

**On Farm Fibre Measurement (OFFM)
Instrument Evaluation Trial**

AWI Project EC397

Final Report

April 2004

TABLE OF CONTENTS

1	EXECUTIVE SUMMARY	1
1.1	<i>Summary of Project.....</i>	1
1.2	<i>Key Outcomes – Phase 1</i>	2
1.2.1	Accuracy	2
1.2.2	Precision	3
1.3	<i>Key Outcomes - Phase 2</i>	4
1.3.1	Clip Preparation	4
1.3.2	Sheep Classing/Selection	5
1.4	<i>Recommendations - Phase 1 and Phase 2</i>	6
2	BACKGROUND.....	7
2.1	<i>Introduction</i>	7
2.2	<i>Objectives of OFFM Instrument Evaluation Project.....</i>	8
2.3	<i>Scope of Trial Design.....</i>	8
2.4	<i>Literature Review</i>	9
2.4.1	Technologies for Fleece Testing – On- and Off-Farm.....	9
2.4.2	Factors Influencing use of OFFM.....	9
2.4.3	Statistical Terms for Instrument Comparisons	10
2.4.4	Statistical Analysis based on IWTO-0	11
2.4.5	Factors Influencing the Performance of OFFM Instruments	11
2.4.6	Between-Instrument Comparisons from Previous Research	12
3	TRIAL DESIGN.....	16
4	PHASE 1 TRIALS.....	16
4.1	<i>Phase 1 Trial Protocol.....</i>	16
4.1.1	OFDA2000 (Method A)	17
4.1.2	Lab OFDA100 (Method-B) and Lab LSN (Method-C)	18
4.1.3	Fleecescan (Method-D).....	19
4.1.4	Whole Fleece (“True Value”) (Method-E).....	19
4.1.5	Data Management.....	20
4.1.6	Data Clean-up	20
4.2	<i>Phase 1 Trial Properties.....</i>	21
4.3	<i>Phase 1 Statistical Analyses</i>	21
4.3.1	On-Farm Measurements	21
4.3.2	Laboratory Measurements	22
4.3.3	Whole Fleece Measurements	22
4.4	<i>Phase 1 Analyses and Results</i>	23
4.4.1	Property-1 (CH - “Chatsworth House”).....	24
4.4.2	Property-2 (CL - “Corea Lane”).....	26
4.4.3	Property-3 (CR - “Corea”)	28
4.4.4	Property-4 (AM - “Bernifay”).....	30
4.4.5	Property-5 (WE - “Wallendbeen East”)	32
4.4.6	Property-6 (AE - “Glenara North”)	34
4.4.7	Property-7 (MP - “Marbarrup”)	36
4.5	<i>Phase 1 – Summary of Findings</i>	38
4.5.1	Properties for Inclusion in Summary	38
4.5.2	Relationship between Whole Fleece Lab LSN and OFDA100 Measurements	38
4.5.3	Relationship between Whole Fleece Average and other Test Instruments.....	42

4.5.4	Analysis of Variance of Site and Whole Fleece Confidence Limits	48
4.5.5	Selection Differentials for MFD	56
4.5.6	Selection Differentials for MFC	57
4.5.7	Sources of Variance for MFD, SDD and MFC Measurement.....	60
4.5.8	Improving the Precision of Measurement.....	65
4.6	<i>Phase 1 – Conclusions and Recommendations</i>	66
4.6.1	Conclusions Related to Accuracy	66
4.6.2	Conclusions Related to Precision	67
4.6.3	Recommendations from Phase 1	68
5	PHASE 2 TRIALS	69
5.1	<i>Phase 2 – Individual Properties</i>	69
5.1.1	Randomisation	70
5.1.2	Comparison of MFD for Fleecescan and OFDA2000 Classed Lines.....	71
5.1.3	OFDA2000	71
5.1.4	Lab LSN and Lab OFDA100	71
5.1.5	Fleecescan.....	71
5.1	Data Clean-up.....	71
5.2.	<i>Phase 2 Trial Properties</i>	72
5.3	<i>Phase 2 Statistical Analyses</i>	72
5.3.1	Fleece Classing.....	72
5.3.2	Sheep Selection	72
5.4	<i>Phase 2 –Analyses and Results</i>	73
5.4.1	Comparison of Fleece Classing Results	73
5.4.2	Comparison of Sheep Selection Results.....	82
5.5.	<i>Phase 2 – Conclusions and Recommendations</i>	86
5.5.1	Conclusions for Fleece Classing Results	86
5.5.2	Conclusions for Sheep Selection Results	87
5.5.3	Recommendations from Phase 2.....	87
6	ACKNOWLEDGEMENTS	88
7	OFFM REFERENCES	90
8	APPENDICES	97
<i>Appendix 8.1</i>	<i>Phase 1 Statistical Analyses for Each Property</i>	97
Appendix 8.1.1	Property-1 (CH - “Chatsworth House”)	97
Appendix 8.1.2	Property-2 (CL - “Corea Lane”).....	105
Appendix 8.1.3	Property-3 (CR - “Corea”)	113
Appendix 8.1.4	Property-4 (AM - “Bernifay”)	121
Appendix 8.1.5	Property-5 (WE - “Wallendbeen East”).....	129
Appendix 8.1.6	Property-6 (AE - “Glenara North”).....	137
Appendix 8.1.7	Property-7 (MP - “Marbarrup”)	145
<i>Appendix 8.2</i>	<i>Phase 2 IWTO-0 Appendix B for Property 1 “Challicum Park” (DO)</i> 153	
Appendix 8.2.1	Validation Property 1 (DO) MFD – Fleecescan / Lab LSN	153
Appendix 8.2.2	Validation Property 1 (DO) MFD – Fleecescan / Lab OFDA100	154
Appendix 8.2.3	Validation Property 1 (DO) MFD – OFDA2000 / Fleecescan	155
Appendix 8.2.4	Validation Property 1 (DO) MFD – OFDA2000 / Lab LSN.....	156
Appendix 8.2.5	Validation Property 1 (DO) MFD – OFDA2000 / Lab OFDA100.....	157
<i>Appendix 8.3</i>	<i>Phase 2 IWTO-0 Appendix B for Property 2 “Mt Sturgeon” (MS) ...</i> 158	
Appendix 8.3.1	Validation Property 2 (MS) MFD – Fleecescan / Lab LSN	158
Appendix 8.3.2	Validation Property 2 (MS) MFD – Fleecescan / Lab OFDA100	159

