

TRIALS AND TRIBULATIONS OF IMPLEMENTING NIRS FOR RAW SUGAR FACTORY LIQUORS

KJSCHÄFFLER

Sugar Milling Research Institute, University of Natal, 4041, Durban

Abstract

Near infrared spectroscopy (NIRS) is capable of determining multi-analytes rapidly and easily. The selection of samples for incorporation into a successful database and the subsequent development of robust equations is much more difficult. The original goals of the project were changed several times. This has led to the discarding of many spectra and laboratory results.

Molasses and back-end samples are now diluted before transmission-NIRS scanning. Factory samples from DAC to C-massecuite are analysed using a 1 mm flowcell. Flushing and filling of the cell has been automated and the throughput is currently 20 samples per hour. Front-end calibration produces rapid estimates for sucrose, pol and brix whilst the back-end calibration yields results for sucrose, pol, brix, dry solids, fructose, glucose and sulphated ash.

Calibrations for sugar juices have been developed from NIRS spectra and laboratory data spanning 4-6 years. Although NIRS predictions on independent samples produced results with excellent fit and precision, the results were often biased. Systematic error varied depending on the datasets used for calibration and validation.

Check solutions, consisting of 10% pure sucrose, were introduced to monitor the effect of instrumental and cell-temperature changes on bias. These check solutions were used to demonstrate that the bias problem was two-pronged. Wavelength (WL) accuracy deviation, together with ambient temperature variations, was shown to be responsible for the systematic bias of NIRS predictions. The implementation of a repeatability file successfully reduced bias.

The investigation of factors responsible for the deviation in wavelength accuracy are on-going. Daily checks are carried out and wavelength adjustments are made each day. The use of repeatability files and cell thermostating is being investigated. Weekly front- and back-end comparisons between NIRS and laboratory methods will continue.

Introduction

Glossy sales brochures for NIRS instruments highlight the procedure's strong points (fast, easy analysis of multicomponents with non-invasive testing). The steep learning curve for the inexperienced NIRS user is often not mentioned. The development of robust calibration equations that will produce accurate predictions on independent datasets is also downplayed or even overlooked. Many of the published articles on the application of NIRS simply demonstrate the concept of NIRS. These investigations normally involve collection of a few hundred samples, scanned by NIRS and compared with conventional laboratory

analyses. The dataset is then randomly split, with one set used for calibration whilst the remaining samples are used for NIRS prediction. These NIRS estimates invariably agree with the laboratory results since all the samples originally came from the same source. In addition the samples are normally scanned over a relatively short period of time under very similar conditions such as controlled laboratory temperature and stable instrument condition.

A comprehensive NIRS investigation has been carried out in the sugarbeet industry by de Bruijn (1997). Transmission NIRS, together with autosampling, was used to determine pol and dry substance in juices, syrups and massecuities.

The current study at the SMRI has been based on the work initiated at CSM (de Bruijn, 1997) and covers:

- the analysis of molasses by transmission NIRS rather than by reflectance, ensuring a single, simple system for all products from direct analysis of cane (DAC) consignments to back-end products (molasses and massecuities)
- the introduction of a flow cell to automate NIRS analyses
- the widening of the calibration database to include all intermediate products
- the problems associated with reference methods and the problems associated with consolidating reference method results and spectral data, storage and retrieval of data over a period of 4-6 years
- the introduction of a check solution to monitor and correct for systematic errors
- the problem of bias and an investigation into the reasons for this bias including instrumental or temperature variations.

Experimental

Samples

Samples for direct cane testing (DAC), mixed juice (MJ) and clear juice (CJ) were scanned undiluted. Syrup samples were diluted (20 g in 100 ml), whilst for molasses (A-, B-, C-) samples, a dilution of 16 g in 100 ml was used. C-massecuite samples were diluted by weighing 13 g and diluting volumetrically to 100 ml. The laboratory's air-conditioning system was on at all times, however temperature and humidity were not monitored.

Analytes

The SMRI test methods (TM) were accredited in 1998 by the National Laboratory Accredited Service (NLA)(Anon, 1998). These accreditations are in accordance with the ISO 025 guideline.

