INTERNATIONAL WOOL TEXTILE ORGANISATION

TECHNOLOGY & STANDARDS COMMITTEE **BARCELONA MEETING**

Raw Wool Group **May 2002**

Chairman: A.C. BOTES (South Africa) **Report No: RWG 02**

The use of NIR to predict residual ash in the IWTO-19 yield test - 2nd trial

By

Jeremy Wear

SGS Wool Testing Services PO Box 15062, Wellington, New Zealand

SUMMARY

This paper is an extension of previous work on the prediction of IWTO-19 Ash residuals presented to the Raw Wool Group at the Nice 2001 meeting (Wear 2001).

150 samples were selected for calibrating and validating an NIR instrument over a much wider range of ash % values than previously used. Both NIR and reference measurements were performed in quadruplicate to improve precision.

The calibration was validated with 45 independent samples and the methods compared using IWTO-0. Initially there was no overall bias but a statistically significant level dependant bias was observed. This was rectified by removing the three highest samples in the validation set as they represented only 0.03% of the natural population. This is a valid option since the instrument software (or laboratory management systems) can be set to intercept any samples with predicted ash values above a specified level, and require these samples to be retested using the reference method.

A comparison was made of precision from the two methods, using IWTO-0, and found no level dependant bias, but overall the differences were statistically significant with the NIR predictions being slightly less precise than the reference measurements. This is contrary to the findings of the initial report. It was surmised that this was due to the lack of spectral diversity in this second trial, which was one of the major differences between the two studies, and this would not be an issue with a commercial calibration which would include significantly more samples.

This report recommends that IWTO-19 Appendix K be edited to allow the use of NIR technology for predicting ash % results.

INTRODUCTION

At the IWTO November Nice meeting in 2001 report RWG 02 was presented concerning the use of NIR technology to predict IWTO-19 residual ash results. Use of NIR predictive technology is allowable under Appendix K of IWTO-19 for determining the ethanol extractives, and if NIR could also be used to replace ash determination there would be significant benefits. Using NIR technology for ash determination would be beneficial by reducing energy utilisation, environmental pollution, and heath and safety risks associated with operating high-temperature furnaces. The findings of this initial report were promising and it recommended the work be repeated with fewer samples of higher precision, and more evenly distributed reference values.

This paper is an extension to the original work, following the recommendations therein and using four replicates per sample to increase precision, and a population of samples structured to cover a wide range of reference values.

APPARATUS AND EXPERIMENTAL

IWTO-19 Appendix D was followed for the reference measurements. NIR Spectra were collected on a NIRS 6500 spectrophotometer using 10g specimens, The same specimens were then used for ash determination using the standard method in Appendix D of IWTO-19. The spectra were collected between 400nm and 2300nm.

Four replicate specimens were measured for NIR spectra and reference determinations on each sample and the data averaged for analysis. This effectively doubles the precision of each determination for the purposes of the comparison.

Analysis software used was WinISI from Infrasoft International, Microsoft Excel, and Statistica from Statsoft International.

