Description and performance of the Agritest Staple Breaker model 2

Peter Baxter
SGS Wool Testing Services, Wellington, New Zealand

Summary

The Agritest Staple Breaker model 2 is a semi-automated instrument for the measurement of staple length, staple strength, and position of break. It uses the measurement principles outlined in the standard test methods AS 2810 and IWTO-30. This paper describes the performance of the instrument in terms of precision and accuracy, and compares the data obtained with data which has been published during development of these test methods. It is concluded that the instrument gives results which comply with the requirements of these standards.

Introduction

The performance of the Agritest Staple Breaker model 1 was described by Vizard et al (1994). The principal criticism of the instrument was the slowness of measurement due to all the procedures being manual. The Agritest Staple Breaker model 2 (SB2) is a much more sophisticated instrument which measures both staple length and staple strength, as well as position of break using a microprocessor-controlled operating system. It is capable of testing upwards of 250 staples per hour. This paper describes trials carried out to assess the repeatability and accuracy of the results from the instrument when performing tests in accordance with AS 2810 or IWTO-30, which are the Australian and IWTO test methods for length and strength respectively.

Description of the Staple Breaker Model 2

The Agritest Staple Breaker model 2 (SB2) is a totally redesigned version of the earlier instrument, the Staple Breaker model 1. The operation of this earlier device and the recording and analysis of results were all performed manually. In addition staple thickness was measured directly using a pneumatically-driven gauge. Despite these differences, Vizard demonstrated good agreement between staple strength results obtained from the Staple Breaker model 1 and the ATLAS device which is described as '*a suitable instrument*' in AS 2810 and IWTO-30. The Staple Breaker model 1 is now used widely in Australasia and overseas for education and research purposes.

The SB2 adopts a gravimetric procedure for estimation of staple thickness as described in AS 2810 and is highly automated. The only manual procedures involve placing each staple between the jaws of the device, and transfer of the tip and base portions of the ruptured staples to the electronic balances. The technology features a microprocessor-controlled electro-optic and pneumatic system which, after self-diagnostic checks, and upon activation, rapidly extends a moveable jaw whilst detecting the ends of the staple, measures the staple length, reliably clamps the ends of the staple, extends the staple to complete rupture, measures the force to rupture, releases the broken portions for manual transfer to electronic balances, weighs the broken portions, and then recycles to the next staple.

The instrument detects and requires corrective action for staples which are misaligned, or of an unmeasurable nature. After all the staples in a sample are broken, the instrument's software, using data input by the operator (either washing yield for flock testing purposes, or core test yield and lot category data for certification tests), calculates the average and variation of staple length, the average and variation of greasy and clean staple strength, and the position of break in terms of the percentage broken in the base, middle and tip. All raw and derived data can be output to a computer.

Calibration

The length measurement carried out by the SB2 must be calibrated against a standard instrument, since this measurement is carried out during the destructive process of strength measurement, and therefore cannot be directly calibrated by using standard gauge blocks. In the Australian Standard AS 2810, an instrument complying with AS 2720 is taken to be the reference. In the latter method, staples are transported on a belt between parallel vertical arrays of 8 light emitting diodes (LEDs) and detectors spaced at 2.5 mm centres. The length of each staple is defined by the belt travel from when the output of any detector falls to 50% of normal, to when the output from every detector has risen above this threshold and has remained there for the equivalent of 5 mm travel. The travel distance less 5 mm is defined as the staple length.

The Agritest Staple Length Meter (SLM) comprises a moving belt with an array of light emitting diodes on one side and an array of detectors on the other. It is calibrated using length blocks (of 50.0, 100.0, and 150.0 mm length) which are traceably calibrated to international standards of metrology. The principles of operation and measurement are exactly as described in AS 2720. The Agritest SLM also has the facility to output all data and derived values to a computer.

The SB2 uses exactly the same LED, detector array and detection criteria as the SLM, and once calibrated for length, (by remeasuring a sufficiently wide range of staples previously measured on the SLM), may be used for both length and strength measurement with only periodic checks on the length measurement system. The other measurement systems used on the SB2 - the force transducer for staple strength measurement, and the electronic balances, can both be checked using certified standard masses.

Assessment of repeatability - experimental design

Two SB2 instruments were used for the repeatability work.

Repeatability is difficult to measure in the length and strength test method because of the very high variances associated with the sampling as compared to the measurement process. The sampling variances cannot be reduced by the commonly-used expedient of blending (which is used to reduce this source of variance in all core test methods), since the between-staples component of variance is extremely high. Additionally, of course, on the SB2, staple measurement is destructive, so the same staple cannot be remeasured.

In order to assess repeatability, two basic procedures were followed:

- Split staple comparisons - this procedure was used to allow direct comparisons to be made with similar work carried out by CSIRO during development of the staple length and strength test method. 10 lots of wool were chosen to cover a wide range of types - 5 were of Australian origin and 5 were from New Zealand. Fleece and oddment samples were used. From each lot, 20 relatively thick staples were selected, and each staple was split into two approximately equal portions. The two portions of the staples were assigned in sequence to two sets, which were either tested on two instruments or on two different days.
- Split lot comparisons - staples from samples of 58 staples each (comprising both fleece and oddment lines, but all of New Zealand wools) were assigned alternately but randomly to the two SB2 instruments, thereby giving ‘subsamples’ of 29 staples each.