- Front-end samples were analysed for sucrose, glucose and fructose by gas chromatography (SMRI TM300), pol (SMRI TM042), brix (SMRI TM005) and sulphated ash (SMRI TM056).
- Back-end samples were analysed for sucrose, glucose and fructose by HPAEC (SMRI TM301). Dry solids were determined by either the vacuum oven drying procedure (SMRI TM035) or by Karl Fischer (SMRI TM030). Other analyses included sulphated ash (SMRI TM057), pol (SMRI TM043) and brix (SMRI TM007)

NIRS

- A Foss NIRSystems 5000 spectrometer, with a scanning range of 1100 to 2500 nm, was fitted with a sample transport module. A 1 mm cuvette was used to scan juice samples in duplicate. This cell was subsequently replaced with a 1 mm flowcell (model NR-70504).
- An automated system developed at the Institute to fill and flush the flowcell (Figure 1) is described below:
 - When not in use the cell remains filled with 3% formaldehyde ensuring that the system remains bacteria free
 - When a new sample arrives the operator presses the start button which activates a solenoid valve and pump so that the cell is flushed with sample (50 ml). The volume pumped is controlled by the size of the pump tube, the speed of the pump (variable) and the timer (Timer 1). The operator then uses the NIRS-PC to identify and scan the sample in the usual way
 - After scanning, the sample transport returns to its home position where a micro-switch re-activates the solenoid valve and pump, flushing the cell with formaldehyde. The system is then ready for the next sample. Currently the system is capable of handling 20-25 samples an hour.
- NSAS (version 3.25) was used to scan samples for the early part of this investigation. InfraSoftInternational's (ISI) NIRS package was then used (1996-1998). More recently ISI's WINISI (version 1.04) has been used for all acquisition, data manipulation, calibration, standardisation and validation. Vision (version 2.22) has also been used, especially for additional diagnostics and graphical views of the spectra.

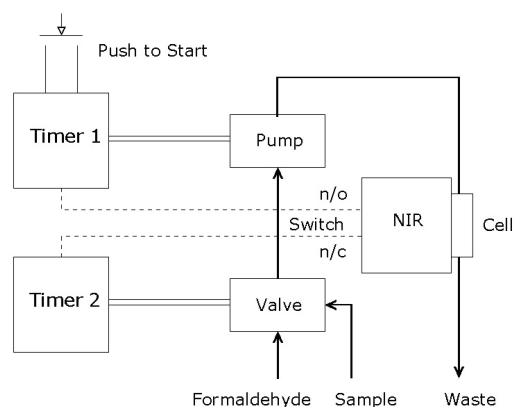


Figure 1. Schematic of automated fill and flush sampling system for NIRS flow-cell.

- Accurate predictions are reliant on the current calibrations being representative of future unknown samples. WINISI's SCORE program was used to convert all spectral data into score files using PCA loadings. Global H values of greater than 3 were considered to be outliers and were deleted. WINISI's SELECT was used to remove redundant spectra from the files prior to calibration. WINISI's "EXPAND product database with new spectra" was used to add new samples to the calibration database. The modified PLS regression method, with 1,4,4,1 mathematics treatment and SNV-detrend correction techniques, was used for all analytes. After calibration, the equations were checked internally with the original file using WINISI's MONITOR routine.

Front-end Calibration Development and Prediction

- The initial mixed juice calibration dataset was obtained from samples scanned using a 1 mm cuvette from 1994 to 1996 (N=562).
- A dataset consisting of 367 samples was obtained from Eston sugar mill in 1997 (cuvette).
- MJ samples from all mills were collected in July 1998 (N=136, cuvette).
- Other products (DAC, MJ, CJ and syrup) were acquired in 1998 to widen the prediction database (N = 438 samples, cuvette).
- The prediction dataset was again expanded in 1999 with 230 front-end samples scanned using the 1 mm flowcell.
- Finally MJ samples (N=334) from all mills (1999 season) were scanned in January and February 2000 using the flowcell.

Back-end Calibration Development and Prediction

- The initial database was obtained by analysing 206 diluted C-molasses samples from all SA sugar mills over the period May 1997 to January 1998 using a 1 mm cuvette.
- During 1998, other back-end products (syrup, B-molasses, C-molasses and C-massecuite) were added to the database (N=317, cuvette).
- In 1999, the calibration was expanded with 245 samples (syrup, B-molasses, C-molasses and C-massecuite) and scanned using the 1 mm flowcell.
- Finally a prediction set, consisting of molasses samples (N=110) from all mills (1999 season) scanned in January and May 2000 using the flowcell, was developed.

Check Solution Development and Application

Nine solutions of pure sucrose (Hulett's Superwhite) were prepared (5 to 13% (g/100 g)). The samples were scanned by NIRS using the automated flowcell procedure. A calibration equation was developed in the usual way. Check solutions (10 g/100 g±0.1 g) were prepared in bulk (2 000 g). Aliquots were then sealed in plastic sachets and deep-frozen. Each time samples were analysed, a check solution was scanned to determine temperature and instrumental bias effects.