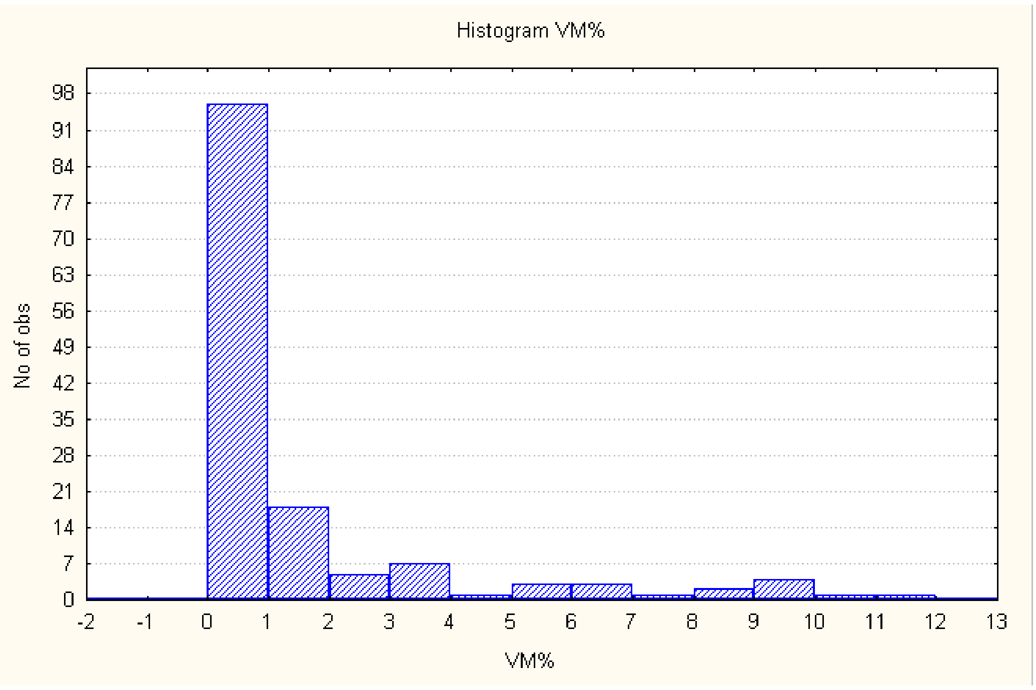
RESULTS AND DISCUSSION

SAMPLE POPULATION STRUCTURING

150 samples were selected from the last 6 months of commercial data. Due to some reserve samples being too small only 145 samples were used for the trial. The samples covered woolbase from 41.6 % to 86.2 % (mean 59.0 %), and airflow diameter from 14.8 µm to 41.4 µm (mean 27.3 µm).

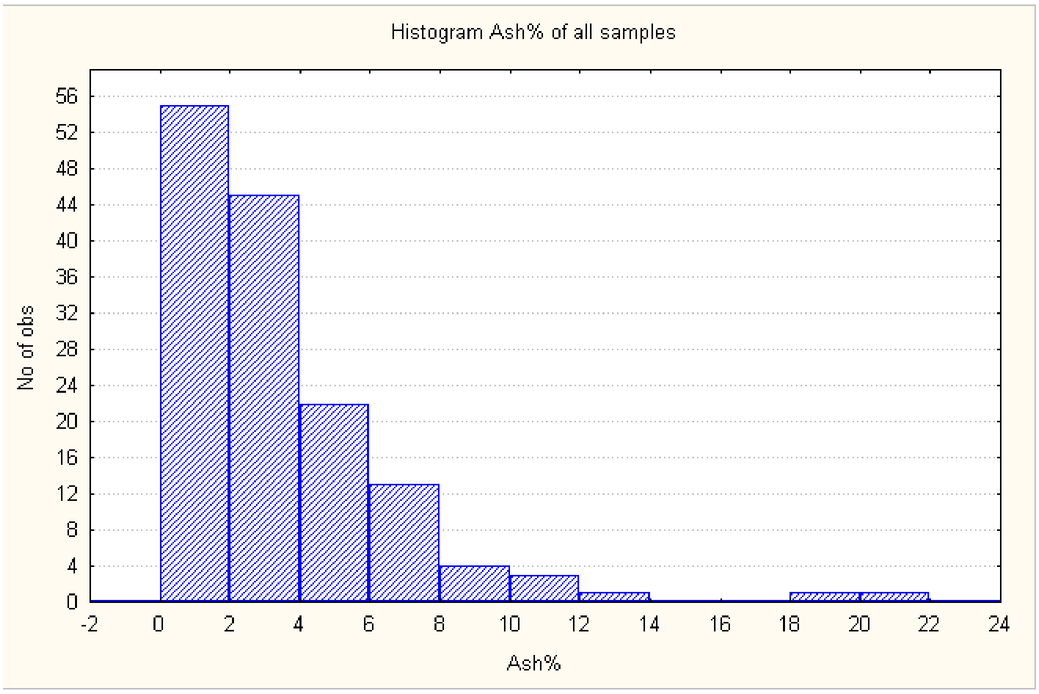
Fifteen samples were selected for exceptionally high vegetable matter content, this is because VM is a contaminant which is heterogeneous and spectrally very different to wool. These were included as samples which may cause difficulty for the predictive algorithms. If the calibration system performs well with these samples included it is unlikely high VM levels will cause errors with commercial measurements. The high VM samples are apparent in histogram shown as figure 1.

Figure 1: Histogram of VMB%



The remaining samples were selected on the basis of their Ash% results. As many high ash results as possible were included and the population was structured to be as diverse as possible - the distribution can be seen in Figure 2 below. In a natural population most samples of New Zealand wool produce ash results after laboratory scouring of less than 2%. A natural population was used in the previous paper and all of the 1200 samples in that trial had Ash% results less than 3%.

