INTERNATIONAL WOOL TEXTILE ORGANISATION

TECHNOLOGY & STANDARDS COMMITTEE **BUENOS AIRES MEETING**

Raw Wool Group **May 2003**

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

Measurement of IWTO-19 Ash Content by Near Infrared Reflectance (NIR) Analysis

By

D. J. Petrie, J. J. Lidgard, J. W. Marler and A. H. M. Ireland

Australian Wool Testing Authority Ltd PO Box 190, Guildford, NSW 2161, Australia

SUMMARY

The prediction of ash content of laboratory-scoured core samples utilising Near Infrared Reflectance Analysis (NIRA) has been investigated. Modified Partial Least Squares (MPLS) Regression was found to underestimate ash content when the sample being tested contained significant quantities of dag. The underestimation was not a consequence of saturation of the NIRA detector but rather appeared to be due to an inability of the MPLS technique to adequately account for dag which was present in the sample but masked by wool.

Application of Artificial Neural Networks (ANN) Regression to the calibration data set produced improved results. The underestimation at higher ash levels was not as evident, indicating that ANN is better able to utilise the spectral information to predict total ash content.

High levels of dag were found to adversely affect the repeatability of the IWTO-19 method for determining ash content. Uneven distribution of dag within samples was believed to be responsible. This finding has implications for NIRA, as any method of prediction can only be as good as the reference method to which it is calibrated.

INTRODUCTION

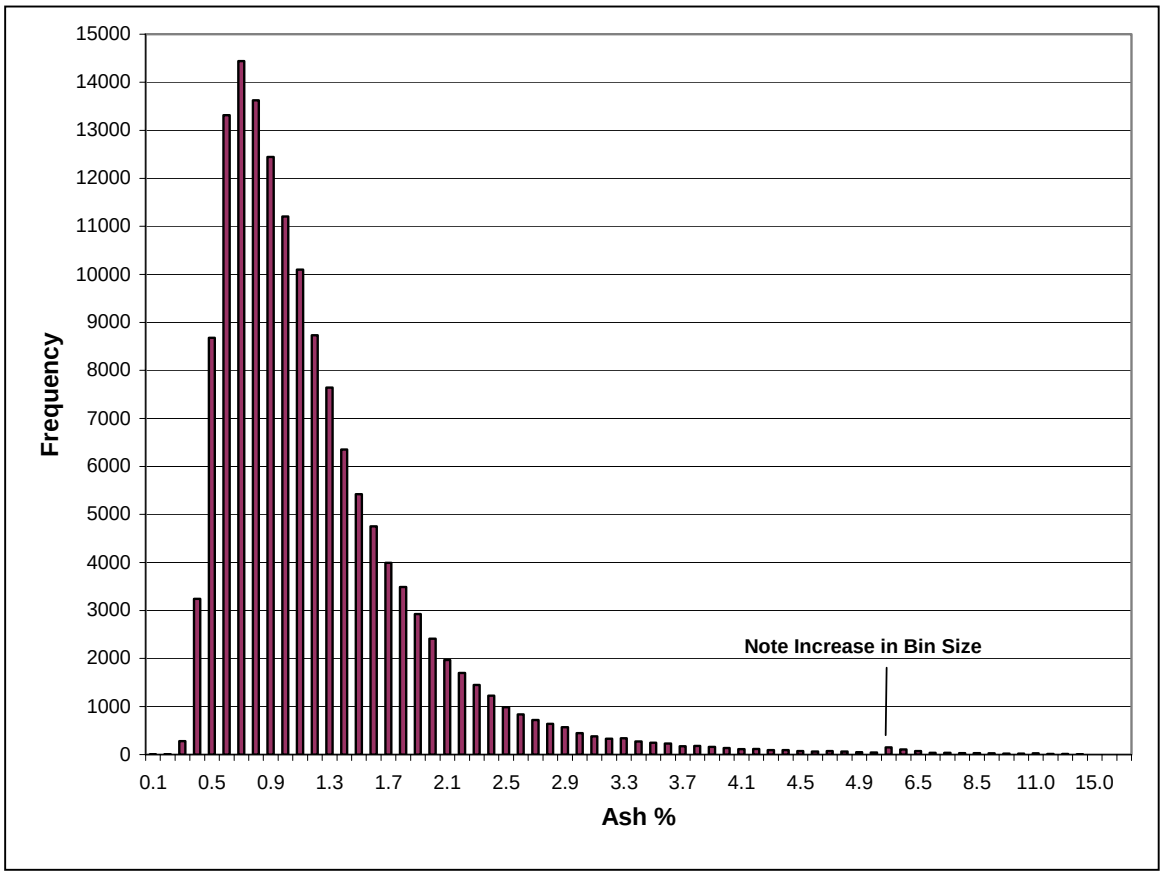
Calibrations that satisfy the requirements of IWTO-19 Appendix K have been successfully developed for the measurement of ethyl alcohol extractable matter using Near Infrared Reflectance Analysis (NIRA). This methodology is now in routine commercial use [(1)]. NIRA has been successfully applied to predict the ash content of flour mill streams [(2)].

A series of papers investigating the extension of this technology to IWTO-19 ash content of New Zealand wools have been presented at recent IWTO conferences [(3,4,5)]. Though the technology is promising and would provide considerable benefits over the conventional method for determining the ash content, concerns were raised about the apparent non-linearity of the relationship between ash content determined by the current laboratory and NIRA methods that necessitates a cut-off point. The data reported also displayed variable precision that warrants further investigation.

RANGE OF ASH CONTENTS FOUND IN NORMAL COMMERCIAL PRACTICE

An analysis of one years commercial laboratory IWTO-19 ash results, (LABash), showed 95% of results were within 0.0-2.5%, 99% of results were within 0-4% LABash and the remaining 1% of samples could measure from 4% up to 15% (see Figure 1).

