INTERNATIONAL WOOL TEXTILE ORGANISATION

TECHNOLOGY & STANDARDS COMMITTEE **NICE MEETING**

Raw Wool Group **December 2002**

Chairman: A.C. BOTES (South Africa) **Report No: RWG 06**

The use of NIR technology for predicting IWTO-19 residual ash in a commercial laboratory

By

J.L.Wear

SGS Wool Testing Services P.O.Box 15062, Kilbirnie, Wellington, New Zealand

SUMMARY

At the IWTO Barcelona meeting 2001, and at Nice 2002, separate papers were presented concerning the use of NIR technology to predict IWTO-19 residual ash content.

This paper is a continuation of that work with the objective of investigating the robustness and stability of the system now that precision and accuracy have been established in the previous papers.

The full commercial population of woolbase results at SGS Wool Testing Services for a 5-month period was analysed, and performance was monitored as per the proposed appendix K for IWTO-19.

It was found that the NIR prediction system performed well against the reference method and a new technique was introduced to objectively calculate the “high ash” cutoff point above which measurements must be performed using the reference method in Appendix D.

It is concluded that, by combining the evidence in this report with the information presented in Wear (2001) and Wear (2002), the NIR prediction method for ash is accurate, precise, stable, and robust and is a suitable alternative to the reference method for a commercial laboratory able to demonstrate competence in its use.

INTRODUCTION

At the IWTO Nice 2001 and Barcelona 2002 meetings, separate papers were presented concerning the use of NIR technology to predict IWTO-19 residual ash content.

The first paper (Wear 2001) demonstrated how a calibration using commercial data (a natural population) can be used to accurately predict ash results up to 1.5% at a precision at least equivalent to the reference method. The model was less effective at higher ash levels. It was concluded that the poor precision and scarcity of these high ash results in the calibration population were the major contributing factors.

The second paper (Wear 2002) examined a calibration which was more evenly structured and included replicate measurements to increase precision. That study showed equivalency by IWTO-0 in both accuracy and precision up to ash levels of 7%. On this basis, a submission was made to alter Appendix K of IWTO-19 to allow for prediction of ash content by NIR. Due to some late proposals to change the wording, consideration of the submission was delayed until the IWTO 2002 Nice meeting.

This paper details the use of an ash prediction equation on commercial laboratory data over a medium term period, and the performance of the system when monitored as per the proposed Appendix K text.

The major objective of this work was to investigate the robustness and stability of the system now that precision and accuracy have been established in work reported in the previous papers.

DATA COLLECTION

For a 5-month period, commercial samples measured for ash by IWTO-19 were also measured on an NIRS 6500 spectrophotometer. Over this period over 15,000 samples were measured. Ash results for these samples were predicted using a commercial calibration and the results were compared.

ANALYSIS OF PERFORMANCE

Figure 1 shows the distribution of mean fibre diameters of the samples tested in the data collection period. This is indicative of the population of New Zealand wool types during the winter months.

