

Spectral matching two different brands of NIR whole grain analysers

P.J. Clancy

NIR Technology Australia, 56 Kitchener Parade, Bankstown NSW 2200, Australia.

Introduction

NIT Whole Grain Analysers have become the standard instrument for measuring protein and moisture in wheat and barley throughout the world. The Foss Infratec has dominated the industry for many years and many organisations have developed large NIT spectral databases for wheat and barley. However there are many applications for whole grain analysis where the Infratec is either too big or too expensive. As such, there is need in the market for lower cost and portable NIT Whole Grain Analysers. NIR Technology Australia has developed the Cropscan 2000 series of NIT analysers to fill this market need.

This paper shows how the Cropscan and the Infratec can be spectrally matched and as such, how these large databases of NIT spectra can be used to develop calibrations which can be used in the Cropscan instruments.

Background

NIR Technology Australia was first commissioned to design a NIT whole grain analyser in 1996. The design criteria were as follows;

- Comparable to the Tecator Infratec
- same wavelength range, ie, 850-1050nm
- transmission technology
- PLS calibrations
- portable, <11kg, battery powered
- rugged - farmer proof
- measure protein <0.35%, moisture <0.25%

In conjunction with a panel of 30 farmers from Western Australia, we developed Cropscan 2000G(see figure 1.) which was released in July 2000. Feedback from Europe and North America led to the development of the Cropscan 2000B(see figure 2.) in 2001. These instruments use the same optical bench, however the 2000B includes a Sample Transport Module which moves a sample cell in front of the light beam and up to nine scans are collected over a window approximately 100mm by 45 mm. This allows averaging of the spectra to improve the accuracy.

Light from a tungsten halogen lamp, illuminates a sample of grains held in the Sample Cell. The light is absorbed by the protein, moisture, sugar and other components in the grains. A small proportion of the original light energy penetrates through the sample and reaches the other side of the sample cell. The emerging light is focused onto the slit of the Spectrograph, where it is dispersed into the frequency range, 720-1100nm. The separated light is projected onto a linear Diode Array Detector. Each detector element, called a pixel, senses the intensity of the light hitting it and produces a current. The amount of light hitting the pixel is proportional to the current generated. This current is amplified, digitised and sent to the computer. By switching each pixel in line with



Figure 1. Cropscan 2000G.



Figure 2. Cropscan 2000B

the amplifier, a scan of the detector is made. The NIT spectrum is computed using the following expression:

$$\text{Absorbance } I = \text{Log } (I_{oi}/I_{si}) \tag{1}$$

Where,

- I_{oi} is the original light energy at frequency i
- I_{si} is the light energy leaving the sample cell at frequency i

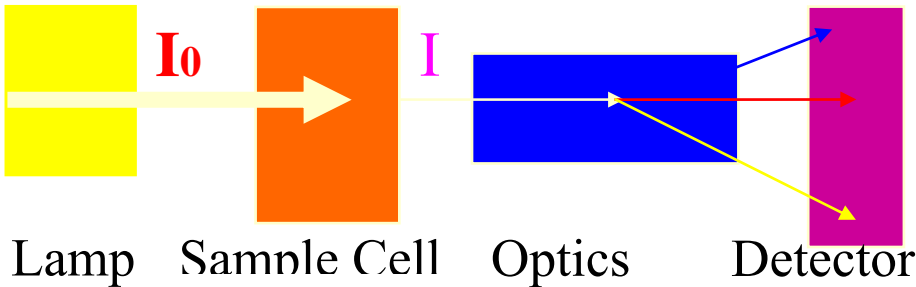


Figure 3. Schematic of the optics and principle of measurement used in the Cropscan 2000 series.

A typical NIT spectrum of wheat is shown in figure 4.

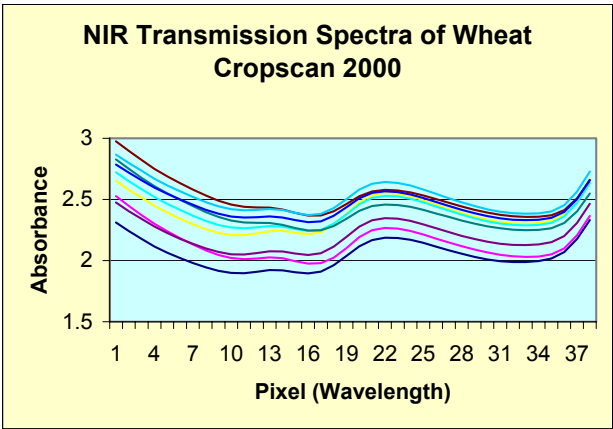


Figure 4. NIT spectrum of a typical wheat.

