INTERNATIONAL WOOL TEXTILE ORGANISATION

TECHNOLOGY & STANDARDS COMMITTEE | NICE MEETING

Commercial Technology Forum | November 2001

Chairman: A.G.DE BOOS (Australia) | Report No: CTF 04
Technical Coordinator: ( )

The Impact of Test Result Precision on On-Farm Decisions

By

J.W. Marler

Australian Wool Testing Authority Limited PO Box 190 Guildford NSW 2161 Australia

SUMMARY

The increasing use by wool producers of diameter measurement On-Farm has highlighted that information on the impact of the precision of the test result for a single sheep on any decisions that are made is lacking in the scientific literature. This paper provides a mathematical study which highlights the key issues that wool producers using either, the laboratory testing of mid-side samples, the FLEECESCAN or the OFDA2000 in-shed, in-store or in-race need to consider.

INTRODUCTION

The number of wool producers using On-Farm testing of Mean Fibre Diameter (MFD) has been increasing. Wool producers need to ask themselves the question as to why they are prepared to have tests performed on individual fleeces. The most common applications are to gather information:

- for use in genetic improvement through a specific breeding program;
- for culling inferior animals;
- for use in classing the flock into lines of different Mean Fibre Diameters for sale; or
- for all of the above.

Each will have different outcomes in the interpretation of test data taken On-Farm or in-shed.

In the first two cases, it is the ranking of an individual sheep with respect to the other sheep in the mob that is the important parameter, whereas in the third case, it is important that fleeces are assigned to the correct line based on the measured MFD.

The precision, or more simply the repeatability, of the measurement will have an impact on all of the applications. For example, the more precise (ie. a lower Confidence Limit) is the measured MFD the smaller the number of incorrect decisions that will be made. It is also important to appreciate that it is the precision of the estimate of the fleece that is important not the precision of the estimate at a single site such as the hip-bone or the mid-side (Marler and Couchman (2001)).

The impact of the precision of measurement on both sheep selection and fleece classing can be evaluated from a simple mathematical approach.

METHOD

Using Microsoft EXCEL it is possible to generate a mob of sheep with a known average Mean Fibre Diameter and known Between-Sheep Standard Deviation (SD) of Diameter. For the purpose of this exercise a mob of 2000 sheep was created having a mob average of 19.0µm with a between sheep Standard Deviation of 1.5µm using the EXCEL function "ROUND(NORMINV(RAND(),19.0,1.5),1)". This function calculates a random value from a population that is normally distributed with a Mean of 19.0µm and a Standard Deviation of 1.5µm. By copying this function to 2000 cells in a single column we now have our 2000 sheep. These diameters can now be defined as the "True" diameters for each sheep.

We can use the same function in EXCEL to simulate testing of these sheep. This is achieved by assuming an overall Confidence Limit for the sampling and test procedures and converting this back to a Standard Deviation representing the total variation (SD_{Total}). The expected measured result for each of the 2000 sheep can be calculated by the "True Diameter + "ROUND(NORMINV(RAND(),0, SD_{Total}),1". The equation assumes that the testing is unbiased.

Using the above processes three different mobs of 2000 sheep each were generated with the same mob average MFD (19.0µm) but differing in uniformity of MFD (ie Between-Sheep SDs in the mobs of 1.0µm, 1.5µm and 2.0µm).

Having generated the data for a range of different Measurement Confidence Limits it is possible to simulate the effects of ranking the sheep from finest to coarsest based on the measured values and comparing these to the True Ranking based on the rank for the True Diameter of each sheep.

A sheep classing exercise has also been performed. Using the measured values for each of the different Confidence Limits, the wool was classed into four lines based on the measured MFD. The cut-off limits for the four lines were derived using the "CLASSER" software (Vizard and Williams(1993)). The MFD for each line was calculated as the arithmetic average of the MFD for each sheep in the classed line. The "True MFD of the Classed Line" was also calculated and compared to the expected values based on the measured data.

RESULTS AND DISCUSSION

1. The Impact of Test Result Precision on Sheep Selection

In order to select the finest sheep in a mob the sheep are listed in order from the finest to the coarsest. The finest is allocated a number (rank) one and the rest numbered in order up to 2000 for the coarsest. The impact of the precision of the test results can be shown by comparing the rank of each sheep based on its measured value compared to the true rank. The Confidence Limit for Fleece Testing has been typically reported as ±0.6µm to ±1.6µm (Morgan (1990), Cottle et al (1996), Baxter (2000) and Baxter (2001)). These reported Confidence Limits generally underestimate the real Confidence Limits as the values typically refer to a single site rather than the entire fleece. For the purpose of illustration three different Confidence Limits (±0.4µm, ±1.0µm and ±1.6µm) have been compared for each of the three different mobs.

From an examination of the graphs (Figures 1 to 6) a number of conclusions can be drawn:

- The greatest error in ranking occurs at the average MFD for the mob. This is evidenced by the ballooning effect in all the figures around the middle. This occurs as a consequence of the normal distribution. Small differences in the measured MFD can move the rank of a given sheep over a large number of other sheep compared to the both extremes where the same difference in measurement will have a smaller impact as there are less sheep numbers to move over. The consequences to sheep selection are minor, as it is normally at the extremes that "On-Farm" decisions are being made (eg the coarsest to cull or the finest, in the case of rams, to keep).

- The lower the Confidence Limit of the test result (ie the more precise) the less the errors in the ranking of the sheep.

- For any given test result precision, the more uniform the MFD between the sheep in the mob the greater the errors in the ranking of the sheep.

Figure 1: Errors in Ranking – Test Confidence Limit **±0.4µm** and Between Sheep SD of Diameter **1.0µm**

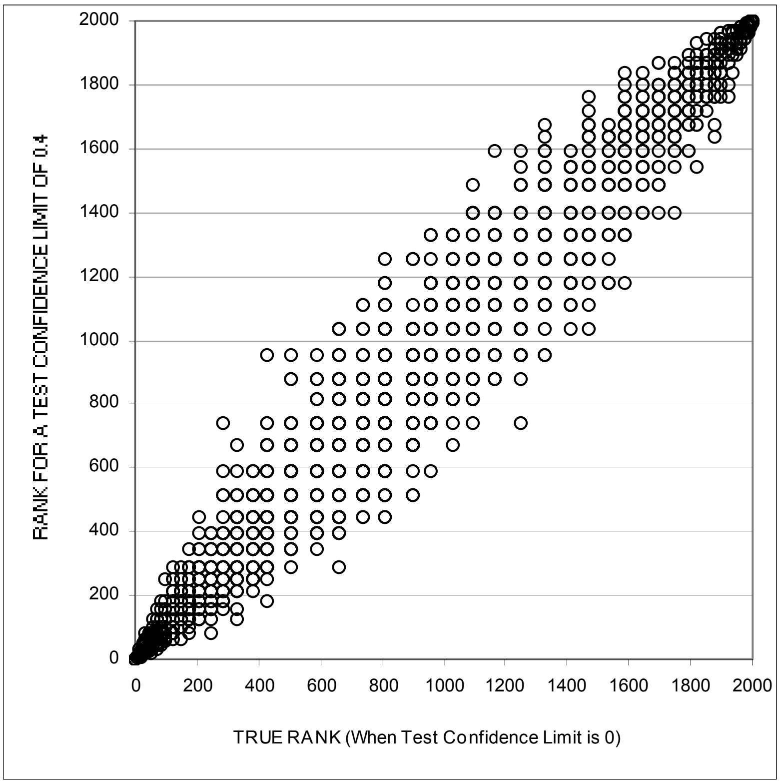


Figure 2: Errors in Ranking – Test Confidence Limit **±1.6µm** and Between Sheep SD of Diameter **1.0µm**