Results and discussion

Mixed Juice calibrations

Initial MJ calibration dataset MJ samples (N=562) were collected over a three year period (1994 to 1996) from all factories (Schaffler and de Gaye, 1997). After using WINISI's SCORE and SELECT, the reduced dataset (N=336) produced the calibration equations shown in Table 1.

Front-end prediction dataset with intermediate liquors (cuvette) Raw sugar factories are traditionally split into front- and back-ends. Frontend dataset contained DAC, mixed juice, clear juice and syrup samples collected in 1998/99.

Expanded front-end prediction dataset with intermediate liquors (flowcell) The application of transmission-based NIRS for sugar liquors can only be considered as a feasible alternative to routine laboratory methods if a flowcell is utilised in place of the standard 1 mm cuvette. Fifty-five CJ samples were used to show that under identical instrument - temperature conditions, the two cells produced very similar results (Anon., 1999a). To date hundreds of samples have been successfully analysed with the flowcell. The only problem that has been experienced was the occasional blockage of the sampling-line due to suspended solids in MJ from factories using mills for extraction. South Africa currently has only four such factories (NB, GH2, DL and EN). For samples from these four mills, in-

Table 1. Summary of Calibration Equations for sucrose, pol and brix in mixed juice (1994 to 1996), 336 samples were selected.

Constituent	Reference		NIRS	
	Mean	SD of mean	SEC	RSQ
Sucrose	9.57	1.27	0.05	0.999
Pol	9.45	1.28	0.04	0.999
Brix	11.42	1.34	0.05	0.999

line filters (plastic filter holder filled with glasswool) ensured trouble-free operation. MJ from diffusers never blocked the lines. The flowcell dataset contained some 230 front-end samples were scanned in the usual way.

Prediction datasets from 2000 Two sets of MJ samples (N=168 and N=166 from all mills taken late in 1999) were also scanned using the flowcell.

NIRS predictions for sucrose, pol and brix in MJ The acid test for MJ NIRS calibrations is to predict sucrose, pol and brix in these independent datasets using the original MJ calibration (1994-1996). The prediction summary is shown in Table 2.

Considering that the prediction sets were scanned one to four years after the calibration equations were developed, the prediction results were very good. The average slope was 0.950. The average RSQ was 0.980. Average standard error of prediction after bias correction (SEP(C)) was also very good at 0.22. The major concern however was that seven of the seventeen predictions were biased. The analytes with high bias have been highlighted in Table 2 (range from -0.30 to +0.40). The bias is therefore variable and was obtained for all three analytes. Two examples of NIRS predictions, one without bias and the other with significant bias, are shown graphically in Figures 2 and 3 respectively.

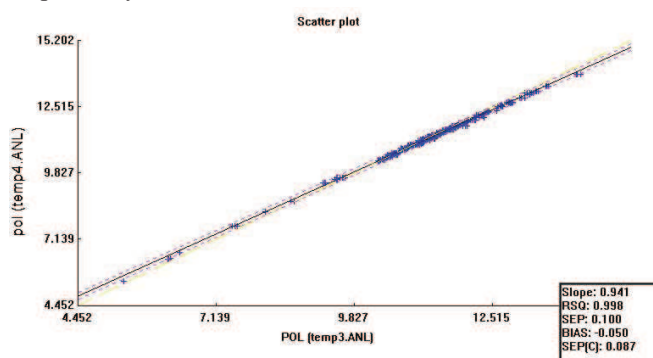


Figure 2. NIRS prediction for pol in MJ versus polarimeter. Equation was from 1994-1996. Prediction set was ESTON97. Bias was very low at -0.05.

Table 2. Summary of NIRS front-end predictions: Calibration equations were derived from 1994-1996 datasets. Predictions were obtained from ESTON97, JULY98, Front-end, FLOWCELL, JAN2000 and FEB2000 datasets.