The Cropscan 2000 series were designed to produce comparable NIT spectra to the Tecator Infratec. The two instrument use slightly different optics. The table below compares the two systems.

Table 1. Cropscan vs Infratec Optics

Component	Cropscan	Infratec
Wavelength Range	720-1100nm	850-1050nm
Grating System	Flat Field Spectrograph	Rotating Grating Monochromator
Detector	Silicon Photo Diode Array	Single Silicon Photo Diode
Sampling System	5-9 sub samples, 45mm diam	10 sub samples, 25mm diam
Calibration Modes	PLS, MLR	PLS, ANN, MLR
Spectral Resolution	8-11nm	7nm
Data Point Resolution	38 pixels	100 data points

Both instruments are optimised to collect NIT spectra between 0 and 5 absorbance units.

Calibration challenge

To develop a NIT calibration for grains requires scanning many hundreds of samples, if not thousands. Don Lawl, showed in a presentation at the Australian NIR Users Group meeting, Horsham, 2000, that the problem with Partial Least Squares(PLS) or Multiple Linear Regression(MLR) calibrations is that there are biases and slopes relevant to each variety of grain or region in which it is grown. Figure 5, illustrates the dilemma of how linear calibrations fail to provide the optimum solution to the problem of NIT calibrations. As such, linear models such as PLS or MLR try to fit a straight line between the data points to give the best average estimate of the predicted parameter, eg, protein and moisture. It would be better to have separate calibrations for

each variety and region rather than bundling all samples into one calibration set and trying to estimate the best predictors for all samples. Artificial Neural Network methods are non linear and can resolve this problem by using pattern recognition to identify the sample spectrum and match it to similar sample spectra. Likewise LOCAL methods overcome the problem by matching the sample spectrum to spectra which are similar and then develop a PLS or MLR calibration for those samples and predict the unknown sample.

As such, ANN and LOCAL methodologies have been demonstrated to be superior to PLS and MLR calibrations for whole wheat and barley. However these methodologies require a large number of sample spectra, ie, ANN 10000 to 30000. These spectra should contain as great variability in variety, environment, region and composition as possible. The spectra should also be collected over many years.

The problem is that it may not be possible to access such a large number of samples, especially those which are several year old. The solution is to therefore develop a means of using the spectral databases collected over many years and to transfer the spectra or calibrations to a lower cost instrument, such as the Cropscan 2000.

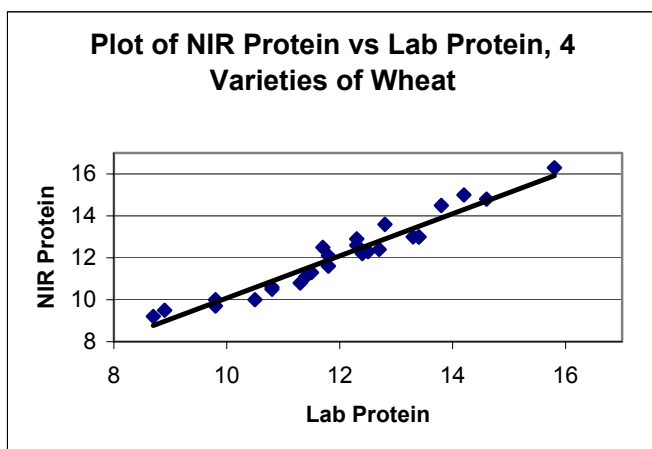


Figure 5. Failure of linear calibration methods.

Solution

We have been exploring the means of using existing databases of Infratec spectra with the Cropscan 2000 series. The first step is to match the Infratec and Cropscan spectra. Then convert the Infratec spectra into Cropscan equivalent spectra. The next step has been to use PLS to develop a calibration based on the converted Infratec spectra, and then use the calibration model on the Cropscan to predict protein and moisture in wheat and barley.