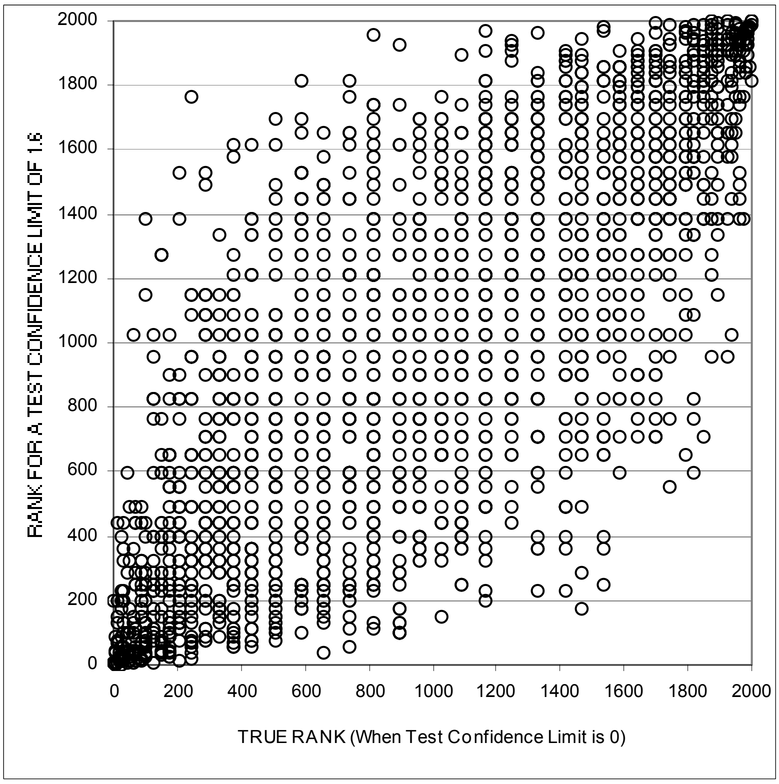


Figure 3: Errors in Ranking – Test Confidence Limit **±0.4µm** and Between Sheep SD of Diameter **1.5µm**

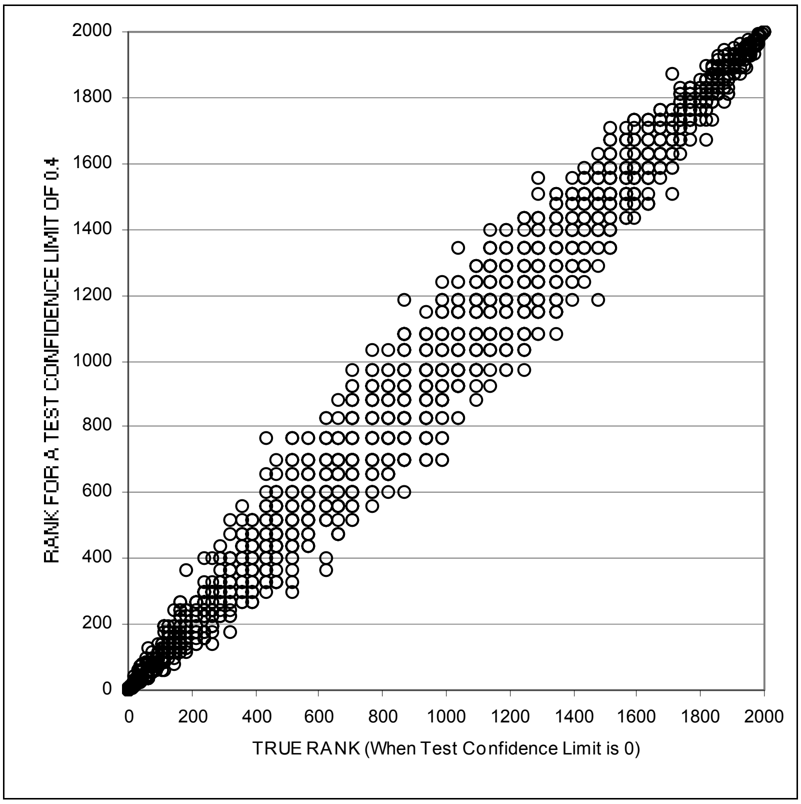


Figure 4: Errors in Ranking – Test Confidence Limit **±1.6µm** and Between Sheep SD of Diameter **1.5µm**

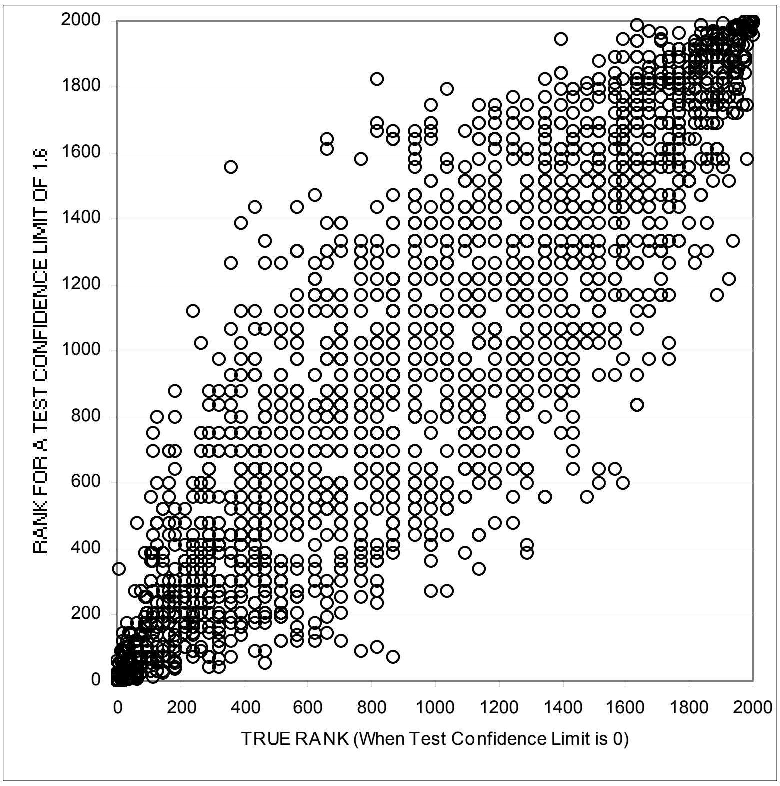


Figure 5: Errors in Ranking – Test Confidence Limit **±0.4µm** and Between Sheep SD of Diameter **2.0µm**