Figure 1. Distribution of LABash results over one year



SAMPLE SELECTION FOR CALIBRATION

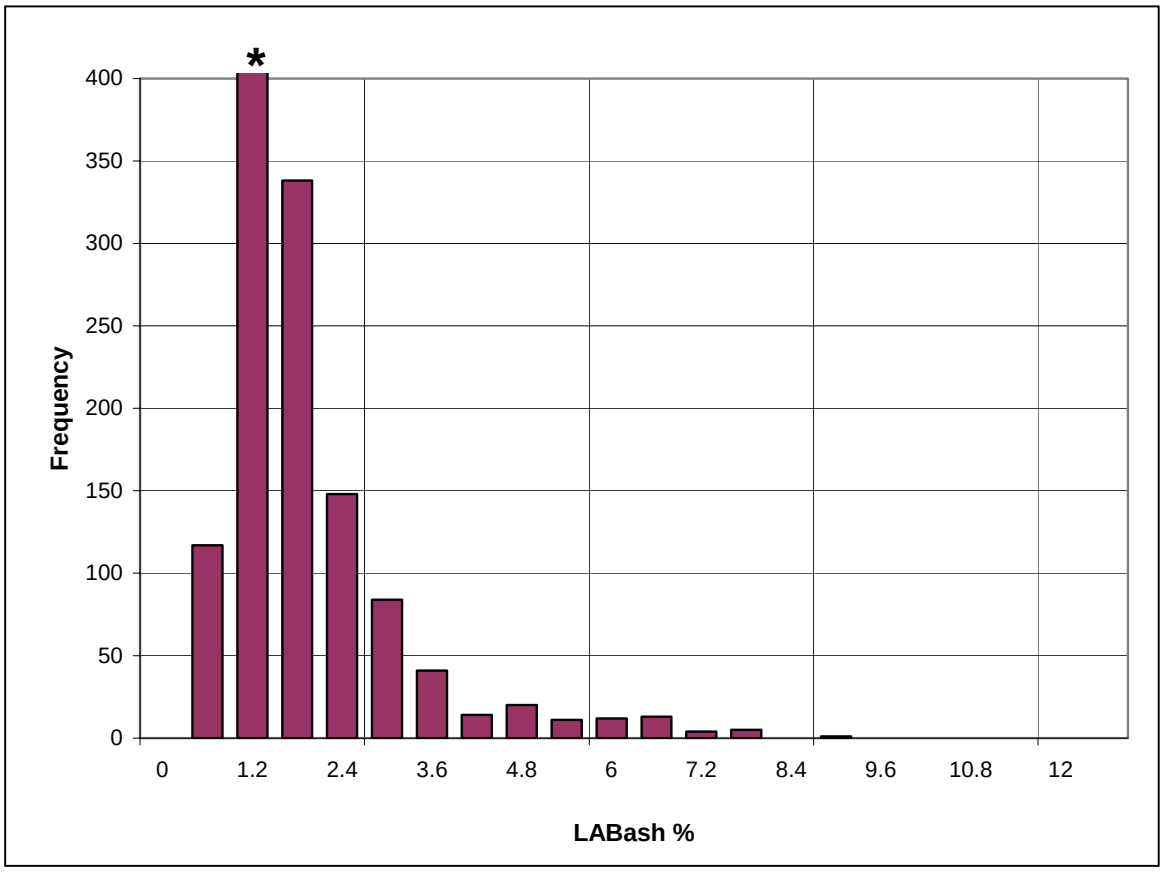
The general approach to calibrating NIR instruments has been to select samples that cover the normal measurement in a manner where they are evenly distributed over the range (i.e. the same number of samples at 1%, 5% and 10%). This is commonly referred to as a “boxcar” distribution.

For this report the population used for the calibration was selectively compiled to distribute samples evenly over the LABash range so that low and high ash percentages were equally represented as best as possible. Different properties such as the percentage of Vegetable Matter (VM), Mean Fibre Diameter (MFD) and the percentage of dag were also evenly distributed. Samples with unusually high LABash, VM, MFD and dag were selected from keepers to fill the high end of the boxcar.

With so many parameters to cover in the calibration set, it became clear that it was not possible to maintain the boxcar distribution for any one parameter. The final selection was based on ensuring there was a wide range of all the parameters listed earlier while maintaining as wide a range in LABash but relaxing the requirement for a true boxcar distribution.

The distribution of LABash results for the calibration set is shown in Figure 2.

Figure 2. Distribution of LABash results for the calibration set



Note: * Actual total for this column was 688

A second population was created from one week of commercial testing, and is referred to as the commercial set. The distribution of all parameters within this set reflected what would best be considered as the "normal" weighting across all the measured parameters.

The NIR calibration can only be as good as the conventional test data that is used as a reference value. In order to limit the impact of outlier results on the calibration data set, samples with large between subsample differences for LABash were removed from the population. Three LABash intervals and corresponding tolerances, that reflected increased variance at higher LABash, were used as follows:

LABash Interval	Maximum Difference
Less than 2%	0.5%
Between 2% and 4%	1.0%
Greater than 4%	1.5%

Table 1 displays the final population ranges for each of the main wool characteristics.

Table 1: Population Ranges

Data Set	Samples	Ash%	MFD	VM%	Dag%	Grease%
Calibration	1496	0.3-8.8	15.6-40.0	0.1-24.8	0-23.4	0.6-3.3
Commercial	1971	0.3-5.1	14.7-34.5	0.1-22.7	0-19.9	0.6-2.0

CALIBRATION BY MODIFIED PARTIAL LEAST SQUARES REGRESSION (MPLS)