Appendix 8.3.3	Validation Property 2 (MS) MFD – OFDA2000 / Fleecescan	160
Appendix 8.3.4	Validation Property 2 (MS) MFD – OFDA2000 / Lab LSN.....	161
Appendix 8.3.5	Validation Property 2 (MS) MFD – OFDA2000 / Lab OFDA.....	162
<i>Appendix 8.4</i>	<i>Phase 2 IWTO-0 Appendix B for Property 3 “Carlyle” (JM)</i>	<i>163</i>
Appendix 8.4.1	Validation Property 3 (JM) MFD – Fleecescan / Lab LSN.....	163
Appendix 8.4.2	Validation Property 3 (JM) MFD – Fleecescan / Lab OFDA100.....	164
Appendix 8.4.3	Validation Property 3 (JM) MFD – OFDA2000 / Fleecescan.....	165
Appendix 8.4.4	Validation Property 3 (JM) MFD – OFDA2000 / Lab LSN	166
Appendix 8.4.5	Validation Property 3 (JM) MFD – OFDA2000 / Lab OFDA100	167
<i>Appendix 8.5</i>	<i>Phase 2 Potential Returns from OFFM Clip Preparation</i>	<i>168</i>
<i>Appendix 8.6</i>	<i>Methods assessed for use in Statistical Analyses</i>	<i>171</i>
<i>Appendix 8.7</i>	<i>Implications of Confidence Limits on On-Farm Decisions.....</i>	<i>176</i>
<i>Appendix 8.8</i>	<i>“Micron Creep” or “Diameter Shrinkage”</i>	<i>179</i>
<i>Appendix 8.9</i>	<i>Identification Errors in Trial.....</i>	<i>180</i>
Appendix 8.9.1	Opportunities for Identification Errors	180
Appendix 8.9.2	Resolution of Identification Errors.....	180
<i>Appendix 8.10</i>	<i>Fleecescan and OFDA2000 Operators.....</i>	<i>182</i>
<i>Appendix 8.11</i>	<i>Resources available from AWI Project EC 397</i>	<i>183</i>
Appendix 8.11.1	Wool Samples.....	183
Appendix 8.11.2	Database and Raw Data.....	184
<i>Appendix 8.12</i>	<i>IWTO-0-01 Appendix-B</i>	<i>185</i>
<i>Appendix 8.13</i>	<i>IWTO-31-02 Appendix C, Clause 6.2.2</i>	<i>198</i>
<i>Appendix 8.14</i>	<i>Original Trial Protocol</i>	<i>201</i>
<i>Appendix 8.15</i>	<i>Operating Procedures for OFDA2000.....</i>	<i>223</i>
<i>Appendix 8.16</i>	<i>Operating Procedures for Fleecescan</i>	<i>262</i>

1 EXECUTIVE SUMMARY

1.1 Summary of Project

The 2003 Australian Wool Innovation Ltd (AWI) On-Farm Fibre Measurement (OFFM) Instrument Evaluation Trial was a direct outcome of market research commissioned by AWI, which indicated an increasing need by wool producers for an independent and objective assessment of the commercial application of the two current OFFM technologies (OFDA2000 and Sirolan Fleecescan™). AWI established an Expert Advisory Group to provide assistance and guidance with the implementation of the trial and the analysis and interpretation of the trial results.

Rather than simply compare the OFDA2000 (incorporating both pinbone and midside sampling) and the Sirolan Fleecescan™ on-farm, the trial included comparisons with laboratory fleece measurement systems (at an AWTA Ltd Laboratory). Hence, both midside sampling and pinbone sampling followed by laboratory preparation and measurement on both a Sirolan Laserscan and an OFDA100 were part of the trial protocol.

The key objectives of the instrument evaluation project (the trial) were to quantify:

- (1) The precision and accuracy of the commercially available sampling/measurement systems, with respect to an entire fleece; for the measured characteristics, principally Mean Fibre Diameter (MFD); but including: Standard Deviation of Fibre Diameter (SDD), Coefficient of Variation of Fibre Diameter (CVD), Comfort Factor (CFR) and Mean Fibre Curvature (MFC).
- (2) The ability of the sampling/measurement systems, principally the Fleecescan and the OFDA2000 (with midside sampling) to class a single mob wool clip into lines of distinctive MFD categories (minimum category of one micrometre (μm)); and
- (3) The ability of the sampling/measurement systems, principally the Fleecescan and the OFDA2000 (with midside sampling) to rank individual sheep on the basis of MFD.

A key element of Phase 1 of this trial was to establish the “true” Whole Fleece values for each individual fleece. For the first time in any study of this type, each individual fleece was core sampled and tested using procedures that reflect those used for Presale Certification of sale lots. These values became the benchmark against which all OFFM systems were compared.

Objective (1) was principally investigated in Phase 1 of the trials, while Objectives (2) and (3) were examined in Phase 2 of the trials. The trials were specifically directed towards providing knowledge that could be used by woolgrowers (or their sheep and wool customers) to make informed decisions with respect to the use of these OFFM instruments for clip preparation, sheep breeding and/or flock management.

The trial provided an opportunity to benchmark one instrument against the other; and against the other commercially accepted method of fleece measurement using on-farm midside sampling and off-farm laboratory measurement; and against IWTO Certified results for on-farm prepared sale lots.

The trial also provided an opportunity to assess variability between operators and their instruments. As each operator had a separate instrument, these two components could not be separated in this trial; hence the variance for the operator-instrument effect was combined. It was expected that the majority of this variance would be due to operators, since experience suggests that there is little variation between instruments when used for diameter distribution measurement.