Results for within-instrument repeatability

Assessment of within-instrument repeatability was carried out using the split staple method, and was only carried out on one instrument since it became clear that the results were dominated by the variances between the two halves of the staples rather than between the two days of measurement. Table 1 shows the overall means and standard deviations of the lot differences over the 10 lots of 20 split staples.

Table 1: Within-instrument differences (day 1 to day 2)

	Day 1 to day 2 differences		CSIRO data	
Measurement	Average	SD	Average	SD
SLM staple length	-0.3	1.8		
SB2 staple length	0.0	1.4	-0.3	0.6
CoV staple length	-0.3	1.0		
Staple strength	0.3	2.0	0.4	0.6
Position of break	-0.8	1.3	-0.2	0.6

Comparison data to the right of the table (headed CSIRO) refers to ATLAS Repeatability trials reported to the Australian Standards TX/12 committee in document TX/12/84-47 (Standards Australia 1984c). Their results were obtained on 12 lots of 20 staples each. Overall none of the average differences are significantly different from zero. The standard deviations of the differences on the left hand side of the table are all significantly larger than those shown on the right, but this is almost certainly due to the nature of the samples rather than the instruments. It will be seen when we look at between-instruments that the variances are statistically indistinguishable from the above.

Analyses of variance were carried out for each parameter:

Table 2 - ANOVA - Staple length

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	69252.926	9	7694.770	42.735	0.0000
Day	0.206	1	0.206	0.001	0.9731
Sample x day	176.219	9	19.580	0.109	0.9995
Error	67341.342	374	180.057		
Total	136770.693	393	348.017		

Table 3 - ANOVA - Staple strength

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	51510.075	9	5723.342	53.652	0.0000
Day	7.454	1	7.454	0.0700	0.7917
Sample x day	334.229	9	37.137	0.3480	0.9581
Error	39896.459	374	106.675		
Total	91748.218	393	233.456		

Table 4 - ANOVA - Position of break %

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	14973.200	9	1663.689	12.567	0.0000
Day	72.709	1	72.709	0.549	0.4591
Sample x day	157.100	9	17.456	0.132	0.9988
Error	49510.649	374	132.381		
Total	64713.658	393	164.666		

Table 5 - ANOVA - Staple length measured by SLM

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	59736.186	9	6637.354	40.555	0.0000
Day	10.723	1	10.723	0.066	0.7981
Sample x day	283.911	9	31.546	0.193	0.9949
Error	61209.314	374	163.661		
Total	121240.135	393	308.499		

Several points are worth noting in Tables 2 through 5. In the model used, 'day' is not a significant factor, nor are there significant interactions between 'day' and 'sample'. In each case 'sample' is the only factor affecting the results, which is exactly how it should be. The error terms in the analyses represent residual variances which cannot be explained by the models - in this case the variance between measurements on the staple halves. It is worth noting that measurements on the SLM of the two staple halves gave an error term which is similar (statistically indistinguishable) from the error term in the analysis of the SB2 length data. In other words, the measurement of staple length on the SB2 gave rise to no additional variance to that which was evidenced when using the standard instrument.

Between-instrument repeatability - split staples

Between-instrument repeatability was assessed using the split-staple method on 3 separate days but using the same lots and the same staple preparation operator. Table 6 shows the results compared against identical trials carried out by CSIRO on early ATLAS instruments. In this case there are no significant differences in variance between the left and right hand sides of the table.

Table 6 - Between-instrument differences

	Between instruments: day 1		Between instruments: day 2		Between instruments: day 3		CSIRO data	
Measurement	Average	SD	Average	SD	Average	SD	Average	SD
SLM* staple length	0.1	1.2	0.0	1.2	0.0	1.4		
SB2 staple length	-0.4	1.8	-0.3	1.0	-0.6	1.1	0.0	1.0
CoV staple length	0.4	1.0	0.4	1.0	1.2	1.0		
Staple strength	1.0	1.6	1.6	2.4	0.7	1.4	-1.1	1.6
Position of break	-0.2	1.4	-0.2	1.5	0.4	1.8	-2.7	2.4

It should be noted that in this case the figures for SLM are again the between-staple-halves differences, since all SLM measurements were carried out on the same instrument. The data in this column provides a check against the reference method (AS 2720).

Only in one case (CoV SL day 3) was the average difference between instruments significant at the 0.05 level. In all cases the variances (SD^2) are not significantly different from the CSIRO data. In all cases also, there is no significant difference at the 0.05 level between the variances of the differences for the SLM and SB2.

The above data confirms that the SB2 has a similar level of between-instrument repeatability as the ATLAS, and in length measurement, a similar level to the reference method AS 2710.