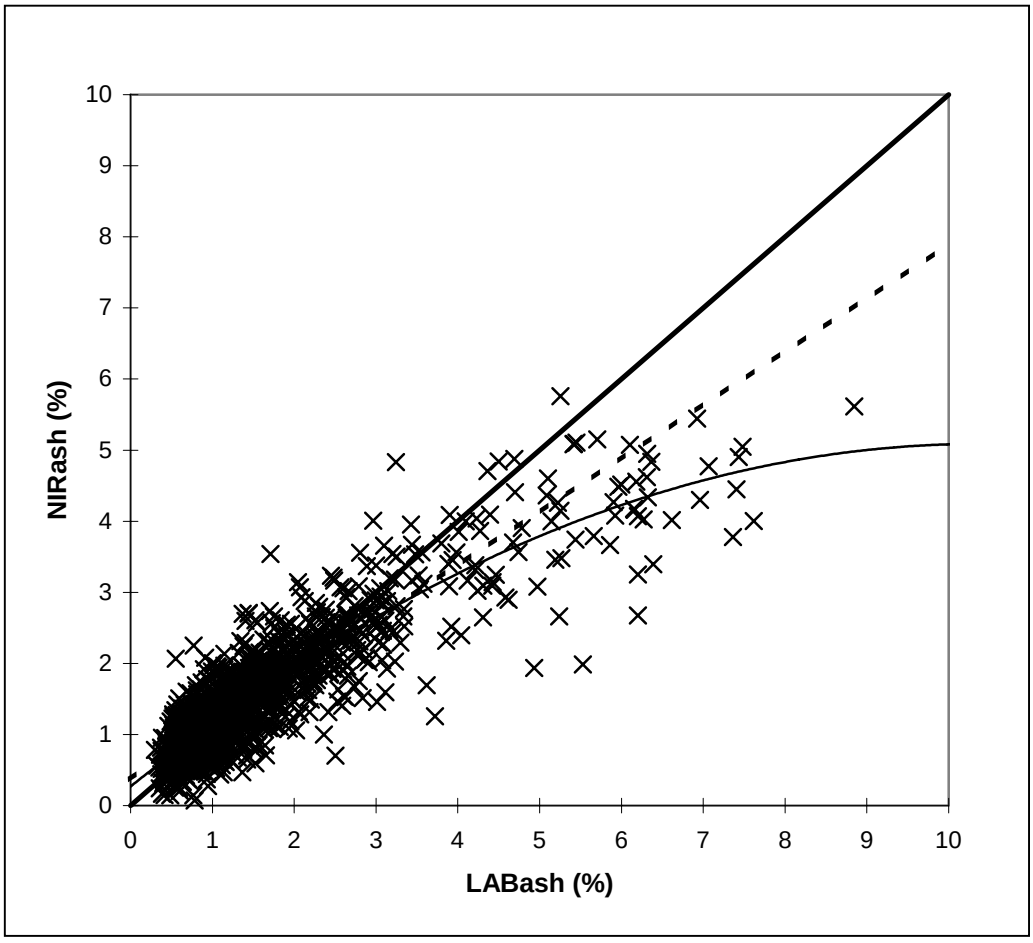
Calibration equations were created using WinISI software version 1.02 by MPLS analysis. Zeroth, 1st and 2nd derivative calibration mathematics were compared using wavelengths between 1108nm–2200nm at intervals of 8nm. The default settings were used as recommended by WinISI. Equations using zero, one and two outlier passes were compared.

An equation's performance was judged by its ability to predict both the calibration data set (i.e. the population that was used to derive the calibration) and an independent validation set (i.e. the Commercial data set). Geometric mean regression procedures as described in IWTO-0 were carried out for the comparisons of ash content measured by NIRA (NIRash) and LABash.

Equations based on 2nd derivative calibration mathematics were often found to perform better than 1st or 0th derivative equations. The correlation and standard errors were similar between the different mathematical treatments. The slope of the IWTO-0 trend lines for the NIRash plotted against LABash data was usually found to be between 0.7-0.9, with an offset between 0.3-0.4.

Figure 3 displays the best equation created from the boxcar population, based on 2nd derivative mathematics, predicting the population used for calibration. As LABash increased NIRash was increasingly underestimated and precision became poorer. Plotting a quadratic trend line through the data highlighted the curvilinear nature of the relationship. The curvilinear relationship evident in Figure 3 is consistent with previously published data [5].

Figure 3. LABash versus NIRash – Uncorrected equation

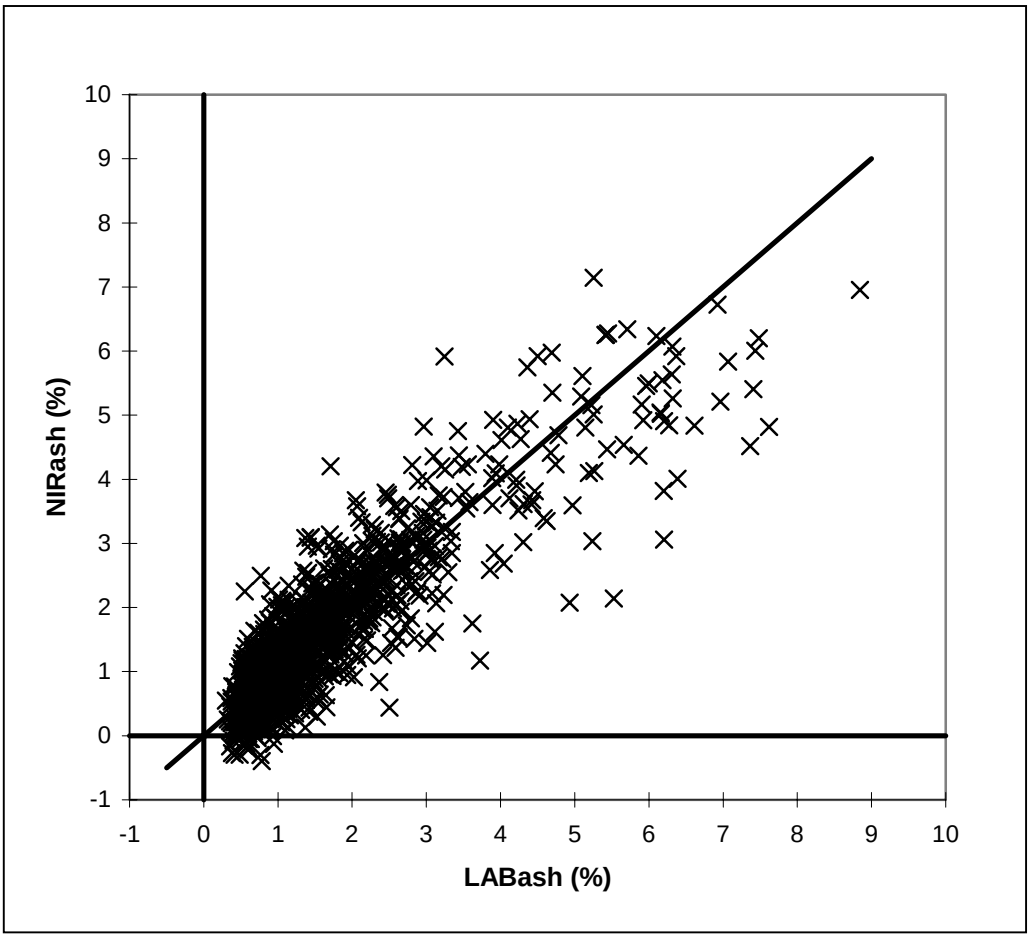


All treatments examined exhibited the same trend as presented in Figure 3.