1.2 Key Outcomes – Phase 1

1.2.1 Accuracy

The overall means and biases (referenced to the Whole Fleece Average Value) are presented below (**Note:** for a given parameter values with the same letter are considered to be statistically equivalent ($p>0.05$)):

Parameter	Whole Fleece Avg	Pinbone Samples			Fleece	Midside Samples		
		Lab LSN	Lab OFDA100	OFDA-2000	Fleece-scan	OFDA-2000	Lab OFDA100	Lab LSN
MFD	19.5	20.1cd	20.2d	20.3d	19.7bc	19.3ab	19.3ab	19.2a
Diff Wh Flc		0.5	0.7	0.7	0.2	-0.2	-0.2	-0.4
SDD	4.1	4.0cd	4.0d	3.7b	4.0cd	3.6a	3.9c	3.8b
Diff Wh Flc		-0.1	-0.1	-0.3	-0.1	-0.5	-0.2	-0.3
CVD	20.8	19.7b	19.7b	18.3a	20.1b	18.5a	20.1b	19.6b
Diff Wh Flc		-1.0	-1.1	-2.4	-0.7	-2.2	-0.7	-1.2
CFR	98.1	97.6ab	97.4a	97.6ab	98.0abc	98.8c	98.6bc	98.6bc
Diff Wh Flc		-0.6	-0.7	-0.6	-0.2	0.6	0.4	0.4
MFC	93	94c	84ab	80a	99c	82a	88b	98c
Diff Wh Flc		2	-9	-13	6	-11	-4	5

The following are the key conclusions that can be drawn from the analysis of the MFD results:

- (a) For on-farm MFD measurements, OFDA2000 measured on midsides and Fleecescan measured on fleeces gave equivalent accuracy compared with the whole fleece average.
- (b) For MFD, the pinbone samples measured on-farm or in a laboratory provided significantly higher results (average +0.6 μm) than the whole fleece average. The pinbone as a sampling site does not give as accurate a measure of MFD for the whole fleece.
- (c) For MFD, the midside samples measured on-farm or in a laboratory tended to produce lower results (average -0.3 μm) than the whole fleece average, but they were not always significantly different.
- (d) For MFD, on an individual property basis, it is unlikely that on-farm results from OFDA2000 midside and Fleecescan will exactly match the Certified Test results for a property.

In general, the accuracy of measurement for the four other parameters by the seven instruments/methods used was more variable than for MFD. The following conclusions apply:

- (e) SDD and CVD measured using OFDA2000 exhibited the greatest divergence from the whole fleece average (-0.3 μm to -0.5 μm for SDD and -2.2% to -2.7% for CVD) and were probably influenced by the “distribution trimming” tool used in the OFDA2000 software. It is recommended that the manufacturers further examine this.
- (f) Compared to the whole fleece average, the CFR results for pinbone samples were, averaged across all instruments, significantly lower (-0.7%), while the midside results were overall significantly higher (+0.6%).

(g) It is generally acknowledged that the measurement of MFC is greatly influenced by the sample preparation for each instrument. The whole fleece Lab OFDA100 and Lab LSN results differed on average by 10 degrees/mm. When whole fleece Laserscan was used as the basis for comparisons for Fleecescan and Lab LSN; and when the whole fleece OFDA100 was used as the basis for comparisons for Lab OFDA100 and OFDA2000 the following conclusions can be drawn for the circumstances prevailing in this trial:

- midside sampling followed by laboratory testing produced MFC results consistent with the whole fleece values;
- pinbone sampling produced MFC results that were lower than the whole fleece values by 4 degrees/mm;
- Fleecescan produced MFC results consistent with the whole fleece values; and
- OFDA2000 produced MFC results that were 7 degrees/mm lower than the corresponding whole fleece OFDA100 values, whether the sample was drawn from the midside or the pinbone. The difference is likely to be related to the totally different preparation systems used for the OFDA100 and the OFDA2000.

It is probable that calibration systems for MFC will shortly be introduced. It is to be expected that the biases between instruments/preparation systems found in this trial will be significantly reduced when such systems are implemented.

1.2.2 Precision

The Overall Confidence Limits are presented below (**Note:** for a given parameter, values with the same letter are considered to be statistically equivalent ($p > 0.05$)):

Parameter	Fleece	Pinbone Samples			Fleece	Midside Samples		
	Whole Flc Avg	Lab LSN	Lab OFDA100	OFDA -2000	Fleecescan	OFDA -2000	Lab OFDA100	Lab LSN
MFD	19.5	±1.19c	±1.25c	±1.39d	±1.17b	±1.24c	±1.05ab	±1.04a
SDD	4.1	±0.7a	±0.5a	±0.5a	±0.7a	±0.5a	±0.5a	±0.6a
CVD	20.8	±3.3c	±2.4a	±2.2a	±3.4c	±2.2a	±2.4a	±2.9b
CFR	98.1	±2.3a	±2.6a	±2.6a	±1.8a	±1.6a	±1.5a	±1.6a
MFC	93	±15b	±12ab	±13b	±13ab	±12ab	±10a	±13ab

The main conclusions related to analysis of the MFD results are as follows:

- (a) To be meaningful for sheep selection and fleece classing, precision comparisons can only be made for estimates relevant to the whole fleece as those relevant to a site do not include all the sources of variation present over an animal.
- (b) The Confidence Limit estimates for MFD measurements varied from $\pm 1.04 \mu\text{m}$ to $\pm 1.39 \mu\text{m}$ across the different systems evaluated in the trial. The following comments are relevant to MFD:
 - midside sampling and laboratory testing in accordance with AS/NZS 4492.4:2000 (Laserscan) and AS/NZS 4492.5:2000 (OFDA100) gave the lowest (best) and equivalent Confidence Limits of $\pm 1.04 \mu\text{m}$ for Laserscan and $\pm 1.05 \mu\text{m}$ OFDA100;
 - the Fleecescan ($\pm 1.17 \mu\text{m}$) and OFDA2000 midside sampling ($\pm 1.24 \mu\text{m}$) were considered to be equivalent in precision;
 - the pinbone sampling produced results which had poorer (higher) Confidence Limits than midside sampling;
 - the pinbone sampling and OFDA2000 measurement produced the (worst) highest Confidence Limit ($\pm 1.39 \mu\text{m}$) of all the methods evaluated in this trial; and

- the 95% confidence limits increased with the average MFD of the mobs, reflecting the normal level-dependency observed for laboratory-based diameter measurement systems.