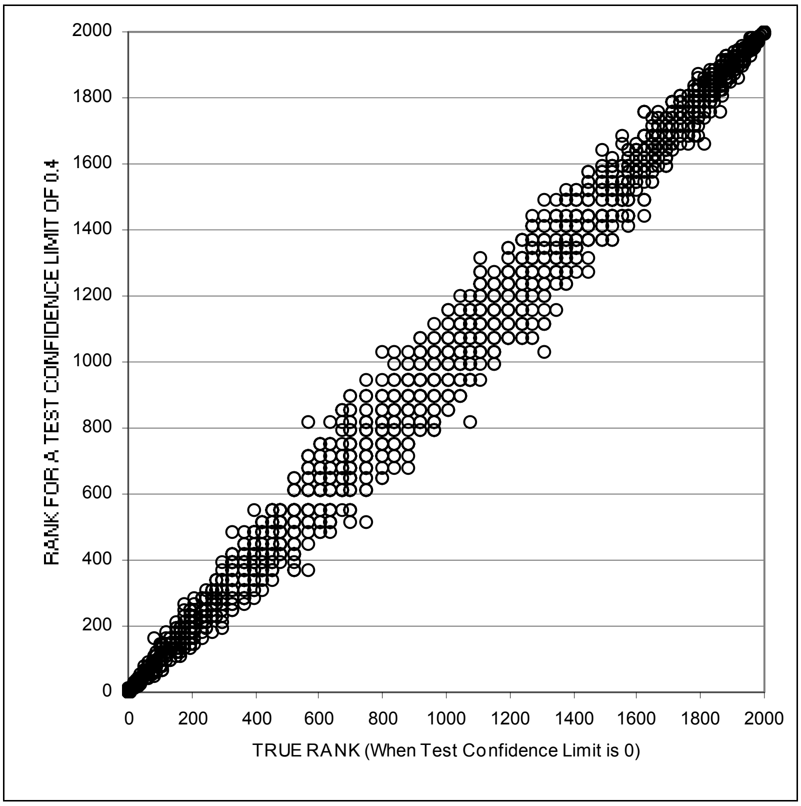
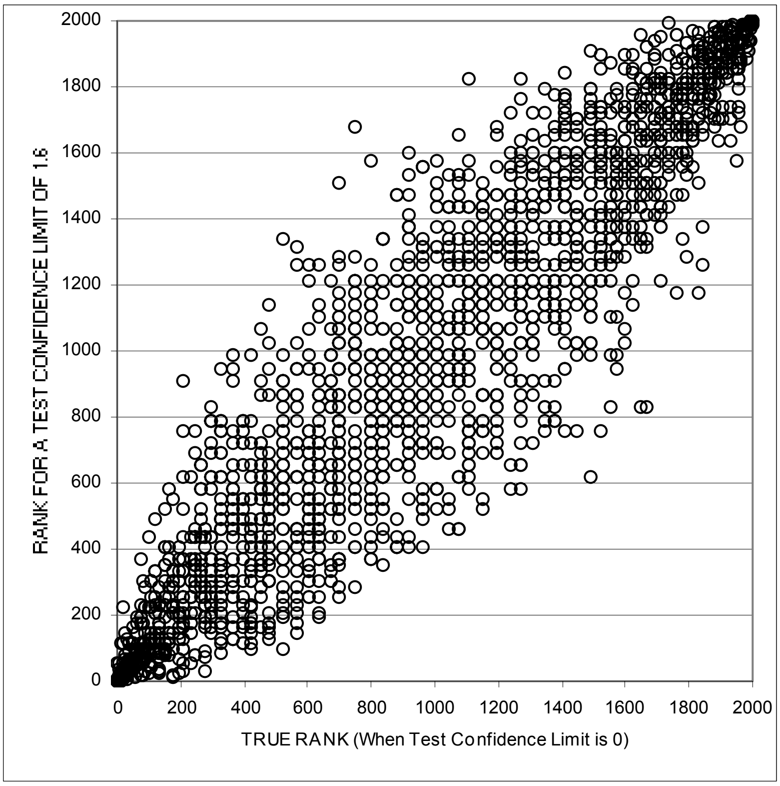


Figure 6: Errors in Ranking – Test Confidence Limit **±1.6µm** and Between Sheep SD of Diameter **2.0µm**



The influence of the incorrect selection of sheep to remain on the farm as a percentage is shown to be dependent on the uniformity of the MFD in the mob and the precision of the test is shown in Table 1.

Table1: Percentage of Incorrect Selections

Standard Deviation of MFD of the Mob	95% Confidence Limit of Test Result	35% Cull Rate (ie 700 out of 2000)	95% Cull Rate (ie 1900 out of 2000)
1.0μm	±0.4μm	4%	12%
	±1.0μm	11%	27%
	±1.6μm	15%	47%
1.5μm	±0.4μm	3%	13%
	±1.0μm	7%	19%
	±1.6μm	11%	33%
2.0μm	±0.4μm	2%	8%
	±1.0μm	5%	17%
	±1.6μm	9%	31%

Peterson and Gherardi (2001) found the same effects, which they presented in graphical form, but did not indicate the trends they found were influenced by the diameter uniformity of the mob. The following series of figures show the relationships for each of the three mobs in the same format that they used.

Figure 7: The Impact of the Precision of the Test on the Percentage of Animals Incorrectly Selected to Remain on the Property for Different Culling Rates (Between-Sheep SD = 1.0 μm)

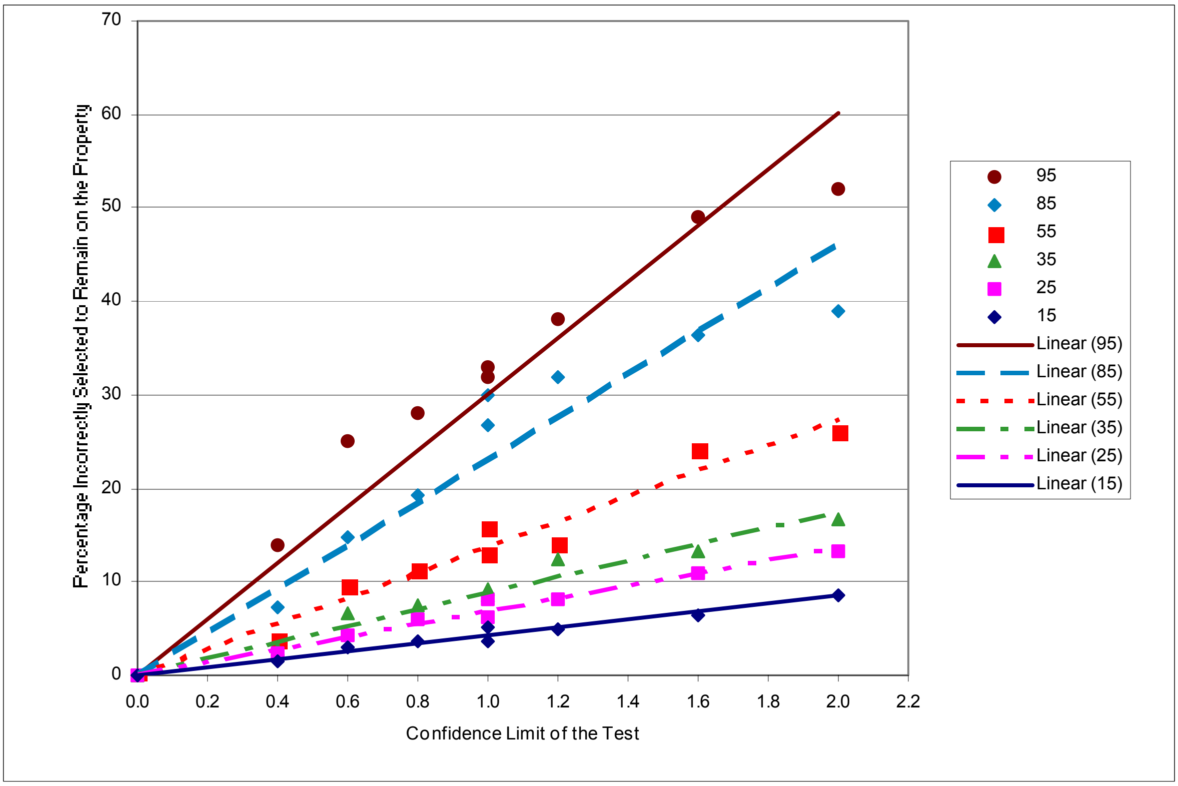


Figure 8: The Impact of the Precision of the Test on the Percentage of Animals Incorrectly Selected to Remain on the Property for Different Culling Rates (Between-Sheep SD = 1.5 μm)

