

University of Wales Trinity Saint David

Wales institute of Science and Arts



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

Ultrasonic attenuation coefficient determination

Structural integrity & material evaluation ACEA7013

Joseph Britton

2104278

December/2024

Table of Contents

Table of figures	iii
Table of tables.....	iii
Nomenclature	iv
Abstract.....	v
Apparatus.....	v
1. Introduction	1
1.1. Method	1
2. Measurements and graphs of raw measurements	4
3. Calculations and error estimates	7
3.1. Attenuation coefficient.....	7
3.2. Standard deviation & confidence limits	8
3.3. Normal distribution.....	9
3.4. Near field distance	9
4. Results and discussion	11
4.1. Mild steel block.....	11
4.2. Brake disc	14
5. Conclusion.....	17
References	18

Table of figures

Figure 1 - Apparatus list	vi
Figure 2 - Wave amplitudes for mild steel block with annotations.....	2
Figure 3 - Mild steel block width measurement	2
Figure 4 - Scan of brake disc showing 'noise' in data	3
Figure 5 - Direction labels for the mild steel block	4
Figure 6 – A) Raw data for the mild steel block width B) Raw data for the mild steel block length	5
Figure 7 - Sidewall effects for testing on mild steel block length.....	5
Figure 8 - raw data for brake disc sample represented graphically	6
Figure 9 - Calibration date for ultrasonic scanner	7
Figure 10 – A) Attenuation coefficients for the brake disc sample B) Attenuation coefficients for the mild steel block sample for the width C) Attenuation coefficients for the mild steel block sample for the length	8
Figure 11 - Statistical analysis of multiple echoes for the mild steel sample width.....	12
Figure 12 - Comparison of normal distributions for the mild steel block sample	12
Figure 13 - Illustrating 'noise' in data for third echo	13
Figure 14 - Statistical analysis of multiple echoes for the mild steel sample length.....	14
Figure 15 - Comparison of normal distributions for multiple echoes on brake disc	16
Figure 16 - Statistical analysis of multiple echoes for the brake disc sample	16

Table of tables

Table 1 - Raw results for mild steel sample	4
Table 2 - Raw data for the brake disc sample.....	6
Table 3 - Z table values for various confidence levels (Hazra, 2017).....	9
Table 4 - Attenuation coefficients for mild steel block sample	11
Table 5 - Attenuation coefficients for the brake disc sample.....	15

Nomenclature

Symbol/Term	Definition	Units
a	Attenuation coefficient	dB/mm
A_0	Initial wave amplitude	dB
A	Later wave amplitude	dB
z	Distance travelled by wave	mm
σ	Population standard deviation	Unitless
μ	Mean	Unitless
N	The size of the population	Unitless
x_i	Each value of the population	Unitless
n	The total number of values	Unitless
Z	Z-value	Unitless
λ	Wavelength	mm
NF	Near field distance	mm
D	Transducer diameter	mm

Abstract

Ultrasonic testing (UT) is one of the major Non-destructive testing (NDT) techniques to measure material properties by means of attenuation coefficient, which measures the loss in amplitude of ultrasonic waves while passing through a material. In the present experiment, attenuation coefficients for a mild steel block and a brake disc were measured using a 5-MHz zero-degree probe.

For the mild steel block, the reported values were 0.11 dB/mm as the attenuation coefficient for width and 0.07dB/mm due to its homogeneous structure; whereas for the brake disc, higher variability was attained, with attenuation coefficients as high as 0.41dB/mm, driven by surface irregularities, a heterogeneous composition, and interference caused by the near field. The reported values of standard deviation were also ± 0.05 dB/mm for steel and ± 0.15 dB/mm for the brake disc, while the confidence intervals showed the accuracy within ± 0.01 mm. This study emphasizes the importance of surface preparation and transducer alignment in UT accuracy, especially for irregular materials. While effective in the determination of attenuation coefficients, the reliability of UT diminishes under poor surface conditions and irregular geometries, hence signalling the need for an improved testing practice.

Apparatus

For images of the apparatus used for this experiment, see figure 1. Descriptions of the apparatus used is shown below.

- A. 5MHz transducer
- B. Mild steel block sample
- C. Couplant gel
- D. Vernier calliper
- E. Brake disc sample
- F. Silverwing D-Scan 705 ultrasonic flaw detector



1. Introduction

Ultrasonic testing is a widely used non-destructive testing (NDT) technique for detecting internal flaws of a component. One key parameter for ultrasonic testing is the attenuation coefficient of a material, this quantifies the loss of ultrasonic wave amplitude as it propagates through a material. This is critical for assessing material properties where an engineer may be unsure of material composition.

In this experiment, two materials will be analysed with an ultrasonic flaw detector and 5MHz zero-degree probe to measure amplitude decay between waves. These measurements are then used to calculate the attenuation coefficients of a mild steel rectangular block and a brake disc.