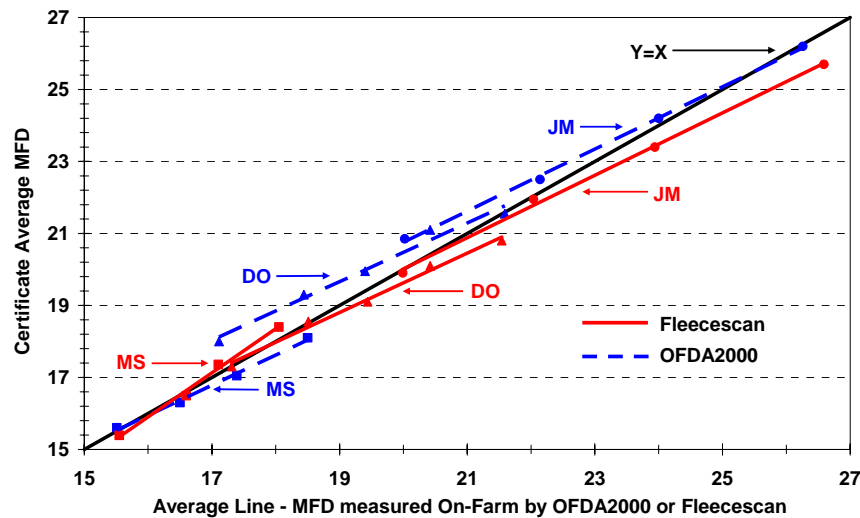
For the other parameters the following conclusions related to precision can be drawn:

- (c) The Confidence Limit estimates for SDD measurements varied from $\pm 0.5\mu\text{m}$ to $\pm 0.7\mu\text{m}$ and with all methods of measurement being considered equivalent.
- (d) The Confidence Limit estimates for CVD measurements varied from $\pm 2.2\%$ to $\pm 3.4\%$. The OFDA2000 with mid-side sampling produced the best results ($\pm 2.2\%$); but this needs to be considered in light of the “accuracy” results shown above, in that that the OFDA2000 also produces lower CVD values by 2% to 3%. Fleecescan and pinbone sampling with Lab LSN measurement gave the worst Confidence Limit of $\pm 3.4\%$ and $\pm 3.3\%$, respectively.
- (e) The Confidence Limit estimates for CFR measurements varied from $\pm 1.5\%$ to $\pm 2.6\%$ with all methods of measurement being considered equivalent. There was a tendency for the pinbone samples to produce poorer (higher) Confidence Limit than the midside samples.
- (e) The Confidence Limit estimates for MFC measurements varied from ± 10 degrees/mm to ± 14 degrees/mm.
- (f) Ideally, measurements on high value animals (such as rams) should be performed using the best possible precision (the lowest Confidence Limit) as there is greater selection emphasis placed on these animals and more reliance placed on the test result.
- (g) Laboratory measurements in accordance with AS/NZS 4492.4:2000 (Laserscan) and AS/NZS 4492.5:2000 (OFDA100) offer slightly better precision (lower Confidence Limits) and correlations with whole fleece measurements, ultimately resulting in slightly better selection differentials than both OFDA2000 and Fleecescan on-farm methods. However, there is less than $0.05\ \mu\text{m}$ at 80% and $0.07\ \mu\text{m}$ at 20% in selection differential between on-farm and laboratory measurement methods.
- (h) These differences in precision are unlikely to have a large impact on the genetic and economic gain that can be derived from animal selections using MFD. Thus, wool producers need to weigh up the convenience benefits of on farm testing against any small loss of precision when considering which test procedure to use for their flock and circumstances. Other service factors such as testing costs, labour efficiency and yard or shed design may be more important.

1.3 Key Outcomes – Phase 2

1.3.1 Clip Preparation

The figure below shows the strong relationship between the on-farm Fleecescan and OFDA2000 results for each bin line and the Certified Test results for each bin. The 1:1 line ($Y=X$) is also presented. Values above the $Y=X$ line indicate OFFM results are finer than the Certified results and values below the line indicate OFFM results are coarser than the Certified results.



The following conclusions were drawn based on the results presented for the three properties:

- (a) Laboratory measurements on midside samples were slightly more precise than those made on-farm.
- (b) There was a strong relationship between the two on-farm measurement technologies and their ability to measure MFD compared to Certificate Testing.
- (c) When used for fleece classing, both the OFDA2000 (when measuring midside samples) and the Fleecescan (when measuring skirted fleeces) produced lines of wool of different diameters in the direction to be expected (i.e. the line that was expected to be the finest was the finest, and so on).
- (d) The greater the range in diameter within the flock and the more precise the measurements, the more likely it is that measurement assisted classed lines and animal selections will be consistent, regardless of the instrument used.
- (e) From the fleece classing and sheep selection results it can be concluded that the precisions of Fleecescan and OFDA2000 midside sampling are near equal.
- (f) Biases were observed between the OFFM systems and the Certified Test results, and these varied from property to property, ranging overall on individual lots from $-1.1 \mu\text{m}$ to $+0.9 \mu\text{m}$. A portion of these biases could be attributed to the midside sampling site ($-0.5 \mu\text{m}$ to $-0.2 \mu\text{m}$ based on the averages of the combined Lab results on the three properties), and the remainder to the individual OFFM instruments themselves (OFDA2000 midside $-0.3 \mu\text{m}$ to $+0.5 \mu\text{m}$, and Fleecescan $+0.1 \mu\text{m}$ to $+0.4 \mu\text{m}$).
- (g) “Micron Creep” was observed for both measurement technologies used on farm. It would also exist for Laboratory measurements if they were to be used for classing the fleeces.

1.3.2 Sheep Classing/Selection

- (a) Whilst the OFDA2000 showed slightly higher correlation coefficients with the Lab OFDA100 and Lab LSN than the Fleecescan, this is to be expected since the OFDA2000, Lab OFDA100 and Lab LSN were all using samples obtained from the midside (they are comparing like with like), whereas the Fleecescan uses a random selection of fibres obtained from the skirted fleece.
- (b) For all three properties the highest correlation between the measurements was between Lab-OFDA (midside) and Lab-LSN (midside), as would be expected given the slightly better precision of the laboratory measurements as shown in Phase 1 of the project.

- (c) The percentages of co-selected sheep vary with the ratios of the measurement precision of the two measurement systems to the range of MFD values in the selected group. Thus, for example, the percentage of co-selected sheep is higher in all cases for the 80% of sheep selected with the lowest MFD than it is for the 20% of sheep selected with the lowest MFD. This is to be expected since the ratio of the range to the measurement precision of MFD values in the 80% selection situation is significantly higher than for the 20% situation.
- (d) In order to improve the agreement between measurement systems in the selection of the "top" 20% of a flock, it would be necessary to improve the precision of measurement so that the ratio of the range to the precision increases. This confirms the observation made in Phase 1, that growers using the technologies for ram selection should compare methods and costs of improving the precision of the measurement against the benefits to be obtained.
- (e) Similarly, the percentages in any category of comparison are higher for Property 3 than for Property 2. The former had a wider range of MFD values than the latter (approximately 15 μm compared to 6 μm). The results for Property 1 fell somewhere between the corresponding values for the other two properties, as did the range of MFD values (approximately 9 μm).
- (f) The percentages of co-selected sheep were highest when comparing the two laboratory midside measurements, and lowest when comparing the two on-farm technologies, with the corresponding comparisons between an on-farm technology and a laboratory measurement falling somewhere between. These relationships mirror those of the correlation coefficients.
- (g) It can therefore be concluded that the greater the range in diameter within the flock, and the more precise the measurement, the more likely it is that animal selections made using different instruments will be consistent.