Dataset	Analyte	SEP	Bias	SEP(C)	Slope	RSQ
ESTON97	Sucrose	NA	NA	NA	NA	NA
	Pol	0.10	-0.05	0.09	0.941	0.998
	Brix	0.23	-0.22	0.08	0.982	0.996
JULY98	Sucrose	0.32	-0.29	0.13	0.872	0.991
	Pol	0.18	0.07	0.17	0.851	0.984
	Brix	0.30	0.20	0.23	0.914	0.940
Front-end	Sucrose	0.23	0.01	0.23	0.989	0.994
	Pol	0.32	0.19	0.26	0.986	0.992
	Brix	0.51	0.43	0.26	1.012	0.994
FLOWCELL	Sucrose	0.34	0.04	0.34	1.010	0.985
	Pol	0.31	-0.01	0.31	1.000	0.986
	Brix	0.43	0.21	0.38	1.025	0.986
JAN2000	Sucrose	0.14	-0.01	0.14	0.935	0.988
	Pol	0.37	0.34	0.14	0.923	0.988
	Brix	0.21	0.08	0.19	0.956	0.971
FEB2000	Sucrose	0.20	0.08	0.18	0.916	0.973
	Pol	0.42	0.37	0.21	0.897	0.964
	Brix	0.34	-0.10	0.32	0.871	0.924
Average				0.22	0.950	0.980

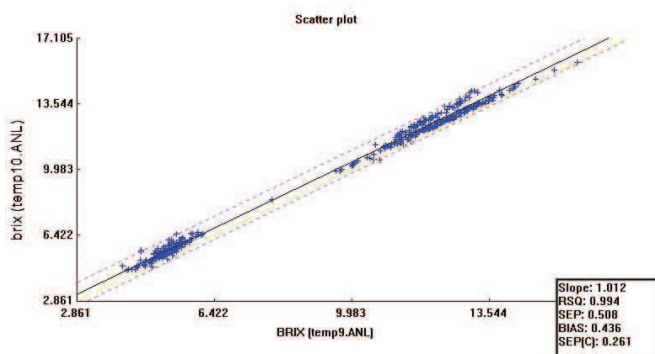


Figure 3. NIRS prediction for brix in MJ versus refractometer. Equation was from 1994-1996. Prediction set was Front-end98. Bias was high at 0.44.

This bias could be removed or reduced by adding the first 10 samples of the particular prediction dataset to the original calibration dataset. As slope and RSQ parameters were excellent, this type of systematic error could be due to changes in :

- the reference methods
- sample processing
- sample temperature
- the instrument

At this point, the concept of a stable check solution to monitor temperature and instrument changes on NIRS predictions was initiated (see check solutions).

Molasses Calibrations and Testing

Dilution of high brix liquors prior to NIRS The original brief on the use of NIRS for routine prediction of analytes in factory products was to analyse samples with as little sample preparation as possible. For this reason high brix material was poured

Table 3. Summary of Calibration Equations for seven constituents in diluted back-end products (1997 to 1999), 558 samples were selected for calibration.

Analyte	Reference		NIRS	
	mean	SD of mean	SEC	RSQ
Fructose	1.07	0.30	0.02	0.995
Glucose	0.67	0.18	0.02	0.984
Sucrose	5.96	1.89	0.08	0.998
Pol	5.48	1.99	0.04	0.999
Brix	12.78	0.50	0.06	0.987
Ash	1.79	0.41	0.03	0.994
Dry Solids	12.18	0.45	0.05	0.986

directly into a 30 mm cuvette and the samples were scanned by reflectance. Excellent results were obtained (Schäffler and de Gaye, 1997). However this approach has three disadvantages:

- in the routine laboratory, all products must be analysed using the same configuration. Therefore transmission NIRS in a narrow cuvette (1 mm) was chosen
- pouring highly viscous masseccutes into a 30 mm cuvette is extremely difficult. Breakage costs are also prohibitive
- dilution prior to analysis with a flowcell, is preferable as sampling and scanning can be easily automated.

For this reason, C-molasses samples (206 from May 1997 to January 1998) were diluted (10-20% solids) and analysed by transmission NIRS in the usual way. WINISI's SCORE programme was used to eliminate outliers.

Expanded back-end calibration dataset with intermediate liquors The initial database was extended with WINISI's EXPAND PRODUCT WITH NEW SPECTRA. Details of the new samples were:

- Back-end98: syrup, B-molasses, C-molasses and C-masseccute samples (cuvette)
- FLOWCELL99: syrup, B-molasses, C-molasses and C-masseccute samples (flowcell)

The new calibration file contained 558 samples. PLS calibration for seven constituents yielded the results shown in Table 3. The excellent correlations and low SEC results obtained indicate that there is an excellent relationship between the spectra and the laboratory results.

NIRS predictions for seven analytes in back-end products. These equations were evaluated by predicting the seven analytes in two datasets collected in January and May 2000. The prediction summary is shown in Table 4.