Tables 7 through 10 show the results of analyses of variance on the relevant factors and interactions:

Table 7 - ANOVA - Staple length

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	200990.637	9	22332.293	162.970	0.0000
Instrument	72.625	1	72.625	0.530	0.4668
Day	600.164	2	300.082	2.190	0.1124
Sample x Inst	296.117	9	32.902	0.240	0.9886
Sample x Day	7590.662	18	421.703	3.077	0.0000
Inst x Day	17.329	2	8.665	0.063	0.9387
Sample x Inst x Day	208.313	18	11.573	0.084	1.0000
Error	150462.246	1098	137.033		
Total	360232.898	1157	311.351		

Table 8 - ANOVA - Staple strength

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	131969.213	9	14663.246	131.578	0.0000
Instrument	321.782	1	321.782	2.887	0.0896
Day	114.778	2	57.389	0.515	0.5977
Sample x Inst	306.552	9	34.061	0.306	0.9732
Sample x Day	4340.980	18	241.166	2.164	0.0033
Inst x Day	26.686	2	13.343	0.120	0.8872
Sample x Inst x Day	496.018	18	27.557	0.247	0.9995
Error	122362.563	1098	111.441		
Total	259965.153	1157	224.688		

Table 9 - ANOVA - Position of break %

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	64181.496	9	7131.277	62.507	0.0000
Instrument	2.292	1	2.292	0.020	0.8873
Day	85.879	2	42.939	0.376	0.6864
Sample x Inst	414.473	9	46.053	0.404	0.9336

Sample x Day	5178.872	18	287.715	2.522	0.0004
Inst x Day	39.582	2	19.791	0.173	0.8408
Sample x Inst x Day	329.561	18	18.309	0.160	1.0000
Error	125268.862	1098	114.088		
Total	195543.209	1157	169.009		

Table 10 - ANOVA - Staple length measured by SLM

Due to	Sum squares	DoF	Mean square	F-stat	Significance
Sample	173181.028	9	19242.336	136.778	0.0000
Instrument	0.352	1	0.352	0.003	0.9601
Day	509.918	2	254.959	1.812	0.1640
Sample x Inst	124.606	9	13.845	0.098	0.9997
Sample x Day	9694.581	18	538.588	3.828	0.0000
Inst x Day	2.409	2	1.204	0.009	0.9915
Sample x Inst x Day	375.178	18	20.843	0.148	1.0000
Error	154470.069	1098	140.683		
Total	338347.787	1157	292.435		

These analyses of variance indicate that ‘sample’ is again the only significant variable, and that ‘instrument’ and ‘day’ had no discernible effect on the observed variation. In each case the interaction ‘sample x day’ was highly significant, but this is a reflection of the fact that the sample sets were prepared separately for each day. The error terms are similar to those shown in tables 2 through 5, and are all high, suggesting that the inherent between-staple-portions components of variance are too high for there to be any significant benefit in using this experimental technique as against the alternative of split lot tests.

Considering the entire set of staple by staple comparisons allows us to detect very small differences which are not apparent in the tables above. Whilst there was no significant difference in PoB % between the two instruments, the average difference in staple length was 0.5 (sd 5.2) mm and the average difference in staple strength was 1.0 (sd 6.7) N/ktex. These levels of differences between instruments were considered acceptable in all the published reports on ATLAS (Standards Australia 1984b,c, 1985a,b, Thompson et al 1988, Marler 1989, Jackson and Steer, Stubbs et al 1991).

Between-instrument repeatability - split lot comparisons

Assessment of between-instrument repeatability was also assessed using the ‘split-lot’ comparison method outlined above, on 69 samples. Whilst the subsamples each only comprised 29 staples, this was adequate over the range of samples available to evaluate the differences precisely. The measurements are summarised in figures 1 through 5. In these and the following figures, the ‘a’ plot shows the results from one instrument plotted against the other, whilst the ‘b’ plot shows the differences plotted against the means. The latter type of plot is recommended where the results from independent methods or instruments are to be compared, since it unambiguously highlights level-dependent differences when the data shows non-zero regression constants (see Baxter(1996) for a general discussion of this issue, and Altman and Bland (1983) as a source reference).