1.4 Recommendations – Phase 1 and Phase 2

There is no reason to suppose that the general conclusions obtained on these properties are not widely applicable. The magnitude of the effects on individual properties, and with different operators and instruments, will vary depending on the circumstances prevailing at the time. Recommendations arising from the project are as follows:

- (a) Whilst the study has highlighted the differences in accuracy and precision between the various measurement systems, it has also shown that the practical impacts of these differences on sheep selection are relatively small whereas their impact on clip preparation are much larger and significant. Wool producers should weigh up the convenience/benefits of on-farm testing against any loss of precision when considering which test procedure to use for their flock and circumstances. Other service factors such as testing costs, labour efficiency and yard or shed design may be more important. **Note:** For clip preparation, the poorer the precision of a measured parameter, the more the classed lines may vary from the Certified test results for that parameter.
- (b) Compared to the midside site, the results from pinbone sampling have poorer precision (ie. are more variable) and have larger biases from the whole fleece values (ie. poorer agreement with Certified Results when used for preparing classed lines in a shed). Based on the results from the OFFM Trial, EAG Members recommend that the midside site be used in preference to the pinbone. There may still be users who wish to use the pinbone site for greater convenience and ease of location. They should be made aware of the shortcomings, which should also be presented to the wool producer client before testing commences.

- (c) Measurements on high value animals (such as rams) should be performed using the highest precision (lowest Confidence Limit) available, if this is cost-effective.
- (d) For commercial selection procedures where the animals are of less value, the benefits and convenience of on-farm testing over midside sampling followed by laboratory testing may outweigh any loss (15% to 34%) in precision (**Note:** this equated to less than 0.05 μm at 80% and 0.07 μm at 20% in selection differential between on-farm and laboratory measurement methods).
- (e) The OFDA2000 manufacturer should re-examine the application of the “distribution trimming tool” as the measured values for SDD and CVD were lower than other methods used in the trial.
- (f) As it is relatively imprecise, the use of CFR in sheep selection, using the current commercial sampling/testing procedures, is not recommended. CFR is closely correlated with MFD and consequently it is better to use MFD.
- (g) The precision of the measurement of MFC was found to be relatively poor compared with other parameters, and large biases were found between instruments/preparation systems; thus, until calibration systems are available and implemented, considerable care should be exercised in the use of MFC in fleece measurement.
- (h) The use of SDD and CVD as selection parameters should only be undertaken with the knowledge that the precision of SDD ($\pm 0.6\mu\text{m}$) and CVD ($\pm 2.4\%$ to $\pm 3.4\%$) will limit the expected benefits.
- (i) To enable the most consistent comparisons of all parameters across years, it is recommended that the same instrument/method be used each time.

2 BACKGROUND

2.1 Introduction

Since the late 1990s, both OFDA2000 (measurement based on imaging technology) and Sirolan Fleecescan™/Sirolan Laserscan™ (measurement based on laser technology) instruments have been regularly used to measure mean fibre diameter (MFD), but also other wool fibre traits, on samples taken from fleeces on-farm; in a race pre-shearing and in shearing sheds at shearing time.

In September 2001, Australian Wool Innovation (AWI) initiated a series of research projects to gain a greater understanding of the issues involved with the development and application of the on-farm fibre measurement (OFFM) technologies and systems. Nine research projects were initiated with respect to OFFM covering: market research, a world-wide search for new technologies, an assessment of potential enhancements to existing technologies, scoping a Quality Assurance (QA) program, scoping a wool grower extension program, a benefit cost analysis, an assessment of the potential for electronic identification, an assessment of existing and future decision support systems and a comparison of existing technologies.

The market research commissioned by AWI (TQA 2002), which was completed in June 2002, indicated an increasing need by woolgrowers for an independent and objective assessment of the commercial application of the two current OFFM technologies.

At that time, there had been little independent published information on the performance of the two instruments with respect to their use by woolgrowers for clip preparation, sheep breeding and/or flock management. There was also a scarcity of published data that used recognised and industry accepted statistical methods to determine the precision limits of each instrument and/or to compare their outcomes and their application methods.

By June 2002, AWI had established a draft protocol for an extensive trial that would address the above issues. AWI also established an OFFM Expert Advisory Group (EAG) to provide assistance and guidance with the implementation of the trial and the analysis and interpretation of the trial results. The OFFM EAG comprised the following:

Mr Peter Baxter	SGS Wool Testing Services (New Zealand)
Dr Ralph Behrendt	Department of Primary Industries (Victoria)
Dr Kerry Hansford	Teckel Consulting Pty Ltd
Dr Bill Humphries	CSIRO Textile and Fibre Technology
Dr John James	Consultant
Mr Jim Marler	AWTA Ltd
Dr Paul Swan	AWI Ltd
Mr George Waldthausen	AWI Ltd

Dr John James was appointed to provide the primary analysis and interpretation of the results of the trial with advice and guidance from other EAG members. Dr Kerry Hansford was appointed to undertake the primary preparation the Final Report in consultation with other members of the EAG and project staff.

This project was initiated within AWI's Fibre Specification and Logistics program, for which Mr George Waldthausen is the Program Manager. Mr Gary Macfarlane (GMAC Consulting Pty Ltd) was AWI's Project Manager responsible for the implementation of all aspects of the trial. Mr Russell Pattinson (Miracle Dog Pty Ltd) provided on-ground detailed management of each trial site.

2.2 Objectives of OFFM Instrument Evaluation Project

The key objectives of the instrument evaluation project (trial) were to quantify:

- The precision/repeatability and accuracy of both instruments, with respect to an entire fleece; for the measured characteristics principally, MFD; but including: standard deviation of fibre diameter (SDD), coefficient of variation of fibre diameter (CVD), mean fibre curvature (MFC) and comfort factor (CF).
- The ability of both instruments to class a single mob wool clip into lines of distinctive MFD categories (minimum category of one micrometre (μm)); and
- The ability of both instruments to rank individual sheep on the basis of MFD.