Figure 1 Distribution of Mean Fibre Diameters (Airflow) in the study population

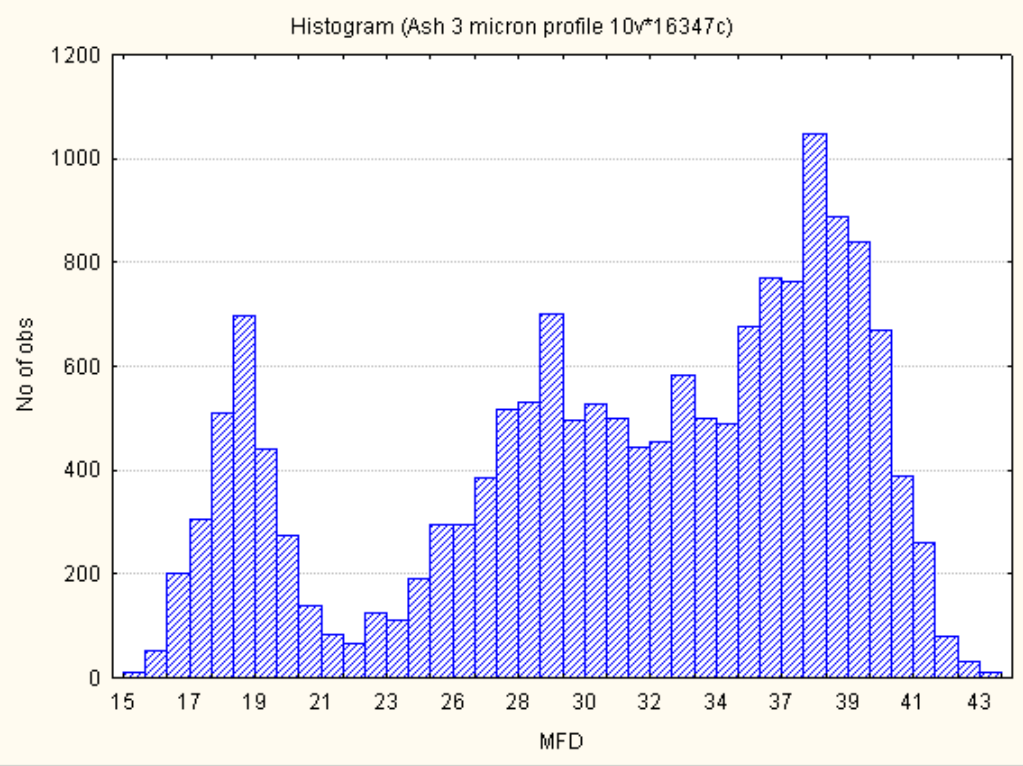


Figure 2 shows the distribution of ash results during the period of the study.

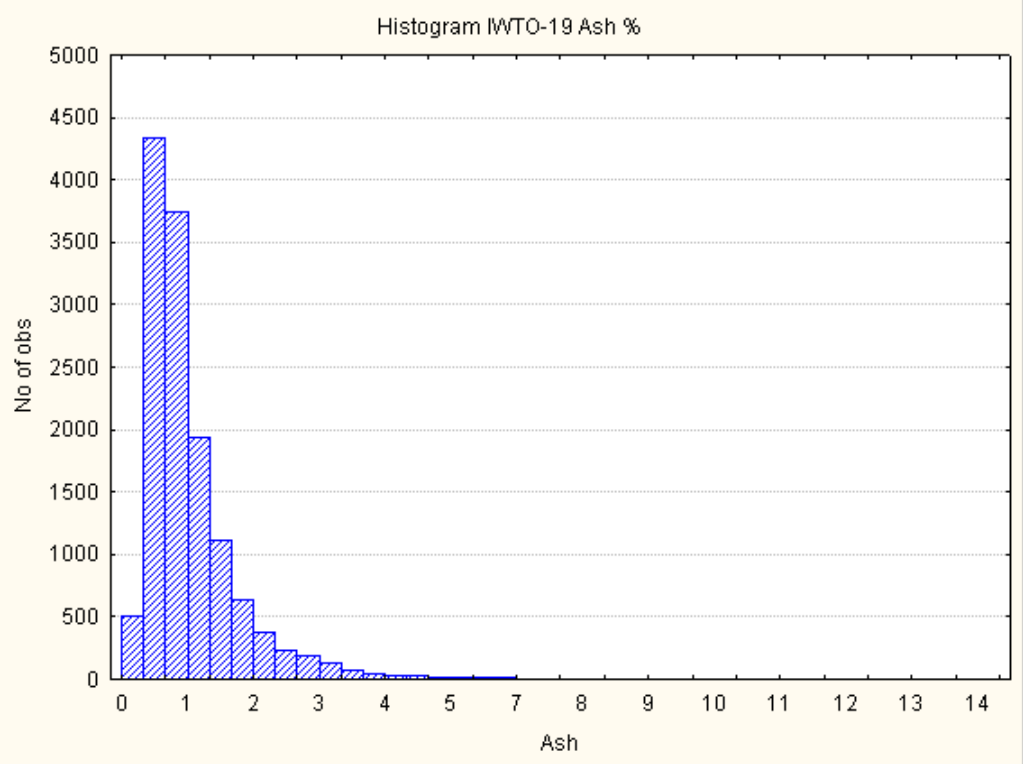
Whilst there are some results up to 14%, the vast majority of data are below 3% ash content. A comparison was performed between the commercial (IWTO-19 App. D) and NIR data, with the results summarised in table 1.

Table 1: Difference between measured and predicted results

Mean Difference	95% CL Lower	95%CL Upper	Standard Error of Prediction
0.0081	-0.0015	0.0177	0.59