Figure 2: Histogram of ash% reference measurements

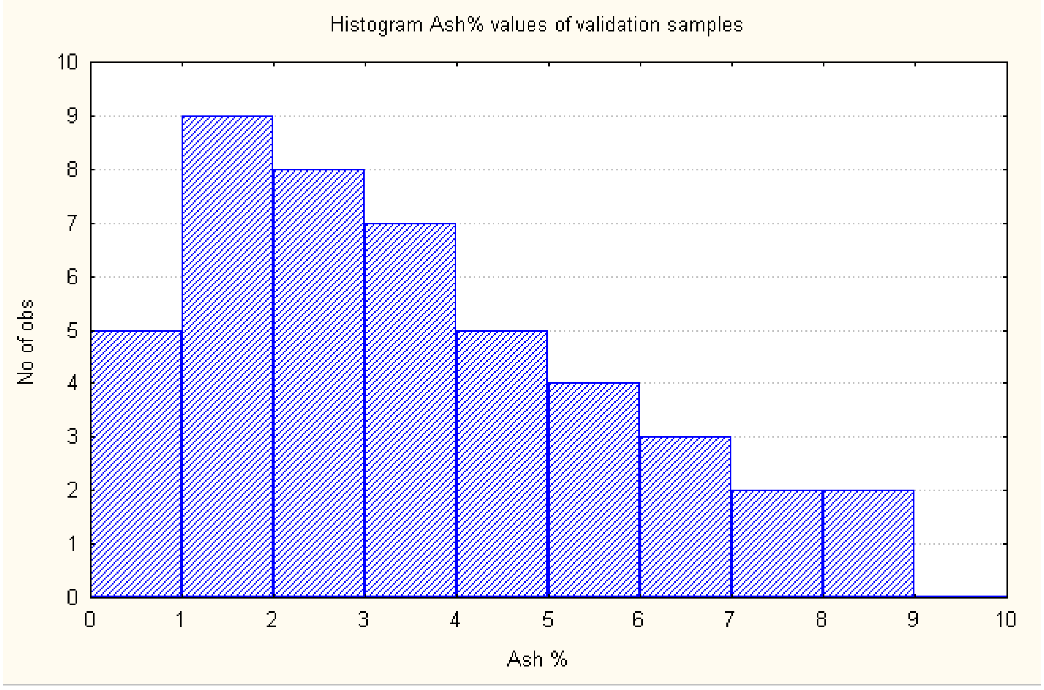


45 samples were selected for the independent validation, leaving 100 for calibration. The validation samples were selected on the following criteria:

1. Covering the range of ash% values as evenly as possible without depleting the calibration population.
2. Not being significantly spectrally different to the samples in the calibration.
3. Being within the boundaries of the bulk of the calibration population - a maximum value of 10% ash was selected for this purpose.

The distribution of ash% values of the validation set can be seen in figure 3.

Figure 3: Ash % distribution for validation set



CALIBRATION

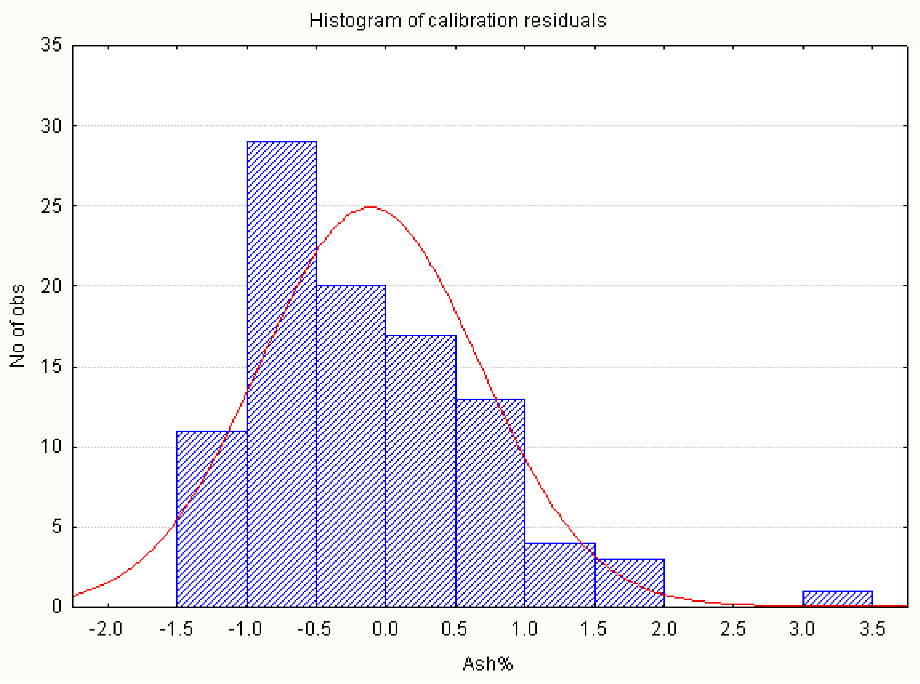
The calibration sample data was subject to partial least squares regression on principle components, and the same mathematical treatment was used as in the first report.