Slope and bias corrections to calibration equations are commonly used in NIR calibration models to adjust for deficiencies in the calibration process. A linear correction was applied to correct for the slope and bias offsets.

Figure 4 displays the equation corrected for slope and bias. Note that the IWTO-0 trendline has been forced onto the ideal line (Y=X).

Figure 4. LABash versus NIRash – Corrected equation



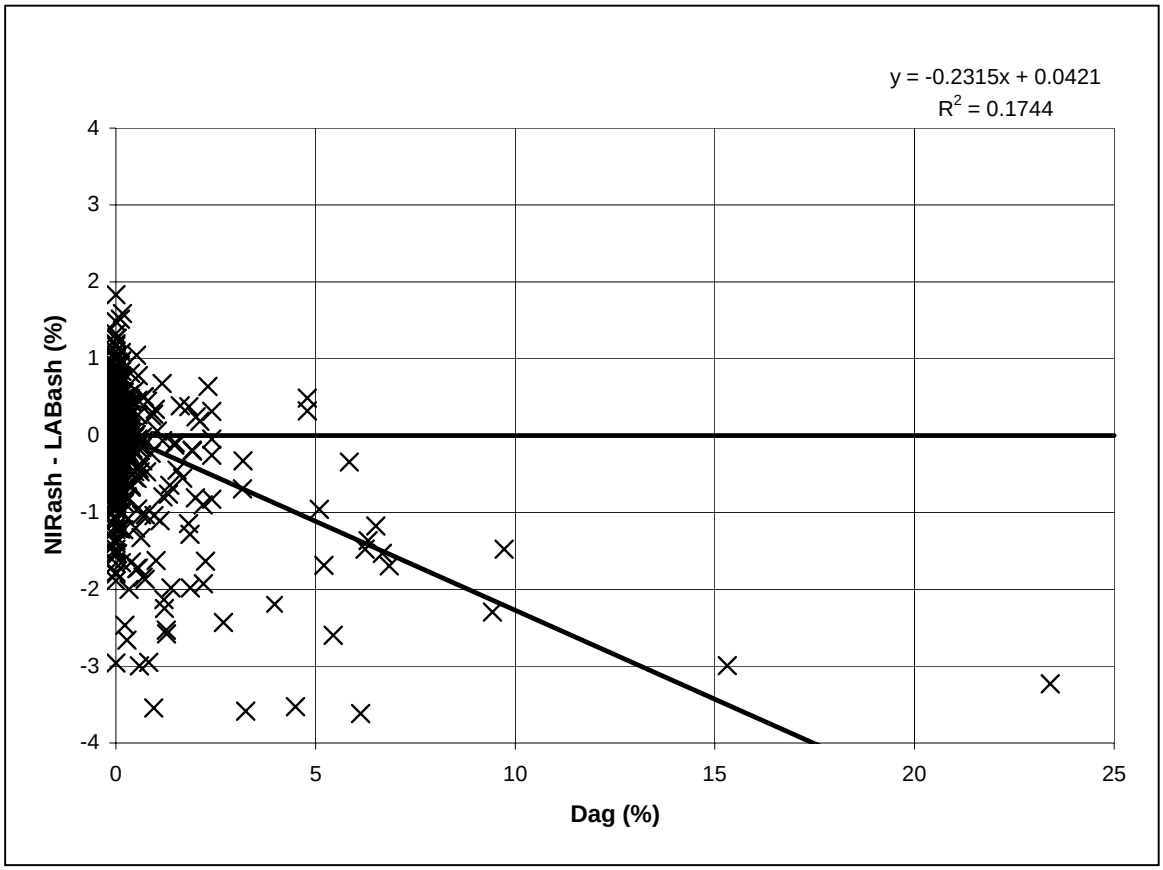
However, on closer examination of Figure 4 one can see that the corrected calibration equation has a number of shortcomings. A number of samples have NIRash values below zero. There is also a tendency to under predict samples with high ash values.

Using linear slope and bias corrections did not resolve the calibration problems in this case.

THE IMPACT OF DAG ON THE MPLS CALIBRATION

Plotting the difference between NIRash and LABash, for the uncorrected calibration model, against a number of different sample characteristics showed that dag percentage was the only parameter significantly correlated with the underestimation of NIRash (see figure 5).

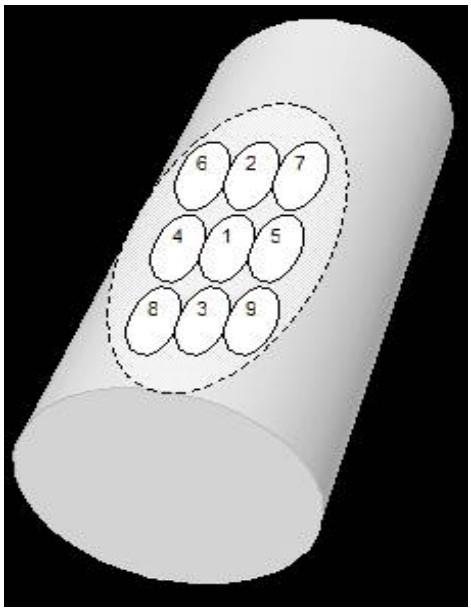
Figure 5. NIRash – LABash Residuals versus Dag Content



RESPONSE OF NIRA TO DAG

One possible cause of the curvilinear underestimation of ash by NIRA was postulated to be a detector saturation problem at high levels of dag. To assess this possibility a 150g subsample of greasy wool, containing a low level of VM and no dag, was scoured and dried to produce a clean sample in the shape of a log.