The above objectives relate to the instruments being a part of an OFFM system operated in an on-farm environment. **Note:** for the purposes of this document an OFFM system, which comprises the measurement instrument, any associated apparatus and application procedures, will be referred to as an "instrument".

The trial was specifically directed towards providing knowledge that could be used by woolgrowers (or their sheep and wool customers) to make informed decisions with respect to the use of these OFFM instruments for clip preparation, sheep breeding and/or flock management.

The trial provided an opportunity to benchmark one instrument against the other; and against the commercially accepted method of fleece measurement using on-farm midside sampling and off-farm laboratory measurement (at an AWTA Ltd laboratory); and against IWTO certified results for on-farm prepared sale lots. In addition, it has provided a benchmark against which any OFFM systems that measure diameter and/or other fibre parameters may, in the future, be compared.

The trial also afforded an opportunity to assess variability between operators and their instruments. As each operator had a separate instrument, these two components could not be separated in this trial; hence the variance for the operator-instrument effect was combined. It was expected that the majority of this variance would be due to operators, since experience suggests that there is little variation within an instrument when used for diameter distribution measurement.

2.3 Scope of Trial Design

The trial was limited to providing information to woolgrowers specifically in relation to the classing and the ranking of the measures defined above. The trial did not assess:

- The other wool fibre measurements that may be provided by either instrument;
- The performance of either instrument in an in-store environment; or
- The two instruments in relation to issues such as ease of use, robustness, Occupational Health and Safety, after sales support, quality assurance, pricing, etc.

2.4 Literature Review

2.4.1 Technologies for Fleece Testing – On and Off-Farm

In an Australian Wool Testing Authority Ltd (AWTA Ltd) Newsletter, Sommerville (2001b) summarised the instruments and methodologies that may be used to measure the mean fibre diameter (MFD) and other fibre characteristics of individual fleeces. Subsequent AWTA Ltd Newsletters have provided more detail on each technology.

Until recently, fleece measurement involved the testing of sub-samples, often sourced from the midside, within a laboratory environment (Morgan 1990; Cottle *et al.* 1996). Instruments such as Airflow (Anderson 1954), Sirolan-Laserscan (Lynch and Michie 1976; Dabbs and Glass 1992; Glass and Dabbs 1992) and OFDA100 (Baxter *et al.* 1991) are those commonly used to estimate the fibre characteristics in a laboratory environment.

A new generation of testing technologies has allowed fleece testing to be conducted either on-farm or in a wool store. The OFDA2000 (Brims *et al.* 1999; Baxter 2001; Petersen and Gherardi 2001) may be used either pre-shearing, with the animals being sampled in a race, or in the shed using samples taken from the fleece immediately after shearing. The Sirolan Fleecescan™/Sirolan Laserscan™ system, hereafter called the Fleecescan (Hansford 1999; Knowles 2000; Humphries *et al.* 2001), is usually operated at shearing in the shed, but may also be operated post-shearing at a central store. As they are the most commonly used on-farm fibre measurement (OFFM) technologies, it is these two instruments, viz. OFDA2000 and Fleecescan, that are the subject of the trial documented in this report.

Note: since commencing this trial, an additional instrument known as “Woolview 20:20” has been developed. No independent data is available to evaluate this new instrument, and therefore, its performance cannot be compared with the results from this trial. Similarly, results reported herein cannot be implied to represent what could be achieved using the “Woolview 20:20”.

2.4.2 Factors Influencing Use of OFFM

In wool production systems, OFFM technologies are primarily used for:

- clip preparation, specifically preparing lines of wool based on MFD to capture premiums in the diameter by price relationship, and
- sheep breeding, in the selection and culling decisions for breeding and ongoing wool production.

There are many factors that can influence the successful application of OFFM technologies when used for the above purposes. Some of these factors include:

- The market price-premium relationship for mean fibre diameter; now and in the future.
- The position of the average diameter of the flock relative to the market price-premium relationship for MFD.
- The cost of the OFFM technology/system including labour, time and compatibility with other farming operations and systems.
- The culling rate that can be applied, which is heavily dependent on and thus often restricted by lambing percentages.
- The variation in MFD between sheep within a flock.
- The use of fleece weight and other wool/sheep traits in sheep selection decisions.
- The age of the sheep at testing and the number of seasons that the test information will be used.
- The accuracy and precision of the OFFM technology.
- The relationship between site measurements and fleece measurements.
- The on-farm infrastructure and logistics able to support the application of application of one or other of the OFFM instruments.

Some of these factors have a major influence on the economic returns from the use OFFM technologies. The relative importance of these factors for an individual woolgrower will depend on the grower's objectives relative to the use of OFFM. Prior to use, wool producers should consider their objectives for OFFM and determine the benefit:cost of the measurement program. Producers also need to consider practical issues such as shed space, animal handling facilities and extra labour requirements before using OFFM technology.

This project investigates the performance of two currently available OFFM technologies against industry accepted technical criteria. It does not specifically assess the practicality or benefit:cost of using each technology for different farm situations.

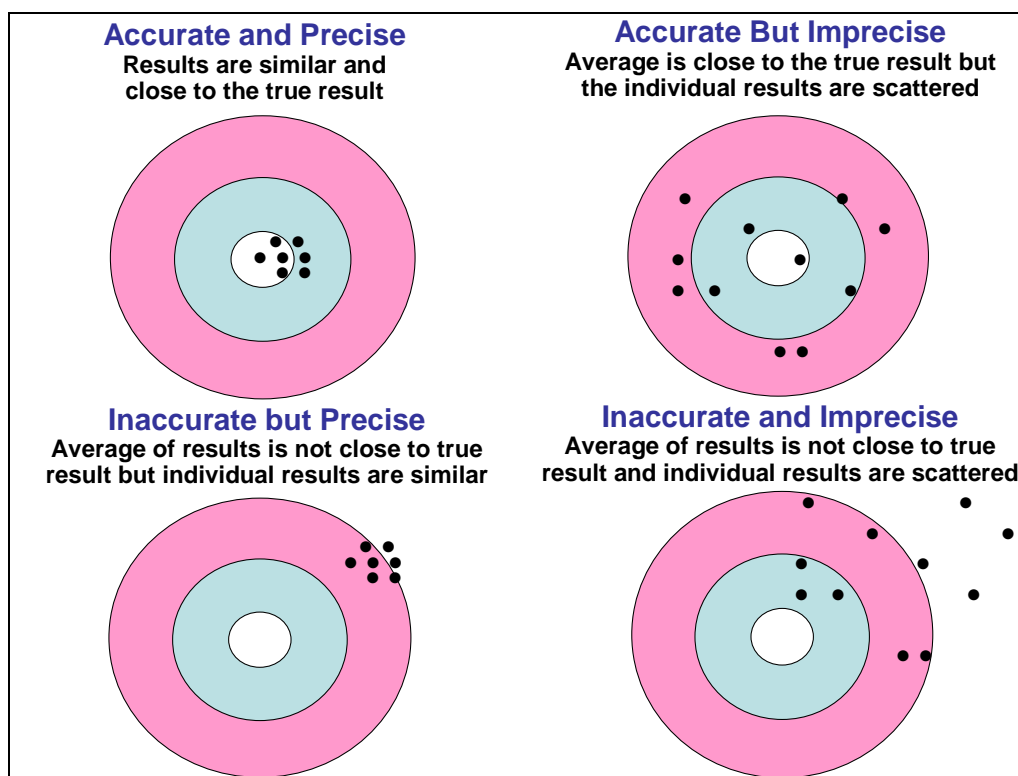
2.4.3 Statistical Terms for Instrument Comparisons