For front-end calibrations the purity of the samples ranged from 76 to 90%. The purity of the back-end samples tested in the two prediction sets was only 34 to 44%. As expected the prediction statistics were not as good as those for the front-end. Nevertheless, the prediction data are remarkably good. The SEP(C) was excellent (0.04 to 0.14). Average slope and RSQ statistics were 0.97 and 0.85 respectively. Once again the major concern was that significant bias was obtained for many of the predictions. Predictions with high bias have been highlighted in Table 4 (range from -0.27 to +0.30). The bias is variable and was obtained for most of the analytes. This bias could be removed or reduced by adding the first 10 samples of the prediction dataset to the original calibration dataset. This bias is similar to that experienced for the front-end juices. Again the concept of a stable check solution to monitor temperature and instrument changes on NIRS predictions was initiated (see check solutions).

Laboratory Results and Consolidation of Laboratory Data and Spectra.

NIRS predictions can be produced extremely quickly (in minutes). By contrast, conventional laboratory methods (moisture, sucrose by hplc, dry solids and sulphated ash) are normally

Table 4. Summary of NIRS back-end predictions: Calibration equations were from 1997-1999 datasets. Predictions were obtained from JAN2000 and May2000 datasets.

Dataset	Analyte	SEP	Bias	SEP(C)	Slope	RSQ
JAN2000	Fructose	0.13	-0.11	0.04	1.131	0.898
	Glucose	0.08	-0.05	0.06	0.922	0.868
	Sucrose	0.29	-0.27	0.13	0.698	0.672
	Pol	0.16	0.11	0.11	0.928	0.677
	Brix	0.30	0.28	0.10	0.949	0.921
	Ash	0.15	0.13	0.08	1.027	0.900
	Dry solids	0.22	0.20	0.09	0.955	0.928
MAY2000	Fructose	0.15	-0.15	0.05	1.195	0.880
	Glucose	0.07	-0.05	0.05	0.914	0.874
	Sucrose	0.20	-0.15	0.13	0.823	0.814
	Pol	0.25	0.21	0.14	1.108	0.897
	Brix	0.31	0.30	0.10	0.999	0.948
	Ash	0.21	0.18	0.11	0.885	0.753
	Dry solids	0.26	0.23	0.11	1.063	0.920
Average				0.09	0.97	0.85

much more time-consuming (2-5 hours). On the other hand, laboratory methods normally have very simple calibration procedures. Method development time is relatively short.

NIRS calibration is totally dependent on laboratory (reference) results. These data are obtained from different analysts and different laboratories over a considerable time period. For back-end results, incorporating seven constituents and hundreds of samples, the problem is compounded. Collating the data from different analysts in different laboratories is a time-consuming task. Data is normally consolidated in spreadsheets. The matching spectra must then be sorted and melded with the laboratory data. This process is time-consuming and prone to error. Preparation of spectra and laboratory data for calibration or validation is a serious drawback to NIRS and is often downplayed. Current NIRS software still makes use of DOS's 8 character file-naming protocol. As NIRS investigations result in the generation of many files, this limitation is a serious disadvantage.

Arguments that calibrations are generic and can be used internationally are also misleading as laboratory methods are normally not identical. The uncertainty in standardising instruments with high moisture samples is also an open question.

This study has shown that large NIRS libraries of samples, collected over several years, can produce biased results mainly due to temperature or instrument changes (see check solutions). Smaller libraries, collected over a much shorter time period, will be easier to maintain. As cane quality changes slowly with time, the smaller dataset can be updated each week with the previous week's controls.

Check Solutions

Calibrations and predictions Monitoring the accuracy of NIRS estimates is unfortunately a time-consuming exercise. Samples must be sent for conventional laboratory analyses. These results are then compared to the NIRS predictions. Differences between NIRS and conventional analyses can be due to:

- New samples with spectral differences from those used for the calibration. Here additional samples must be taken to widen the calibration database. The front-end and back-end

calibrations described in this paper have been developed from databases covering many years.

- Changes in the reference or laboratory methods. This may be due to a modification in the method, a new analyst or the use of a different laboratory. As the methods used at the SMRI are accredited and include regular inter-laboratory tests, this type of error is unlikely.
- Instrument and temperature drift. This is of concern as it is extremely difficult to monitor sudden changes in the NIRS instrument.