Nine large pieces of dag were selected from high dag samples and weighed. The area covered by these nine pieces was roughly equal to the area covered by the brightest area of the NIR light (see diagram below).

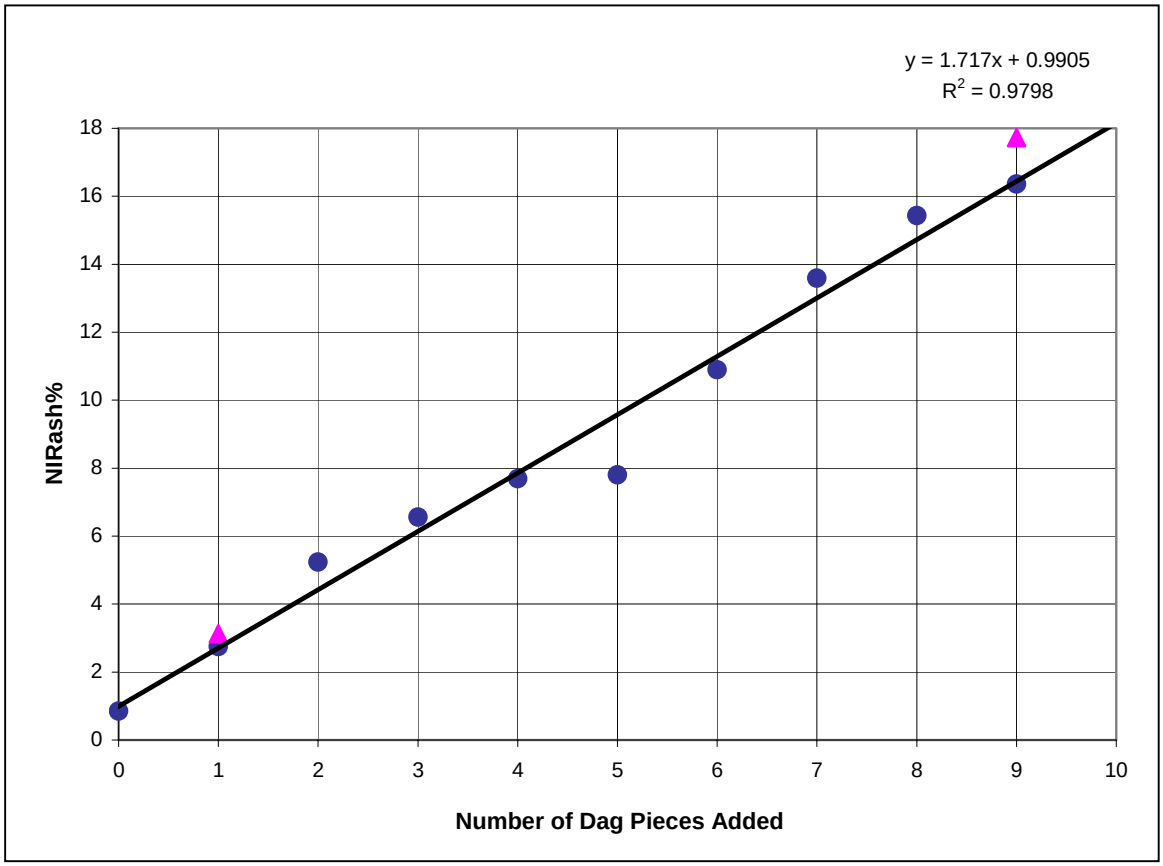


An NIR measurement was taken of the clean sample. Dag pieces were then placed onto the log one at a time starting in the centre of the bright NIR region. An NIR measurement was taken as each dag was placed onto the log. A repeat NIR measurement was taken as the first and last piece of dag was in place.

All dag was removed from the log, recording their numbers and positioning, and the clean sample re-measured. Each dag was then crushed and spread out over the bright NIR region. An NIR measurement was taken as each dag was spread over the log. A repeat measurement was taken on the first and last increments.

The effect of adding dag in pieces is shown in Figure 6. The NIRash result increased proportionately with the amount of dag added. The initial measurement of the clean log was included in the regression, shown as the 0th piece.

Figure 6. NIRash vs Number of Dag Pieces Added



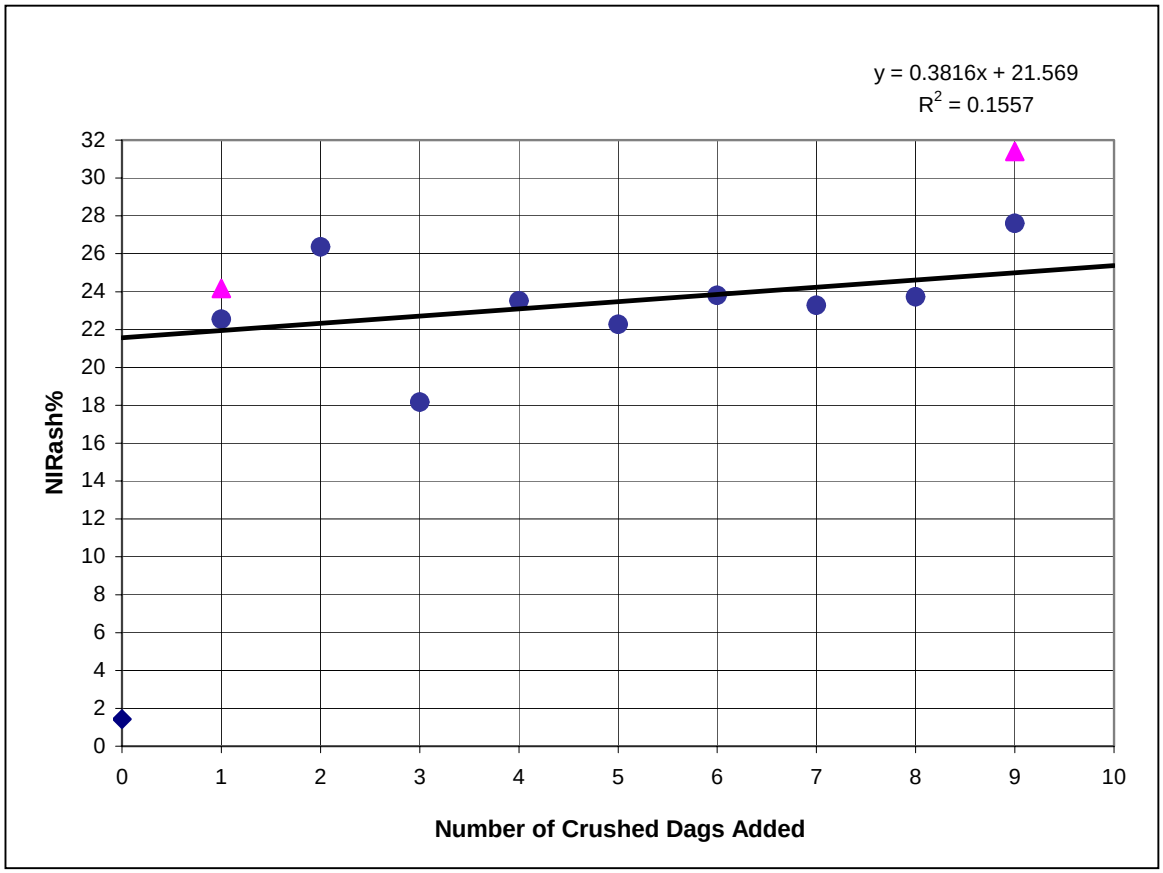
Note: Triangle points indicate repeat measurement.