A number of statistical criteria are used to determine and compare the performance of the various measurement technologies. Bow (1988) and Sommerville (2002c) described the main criteria in the context of both the technical and commercial requirements of wool testing. Two statistical criteria, accuracy and precision (see Figure 1), are most commonly used.

Accuracy describes the correctness of a result. It is a measure of the closeness of the test result to the true result. Technically, the only measurement that can be described as completely accurate is one that involves counting objects. Therefore, to facilitate the determination of the accuracy of an instrument, the "true" result must be defined. In the case of this trial, the "true" value of each skirted fleece was estimated by multiple core sampling to produce two 70 g core samples followed by testing based on the IWTO Certified Method for OFDA100 (IWTO 2002c) and Laserscan (IWTO 2004).

Precision is the repeatability (or reproducibility) of a measurement. It describes the amount of random variation about the mean as a result of natural variability in the material being measured (due to sampling), or variability caused by random errors in the measurement system. Precision can be changed by increasing or decreasing the amount of sampling and measurement.

Figure 1 Diagrammatic “bulls eye” representation of accuracy and precision



Two types of errors are important in the consideration of accuracy and precision. They are random or indeterminate errors and systematic or determinate errors (bias), which combine to produce the error in the mean of a number of replicate measurements. Random or indeterminate errors impact upon precision whereas bias has little impact on precision but a significant effect on accuracy.

2.4.4 Statistical Analysis based on IWTO-0

Within the wool industry, IWTO-0 Appendix B (IWTO 2002a) formalises the statistical methods that are to be used to assess the equivalence or otherwise of two test methods or instruments used to issue IWTO certificates or reports. The Standard includes detailed techniques for comparing both bias and precision. Since the on-farm measurements considered in this report are not intended to be used to issue certificates, the level of analytical detail covered in this Standard is more rigorous than is required for most of the comparisons in this project. However, some of the techniques mentioned in the Standard were used the analyses.

2.4.5 Factors Influencing the Performance of OFFM Instruments

Valid comparisons between OFFM instruments are difficult as the samples tested vary greatly. For example, usually the test specimen measured by OFDA2000 is sourced from a single staple, for the Lab LSN and Lab OFDA100 it is sourced from a single site sample and for the Fleecescan it is sourced from an entire fleece. For this reason, the precision or repeatability of each instrument may not provide meaningful results in terms of the “best” instrument.

In addition, there are a number of other factors that influence the performance (precision and accuracy) of the instruments, which can subsequently affect any comparisons made. In terms of the equipment evaluated in this project, such factors include:

- Number of samplers ie.
 - OFDA2000 – different samplers take staple from sampling site
 - OFDA2000 – different operators prepare micro-staple from staple

- Lab OFDA100 and LAB LSN – different samplers take sample from site
- Number of samples/test specimens ie.
 - OFDA2000 – test more than one micro-staple per sample (eg. two test specimens from midside)
 - OFDA2000 – test more than one site (eg. one test specimen from both midside and pin-bone)
 - Lab OFDA100 and LAB LSN – test more than one test specimen (eg. two test specimens from midside)
 - Lab OFDA100 and LAB LSN – test more than one site (eg. one test specimen from both midside and pin-bone)
 - Fleecescan – test more than one test specimen from the same fleece (eg. two test specimens (separate corings) from one fleece)
- Number of measurements ie.
 - OFDA2000 – test more than one micro-staple per staple (eg. two micro-staples from midside)
 - Lab OFDA100 – test more than one slide per sample (eg. two slides from each sample)
 - LAB LSN – test more test specimens per sample (eg. two test specimens of 100 snippets each)
 - LAB LSN – test more snippets per test specimen (eg. 600 to 1000 snippets)
 - Fleecescan – test more snippets per test specimen (eg. 600 to 1000 snippets)

The above list is not complete; however, it indicates the complexity of designing trials aimed at comparing different OFFM instruments.

2.4.6 Between-Instrument Comparisons from Previous Research

Previous research conducted to evaluate and compare the performance of the different OFFM technologies primarily focused on the precision of the instruments/methods and the correlation between them. Tables 1 and 2 provide a summary of the results of some of these trials¹.

Table 1 (see Page 14) summarises previous research that determined correlations between the various fleece testing instruments/methods for fleece parameters of Merino wool. For these trials, OFDA2000 was operated using standard commercial procedures, with testing based on one greasy midside micro-staple. Fleecescan measurement was based on a single coring of each fleece, with the assumption that between 600 and 1000 snippets were measured². Laboratory testing by either Laserscan or OFDA100 was conducted using conventional mid-side methods. Although it was not detailed in all references, it is generally assumed that the Australian and New Zealand standards for fleece testing (AS/NZS 2002a and AS/NZS 2002b) were followed.

Note: It should be noted that correlation coefficient values are influenced by the range of data included in the comparisons. For a specific level of imprecision (as measured by, for example, the standard error of a regression), the correlation coefficient will increase as the range of data is increased. It is therefore difficult to assign significance to differing R values from different trials.