Figure 1a - Staple length comparisons - two SB2s

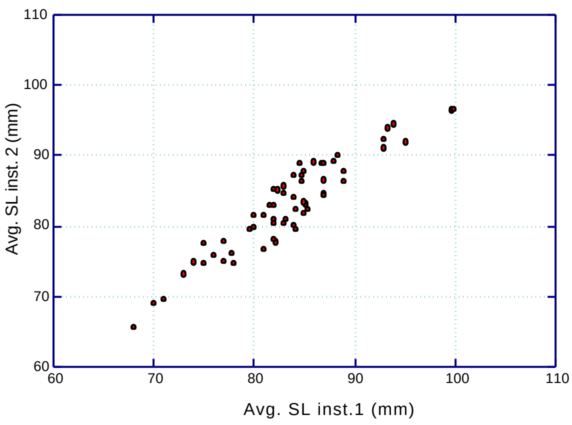


Figure 1b - Staple length differences between two SB2s

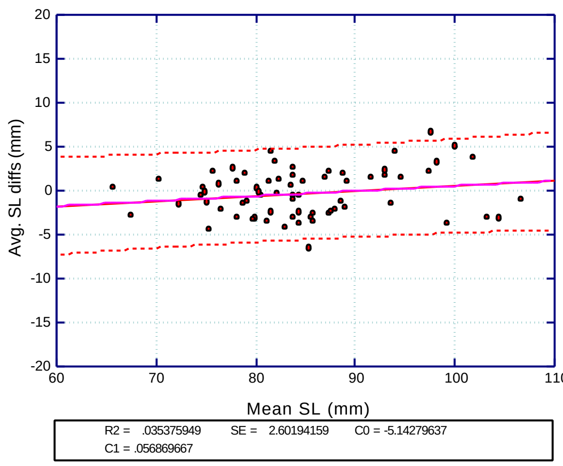


Figure 2a - Staple strength comparisons - two SB2s

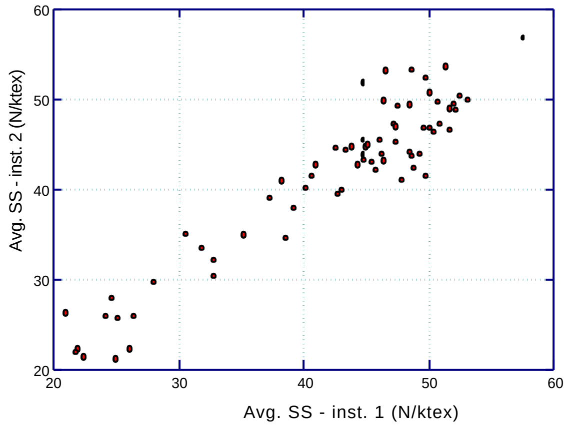


Figure 2b - Staple strength differences between two SB2s

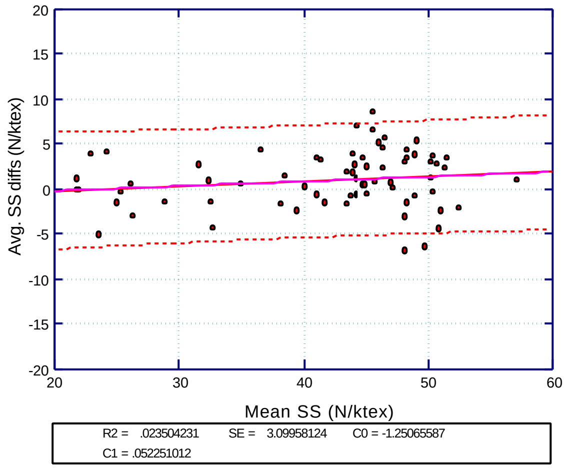


Figure 3a - CoV staple length comparisons - two SB2s

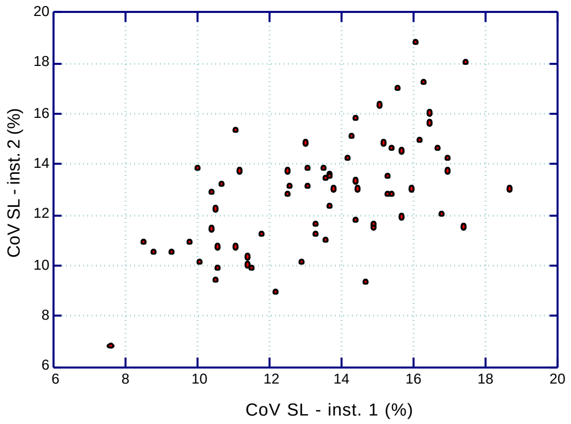
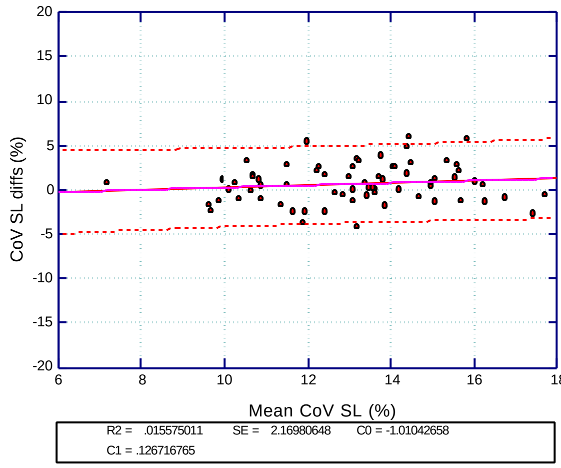