This result demonstrated that the instrument detector does not suffer a 'detector saturation' problem in its current position and height above the sample.

The effect of adding crushed dag is shown in Figure 7. Measuring the first crushed dag gave a very high NIRash result. Following measurements at increased dag levels did not significantly increase this result, indicated by the relatively small slope of the trendline, which excludes the initial clean measurement.

As most of the bright light region was covered after the first crushed dag was added, the additional crushed dags only increased the thickness of the crushed layer of dag. As the measurement did not significantly increase, this would indicate that the NIR is unable to measure through dag.

Figure 7. NIRash vs. Number of Crushed Dag Pieces Added



Note: Triangle points indicate repeat measurement.

One can conclude from these small trials that the curvilinear trends observed in earlier studies were not arising from a detector saturation problem. The NIR instrument was capable of detecting dag as it was added to the surface of a relatively clean wool sample. The results for the crushed dag samples indicate that the material that constitutes dag is opaque to the NIR region of the spectra.

IMPACT OF DAG ON LABASH

The adverse impact of dag on NIRash measurement, using an MPLS calibration, has been clearly established. In order to determine the impact of dag on the LABash method, 12 samples ranging from 0.3% to 12% ash were selected. Two 150g subsamples were taken from each subsample and scoured and dried in accordance with IWTO-19. Each subsample was then subdivided into as many 10g specimens as possible for LABash measurement. The number of specimens ranged between 6 and 10 depending on the subsample mass. The Standard Deviation (SD) of LABash and average LABash were calculated for each subsample.

The Standard Deviation (SD) of LABash increased as the average LABash increased, as indicated by the trend line in Figure 8