Table 1 summarises the calibration results, and the residuals can be seen graphically in figure 4.

Table 1: Calibration result

	Calibration result	**Result RWG02 2001**
Standard error	0.793%	0.235
R-squared %	92%	65%
Number of terms	9	9

Figure 4: Calibration residuals



The residuals are approximately normally distributed indicating that the calibration errors are random.

The standard error of calibration in the previous report was significantly less than the result of this calibration. The reason for this is the much wider range of samples used for this analysis, which has an average reference level of 3.3 % whereas the original work had an average of 0.8 %. The standard error of calibration in this 2nd trial is 24% of the mean value, whereas in the initial report it was 30%. In that report it was also shown that the precision of ash% measurement is proportional to the level of ash in the sample, so these differences are to be expected. The improvement in R-squared is also a result of the extended data range.

VALIDATION

The 45 validation samples were predicted using the calibration equation derived above and the results were analysed as per IWTO-0 Appendix B.

Overall Bias Calculations

The difference of the means can be compared simply in the form of a paired t-test in table 2 below.

Table 2: Paired comparisons - averages and differences

Statistic	Reference	NIR	Average	Difference
Average	3.428	3.475	3.452	0.047
SD	2.206	1.970	2.068	0.623
t-Val				0.509
p-Val				0.613
Significance				ns

Table 2 indicates that the means obtained were statistically similar, and there was no statistically significant overall bias.

Level dependant Bias calculations

Table 3 and Figures 5 and 6 summarise the results of the Geometric Mean and Difference versus Average comparisons of the NIR and reference methods.

Table 3: Regression summaries

Statistic		**GM**	**DVA**
Estimated Slope		0.893	0.111
Standard Error of Slope		0.037	0.045
Significance of Slope	t-Value	2.861	-2.440
	p-Value	0.007	0.019
	Significance	**	*
Significance of Correlation	R-Value	0.962	-0.386
	t-Value	23.009	2.707
	p-Value	0.000	0.010
	Significance	***	**