Two key conclusions can be drawn from the data presented in Table 1:

- Considering all fibre parameters, the correlations between any instrument/method is the highest for MFD. This is not surprising as the focus of the development and calibration of all instruments has been for the measurement MFD.
- The relationship between measurements of any fibre parameter is strongest for like instruments (viz. Laserscan and Fleecescan; OFDA100 and OFDA2000). This would be expected as the measurement principles used are similar for each pair of instruments (ie. laser and optical microscopy, respectively).

Table 2 (see Page 15) summarises previous research estimates of the precision of fibre parameters of Merino wool measured using different instruments/methods. Additional information related to the interpretation is provided, for example, the sample tested and the source of the variation represented in the Confidence Limit estimate.

Note: The precision of MFD and SDD is normally level-dependent. In most cases this has not been taken into account in comparing precision estimates for fleece testing because of the normally relatively limited range of MFD values involved in such trials.

Table 2 highlights two important issues involved in the estimating confidence or precision limits for test instruments:

- The source of the sample, and hence the variation measured, is a major determinant of test instrument precision. It is important that comparisons between instruments be made with this fact in mind.
- Taking the source of the sample into account, the precision of MFD is high relative to that of SDD, CVD, CFR and particularly MFC. As with correlations between instruments, the focus of the development and calibration of all instruments has been for the measurement MFD.

¹ This is an indicative list. See Section 7 (OFFM References) for a more extensive listing of relevant research.

² On the basis of CL reported by Marler *et al.* 2002, (viz. ± 1.02 and ± 1.04 μm for 1000 and 600 snippets respectively), the difference in snippets measured is not expected to have a large impact on reported correlations.

Table 1 Correlations (r) for fleece parameters measured using various fleece testing instruments/methods

Reference	Method	Instrument 1	Instrument 2	Parameter	Corr (r)
Brien <i>et al.</i>	Pearson	OFDA2000	Lab OFDA100	MFD	0.86
Petersen and Gherardi,	Linear Regress	OFDA2000	Lab OFDA100	MFD	0.83
Baxter 2001	Linear Regress	OFDA2000	Lab OFDA100	MFD	1.00
		OFDA2000	Lab OFDA100	MFC	0.89
		OFDA2000	Lab OFDA100	MFD	0.98
Behrendt <i>et al.</i> 2001	Pearson	OFDA2000	Lab OFDA100	MFD	0.94
				CVD	0.89
				MFC	0.95
		OFDA2000 Note: Pinbone sample	Lab OFDA100	MFD	0.89
				CVD	0.83
MFC	0.92				
Behrendt <i>et al.</i> 2002	Linear Regress	OFDA2000 -Mid	Lab LSN and OFDA100	MFD	0.94
				CVD	0.75
				MFC	0.77
Petersen and Gherardi 2002	Linear Regress	OFDA2000 Note: Pinbone sample	Lab LSN Note: 8 fleece sites sampled	MFD	0.89
		Fleecescan	Lab LSN Note: 8 fleece sites sampled	MFD	0.91
Knowles 2002	Pearson	OFDA2000	Fleecescan	MFD	0.84
		OFDA2000	Lab Laserscan	MFD	0.81
		Fleecescan	Lab Laserscan	MFD	0.81
Hansford <i>et al.</i> 2002	Linear Regress	OFDA2000	Fleecescan	MFD	0.80
		OFDA2000	Lab Laserscan	MFD	0.80
		Fleecescan	Lab Laserscan	MFD	0.87
		OFDA2000	Fleecescan	SDD	0.63
		OFDA2000	Lab Laserscan	SDD	0.64
		Fleecescan	Lab Laserscan	SDD	0.68
		OFDA2000	Fleecescan	CVD	0.68
		OFDA2000	Lab Laserscan	CVD	0.69
		Fleecescan	Lab Laserscan	CVD	0.72
		OFDA2000	Fleecescan	CFR	0.63
		OFDA2000	Lab Laserscan	CFR	0.66
		Fleecescan	Lab Laserscan	CFR	0.71
		OFDA2000	Fleecescan	MFC	0.62
OFDA2000	Lab Laserscan	MFC	0.61		
Fleecescan	Lab Laserscan	MFC	0.68		

Table 2 Estimates of 95% Confidence Limit for fleece parameters measured using various fleece testing instruments/methods

Reference	Instrument	Sample	Parameter	Source of Variation	95% CL
Morgan 1990	Airflow	Blended Fleeces	MFD	Blended Fleeces	±1.2
Cottle <i>et al.</i> 1996	Laboratory OFDA100	Midside (4 slides)	MFD	Midside Site	±0.60
			SDD	" "	±0.41
			CVD	" "	+1.73
Baxter 2000	OFDA2000 (precision model from variance components)	Midside (1 greasy micro-staple)	MFD	Midside Site	±1.0 @ 15µm to ±2.0@ 42 µm
Baxter 2001	OFDA2000 On-Farm	Midside (1 greasy micro-staple)	MFD	Midside Site	±0.8
			SDD	" "	±0.3
			MFC	" "	±12
Baxter and Johnston 2002	OFDA2000 On-Farm	Midside (1 greasy micro-staple)	MFD	Midside Site	±1.1 to 1.2
			CVD	" "	±2
			MFC	" "	±9
Behrendt <i>et al.</i> 2002	OFDA2000 On-farm	Midside (1 greasy micro-	MFD	Midside Site	±0.7
			CVD	" "	±1.4
Marler <i>et al.</i> 2002a and 2002b	OFDA2000 On-farm	Midside (1 greasy micro-staple)	MFD	Fleece	±1.28
			SDD	" "	±0.60*
			CVD	" "	±2.20*
			MFC	" "	±10.9
	OFDA2000 On-farm	Midside (1 greasy micro-staple)	MFD	Fleece	±1.28
			SDD	" "	±0.60*
			CVD	" "	±2.20*
			MFC	" "	±10.9
	OFDA2000 On-farm	Flank (1 greasy micro-staple)	MFD	Fleece	±1.41
			SDD	" "	±0.6*
			CVD	" "	±2.23*
			MFC	" "	10.6
	Fleecescan In-Store	Fleece (1000 snippets)	MFD	Fleece	±1.02
			SDD	" "	±0.8
			CVD	" "	±3.5
			MFC	" "	±7.4
	Laboratory Laserscan	Midside (1000 snippets)	MFD	Fleece	±1.19
			SDD	" "	±0.7
CVD			" "	±3.0	
MFC			" "	12.3	

* OFDA2000 between sheep range of SDD/CVD lower than that for Fleecescan and Lab LSN.