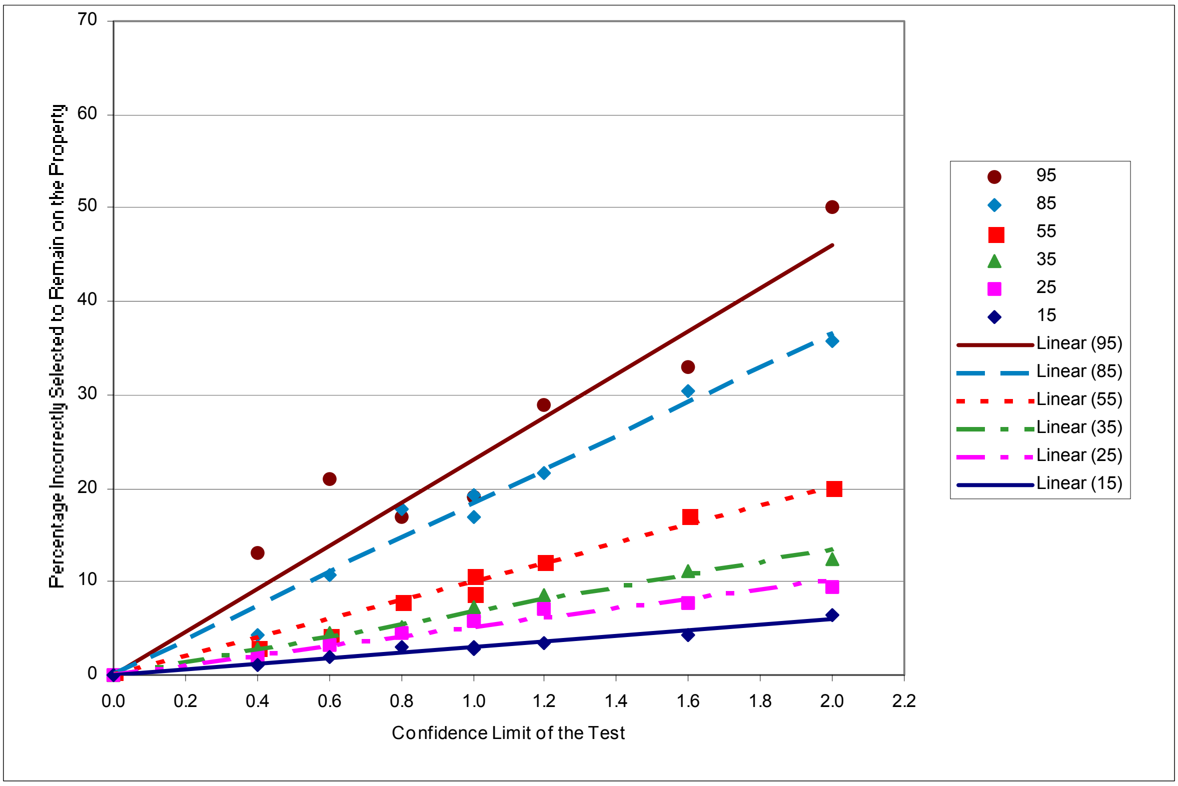
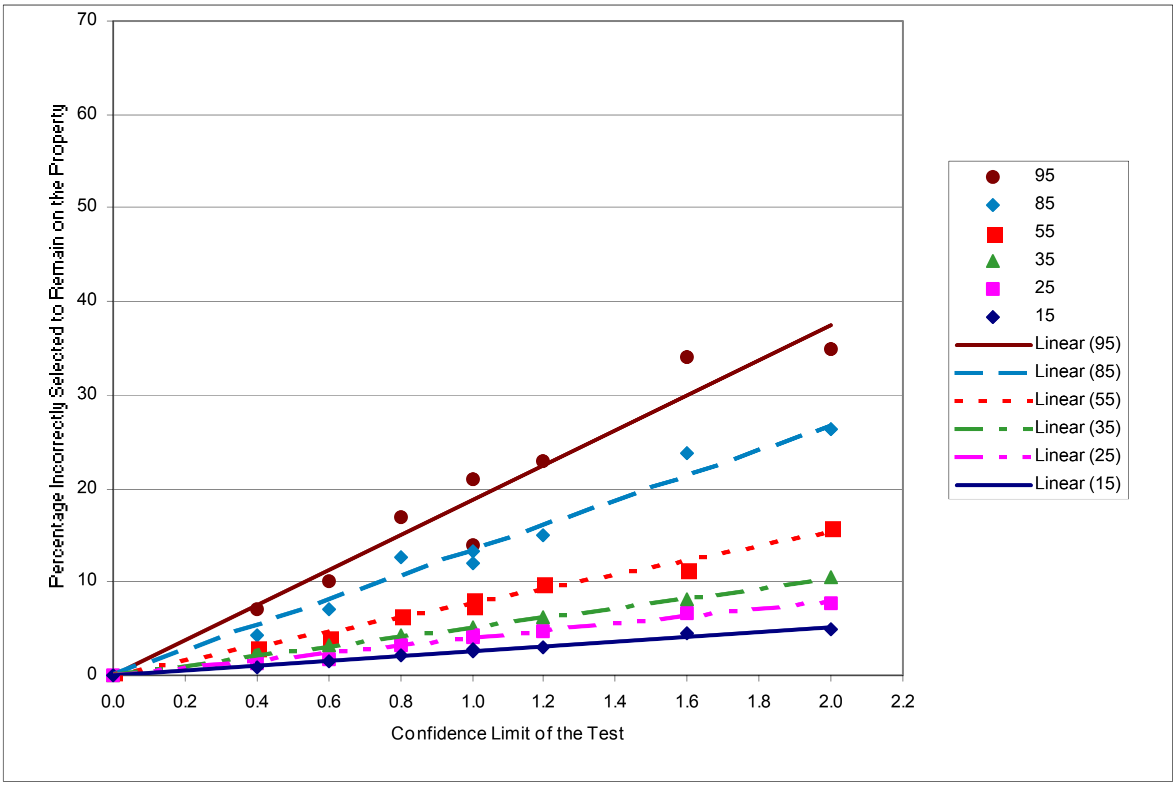
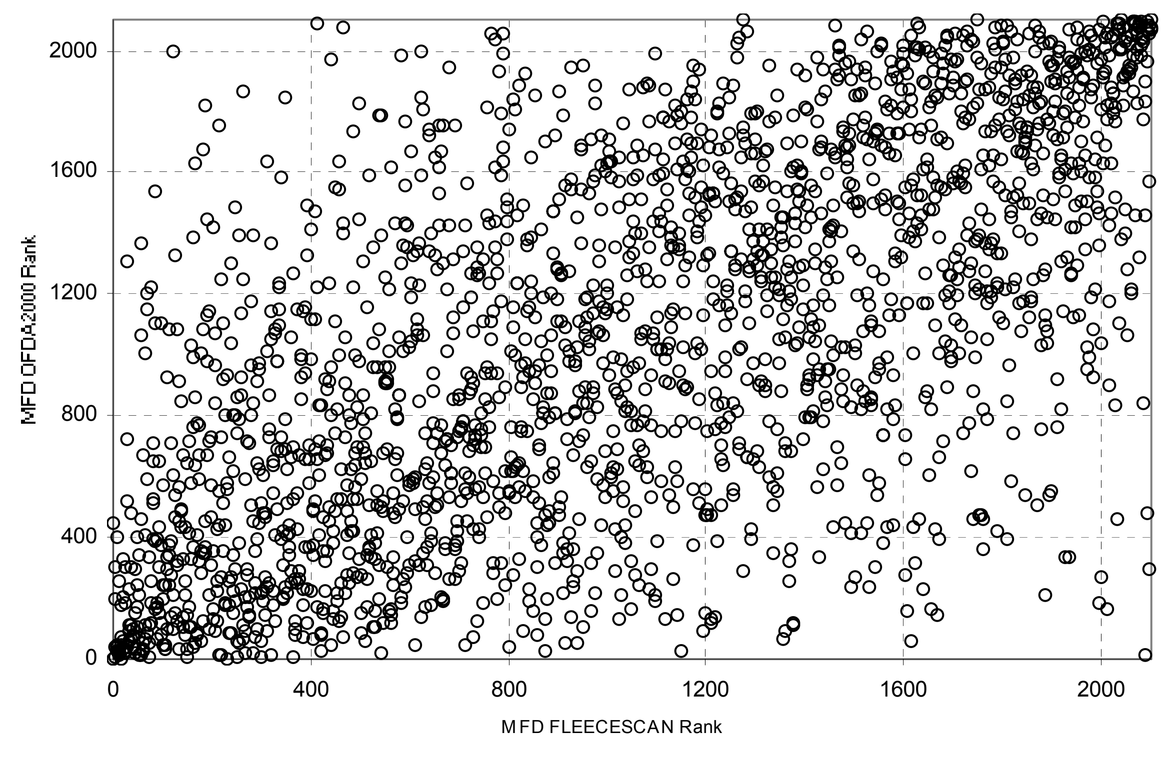