Figure 5: Geometric mean plot showing NIR predictions versus reference results

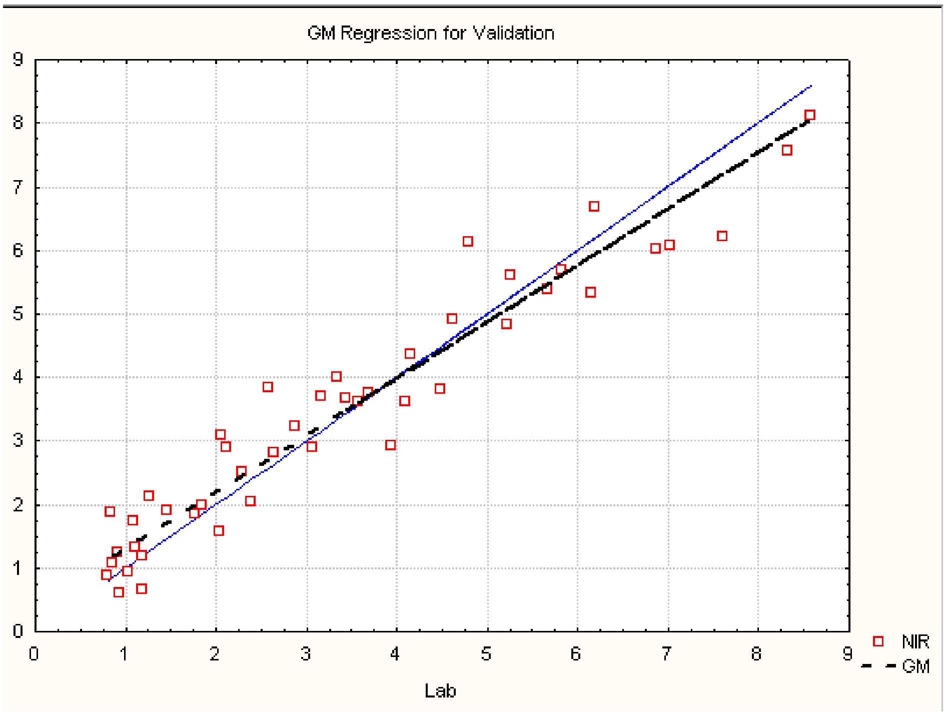
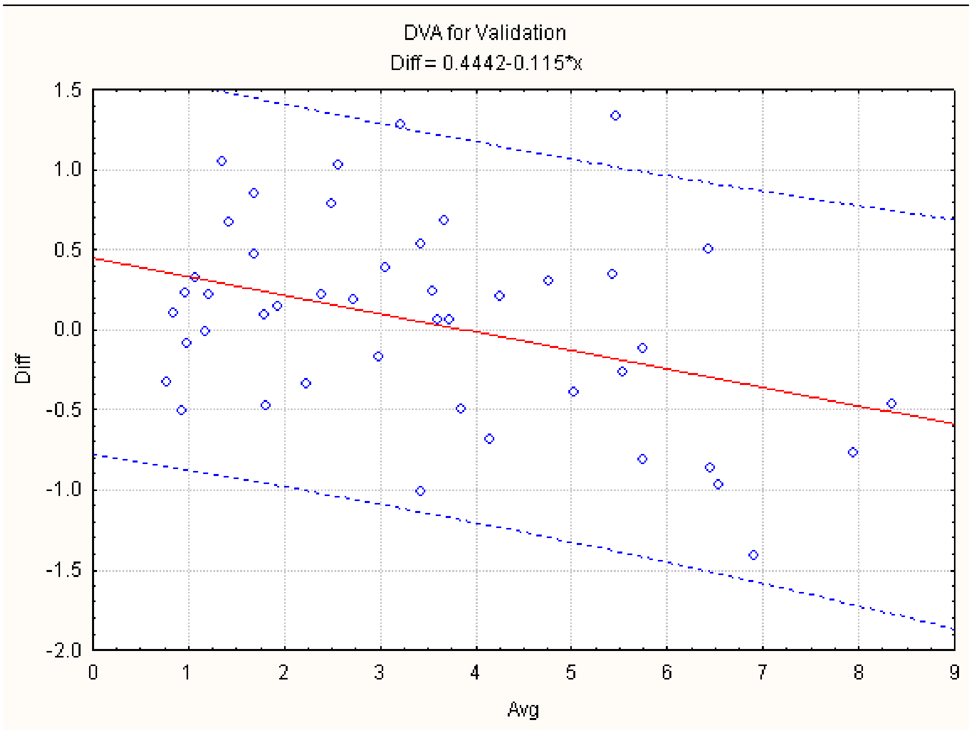


Figure 6: Difference versus average plot



Both the GM and DVA methods indicate there is a statically significant level-dependant bias at the 99% and 95% level respectively. However, the data fit on the GM plot looks adequate, and the level-dependant bias is probably caused by the three samples over 7% ash having slightly lower NIR predictions than the reference measurements. Because these samples are at one end of the regression line, and there are only 45 points, they exhibit a significant amount of leverage on the analysis. It is possible these points are below the line purely by chance, so the analysis was recalculated with these points removed. Of the large population examined for sample selection, less than 0.05% had ash residuals greater than 7%. This would be a very feasible maximum for commercial measurement and could easily be set as a criterion in the instrument software or the laboratory management system.

The correlation between NIR prediction and reference results was significant at the 99.9% level for the GM regression, indicating the relationship is well fitted. The DVA correlation is significant at the 95% level indicating a slight level-dependant bias in the residuals.

VALIDATION FOR SAMPLES BELOW 7% ASH

The analysis was repeated with samples above 7% removed.

Overall Bias Calculations for samples < 7% ash

Table 4: Paired comparisons - averages and differences (<7% ash)

Statistic	Reference	NIR	Average	Difference
Average	3.089	3.203	3.146	0.114
SD	1.856	1.725	1.768	0.581
t-Val				1.267
p-Val				0.212
Significance				ns

Again there is no overall bias.