The aim of this experiment is to determine the attenuation coefficients of both materials and assess this techniques validity in doing so, user input on the system will also play a factor in this validity, hence it will be analysed.

1.1. Method

To begin this experiment, it required each material to be prepared with a couplant material in between the probe and material surface to eliminate any air. This ensures that no loss of beam energy as the acoustic impedance is matched (Xu, 2024). Also, the ultrasonic flaw detector needed to be calibrated for each material prior to use. This involves adjusting beam velocity to ensure distance measurements match real world measurements, taken with a vernier calliper accurate to 0.1mm, as illustrated through figures 2 and 3, velocity values can also be researched for specific materials through an ultrasonic velocity table.

It is essential to adjust the visual range on the x axis, for the mild steel block the width measured at $19\text{mm} \pm 0.1\text{mm}$, hence the x axis is set to 70mm allowing visualisation of 3 echoes. The gate must be moved into an adequate location when taking each reading, meaning adjustment of the x and y locations and width of the gate will need altering for each echo. Once correctly located, the graphical representation can be paused on the screen allowing time to record the data received, see figure 2.



Figure 2 - Wave amplitudes for mild steel block with annotations

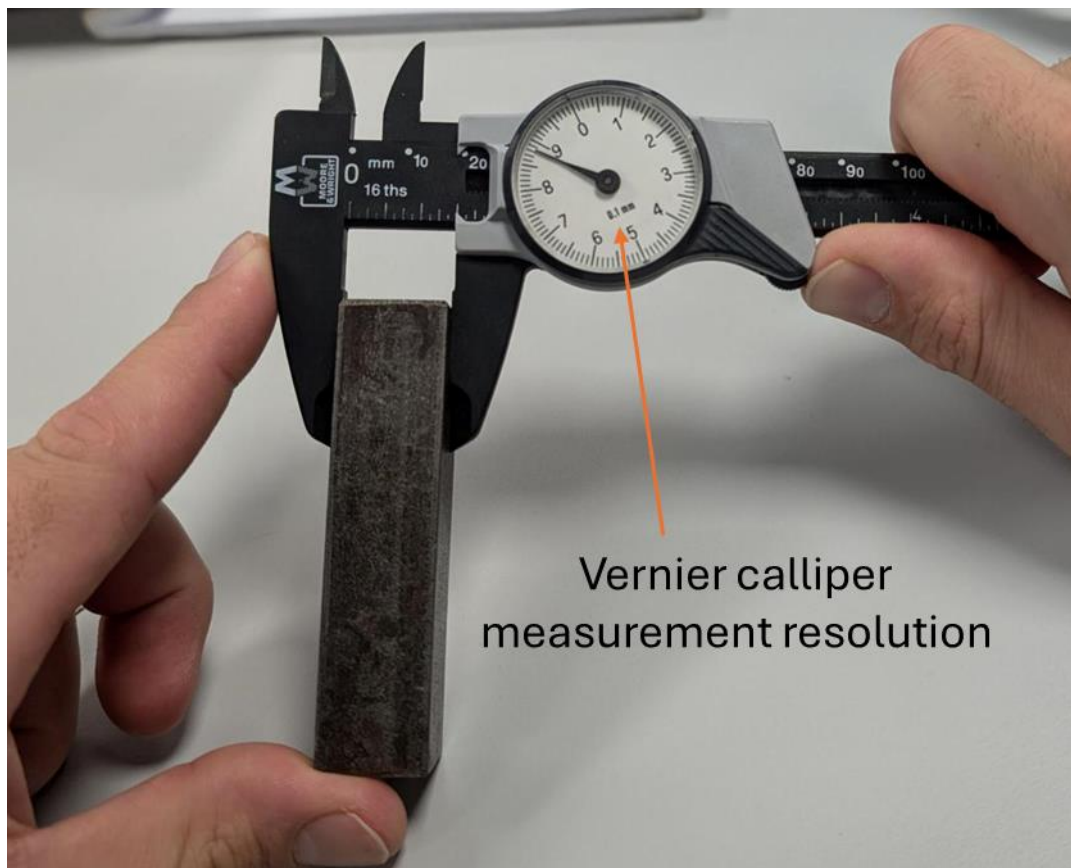


Figure 3 - Mild steel block width measurement

Due to complications caused with the brake disc, the testing was performed on the hub mount. This removed the anomalies in data received when testing the friction surface of the brake disc, caused from surface undulations, meaning the waves were hitting the material at an angle, thus leading to skewed and noisy data. Once more, rust on the opposite surface meant the thickness was not equal at all measuring points (Wang, 2018). The effects of this are shown in figure 4.