First experiment

21 samples of French wheat were scanned on an Infratec 1229 and then on a Cropscan 2000B. The Infratec spectra were reduced to 21 data points, roughly every 5th wavelength reading. The Cropscan spectra were aligned to the Infratec at approximately 910, 950 and 978nm. These points correspond to the peaks and trough in the spectra. Figure 6a and 6b shows the spectra from both

instruments. The correlation between the two sets of spectral data was greater than 0.98. A simple slope and intercept for each pixel to wavelength was computed and applied to the Infratec spectra. This converted the Infratec spectra into Cropscan equivalent spectra as shown in figure 7.

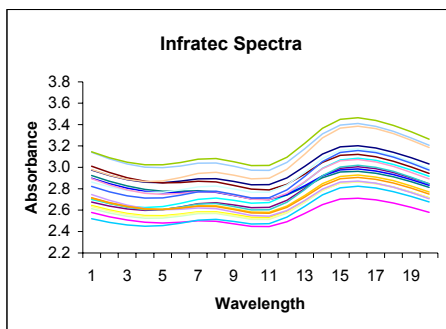


Figure 6a. Spectra of 21 samples of French wheat measured on an Infratec.

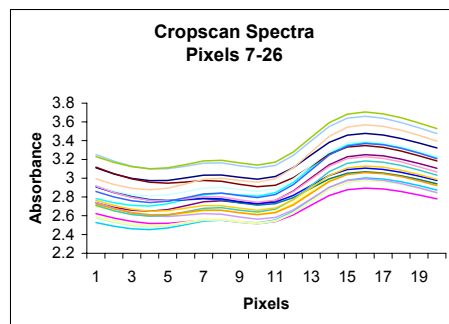


Figure 6b. Spectra of 21 samples of French wheat measured on a Cropscan.

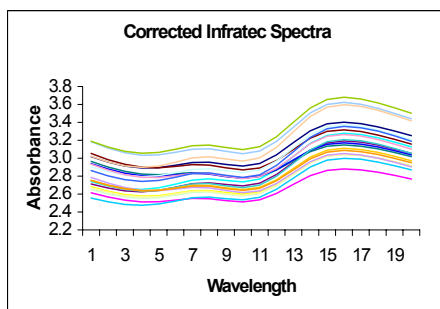


Figure 7. Spectra of 21 samples of French wheat converted from Infratec to Cropscan.

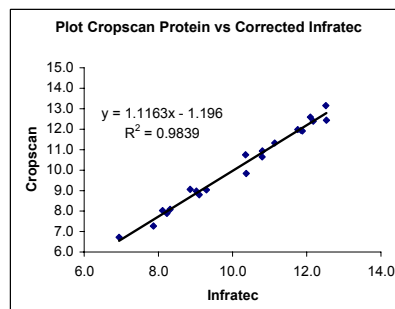


Figure 8. Comparison of results from the two instruments after correction of the Infratec spectra.

A calibration model used in the Cropscan 2000 to measure protein in wheat was applied to the 21 converted Infratec spectra. The SEP for the converted Infratec spectra was 0.34% which compares favourably with the SEP for the calibration set, ie, SEC = 0.33%. Figure 8 shows the comparison between the data from the Cropscan and converted Infratec spectra.

Second experiment

9 Austrian wheat samples were scanned on an Infratec 1229 and a Cropscan 2000B. Using the same method as described above, the Infratec and Cropscan spectra were matched and the slope and intercept computed for each of the 21 pixels to wavelengths. The 319 Infratec spectra of Austrian wheat were then converted to Cropscan equivalent spectra. PLS calibrations were developed for protein and moisture using NTAS(NIR Technology Australia Software). Figure 9 and 10 show the plots of SEC vs number of Principle Components (PC's) used in the calibrations. The Standard Error of Calibration(SEC) were as follows;

Protein = 0.21% #PC's = 12 Moisture = 0.25% #PC's = 9
 The calibration results are shown in Figures 11 and 12.

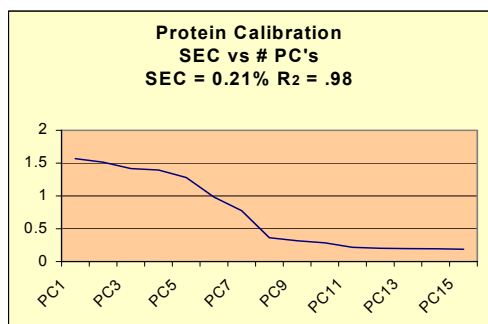


Figure 9. Protein calibration, SEC v PCs.