Figure 3b - CoV SL differences between two SB2s



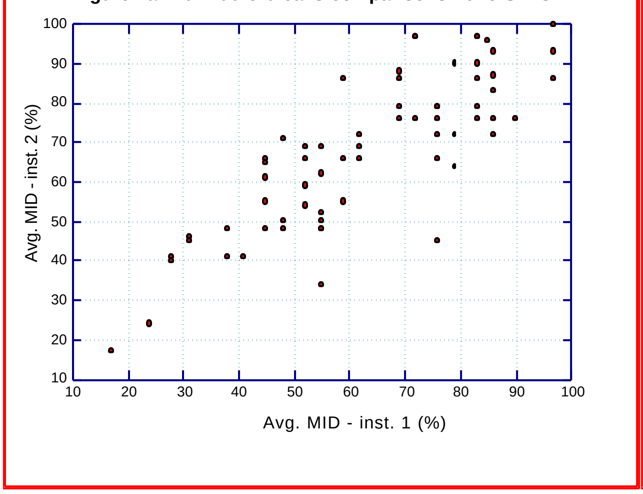


Figure 4a - % Middle breaks comparisons - two SB2s

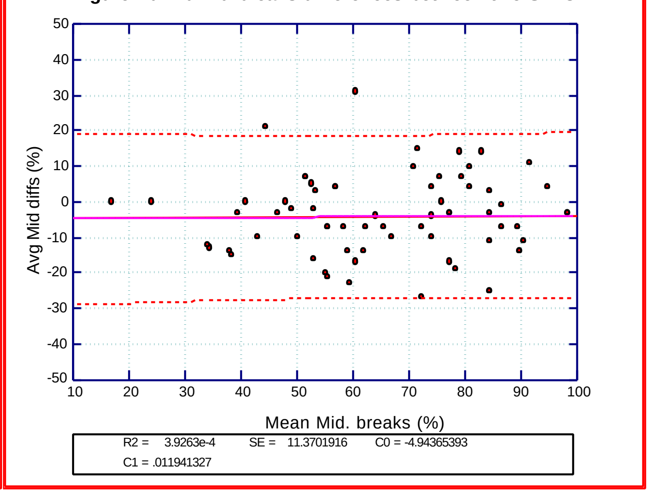


Figure 4b - % Mid breaks differences between two SB2s

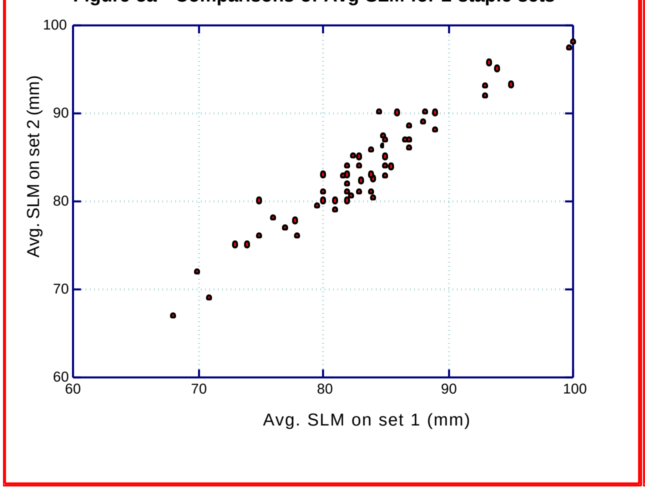


Figure 5a - Comparisons of Avg SLM for 2 staple sets

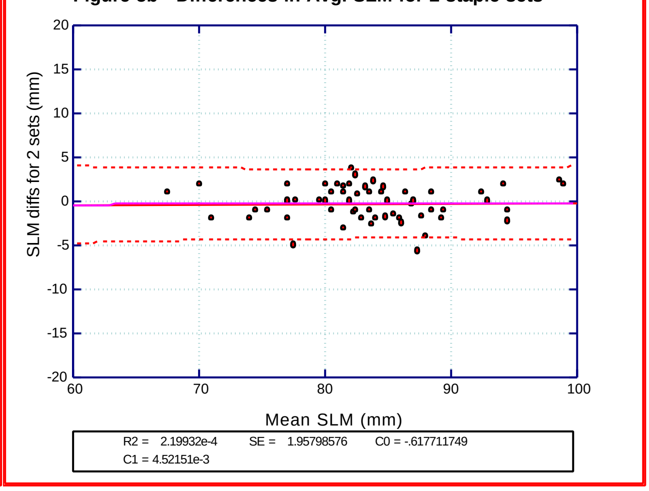


Figure 5b - Differences in Avg. SLM for 2 staple sets

Taking each comparison in turn:

- Staple length - the average difference was 0.3 mm (sd 2.6), which was not significant at the 0.05 level. The standard deviation of the differences is less than would be expected considering only the between-staple variances shown in IWTO-30. The regression coefficients in figure 1b are not significantly different from zero.
- Staple strength - the average difference was 0.9 (sd 3.1) N/ktex, which is significant at the 0.05 level, but within the acceptable range as defined in the literature. The standard deviation of the differences is less than would be expected considering only the between-staple variances shown in IWTO-30. The regression coefficients in figure 2b are not significantly different from zero.
- Coefficient of variation of staple length - the average difference was 0.7 (sd 2.2) %, which was significant at the 0.05 level. There is no information on which to base a judgement as to acceptability, but a review of the expected precision suggests that this is totally acceptable. The standard deviation of the differences is similar to the figure implied by the statements in IWTO-30. The regression coefficients in figure 3b are not significantly different from zero.
- Percentage of middle breaks - it should be noted that this is a measurement with particularly poor precision when only 29 staples are tested because of the way in which the figure is derived (the weight percentage of tip mass is converted to a tip, middle or base category for each staple, and the total counts in each category are converted to a percentage of the total number of staples - the best sensitivity obtainable in this experiment is therefore only 1 in 29 or 3%). The average difference on Mid% was 4.2 (sd 11.3) %, which is significant, but the data presented in

table 9 are perhaps more relevant. When the PoB% was compared on a staple by staple basis there was no significant difference between the instruments. This suggests that the difference illustrated here may be a function of the rounding used in the conversion from PoB% to position of break allocation within one of the instruments, and this is being investigated. There are no data on which to base a judgement of acceptability but common sense suggests that the comparison outcome is satisfactory. The regression coefficients in figure 4b are not significantly different from zero.

- Figures 5 are comparisons of the average staple lengths of the two subsamples from each lot, since only one SLM instrument was used. Only 60 pairs of comparisons were available. The average difference was 0.2 (sd 1.9) mm, which is not significant at the 0.05 level. The regression coefficients in figure 13b are not significantly different from zero.