Table 5 DVA and GM regressions (< 7% ash)

Statistic	**GM**	**DVA**
Estimated Slope	0.930	-0.075
Standard Error of Slope	0.046	0.051
Significance of Slope t-Value	1.529	-1.478
p-Val	0.134	0.147
Significance	ns	ns
Significance of Correlation R-Value	0.950	0.226
t-Value	19.188	1.467
p-Val	0.000	0.150
Significance	***	ns

The removal of the three validation samples with greater than 7% ash greatly improved the results of the comparison. The was now no level-dependant bias on either the GM and DVA. Further evidence to this effect is in the GM correlation being highly significant and the DVA correlation being not significant at the 5% level.

On the basis of this work it can be concluded that with the use of higher precision data an NIR instrument can be calibrated to predict Ash% to an accuracy not statistically different from the reference method.

PRECISION OF THE VALIDATION

It was decided to compare the precision of the NIR determinations and the reference measurements. As the methods are at the sample level (specimen level and paired observations) and replicate measurements are available it was decided that this would be best achieved by repeating the IWTO-0 analysis on the within-sample standard deviations for the two methods. Both the calibration and validation samples were included in this analysis to increase accuracy.

The results can be seen in tables 6 and 7 and figures 7 and 8.

Table 6: Analysis of absolute bias in SD

Statistics	**Reference**	**NIR**	**Average**	**Difference**
Average	0.512	0.775	0.644	0.263
SD	0.525	0.539	0.425	0.641
t-Val				4.951
p-Val				0.000
Significance				***

There was a small but statistically significant bias between the methods, indicating that the NIR predictions were of poorer precision than the reference method.

Table 7: Analysis for level-dependant bias in SD

Statistic		**GM**	**DVA**
Estimated Slope		1.028	0.043
Standard Error of Slope		0.083	0.126
Significance of Slope	t-Value	0.006	0.339
	p-Value	0.995	0.735
	Significance	ns	ns
Significance of Correlation	R-Value	0.276	0.028
	t-Value	3.434	0.339
	p-Value	0.001	0.735
	Significance	***	ns

There was no evidence of a level-dependant bias between the standard deviations of the methods.

Figure 7: GM regression showing the SD of the NIR prediction versus the SD of the reference method