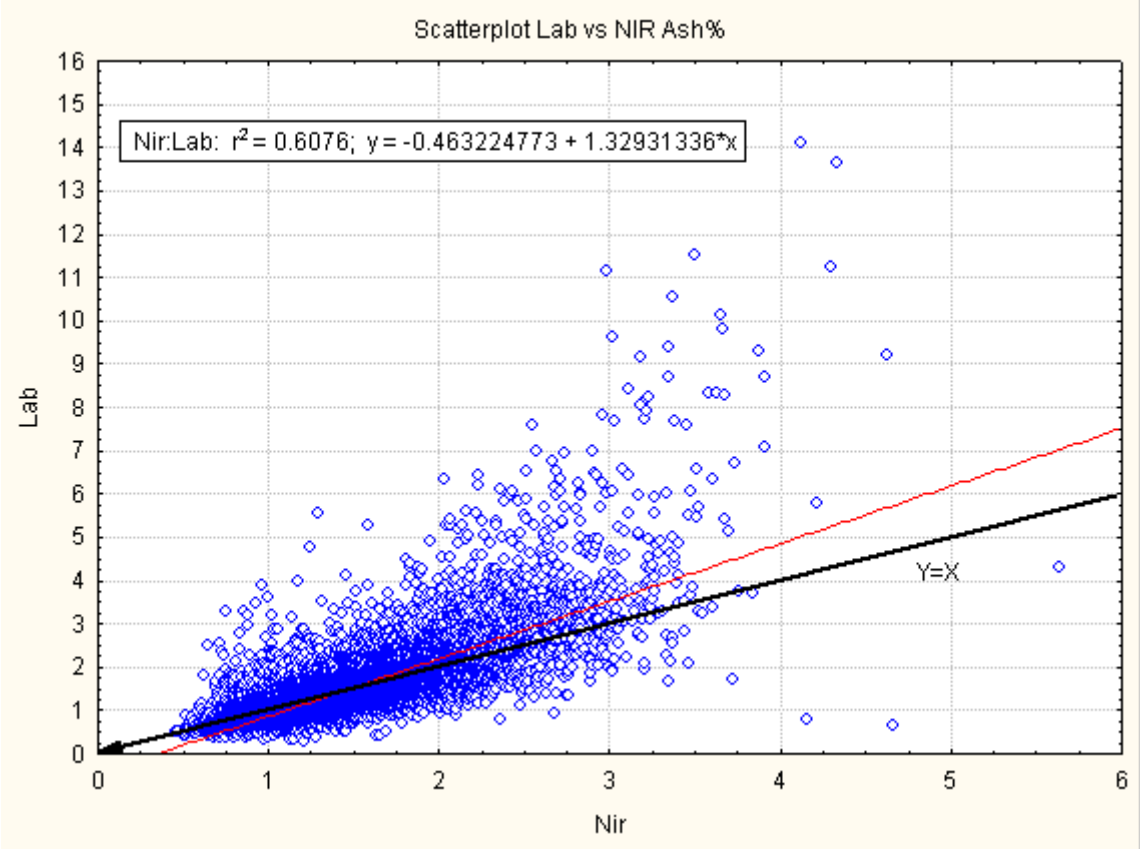
The mean difference was not significantly different to zero and the standard error of prediction was acceptable given the range of results.

Figure 2: Distribution of ash% results determined by IWTO-19, Appendix D



The IWTO-19 Appendix D results are shown plotted against the predicted (NIR) results in figure 3.

Figure 3: IWTO-19 App D versus NIR, all results



From this plot, it can be seen that the vast majority of data is below 2% and the equation predicts well in this region. However, as the ash levels increase, the precision becomes poorer and the NIR instrument tends to underestimate the ash results. This phenomenon had been apparent in the data shown in the previous two papers, and has been managed by truncating the data at a suitable level. The precision of ash measurement is highly level-dependant, as was demonstrated in the first report, where the standard deviation of replicates ranged from 0.15% at 0.8% ash to 0.45% at the 2.2% level.

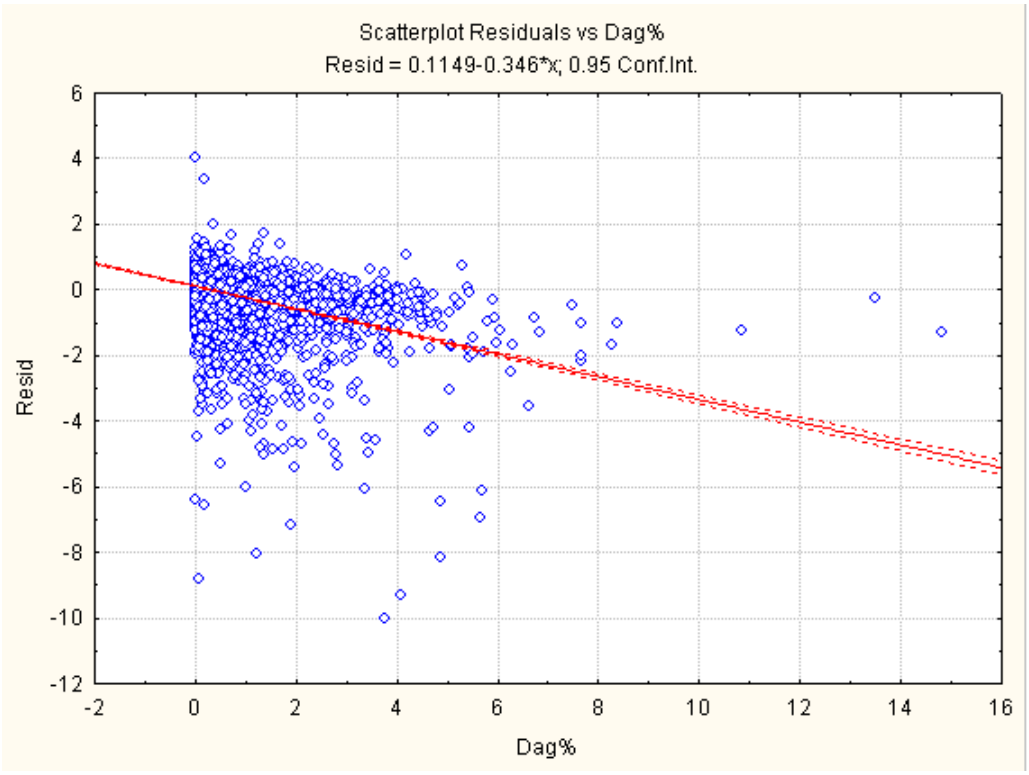
As all the yield component results were available for the data an attempt was made to determine whether any other measurable components might contribute to the error. Table 2 shows the correlation coefficients between the predicted minus measured differences and other residual components.