For cereal and fodder applications, ISI have introduced a check cell. This is a dried feed-concentrate that has been finely ground and hermetically sealed into a special cell. The user receives a spectrum file that has been scanned on the master instrument and a calibration file. These are used to transfer calibrations from one instrument to another. A procedure is included to clone the slave instrument to produce identical results to the master instrument. The use of the check cell, each day, is also an obvious answer to monitoring long-term instrument drift and bias. The ISI checkcell unfortunately cannot be used for SMRI work as:

- It can only be used in the reflectance mode. For current factory applications liquids are pumped through a flowcell and the transmission detector is used.
- The check cell contains dried material with absorbance peaks that are very different to sugar juice samples with their very high water content. In addition, the absorbances of the check cell are significantly lower (0.2 to 0.5) than absorbances of juice samples in a 1 mm cell (0.5 to 2.5).

Nevertheless the concept of a check cell is intriguing and has the potential of ensuring accurate instrument calibration. For both GC and HPLC analysis of sugars, the SMRI has been using check solutions for many years (SMRI TM 300 and 301). Check solutions are synthetic juices accurately prepared for monitoring systematic bias. This concept can be transferred to NIRS as:

- A single check solution (e.g. nominally 10%, target 10.0±0.1 g) can be prepared from high quality sucrose.

- The solution is prepared gravimetrically ensuring both high accuracy and repeatability (spot checks by GC have shown that accuracy is typically ± 0.05 units).
- By preparing the solution in bulk, small aliquots (50 ml in polythene sachets) can be frozen and thawed for testing as required.
- Once the calibration has been developed, the single check solution can be used daily or weekly to check the instrument.

Nine check solutions (5 to 13%) were used to produce a calibration equation for sucrose. The results of the NIRS calibration are shown in Figure 4.

The results are excellent and this calibration is a good example of the precision of NIRS, especially where the solutions are pure and the laboratory result (weighing) is extremely precise (bias=zero, slope=1.000, RSQ=1.000, SEP=0.01). If the instrument is out of the confidence limits, due to a breakdown, lamp change or instrument malfunction, it is now a simple matter to re-standardise the instrument using the master - slave concept (Osborne *et al.*, 1999; Flinn and Saunders, 1995; Park *et al.*, 1999; Berding and Brotherton, 1995).

This calibration equation was then used to monitor a 10% sucrose solution freshly prepared each day over a 10 day period. The ten solutions were scanned to produce 54 NIRS sucrose estimates (Table 5).

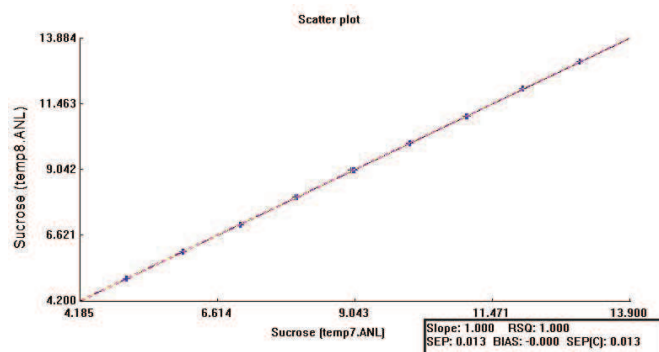


Figure 4. NIRS calibration for sucrose using 9 solutions, December 1999.

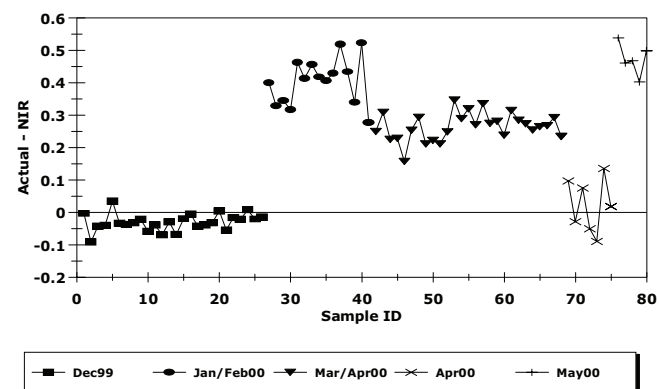


Figure 5. NIR prediction of sucrose in check solutions. Residual (Actual - NIR) for samples scanned from December 1999 to May 2000.

It is clear from Table 5 that the concept of preparing synthetic sucrose solutions for monitoring NIRS instrument accuracy is valid. The average difference between weighing and NIRS was only -0.04 units with excellent upper and lower limits.

Effect of Temperature on NIRS Predictions for check solutions
The check solutions were used over a period of five months. A check solution was run prior to analysing samples. A graph of the difference between actual and NIRS sucrose estimates is shown in Figure 5.