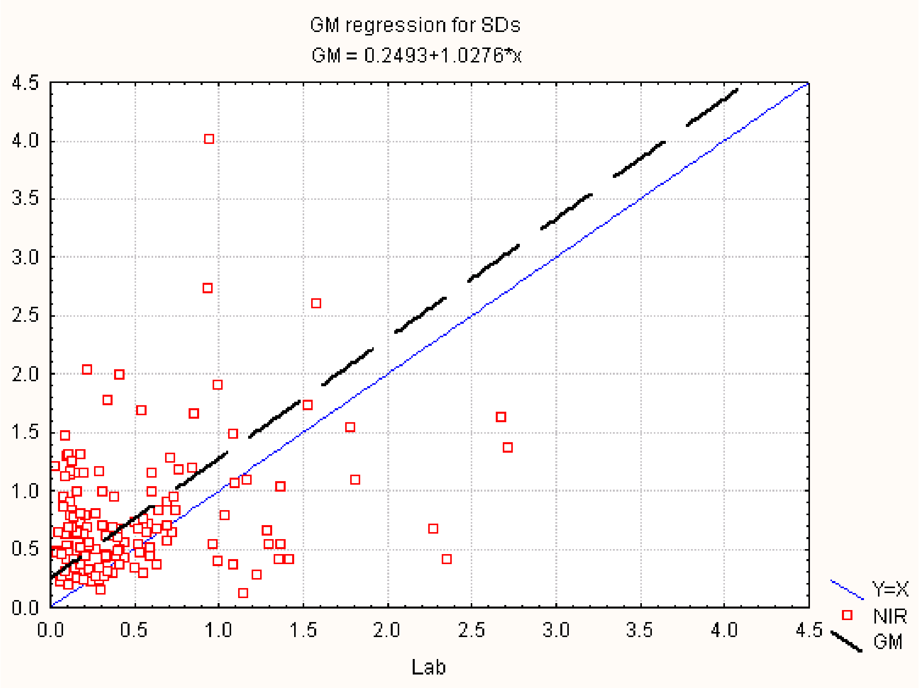
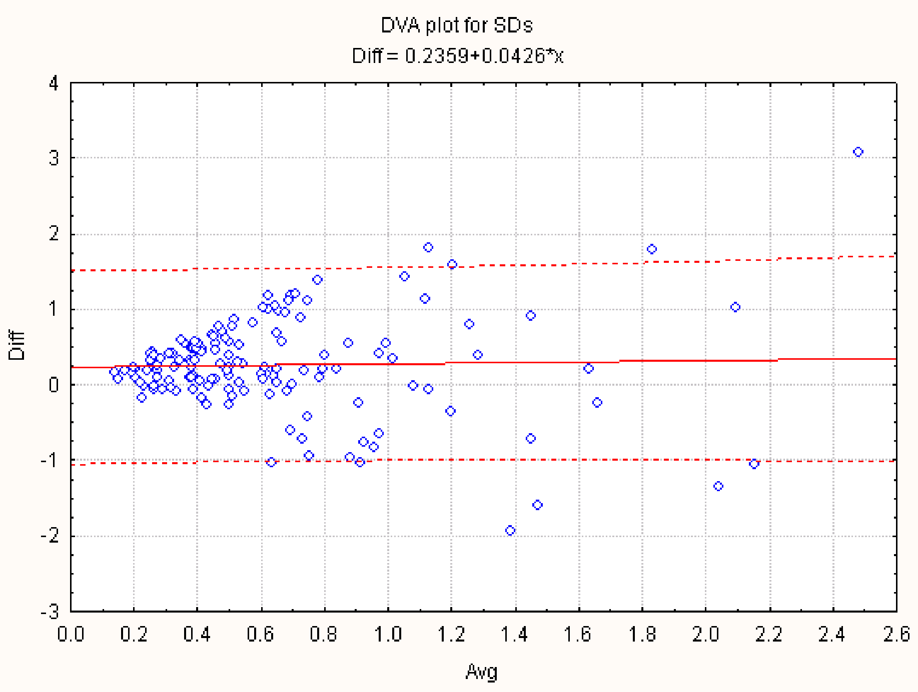


Figure 8: DVA regression for SD



In the previous paper the NIR predictions were of superior precision to the reference method. The reason for this is almost certainly a greater spectral diversity in the first trial. A rule of thumb for NIR calibration is there should be at least ten times as many samples as there are regression terms. There were 9 terms in this current calibration and 100 samples, whereas in the previous calibration there were the same number of terms but 1287 samples. It should be recalled that the samples in the first trial were selected on the basis of spectral diversity, whereas in this trial the samples were selected for diversity of reference results and vegetable matter content. It is anticipated that a commercial calibration would produce results of at least equivalent precision to the reference method, because of the larger calibration population which would be required.

CONCLUSIONS & RECOMMENDATIONS

The results detailed in this and the previous report show that an NIR instrument can be calibrated to provide Ash% results which are statistically equivalent to the results from the reference method at levels less than 7%. This covers 99.97% of New Zealand wools.

The NIR predictions were of slightly worse precision in this trial, but it is anticipated on the basis of the first trial that this would be improved with a larger scale calibration required for commercial measurement.

On the combined findings of the two reports it can be concluded NIR analysis is a suitable alternative to ash determination by combustion.

It is recommended that Appendix K of IWTO-19 be modified to allow the measurement of ash residuals by NIR. It is recommended that the minimum number of calibration samples required by Appendix K be increased to 200 to improve the precision of the calibration for ash. Proposed Text changes are detailed in Appendix 1 (which includes minor change to the text of 6.2.3 of the main text to specifically reference Appendix K).

The quality control monitoring detailed in this appendix should also be applied to the ash prediction results. Appendix K already requires quadruplicate determinations for the calibration and validation, as was used in this report.

REFERENCES

Wear, J.L. (2001) *The use of NIR to predict residual ash in the IWTO-19 yield test*. IWTO T&S Committee, Nice, Report RWG02

IWTO-19-98:*Determination of Wool Base and Vegetable Matter Base of Core Samples of Raw Wool.*

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.

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 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 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) The calibration is validated by at least 100 independent samples 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 %.

d) 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, the result is discarded and the ~~extraction~~ primary method procedure ~~described in Appendix E~~ must be carried out.

e) The calibration/s must be monitored by checking at least 20 test specimens per week. Current records of the quality control cheeks, 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 %.