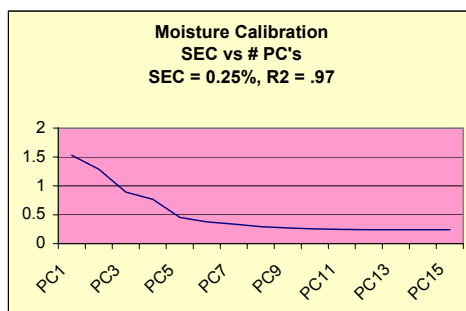


Figure 10. Moisture calibration, SEC v PCs.

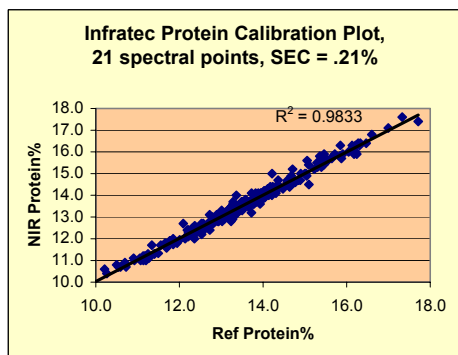


Figure 11. Protein calibration scatter plot.

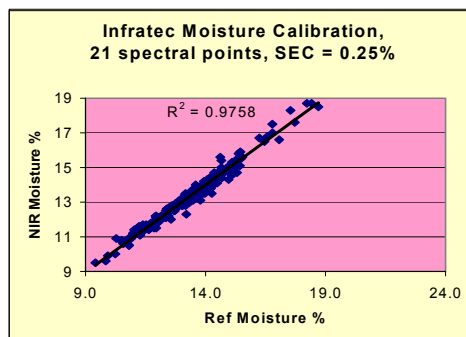


Figure 12. Moisture calibration scatter plot.

A PLS calibration on the original Infracore spectra were run through WinISI to see if there was any loss in accuracy using only 21 wavelengths in the Infracore spectra. The NTAS PLS model was almost exactly the same for protein using 12 PCs.

The PLS models were used to predict the 9 CropScan samples which were used to match the two instruments. Note that these samples were not in the 319 sample calibration set. Unfortunately there were no other samples available as a prediction set. The results of the prediction are shown in figure 13 and 14. Although the linearity for the protein data is linear, the SEP = 0.36% and for moisture with a narrow range, the SEP = 0.28%.

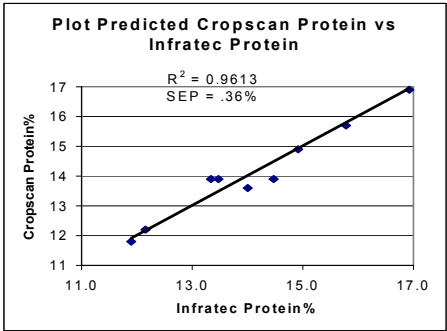


Figure 13. Comparison of protein results from Cropscan and Infratec.

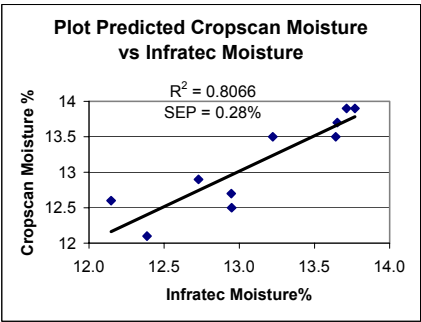


Figure 14. Comparison of moisture results from Cropscan and Infratec.

Conclusion

The data presented in this paper demonstrates the ability to spectral match the Cropscan and the Infratec. It also demonstrates the possibility of using Infratec spectral databases to develop PLS calibrations for use on the Cropscan 2000. However the objective is to use the Infratec databases for ANN and LOCAL methodologies, and thereby reduce the effects of bias and slope based on region, variety and environment. Thus far we have not been able to explore these issues, however it is considered that if the spectra from the Infratec can be converted into Cropscan equivalent spectra then it should be possible to use these databases to perform ANN and LOCAL calibrations on Cropscan 2000 instruments.