Table 2: Correlation matrix for Prediction errors vs. yield residuals.

Correlation Matrix						
	Burr	Hard Head	Dag	Seed	Skin	Ash
Residual NIR - Lab	-0.05	-0.03	-0.39	-0.10	-0.02	-0.83

Statistically-significant correlations were observed between the error residuals and Ash and Dag. As expected, ash was highly-significantly correlated simply due to the level-dependency seen in figure 3. Figure 4 shows the residuals plotted against Dag%:

Figure 4. NIR-Lab ash residuals vs. Dag%



The statistically-significant regression relationship shown in figure 4 indicates that samples containing Dag tend to have a higher ash content when measured by IWTO-19 App. D than when predicted by the NIR calibration. The majority of samples containing more than 2% Dag result in a negative difference (when the reference results are subtracted from the NIR results). This is most likely because Dag is generally discrete within the sample, and there is no guarantee that it will be represented either in the ash crucible or scanned by the NIR. This would also contribute to the poor precision in both reference and NIR measurements when Dag material is present.

The curvilinear relationship suggested by figure 3 is problematic in attempting to determine an appropriate cutt-of point at which the NIR predictions become unreliable. As the NIR calibration tends to under-predict, and commercially only the NIR result would be available to make the decision on whether or not a reference measurement is required, it is essential that the cutoff point is determined accurately. This cutoff point will also probably vary from country to country, depending on the environmental

conditions, animal husbandry methods, nature and degree of contamination, and method of classing of the local wools. For example, in harsh Mediterranean climatic conditions, dag may become baked hard, contain a much higher level of mineral matter, and, compared with a wetter climate, may be retained in the wool in such a manner that it survives laboratory scouring to a much greater extent.

In the earlier reports, the cutoff points were determined by eye based on studying the residual lots. This method is probably not entirely appropriate for IWTO-19 due to its reliance on the judgement of the individual concerned. For this reason a simple objective method has been devised (although undoubtedly other methods could be used). As the departure from linearity is essentially a departure from a mean residual of zero, a suitable procedure would be to determine at what level of NIR result does the mean difference become significantly, and consistently, different from zero. The following method was utilised.

1 The procedure is to be used where there is a noticeable departure in linearity between the two measurement systems.

2 The commercial data was divided into 0.1% groups based on the NIR results.

3 For each 0.1% interval, the mean residual and 95% confidence interval on the mean was calculated (reference minus NIR prediction).

