanced clonal selection (2 reps x 78 plots; 2 trials x 2 reps x 49 plots) and two R&D trials (2 reps x 18 clones x 18 spaced plants; 4 reps x 180 plots). All trials except for the latter involved *Saccharum* spp. hybrid material. The latter was an assessment trial containing predominantly *Saccharum officinarum* L. clones. These trials collectively contained more genetic variation than the trials used in the instrument comparison, were from geographically more diverse locations, but as with the earlier trials, encompassed a restricted temporal sample of seasonal variation, being all late-season harvests. Spectral and RLA data for almost 1,200 samples were available from these trials.

The HSSA operated under ambient tropical conditions. A software limitation meant that only 48 scans of a sample were taken from commencement of an acceptable profile. A set of random samples (n = 103) was recycled through the HSSA so spectral data for duplicates sub-samples were available. Subsequent to scanning, samples were mixed before RLAs were performed. The RLA data collected were as described earlier, with the exception that a true theoretical fibre was determined.[6] Calibrations developed were applied to the population of primary and duplicate spectra. Predicted data were subjected to sub-sampling analyses to determine a standard error for each component that described the predictive precision of the HSSA analyses. Again, duplicate spectra were excluded from the initial calibration set, and all calibrations were developed using the OPUS™ software package.

Results

Instrument evaluation

All component means except moisture were high (Table 1). Moisture, as expected late in the season, was at a low value but in its expected range. There was ample variation for all components except purity, and to a lesser extent, polarization readings (Table 1), as indicated by standard deviation values. There was variation between instruments in number of terms in the calibration equations developed. Most Matrix-E calibrations used one additional term. Coefficient of multiple determination (R^2) values were about equal for calibrations for Brix, fibre, and moisture (Table 1). Calibrations for the B6500 monochromator showed marginally higher values than the Matrix-E for the remaining components. All SECV values for the B6500 were higher than the value obtained

with the Matrix-E. Their relativity (B6500/Matrix-E) ranged from 104 to 127%. Standard error estimates from calibration predictions on duplicate spectra indicated that B6500 values were higher for all components (Table 2). Their relativity (B6500/Matrix-E) ranged from 112% for pol. reading to 168% for purity (Table 2). Predictive precision of both instruments, relative to RLA precision, was less for fibre, press dry matter, and pol. reading but was greater for Brix, moisture, and purity. The instruments differed in their precision relative to the RLA value only for CCS (Table 2).

Table 1. Summary population and cross-validation statistics, for seven quality components of disintegrated sugarcane stalk tissue, obtained from the 2001 instrument comparison using semi-automated presentation, the 2002 assessment of the automated high-speed sugarcane analyzer (HSSA), and data for the Infracana[1], an automated presentation system based on a scanning monochromator.

Component	Statistic	Semi-automated		Automated	
		Matrix-E	B6500 LCM	HSSA	Infracana
Brix in juice (g kg^{-1})	n	502	399	1,087	183
	Mean	235.9	236.0	244.7	183.6
	SD	9.1	9.3	21.1	--
	# terms	9	8	9	--
	R^2	0.948	0.942	0.962	0.966
	SECV	2.09	2.66	4.12	5.08
Commercial cane sugar (g kg^{-1})	n	495	399	1,089	173
	Mean	165.2	165.0	158.0	96.7
	SD	7.8	8.1	25.0	--
	# terms	10	13	10	--
	R2	0.870	0.910	0.949	0.948
	SECV	2.83	2.95	5.66	3.19
Fibre – T (g kg^{-1})	n	502	403	1,090	171
	Mean	160.7	160.7	136.9	129.1
	SD	14.2	14.5	15.5	--
	# terms	9	10	10	--
	R2	0.904	0.910	0.852	0.901
	SECV	4.38	4.92	5.95	6.99
Moisture (g kg^{-1})	n	499	393	1,068	170
	Mean	663.0	663.0	665.3	714.9
	SD	12.5	13.0	23.2	--
	# terms	9	5	9	--
	R2	0.950	0.953	0.963	0.912
	SECV	2.78	3.06	4.45	5.92

(Table 1 cont.)

Component	Statistic	Semi-automated		Automated	
		Matrix-E	B6500 LCM	HSSA	Infracana
Pol. reading (°Z)	n	499	393	1,084	180
	Mean	92.7	92.8	90.2	64.8
	SD	4.2	4.1	12.2	--
	# terms	10	8	10	--
	R2	0.941	0.952	0.967	0.961
	SECV	1.02	1.09	2.21	1.63
Purity (%)	n	501	392	1,064	--
	Mean	93.3	93.3	87.1	--
	SD	1.3	1.3	4.8	--
	# terms	10	8	10	--
	R^2	0.423	0.506	0.874	--
	SECV	1.01	1.17	1.71	--

[1] Data were taken from [7], but modified as necessary to account for differences in units.

Table 2. Estimates of sampling standard errors from variation between duplicate routine laboratory analyses (RLA) and between predictions obtained from application of relevant calibrations to duplicate spectra collected in the comparative instrument evaluation in 2001 and from the high-speed sugar analyzer (HSSA) assessment in 2002.