Assessment of accuracy

Accuracy of staple length can be assessed by comparison with the staple length meter, as this is taken as the reference method in AS 2810. The assessment of staple strength accuracy is a more indeterminate process, since there are no absolute reference levels to work from. The test methods refer to the ATLAS instrument as being acceptable, so it provides a basis for comparison.

During the 1994-95 wool season, the opportunity was taken to obtain tuft samples from grab samples which had already been sampled in New Zealand for certification of length and strength (using ATLAS instruments) as part of the normal auction process. Reference to the auction catalogues subsequently allowed the certificated L&S data to be obtained on those lots which had been presale tested.

Comparison Results

The results of the comparisons undertaken under routine operational conditions are shown in figures 6 through 10. All SB2 values are based on the measurement of 58 staples for both length and strength on one instrument. The certificated values would be the result of measurements split between two ATLAS instruments, but this has little effect on the comparisons. It should be noted that the comparisons include components of variance due to tuft sampling, and between laboratories.

Figures 6 illustrate the relationship for individual staples between measured staple length on the SLM and on the SB2. This is effectively verification of the length calibration, and confirms that, as expected from the design, the systems perform in a totally linear manner.

Figure 6a - SLM and SB2 staple length comparison

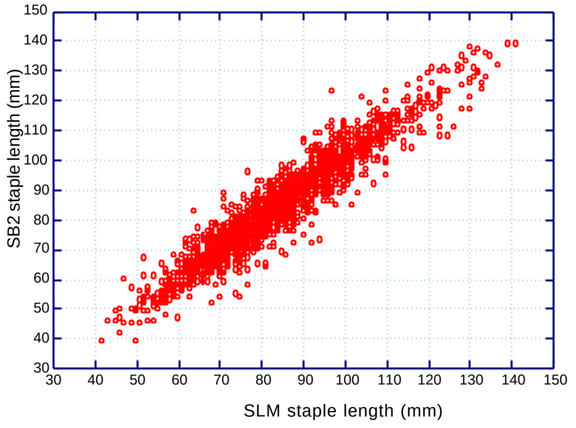
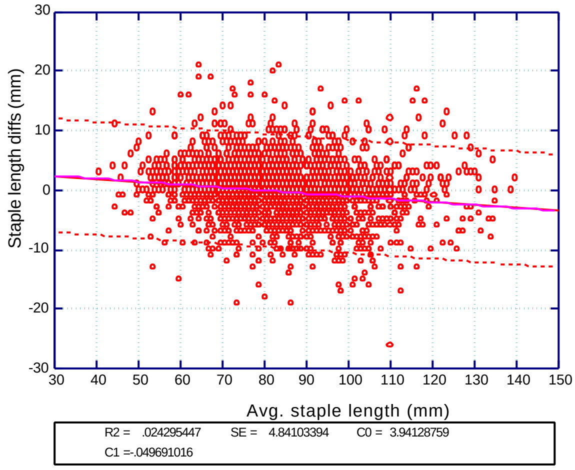


Figure 6b - SLM and SB2 staple length differences



Over the 2346 staples shown in this comparison, the average difference in staple length was 0.2 mm, with a standard deviation of differences of 4.9 mm. The difference is statistically significant but acceptable. The high precision of this comparison allows us to detect that the regression coefficients in figure 6b are both significantly different from zero, but the differences are trivial and of doubtful practical importance, as will be evident when we examine the comparisons between the SB2 and ATLAS-certificated results.

Figures 7 shows the differences between the mean staple lengths from the two instruments:

Figure 7a - Staple length comparison (Certified + SB2)

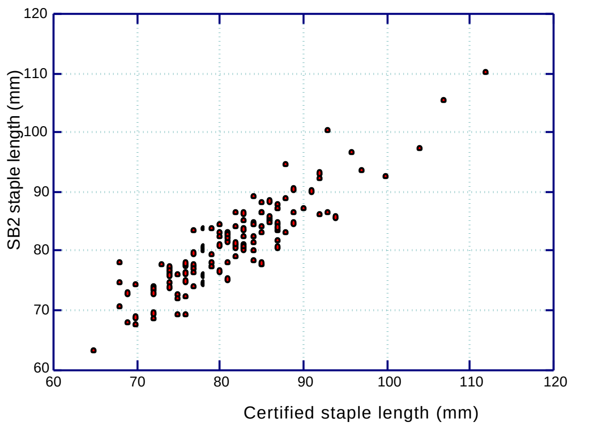
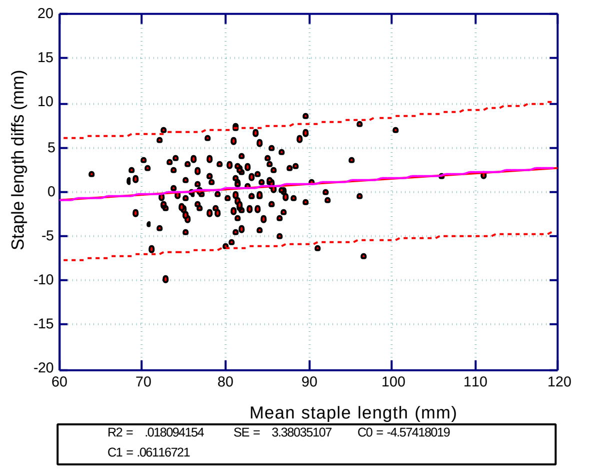


Figure 7b - Staple length differences (Certified + SB2)



The mean difference over the 126 lots is 0.4 mm, with a standard deviation of 3.4 mm, and it is not significantly different from zero. The regression coefficients in figure 7b are not significantly different from zero. The minor slope bias shown in figure 6b may therefore be confirmed as being of no practical importance.