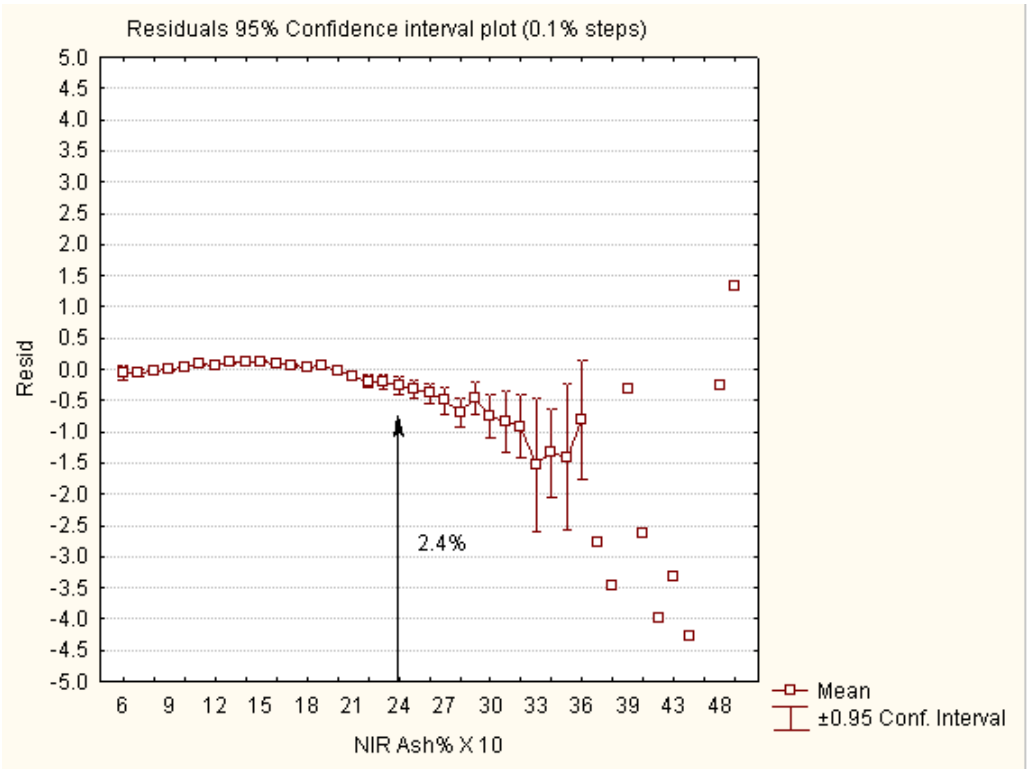
4 These were plotted as points with error bars against the group value and the points connected as illustrated in figure 5 below.

5 The truncation point is selected where there are three sequential points departing with statistical significance from the mean, increasing in one direction, and the departure is accentuated by subsequent points.

6 Significance is determined by the 95% confidence bars not crossing the zero residual line.

This method should be used with an element of common sense as statistical significance can be generated simply by the use of a large amount of data. The essential criterion is the point being determined must be at the beginning of an increasing divergence from zero.

Figure 5: Plot for determining the upper limit of linearity (cut-off level)



Note: The X-axis values were multiplied by 10 for classification purposes.

The application of this technique gives an objective criterion to determine which samples (with NIR ash results greater than the cutoff level) will require measurement by the reference method. The results from these samples can be added to future calibration data sets to improve accuracy.

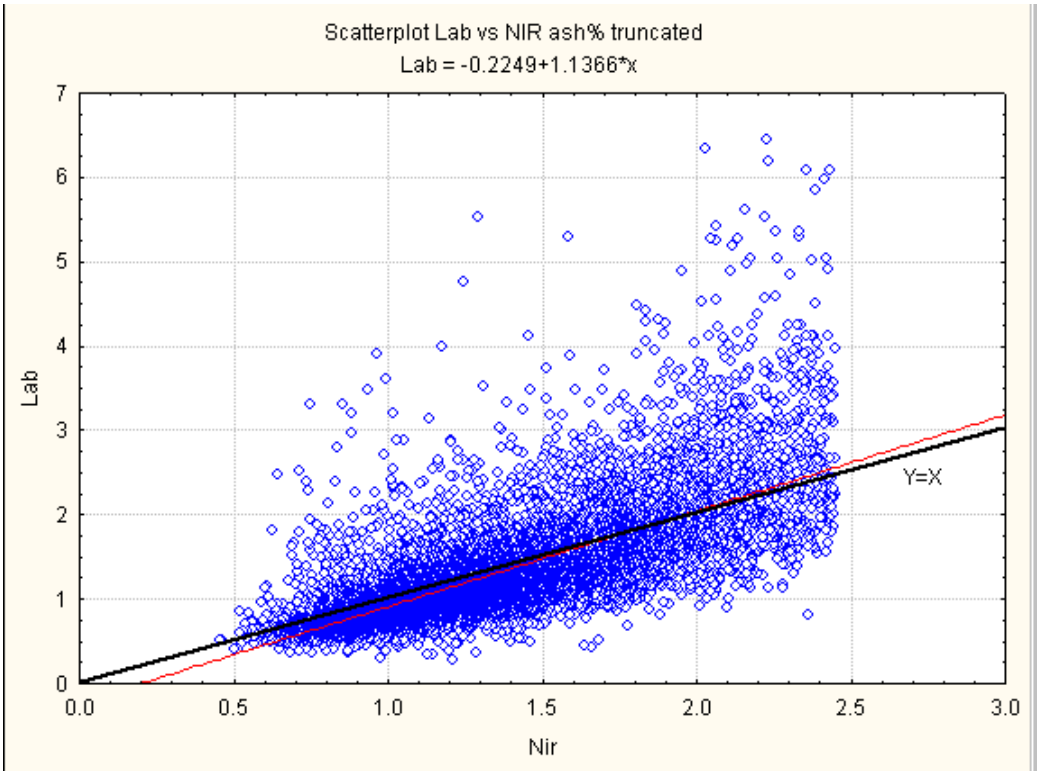
In this case, the cutoff point was 2.4%. Exceeding this limit were 576 cases or 3.7% of the commercial samples in this population. The regression was then recalculated with samples above 2.4% omitted.