Figure 9: The Impact of the Precision of the Test on the Percentage of Animals Incorrectly Selected to Remain on the Property for Different Culling Rates (Between-Sheep SD = 2.0 μm)



A number of wool producers have had their wool tested by both FLEECESCAN and OFDA2000 and have been dismayed at the comparisons. Figure 10 reproduces data from an NZPAC Ltd and NZWTA Ltd Report (2000). There appears to be little to no correlation between the rank that was provided by the two different measurement systems.

Figure 10: Comparison of FLEECESCAN MFD Rank and OFDA2000 MFD Rank



The data created for this report provides an opportunity to compare two measurement systems with either the same or different Confidence Limits. We can consider three different laboratories for the purpose of illustration. Laboratories 1 and 2 have the same measurement capability of ±1.0µm whereas Laboratory 3 has a measurement capability of ±1.6µm. The author believes these hypothetical capabilities exist for commercial operators providing services to growers.

Figure 11 compares two measurement systems both with a capability of ±1.0µm (eg Laboratory 1 compared to Laboratory 2). The mob chosen for this comparison exhibited a high variation in Fibre Diameter (Between-Sheep SD of mob equals 2.0µm).

Figure 12 compares two measurement systems with different capabilities; one system with a measurement capability of ±1.0µm and another system with a measurement capability of ±1.6µm (eg Laboratory 3 compared to Laboratory 2). The mob chosen for this comparison exhibited a high variation in Fibre Diameter (Between-Sheep SD of mob equals 2.0µm).

Figures 13 and 14 show the same comparisons but for a mob with more uniform variation in Fibre Diameter (Between-Sheep SD of mob equals 1.0µm).

Comparing Figures 11 to 14 with the earlier reported data (NZPAC Ltd and NZWTA Ltd) (see Figure 10) would suggest that the observed scatter of results could simply arise from the measurement capabilities of the two sampling and measurement systems and the diameter uniformity of the mob tested.

Hence it is important that the Confidence Limit for both systems is determined for the entire fleece being tested.

Figure 11: Sheep Rank (Finest to Coarsest) Comparisons for two Laboratories with the Same Measurement Capabilities (±1.0μm) for a mob with a Between-Sheep SD of 2.0μm.

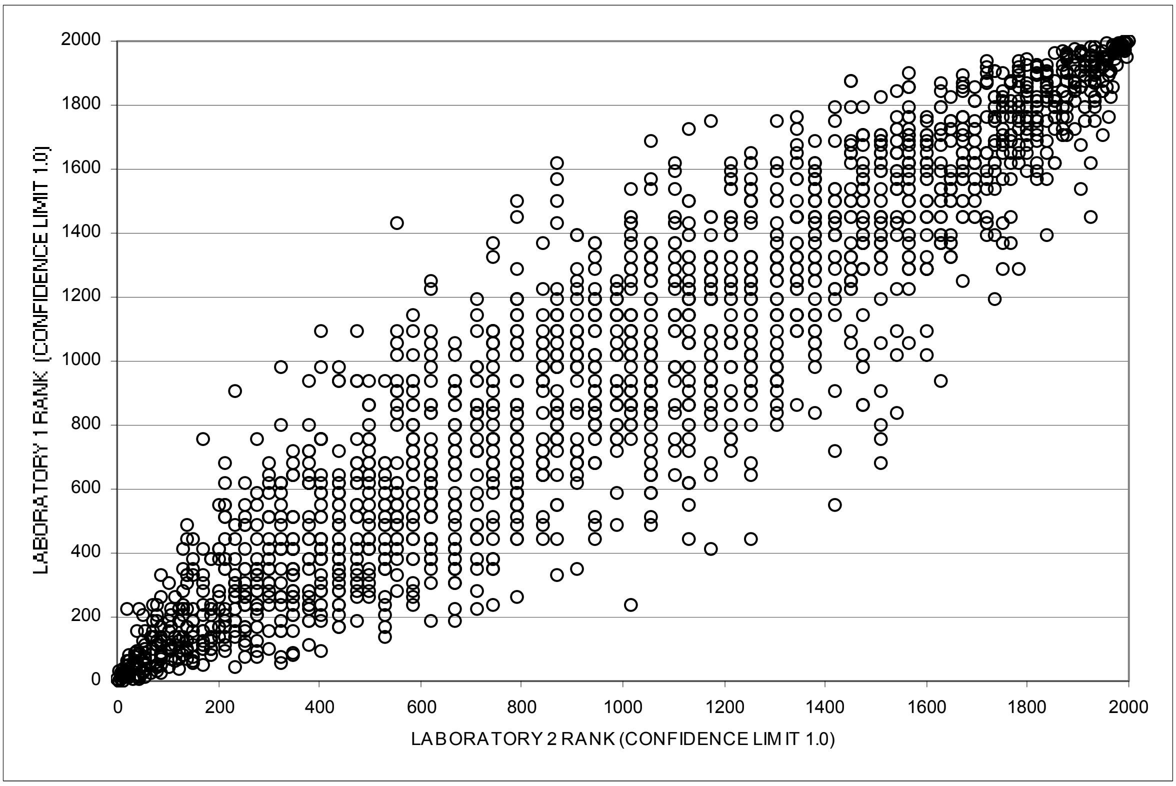


Figure 12: Sheep Rank (Finest to Coarsest) Comparisons for two Laboratories with Different Measurement Capabilities (±1.0μm and ±1.6μm) for a mob with a Between-Sheep SD of 2.0μm.