It can be seen that the agreement between the two systems on mean staple length is excellent. It may also be noted that the standard error of the differences, at 3.4 mm, is below what may be expected from the statistics shown in IWTO-30, where table 2 shows the 95 % confidence limits of mean staple length when measured by 2 operators to be +/- 5.3 mm. The precision of comparisons is therefore expected to be approximately 3.8 mm

This confirms that the ATLAS and SB2 in this comparison performed at least as well as would be expected for two ATLAS instruments.

Figure 8a - Staple strength comparison (Certified + SB2)

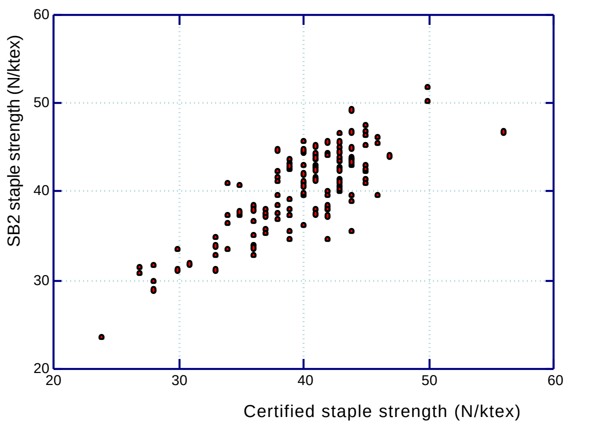
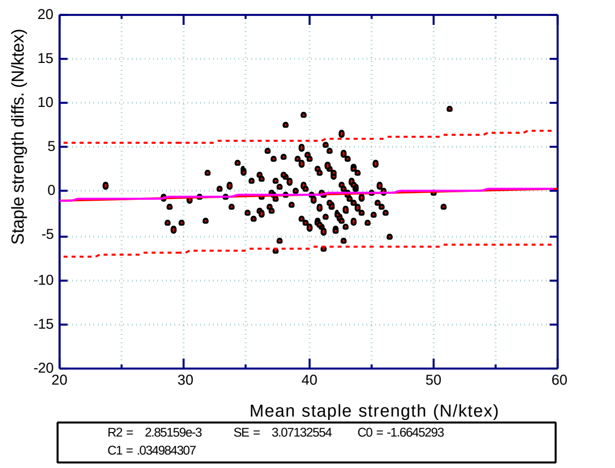


Figure 8b - Staple strength differences (Certified + SB2)



Figures 8 shows the comparison for mean staple strength. It can be seen that the agreement is again excellent: the mean of the paired differences is 0.3 N/ktex, with a standard deviation of 3.1, which again is not significantly different from zero. Figure 8b shows the differences plotted against mean staple strength. The standard deviation of the differences, at 3.1 N/ktex was again quite considerably less than the figure which would be expected from a perusal of IWTO-30, where, based on the quoted 95 % confidence limits of 5.9 N/ktex for 2 operators, we would expect the standard error of the differences to be of the order of 4.3 N/ktex.

The above cover the most important characteristics. For completeness, the comparisons for percentage of middle breaks and coefficient of variation of staple length are also shown in figures 9 and 10.

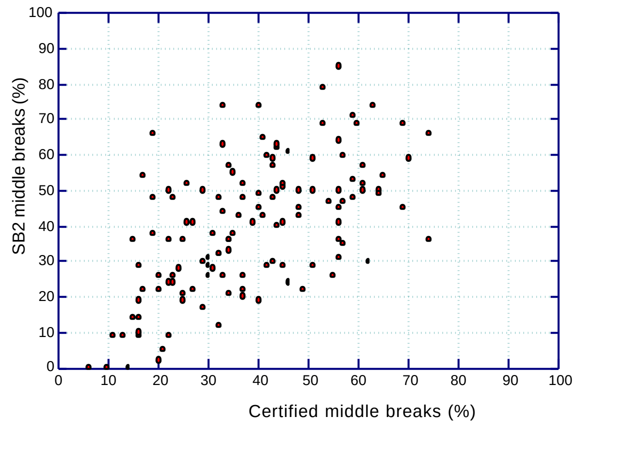


Figure 9a - Middle breaks comparison (Certified + SB2)

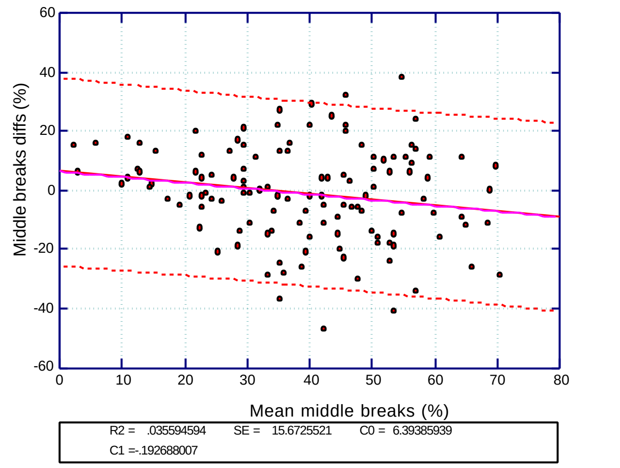


Figure 9b - Middle breaks differences (Certified + SB2)