Figure 8. LABash SD versus Average LABash

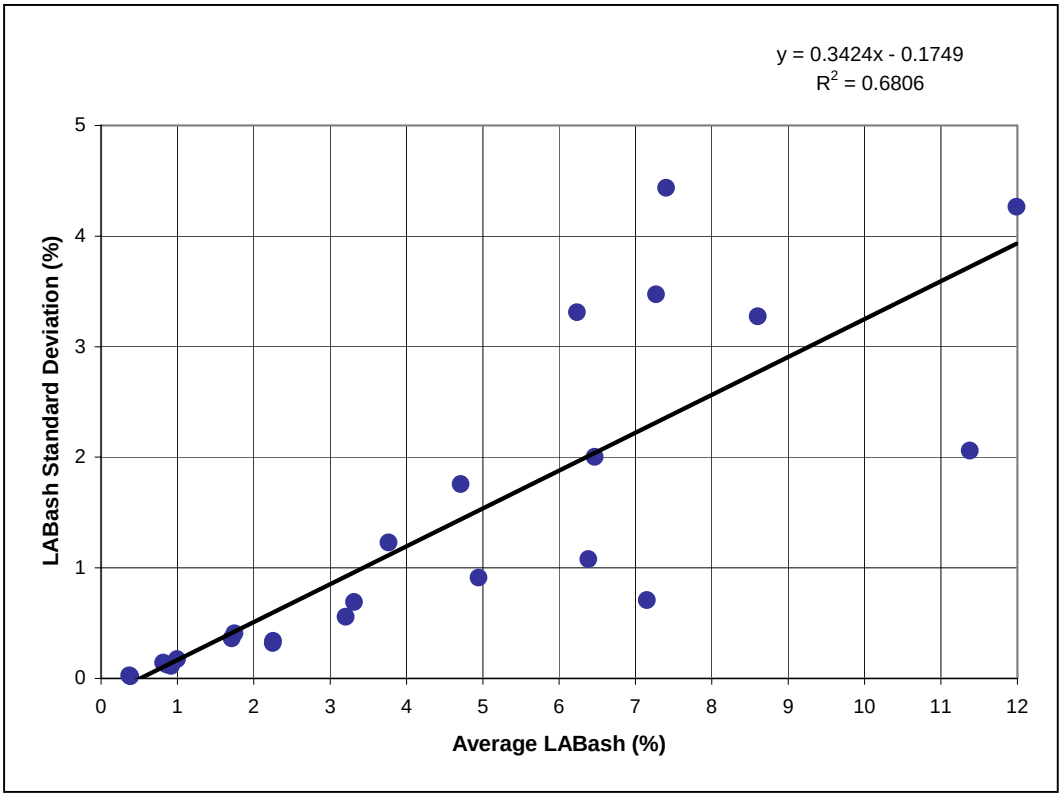


Figure 9. LABash SD % versus Average Dag %

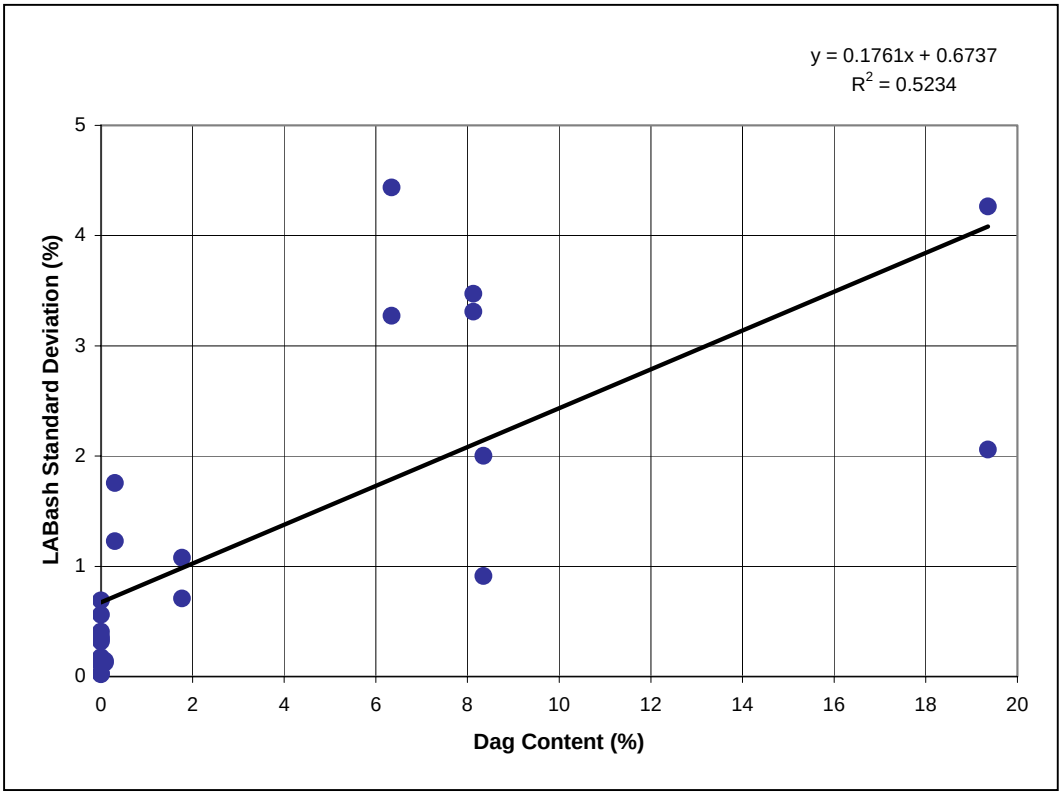


Figure 9 shows that LABash SD increased with increasing levels of dag, leading to the conclusion that dag is a principle cause of variability in LABash test results. The samples with high ash contents were all classed as inferior wool types with low Wool Base. These wools traditionally require testing of additional subsamples because the Wool Base differences between the subsamples is larger than normal, which could in part be due to the impact of the variable LABash content.

This work has shown that the repeatability of both LABash and NIRash measurement is adversely influenced by dag. NIRA will correctly estimate the high ash content of dag if it is visible but dag within the log is often masked by wool. Even when dag is visible at the surface, it may only be a small proportion by area of the total mass of dag which is present. It has been shown that inclusion of high dag samples in the NIRA calibration set reduces the underestimation to some extent, but does not remove it entirely.

CALIBRATION USING ARTIFICIAL NEURAL NETWORKS (ANN)

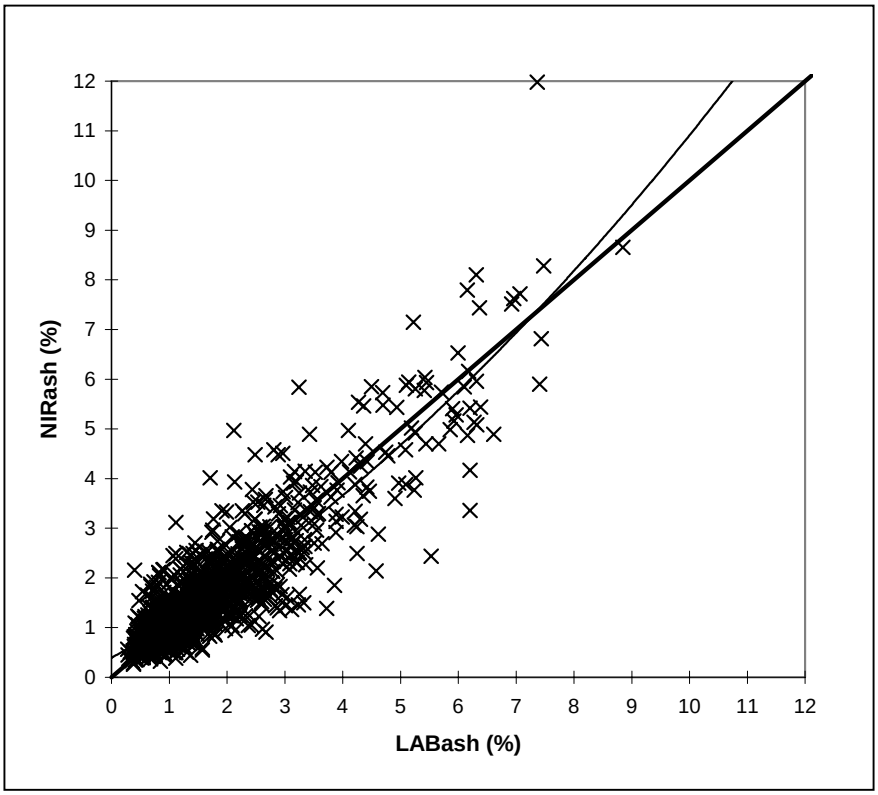
The limitations caused by dag may be an inadequacy in the MPLS regression technique to adequately deal with dag masked by wool. Alternative analytical techniques are available which may better interpret the spectral data available and address the curvilinear relationship that has been evident to date. Artificial Neural Networks [(6)] (ANN) are an adaptive mathematical process that has demonstrated potential to improve the predictive ability when non-linear relationships exist between dependent and independent variables.