It is obvious from Figure 5 that the NIRS estimates for January, February, March and April were badly biased by 0.2 to 0.4 units when compared with the December 1999 results. This type of bias was typical of the accuracy problems that have been experienced over the past few years (Tables 2 and 4).

These findings are additional evidence that the bias encountered for juices and molasses was due to systematic bias probably originating from the instrument and/or temperature factors. Several reports have been published illustrating that temperature changes can influence both water peak spectral intensities and can shift WL positions (Anon., 1999b and Reeves, 1992). One way to counteract laboratory temperature variation is to use a repeatability (REP) file. The aim of a repeatability file is to make the instrument insensitive to environmental changes (*e.g.* instrument or temperature changes) whilst at the same time not reducing the accuracy of the calibration (Anon, 1996; Tillman and Paul, 1998). A repeatability file was produced for the check solution. The file contained spectra of a single check solution scanned at 19, 21, 24, 25, 27 and 29°C. This file was added during WINISI's CALIBRATION development. The addition of the repeatability file had a dramatic effect on the check solution estimates and results were virtually bias-free (Figure 6).

Table 5. NIRS Predictions for sucrose check solutions, ten solutions scanned 54 times over a 10 day period.

Statistic	Sucrose Difference
Average difference (weighing - NIRS)	-0.04
Min difference	-0.10
Max difference	0.05

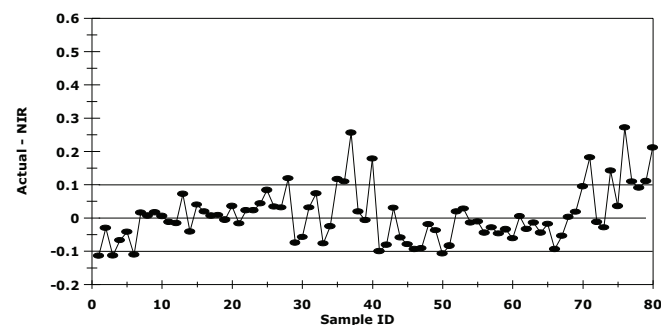


Figure 6. NIR predictions of sucrose in check solutions. Residuals (Actual - NIR) samples from December 1999 to May 2000. Repeatability file (check solution at different temperatures, 19 to 29°C) was used.

Virtually all of the 80 check solutions scanned over the 5 month period were now within ± 0.1 g/100 g target.

Effect of WL Alignment on NIRS Predictions for check solutions Before the NIRS instrument is used for collecting spectra or for routine use, the diagnostics programs should be run. One of these programs is the WL accuracy test. WINISI suggest that WL linearisation constants K and Phi should be altered from current to suggested if the current WL error is more than twice the suggested error. WL errors recorded at the SMRI have on occasions drifted from an ideal of 0.24 to 0.48 nm. A series of tests, using the check solution, was carried out, to determine the effect of changing the WL constants on NIRS bias. K and Phi constants were changed as shown in Table 6.

Changing the WL constants had a dramatic effect. NIRS bias shifted from -0.6 to +0.3. In the same experiment, the detector gain (manual) was changed from 41000 to 57000. Changing the gain resulted in more scatter, but had little effect on NIRS bias.

Effect of WL Alignment on NIRS Predictions for the MJ samples In a further experiment, MJ samples (N=33) were scanned (WL constants: K=3310.78, Phi= 0.31658, error=0.4893). Calibration for sucrose, pol and brix was carried out in the usual way (SEC=0,02, RSQ=0.999). After adjusting the WL constants (K=3309.17, Phi=0.31713, error=0.2326), the same set of MJ samples was again analysed by NIRS. The NIRS predictions of the second set, using the first as calibration equation, are shown in Table 7.

Changing the constants to reduce WL error from 0.4893 to 0.2326 resulted in significant bias for all three analytes. This type of bias is similar to the systematic error observed over the past few seasons. It would appear that these relatively small changes in WL alignment and temperature variations are the major cause of bias. Thus waiting for the actual WL error to reach twice the suggested error, as per the WINISI manual, is not good enough for precise NIRS predictions. The problem of drifting or shifting WL constants has also been reported by de Bruijn and Bout (pers. comm.) and by Berding (pers. comm.). Both groups have noticed that WL drift appears to be a problem with older instruments. It is interesting to note that when a repeatability file, containing check solutions scanned at different temperatures, was used NIRS bias improved. Temperature differences can result in WL shifts which are very similar to instrument WL drift.