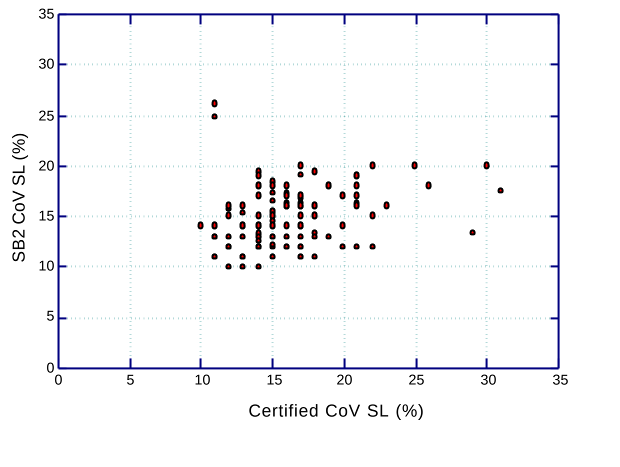


Figure 10a - CoV staple length comparison (Certified + SB2)

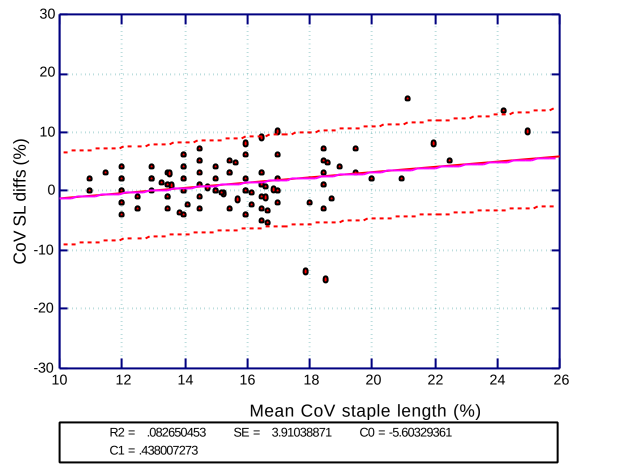


Figure 10b - CoV staple length differences (Certified + SB2)

The means of the paired differences are 1.2 %, standard deviation 15.9 for the percentage of middle breaks (not significant); and 1.2 % standard deviation 4.1 % for CoV of staple length, which is significant. The regression coefficient for the slope in figure 9b is marginally significant, whereas both coefficients are significant in 10b. However, in practical terms, the differences are considered acceptable. The CoV of staple length in particular is highly influenced by the sampling, and the differences indicated here are not considered indicative of any significant differences in the instruments, which would have shown up to a marked extent in the length comparisons. IWTO-30 quotes the precision of CoV staple length to be of the order of +/- 5 % at 90 mm, increasing to +/- 7% at 70 mm.

Conclusions

The work reported here shows quite clearly that the Agritest Staple Breaker model 2 instrument is capable of complying with AS 2810 and IWTO-30 both operationally and in terms of precision and accuracy. The improved performance over the model 1 makes the instrument a much more viable proposition for fleece testing operations.

References

Altman, D.G. and Bland, J.M. (1983), Measurement in medicine: the analysis of method comparison studies, The Statistician, 32, 307-317

Baxter, B.P. (1996), Procedures for comparing a) test methods, and b) different instruments within the same test method, IWTO technical report 17, Capetown

IWTO (1993). IWTO-30: Determination of Staple Length and Staple Strength

Jackson, M A and Steer, A J ( ). Investigations on staple conditioning time, IWTO

Marler, J W, (1989). Between-laboratories staple measurement round trials, IWTO Paris, Dec.

Standards Australia (1984a). AS 2720: Wool - Measurement of mean staple length - Method using the CSIRO staple length meter

Standards Australia (1984b). Comparisons between the CSIRO staple length meter and ATLAS, document TX/12/84-48

Standards Australia (1984c). ATLAS and PERSEUS repeatability trial data, document TX/12/84-47

Standards Australia (1985a).Comparison of ATLAS prototypes for measurement of SL and SS, document TX/12/85-4

Standards Australia (1985b). Staple length data from 3 lots measured via existing system and ATLAS, document TX/12/85-15

Standards Australia (1991). AS 2810: Wool - Determination of mean staple length and mean staple strength

Stubbs, C G, Marler, J W and Bow, M R, (1991). The precision of average staple length and average staple strength measurements, IWTO technical report 9, Nice

Thompson, R L, Andrews, M W, Baumann, A B, Bow, M R, Cantrall, C J, Heuer, P A, and Irvine, P A, (1988). The CSIRO Automatic Tester for Length and Strength (ATLAS), J. Text. Inst., 66 - 78

Vizard, A L, Scrivener, C J and Anderson, G A (1994). The agreement and repeatability of Atlas and Staplebreaker for estimating the staple strength of midside wool samples taken from merino sheep, Wool Tech. Sheep Breed., 42 (2), 112-119.