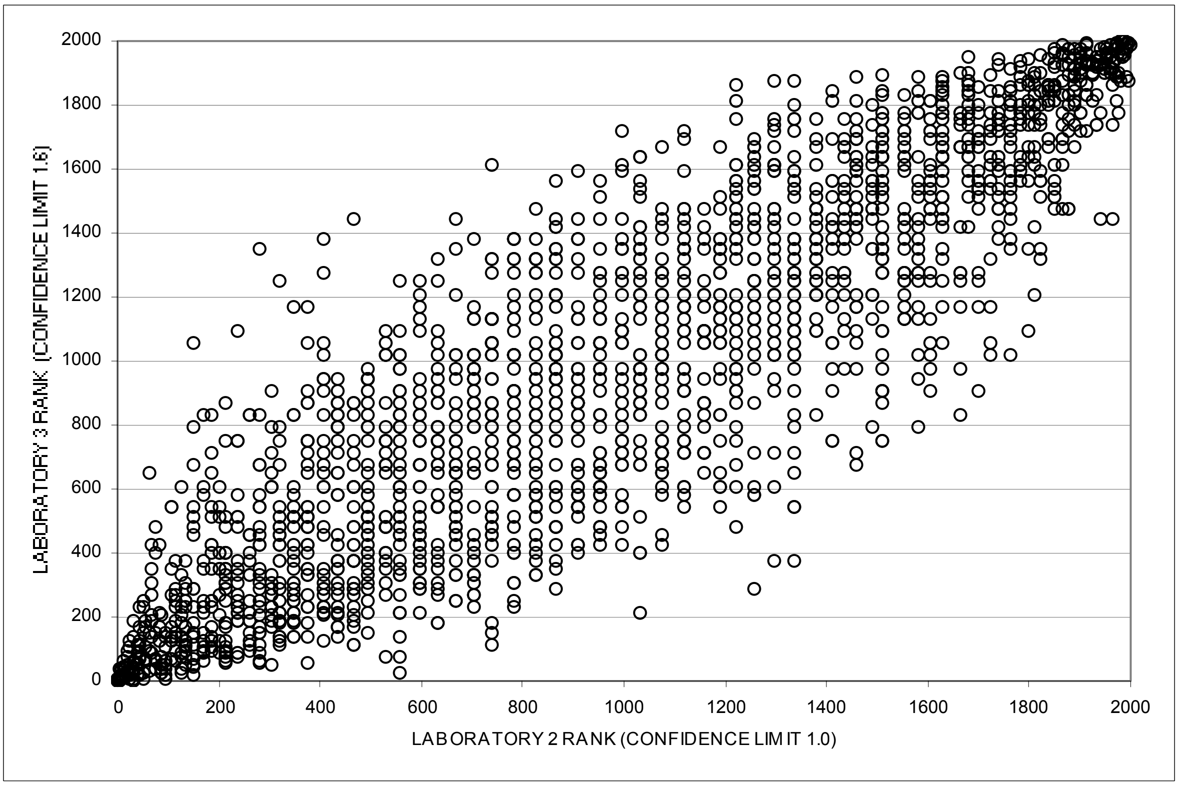


Figure 13: Sheep Rank (Finest to Coarsest) Comparisons for two Laboratories with the Same Measurement Capabilities (±1.0μm) for a mob with a Between-Sheep SD of 1.0μm.

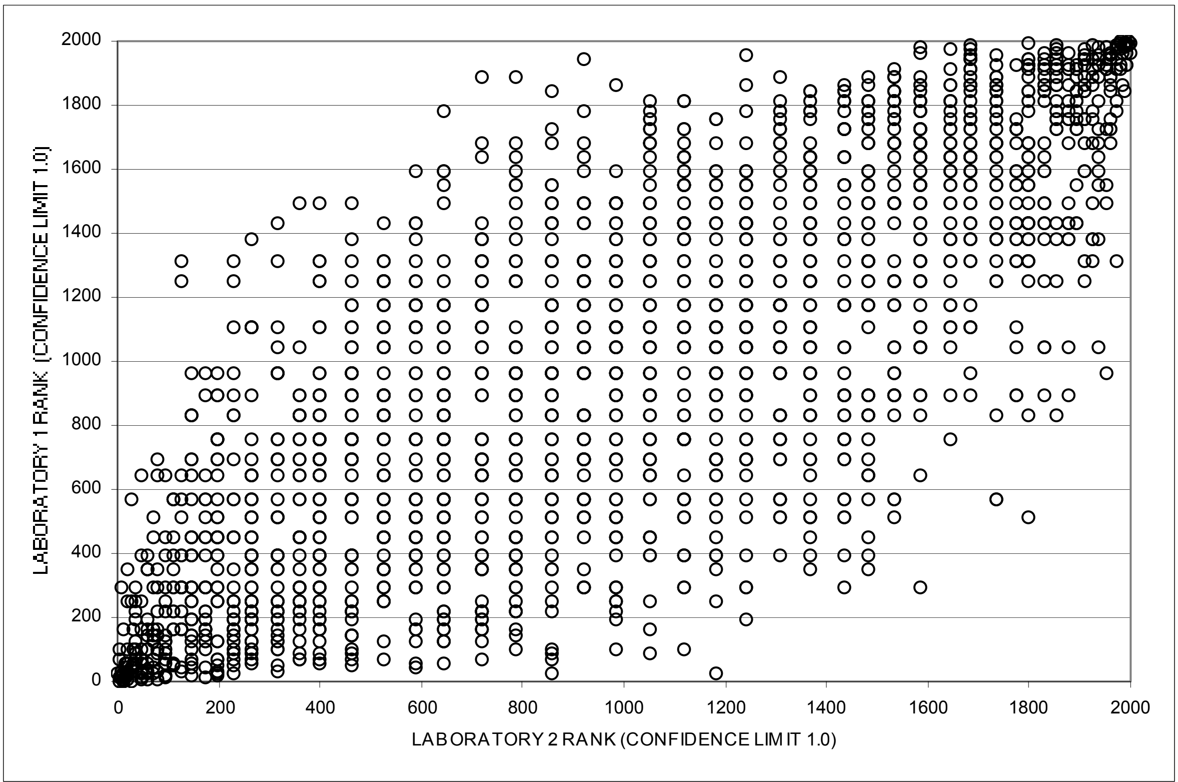
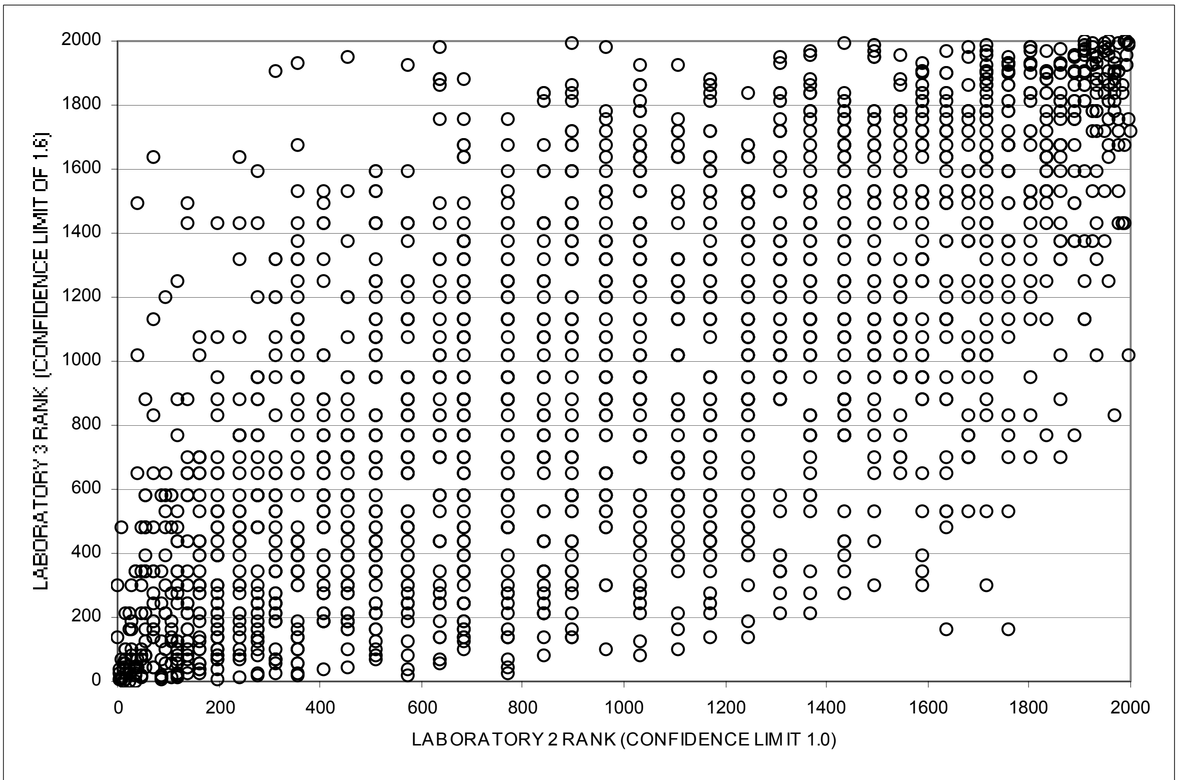


Figure 14: Sheep Rank (Finest to Coarsest) Comparisons for two Laboratories with Different Measurement Capabilities (±1.0μm and ±1.6μm) for a mob with a Between-Sheep SD of 1.0μm.