Table 3: Prediction Statistics for samples <2.4%

Mean	95%CI lower	95%CI Upper	Standard Error of Prediction
0.0455	-0.038	0.053	0.49

Figure 6 shows the improvement in the calibration performance. The few samples remaining with unusually high IWTO-19 App D ash results would normally be detected by between subsample range checks, as all data in this investigation was based on single determinations.

Figure 6 Scatterplot of truncated data



MONITORING

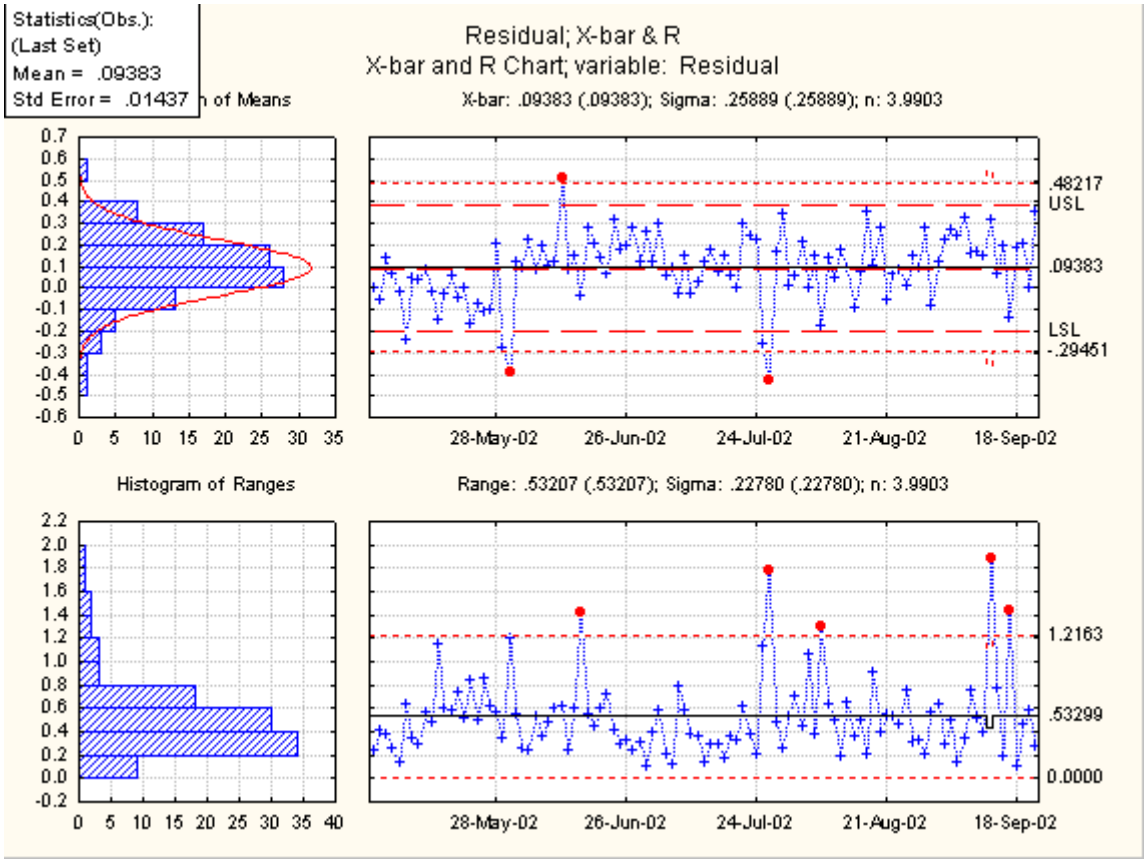
The performance of the system was monitored with the Quality Control samples currently utilised for monitoring Woolbase. This involves passing four samples of different wools through the instrument and reference method on each day. Each sample is a subset of 20 replicates, which are used on 20 consecutive working days, with one new wool being rotated into the working set each week. This allows a variety of wool-types to be used simultaneously, and a continuity of 3 out of 4 wool sets from day to day. This meets the monitoring requirements of the proposed new IWTO-19 Appendix K.

The monitoring system uses standard quality control statistical methods. The differences between reference measurement and NIR prediction were calculated for each sample and averaged for each day, and monitored together with the between-measurement ranges. These are plotted in Figure 7 as an X-bar R chart

X-bar chart. In this chart, the sample *means* are plotted in order to control the mean value of the differences

R chart. In this chart, the sample *ranges* are plotted in order to control the variability of the differences.