Table 6. Effect of changing WL Linearisation constants on NIRS prediction errors for sucrose in check solution.

K	Phi	Error	NIR Error
3309.79	0.31712	0.4362	-0.59
3309.43	0.31708	0.2426	+0.36
3309.79	0.31712	0.4362	-0.50
3309.43	0.31708	0.2426	+0.30

Future steps to avoid the problem.

- A large sample of MJ (20 L) has been split into 400 sachets and frozen. A molasses sample has been diluted and the same procedure adopted.
- One of these sub-samples can be scanned at regular intervals to develop a repeatability file for each product.
- To ensure WL accuracy, the WL constants will be updated daily/weekly.
- If a lamp is replaced or the detector must be rebuilt, the frozen MJ or molasses samples can be used to re-standardise the instrument.
- A newer Foss 6500 NIRS, co-owned by SMRI - SA Sugar Experiment Station, will be tested for WL stability.
- Thermosetting the flowcell will also be considered.
- Less emphasis will be placed on a large representative database. Updating the database with new samples and discarding older samples seems more prudent.
- As a “dress rehearsal” for adoption in routine laboratories, the instrument will be used at the SMRI for the rest of the 2000/2001 season to gain experience in high throughput routine NIRS analysis.

Summary

- Despite the fact that calibration databases for both factory front- and back-ends contained samples collected over many years, NIRS predictions were often biased.
- The bias was variable and depended on which datasets were used for calibration - prediction.
- Incorporation of the first 10 samples of a prediction set into the old calibration dataset and re-calibrating removed the bias.
- A sucrose check solution can be easily prepared, stored and used to monitor instrument and temperature effects on NIRS predictions.
- Alteration of the WL linearisation constants have shown that older NIRS instruments exhibit WL stability problems which in turn result in prediction bias.
- Proposals to counter these systematic errors are suggested.
- The automated flowcell concept is currently being used in “dress-rehearsal” mode to monitor these new proposals.

Table 7. Summary of NIRS predictions: Both calibration and prediction sets contained the same samples. WL constants were adjusted between calibration and scanning the prediction set.

Analyte	SEP	Bias	SEP(C)	Slope	RSQ
Sucrose	0.20	0.18	0.09	0.960	0.980
Pol	0.28	0.25	0.12	0.953	0.977
Brix	0.66	0.65	0.08	0.966	0.993

Acknowledgments

The assistance of Mrs Danielle de Gaye with organisation, preparation and analysis of samples as well as data consolidation is gratefully acknowledged.

REFERENCES

- Anon. (1996) NIRS, version 4.00, Routine operation and calibration development software for NIRS/NIT instruments. Infracsoft International.
- Anon. (1998). SMRI Test Methods.
- Anon. (1999a). Annual Report of Sugar Milling Research Institute, 1998-1999, 12-13.
- Anon. (1999b). Controlling sample temperature. www.winisi.com
- Berding, N and Brotherton, G (1995). Analysis of high-moisture material - fibrated sugarcane. *NIRS The Future Waves: Proc 7th Int Conf on NIRS*, Montreal, 6-11.
- De Bruijn, JM (1997). Development and application of automatic NIRS in factory laboratories. *Zuckerind* 11 : 878-882.
- Flinn, P and Saunders, L (1995). Leaping ahead with NIRS. *NIRSG*, Melbourne 195.
- Osborne, B, Kotwal, Z, Wesley, I, Saunders, L, Dardenne, P and Shenk, J (1999). Optical matching of NIRS reflectance monochromator instruments for the analysis of ground and whole wheat. *J Near Infrared Spectrosc*, 7: 167-178.
- Park R, Agnew R and Barnes, R (1999). The development of NIRS reflectance spectroscopy calibrations for undried grass silage and their transfer to another instrument using multiple and single standardisation. *J Near Infrared Spectrosc* 7: 117-131.
- Reeves, J (1992). Influence of water on NIRS model compounds. *Making light work: Advances in NIRS*, Ed. I Murray, Weinheim, 99.
- Schaffler, KJ and de Gaye, MTD (1997). Rapid NIRS estimation of multi-components in mixed juice and final molasses: the possibility of day-to-day control of raw sugar factories using NIRS. *Proc S Afr Sug Technol Ass* 71: 153-160.
- Tillman, P and Paul, C (1998). The repeatability file - a tool for reducing the sensitivity of NIRS calibrations to moisture variation. *J NIRS Spectrosc* 6: 1-68.