The Calibration and Commercial sets described in Table 1 were used in this investigation to provide direct comparisons with the MPLS method. An ANN relationship was determined for the Calibration data set. A linear correction was applied to correct for the slope and bias offsets. The corrected relationship was validated against the Commercial set. The validation was judged by the geometric mean regression as described in IWTO-0.

The underestimation of ash associated with the high dag samples was not as evident, indicating that ANN is better able to utilise the spectral information to predict total ash content. Plotting the quadratic trend line highlighted the improvement in linearity provided by ANN. This enabled the correction of the calibration for slope and bias without encountering the problem of negative NIRash results being predicted when LABash values are low.

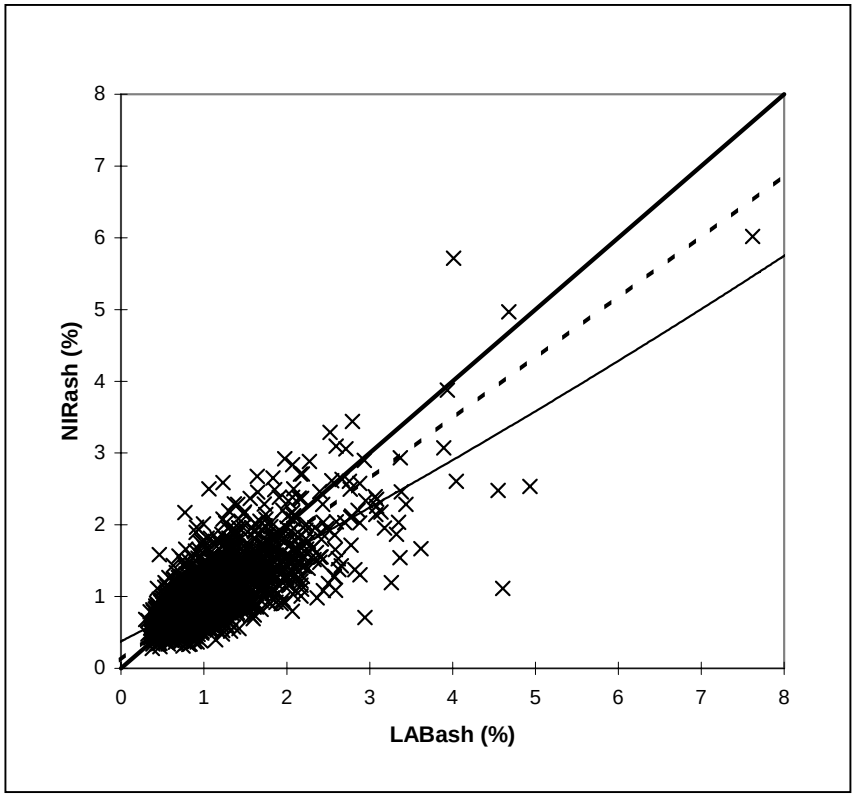
The corrected calibration is shown in Figure 10.

Figure 10. Calibration LABash versus NIRash – Corrected ANN Equation



The validation of the ANN equation using the Commercial set is shown in Figure 11.

Figure 11. LABash versus NIRash – Commercial Set Predicted with Corrected ANN Equation



The average difference between NIRash and LABash was -0.06%. The plotted quadratic trend line confirmed an improved linear relationship between NIRash and LABash.

COMPARISON OF STATISTICAL DATA FOR MPLS AND ANN

Table 2 summarises the statistics of all the calibrations and validations performed.

Table 2: Statistical Data from Calibrations and Validations Using MPLS and ANN

Calibration Technique	Data Set	R	SED	GM Slope	Mean Difference
MPLS Uncorrected	Calibration	0.88	0.54	0.75	0.00
	Validation	0.71	0.46	0.94	0.30
MPLS Corrected	Calibration	0.88	0.54	1.00	0.00
	Validation	0.71	0.46	1.24	0.28
ANN Uncorrected	Calibration	0.89	0.54	0.95	0.05
	Validation	0.74	0.38	0.80	0.02
ANN Corrected	Calibration	0.89	0.54	1.00	0.00
	Validation	0.74	0.38	0.85	-0.06

In all cases the validation regression standard deviation, SED, was lower than the calibration. This was due to the wider distribution of LABash in the calibration set compared to the validation (i.e. the commercial set). There was little difference between the Correlation, R values, between the MPLS and ANN (corrected or uncorrected). The critical difference was in the GM (Geometric Mean) regression slopes in validation and their respective offsets.

For the MPLS there was a large difference in slopes between the calibration set and the validation set, indicating that the MPLS was seeing these two sets differently. Correcting the calibration slope for the MPLS resulted in a large overcorrection when applied to the validation data. The differences were smaller for the ANN.

The validation offsets were lower for the ANN compared to the MPLS.

These trials would indicate there are significant advantages in using ANN over MPLS to derive relationships between NIR spectra and the ash content of wool samples.

CONCLUSIONS

The prediction of ash content of laboratory-scoured core samples has been investigated. Modified Partial Least Squares (MPLS) Regression was found to have difficulties in deriving a suitable calibration to estimate the ash content when the sample being tested contained significant quantities of dag. In contrast, the application of Artificial Neural Networks (ANN) was able to produce much better calibrations for predicting the ash content of samples that were independent of those used to calibrate the NIR instrument.

REFERENCES

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

2. Watson, C.A., Etchevers, G. and Shuey, W.C. (1976) *Relation between Ash and Protein Contents of Flour Mill Streams Determined with the InfraAlyzer and by Standard Approved Methods*. Cereal Chemistry **53**(5): 803-804

3. Wear, J. (2001) *The use of NIR to predict residual ash in the IWTO-19 yield test*. IWTO T&S Committee, Nice, Report RWG 02.

4. Wear, J. (2002) *The use of NIR to predict residual ash in the IWTO-19 yield test*. IWTO T&S Committee, Barcelona, Report RWG 02.

5. Wear, J. (2002) *The use of NIR technology for predicting IWTO-19 residual ash in a commercial laboratory*. IWTO T&S Committee, Nice, Report RWG 06.

6. Wasserman, P.D. (1989) *Neural Computing Theory and Practice,* Van Nostrand Reinhold, New York, ISBN 0-442-20743-3