Component	2001			2002
	RLA	Matrix-E	B6500-LCM	HSSA
Brix in juice (g kg^{-1})	1.735	1.147	1.574	4.722
Commercial cane sugar (g kg^{-1})	1.595	1.468	2.037	5.734
Fibre – T (g kg^{-1})	2.862	3.426	3.953	6.125
Moisture (g kg^{-1})	3.259	2.378	2.726	7.401
Press dry matter (g kg^{-1})	2.869	2.970	3.893	--
Pol. reading (°Z)	0.503	0.566	0.635	2.873
Purity (%)	0.873	0.399	0.669	1.353
# duplicates	107	102	107	103

HSSA evaluation

The population used for the HSSA evaluation, relative to that used for the instrument comparison, had higher mean values for Brix and moisture, and lower values for all other components. All components except fibre had markedly higher standard deviations. This confirmed this population contained greater variability because of genetic sampling. Calibration equations for the same component from the two experiments differed little in the number of terms (Table 1). Equations developed from spectra captured by the Matrix-E from samples presented by the HSSA yielded coefficient of multiple determination (R^2) values comparable to those obtained in the instrument comparison (Table 1) for all components except fibre (lower) and purity (higher).

Relativity (HSSA/Matrix-E), over all components, ranged from 94 to 207%. The SECV values for the HSSA calibrations were higher for all components. Their relativity ranged from 128% for fibre to 217% for pol. reading. The HSSA's performance can be contrasted with those of the Foss Infracana system, an automated, at-line, analytical system using a scanning monochromator[7]. The HSSA produced comparable R^2 values for Brix, CCS, and pol. reading but lower R^2 value for fibre, and a higher R^2 value for moisture (Table 1). The HSSA produced lower SECV values for all components except CCS and pol. reading. In terms of predictive precision, the analysis of predicted values for duplicate spectra revealed that sampling standard errors were inflated relative to those for the instrument comparison, ranging in relativity (HSSA/Matrix-E) from 179% for moisture to 508% for pol. reading.

Discussion

An objective assessment of the Matrix-E FT NIR spectrometer revealed a performance at least equal to the scanning monochromator, a system long regarded as the benchmark instrument for NIR technology. In combination with its inherent physical features, and detector sophistication, the Matrix-E appears an ideal instrument platform on which to base high moisture, at-line, applications. Assessment of the Matrix-E installed in the HSSA revealed performance that was marginally worse than seen in the instrument comparison study due most likely to whole samples not being scanned and operations being conducted under ambient tropical conditions. However, this performance was comparable to that published for an alternative automated system.[7] The reduced performance, as indicated by predictive precision, was disappointing. Can performance of the HSSA be brought close to the upper bound as defined by the Matrix-E in the instrument comparison? Intuitively, this seems possible. Performance of the prototype HSSA was excellent when presented with constant-sized samples, and the target of 400 samples per day would be achievable. In fact, experience in this assessment suggested that, at this rate, a rotation of disintegrator operators would be necessary. Implementation of a new version of the HSSA's PLC code will allow realization of the original objective to scan total samples, not just the initial portion of presentations. Real problems were encountered with scanned entry of bar-code sample identities, particularly for duplicate sub-samples. Statistics reported are for a minimally edited data set (i.e., limited deletion of suspect spectra) despite our knowledge this contained corrupted sample identities. The approach to, and incorporated safeguards for, bar code acquistion contained in the revised PLC code will eliminate this source of error. Trouble-free performance of the HSSA is very dependent on disintegrator performance and maintaining this at a consistently high level will be integral to the HSSA's success. Development of the HSSA was effected within 18 months of seeing the Matrix-E presentation.[5] Overall, development of the HSSA exceeded expectations. Assessment of the prototype was a learning experience, allowing improvement to the unit for the 2003 season. Construction of additional units is scheduled, allowing operation of a networked system of five units operating within BSES under a global calibration. A "low-speed" unit, or LSSA also will be developed for lower investment situations. Application to other at-line, agricultural situations also is envisaged.

Acknowledgements

We acknowledge the ready and willing support of Bruker Optik GmbH, Ettlingen Germany. The generous support of Biolab Technology, NZ, for their investment in the HSSA project, also is gratefully acknowledged. The continuing support of BSES for the initiative embodied in the HSSA has been a key to its success. Collaborative funding from BSES, the Queensland Department of

Primary Industries, and Central Queensland University (Prof. K.B. Walsh) allowed the visit of Prof. W.F. McClure, as a visiting fellow, to participate in, among other activities, exploratory design embodied in the HSSA. The important contribution of the juice laboratory team at BSES Meringa, in particular Ann Rizzo, Kaylene Court, and Rhylee Swain, for delivery of the RLA results, and sufferance of the, at times, disruptive HSSA "experimentation" is acknowledged with real appreciation.

References

1. N. Berding and G.A. Brotherton, in *Proceedings of 2nd International Users Meeting for the Sugar and Alcohol Industries*, 2-3 August 1999 Ribeirao Preto, Sao Paulo Brazil, 16 pp. (Proceedings unpaginated) (1999).
2. N. Berding and G.A. Brotherton, in *Sugarcane: Research towards Efficient and Sustainable Production*, Ed by J.R. Wilson, D.M. Hogarth, J.A. Campbell, and A.L. Garside. CSIRO Division of Tropical Crops and Pastures, Brisbane, p. 57 (1996).
3. N. Berding and G.A. Brotherton, NIR news **7(6)** p. 14 (1996).
4. N. Berding, in *Near Infrared Spectroscopy: Proceedings of the 10th International Conference*, Ed by A.M.C. Davies and R.K. Cho. NIR Publications, Chichester, UK, p. 525 (2002).
5. A. Schmidt, K. Schorb, R Matsch, and U. Nägele, in *Near Infrared Spectroscopy: Proceedings of the 10th International Conference*, Ed by A.M.C. Davies and R.K. Cho. NIR Publications, Chichester, UK, p. 433 (2002).
6. N. Berding and J.S. Pollock, *Inter. Soc. Sugar Cane Technol. Sugarcane Breeders' Newslettter* **44**, 43 (1982).
7. L.R. Madsen II, B.E. White, and P.W. Rein, *J. Proc. Amer. Soc. Sugar Cane Technol*. In press, 14 pp. (2003).