Figure 7 Quality control X-bar R charts for mean and range of measured and predicted ash residuals



In this example period there were three results outside the 95% confidence limits (none consecutive) which is an acceptable outlier rate. Five samples showed excessive variability, and there was no increasing or decreasing trend. The mean difference between the reference and NIR systems was 0.09% overall. The histogram of means shows the differences are normally distributed and therefore probably random errors. It can be concluded that the NIR prediction of ash results has been stable and performing well over the five month period the data was collected.

CONCLUSIONS AND RECOMMENDATIONS

Predictions for five months of commercial results demonstrated that utilising NIR technology was effective for the full array of New Zealand wool types. It was shown that accuracy issues associated with measuring extreme ash results could be addressed by objectively trimming the allowable measurement ranges. This resulted in a maximum allowable prediction value of 2.4%. Less than 4% of the samples measured were above this level and would therefore have required reference measurements to be carried out in a commercial situation. It is recommended that this objective trimming procedure be included as an acceptable technique in appendix K.

High dag content was the only woolbase parameter that could be statistically shown to effect prediction accuracy. Dag is very unevenly dispersed in samples and almost certainly causes just as many repeatability problems with the reference method

Using the monitoring criteria described in the proposed revisions to appendix K and statistical quality control charts the performance of the NIR system was tracked against the reference measurements and found to be stable for the 5 month period.

It can be concluded that, by combining the evidence in this report and the information presented in Wear (2001) and Wear (2002), the NIR prediction method for ash is accurate, precise, stable, and robust and is a suitable alternative to the reference method for a commercial laboratory able to demonstrate competence in its use. It is recommended that the attached proposed Amendments to IWTO-19 be accepted.

REFERENCES

Wear, J. (2001) *The use of NIR to predict residual ash in the IWTO-19 yield test*. IWTO, Nice, France, RWG 02

Wear, J. L. (2002) *The use of NIR to predict residual ash in the IWTO-19 yield test - 2nd trial*. IWTO, Barcelona, Spain, RWG 02

APPENDIX 1

The following are proposed text changes to IWTO-19:

IWTO-19-98 main text

6.2.3 Determination of Extraneous Materials in each Scoured Subsample

Determine the ash, ethyl alcohol extractives and vegetable matter in each scoured subsample by the methods given in Appendices D, E and F, ~~or if appropriate, K.~~

Residual ethyl alcohol extractives and ash may be predicted by the methods detailed in Appendix K, subject to the laboratory demonstrating compliance with the Essential Requirements in K3 for each parameter to the IWTO-approved accreditation authority, and reporting the results to IWTO.

APPENDIX K

Method for Determining the ~~Ethyl Alcohol Extractable Matter~~ Percentage of Residual Materials in the Scoured Subsamples using an NIRA Instrument

KI Scope

This Appendix sets out a suitable method for determining the ethyl alcohol extractable matter and ash content in a scoured subsample of raw wool for determination of Wool Base and Vegetable Matter Base. For the purpose of this appendix the methods for determining the reference measurements will be referred to as the "primary method" and results for all methods will be referred to as "the residuals".

K2 Principle

A Near Infrared Reflectance Analysis (NIRA) instrument calibrated against the primary method may be used to estimate the ~~ethyl alcohol extractable matter~~ residual matter content directly.

In cases where a result is in dispute, the primary method as described in Appendices E and/or D must be used for determining the ~~ethyl alcohol extractable matter~~ residuals.

K3 Essential Requirements

a) The test specimen shall be drawn from the scoured subsample in such a manner as to avoid any change in its ~~ethyl alcohol extractable matter~~ residuals.

b) The NIRA instrument shall be calibrated against the primary method.

c) The calibration samples must be representative of the commercial population of test samples, and exhibit the full range of variation of all wool characteristics to be expected in routine testing, unless the software utilised by the instrument allows "outlier" spectra to be identified during measurements. In the latter situation, the calibration samples must initially encompass as wide a range of variation as is feasible, and the calibration data set must then be extended to widen the sample population over a period of time.

d) The calibration of the instrument must be validated prior to use against samples that have not been used in the calibration, and, when in use, it must be monitored on a regular basis.

K4 Method

K4.1 Apparatus

a) An NIRA instrument is required.

K4.2 Procedure

The procedure is as follows:

a) Test specimens presented to the NIRA instrument must be of sufficient mass to give the appropriate packing density and minimum thickness recommended by the instrument manufacturer.

b) The NIRA instrument is calibrated against the ~~ethyl alcohol extraction~~ primary method, using at least ~~80~~ 200 test specimens covering the probable range and population structure of all wool characteristics to be encountered in routine testing, using a multiple linear regression or another suitable statistical technique.