It is not surprising to the author that many wool producers are confused about the benefits or otherwise of the new technologies when applied "On-Farm". The outcome can clearly be different from wool producer to wool producer. Wool producers starting with a mob that has a relatively high variability in the MFD between their sheep have the most to gain.

The precision of the test result will play an important part in selecting the best animals (eg selecting the best rams, where the culling rate will be high). It is common that to have a test result with higher precision usually means a higher cost as this often entails testing more samples. Hence there will be a commercial balance between the cost of the test and the commercial benefits that can be derived from it. The best strategy for selecting the best rams could well be to test all the ram hoggets using the current fleece testing procedures and based on these results select a smaller number for more intensive sampling and testing before a final decision is made.

It must be remembered that sampling over the animal is an important factor in the overall precision of the result.

Turner (1956) was well aware of the problems of sampling at one site for determining Staple Length as noted in the following quotation from her work:

> *"Turner et al recommend the midside for general purposes, but define two other sites which should be sampled in addition for greater accuracy (as with top stud rams), or when an estimate of uniformity is required. These sites are on a diagonal line running through the midside; one is located near the hip bone and the other just behind the front leg, on the edge of the belly."*

The comment is as equally relevant to Mean Fibre Diameter as it is to Staple Length. Any differences (biases) that exist between sites over the fleece must be taken into account in the calculation of the Confidence Limit for the fleece either by:

- Random sampling over the entire fleece (eg as in the fleece minicorer used in the FLEECESCAN system); or
- Including a component of variance for the between-sites variation in the precision calculation where a single site is sampled and tested (eg mid-side sampling followed by laboratory testing, hip-bone sampling followed by OFDA2000 testing etc).

This will give the correct Confidence Limit for the fleece but doesn't address the issue of the impact of any bias between one particular site MFD and the overall fleece MFD.

Mid-side sampling was developed to aid genetic improvement in breeding programs. The heritability of Mean Fibre Diameter was reported by Taylor and Atkins (1992) to be 0.48 and so a moderate precision was all that was needed to input into the genetic models. The data generated for this report provides an opportunity to examine the impact of the Confidence Limit of a test result on the calculated "Genetic Gain per Annum". For the purpose of this exercise the 35% cull level was considered to represent the ewes and the 95% cull level represents the rams. The average MFD of the mob was assumed to be 19.0µm and the time taken to turnover one generation was considered to be 5 years. Table 2 shows the impact of Test Result Confidence Limit on the Genetic Gain. The data in the table is reported to two decimal places as the figures relate to the averages for populations of sheep and not a single animal.

Table 2: The Impact of Test Result Confidence Limit on Genetic Gain per Annum

Between-Sheep SD (µm)	Confidence Limit (µm)	MFD for Selected Ewes (µm)	MFD Differential (µm)	MFD for Selected Rams (µm)	MFD Differential (µm)	Genetic Gain per Annum (µm)
1.0	0.0	18.46	0.54	16.86	2.14	0.13
	0.4	18.46	0.54	16.94	2.06	0.12
	1.0	18.54	0.46	17.09	1.91	0.11
	1.6	18.59	0.41	17.40	1.60	0.10
	2.0	18.62	0.38	17.50	1.50	0.09
1.5	0.0	18.19	0.81	15.87	3.13	0.19
	0.4	18.21	0.79	15.89	3.11	0.19
	1.0	18.20	0.80	15.97	3.03	0.18
	1.6	18.29	0.71	16.36	2.64	0.16
	2.0	18.34	0.66	16.44	2.56	0.15
2.0	0.0	17.99	1.01	15.09	3.91	0.24
	0.4	17.98	1.02	15.02	3.98	0.24
	1.0	18.00	1.00	15.18	3.82	0.23
	1.6	18.09	0.91	15.30	3.70	0.22
	2.0	18.13	0.87	15.46	3.54	0.21

From Table 2 one can conclude the Confidence Limit of the Test Result has a small impact on the calculated "Genetic Gain per Annum". This is possibly a consequence of the heritability of only 0.48.

In order to make sound economic decisions wool producers need to know the diameter variation within their flock and they should use the sampling and testing procedures that give them the highest precision (ie the lowest Confidence Limit). **The precision must relate to an estimate representing the entire fleece and not the precision relevant to a single site on a sheep.** It is unfortunate that nearly all reported values in the literature relate to the precision at a single site of sampling. Estimates of precision on a single site must be corrected to represent the entire fleece by adding a component of variability due to the variability over the fleece to be relevant.

2. The Impact of Test Result Precision on Fleece Classing

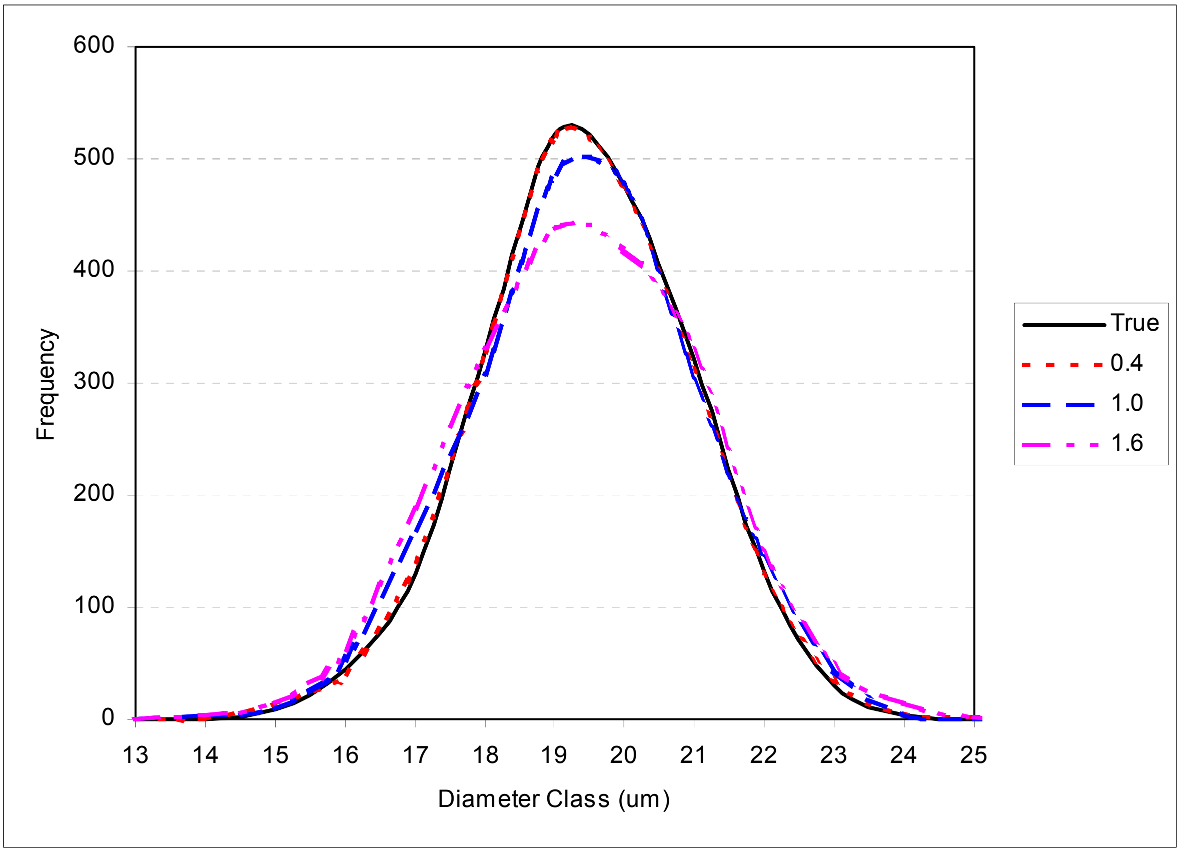
The impact of the test precision on classing the fleeces into different diameter classes in preparation for sale was examined by using the program "Classer" to set the diameter cut-offs for the preparation of four lines (Vizard and Williams, 1993). These were determined for a Mob with an overall average Mean Fibre Diameter of 19.0µm and a Standard Deviation of Mean Fibre Diameter between sheep of 1.5µm. A test precision of ±1.0µm was used for the purpose of setting the limits. Table 3 summarises the limits used.