NOTE: For the purposes of calibration only, the mean of quadruplicate ~~ethyl alcohol extraction~~ residual determinations must be used for each scoured subsample.

c) A calibration undertaken solely for ethanol extract determination shall be validated by at least ~~100~~ 200 independent samples representative of the commercial population to show a relationship between the NIRA method and the ~~ethyl alcohol extraction~~ primary method that has a slope not significantly different from one and a mean difference less than ~~0.1 %~~. 10% of the mean population value.

d) A calibration undertaken for residual ash determination shall be validated by at least 800 independent samples representative of the commercial population to show a relationship between the NIRA method and the primary method that has a slope not significantly different to one and a mean difference less than 10% of the mean population value.

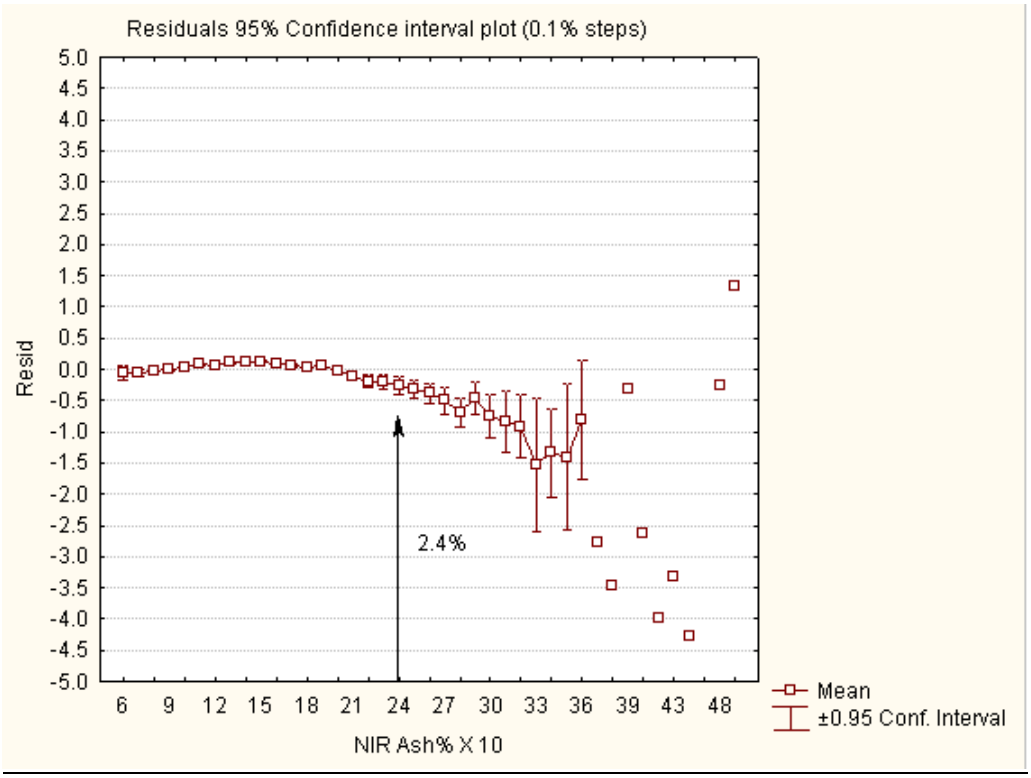
e) In the situation where high residual values lead to a curvilinear relationship between the NIRA and primary method, the data may need to be truncated to ensure that the above conditions can be met. An acceptable method to determine the truncation point is as follows, although any other suitable statistically-based technique may be used:

7 Divide the data into 0.1% ash intervals based on the NIRA results.

8 For each 0.1% interval, calculate the mean difference between the results from the primary method and the NIRA method together with the 95% confidence interval of the mean (i.e. 2 standard errors either side of the mean).

9 Plot the mean differences for each 0.1% interval in NIRA result against the appropriate NIRA level, with confidence interval bars against each mean value, as illustrated in the example below.

10 The truncation point is selected as the NIRA value where there have been three sequential points departing with statistical significance from the mean, increasing in one direction, and with the departure being indicative of an increasing trend of departure from linearity. (Significance is determined by the 95% confidence interval bars not crossing the zero residual line.)



f) The percentage of ~~ethyl alcohol extractable matter~~ residuals in the oven-dry scoured subsample, (E_i or A_i), is calculated to the nearest 0.1 % from the calibration equation.

Where the resultant value for E_i or A_i is greater than the maximum value used in deriving the calibration (i.e. the truncation point referred to in clause e), the result is discarded and the ~~extraction~~ primary method procedure ~~described in Appendix E~~ must be carried out.

g) The calibration/s must be monitored by checking at least 20 test specimens per week. Current records of the quality control checks, including the test data and calculations, must be retained by the laboratory.

K5 Record

Record the following information:

a) The percentage ~~ethyl alcohol extractable~~ residual matter in the oven-dry scoured subsample (E_i and/or A_i) to the nearest 0.1 %.