Table 3: Classing Diameter Limits for the Four Classed Lines

Line Number	Lower Limit (µm)	Upper Limit (µm)
Line 1	0.0	18.2
Line 2	18.3	19.9
Line 3	20.0	20.8
Line 4	20.9	35.0

Figure 15 shows graphically the impact of the Measurement Confidence Limits on the measured diameter distribution for a mob of sheep. The "True" distribution is derived from the statistical programs in EXCEL having an Average MFD of 19.0µm and a Between-Sheep SD of 1.5µm. In the calculation of the measured values it is important to remember that no bias was introduced and so the distributions are centred on the same Average MFD.

Figure 15: The Effect of Measurement Precision on the Measured Diameter Distribution of a Mob of Sheep



From Figure 15 it can be seen that as the precision is worsened (ie the Confidence Limit is increased) larger numbers of sheep are measured finer and coarser than their "True" diameters. The impact this has on the measured diameter distribution is to stretch it in both the fine and coarse directions.

The sheep were then classed based on their "True Value" and their measured values based on three different Confidence Limits (±0.4µm, ±1.0µm and ±1.6µm).

Tables 4 to 6 show the comparisons of the expected Mean Fibre Diameters of the classed lines based on the average of the measured values for each fleece classed into a particular line and the average of the true values for the same fleece.

Table 4: Expected Diameters of Classed Lines Based on a Measurement Confidence Limit of **±0.4µm**

	Expected Value	True Value	**Difference**
Line 1	17.27	17.30	**0.03**
Line 2	19.06	19.06	**0.00**
Line 3	20.35	20.32	**-0.03**
Line 4	21.52	21.48	**-0.04**

Table 5: Expected Diameters of Classed Lines Based on a Measurement Confidence Limit of **±1.0µm**

	Expected Value	True Value	**Difference**
Line 1	17.20	17.43	**0.23**
Line 2	19.08	19.05	**-0.03**
Line 3	20.36	20.23	**-0.13**
Line 4	21.53	21.26	**-0.27**

Table 6: Expected Diameters of Classed Lines Based on a Measurement Confidence Limit of **±1.6µm**

	Expected Value	True Value	**Difference**
Line 1	17.13	17.60	**0.47**
Line 2	19.12	19.08	**-0.04**
Line 3	20.37	20.06	**-0.31**
Line 4	21.70	21.08	**-0.62**

The difference between the "Expected Value" and the "True Value" was greatest for the finest and coarsest lines. As the Confidence Limit increased the magnitude of these differences also increased up to approximately 0.5µm for a Measurement Confidence Limit of ±1.6µm. The above differences have been derived from a random testing model and consequently they would themselves have some uncertainty associated with them.

The lines prepared which were closer to the Mob average exhibited much smaller differences in MFD (eg Line 2) whereas the finest (Line 1) and coarsest (Line 4) prepared lines exhibited the largest differences. The finest prepared line had a true value coarser than what was expected and the coarsest prepared line had a true value finer than what was expected.

These effects have been referred to as "micron creep" or "diameter shrinkage" (Baxter, 2001) and as has been shown above arise from the Confidence Limit of the test and the diameter uniformity of the mob of sheep being classed.

CONCLUSIONS

The factors that have an impact on "On-Farm" decisions related to Mean Fibre Diameter are the precision (ie the 95% Confidence Limit) of the test result and the diameter uniformity of the mob of sheep. The higher the Confidence Limit the greater the adverse effect. Improving the diameter uniformity of the mob of sheep will require sampling and testing procedures of high precision (ie lower Confidence Limits) if economic gains are to be maintained and continued.

Research is needed to quantify the precision limits (ie the 95% Confidence Limits) of test procedures currently used by wool producers to make "On-Farm" decisions. It is important that the precision estimate is relevant to the entire fleece and not simply relevant to a single site on a fleece or sheep which has been the usually reported parameter in the scientific literature. It is equally important that the precision should relate to the exact method being used as deviations in the method can have significant impact on the precision.

REFERENCES

Baxter, B.P. (2000) *Precision of measurement of diameter, and diameter-length profile, of greasy wool staples on-farm, using the OFDA2000 instrument.* 10th International Wool Textile Research Conference, Aachen.

Baxter, B.P. (2001) *On-farm classing of Animals and Fleeces with the OFDA2000*. Wool Tech. Sheep. Breed., **40(2),** 133-155.

Cottle, D.J., Almeida, C.D., Baxter, B.P. and Petrie, D.J. (1996) *Precision of OFDA Fibre Diameter Measurements of Midside Wool Samples.* Wool Tech. Sheep Breed. **44(4):** 295-303.

Marler, J.W. and Couchman, R.C. (2001) *Note: The Ability of the OFDA2000 to Measure Fleeces and Sale Lots On-Farm.* Wool Tech. Sheep Breed. **19(4):** Accepted for publication.

Morgan, P.D. (1990) *Fleece Measurement Interlaboratory Trials – An analysis of the Factors Influencing the Precision of Test Results*. Wool Tech. Sheep Breed. **38(1):** 21-23.

NZPAC Limited and New Zealand Wool Testing Authority Limited. (2000) *Comparison of the OFDA2000 and Sirolan-FleeceScan Measurement Systems*. New Zealand PAC Project Report, September.

Peterson, A.D. and Gherardi, S.G. (2001) *The Ability of the OFDA2000 to Measure Fleeces and Sale Lots On-Farm*. Wool Tech. Sheep Breed. **19(2):** 110-132.

Taylor, P. and Atkins, K. (1992) *Merino Breeding Programs and Fibre Diameter Variability.* Wool Tech. Sheep Breed. **40(3):** 97-101.

Turner, H.N. (1956). *Measurement as an aid to selection in breeding sheep for wool production*. Animal Breeding Abstracts. **24(2):** 87-110.

Vizard, A.L. and Williams, S.H. (1993) *A model to estimate the economic value of using individual fleece diameter measurements to class wool*. Agric. Systems. **41:** 475-486.