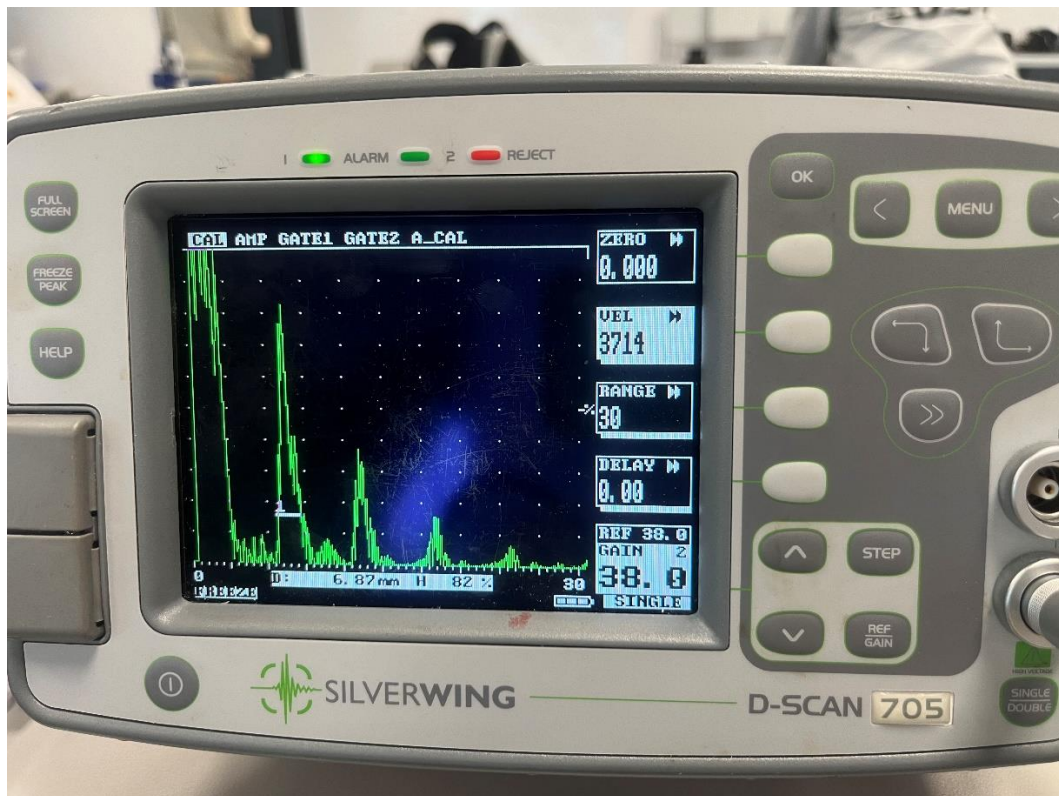


Figure 4 - Scan of brake disc showing 'noise' in data

2. Measurements and graphs of raw measurements

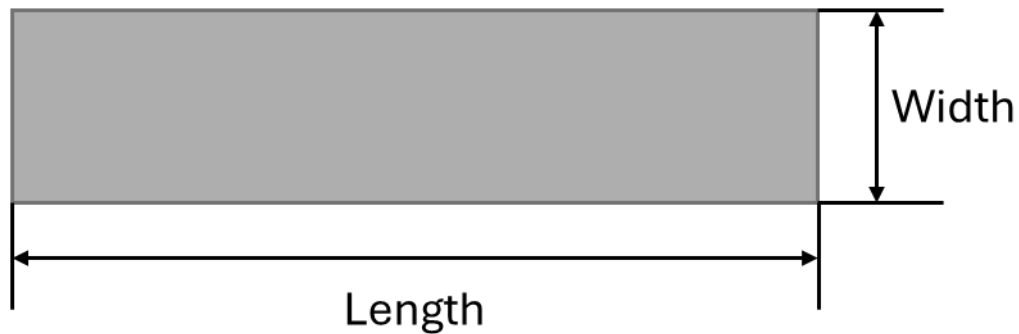


Figure 5 - Direction labels for the mild steel block

Table 1 - Raw results for mild steel sample

Mild steel rectangular block							
Measurement		Echo 1		Echo 2		Echo 3	
		Distance (mm)	Amplitude (%)	Distance (mm)	Amplitude (%)	Distance (mm)	Amplitude (%)
Direction 1 (width)	1	19.44	80.0	37.68	40.0	55.67	20.0
	2	19.44	80.0	37.44	39.0	55.67	20.0
	3	19.45	91.0	37.44	45.0	55.67	26.0
	4	19.47	85.0	37.44	41.0	55.67	23.0
	5	19.47	81.0	37.44	39.0	55.67	22.0
	6	19.46	90.0	37.44	48.0	55.67	38.0
	7	19.47	81.0	37.44	41.0	55.67	27.0
	8	19.47	78.0	37.44	39.0	55.67	30.0
	9	19.50	78.0	37.68	40.0	55.67	26.0
	10	19.50	73.0	37.68	38.0	55.67	23.0
Mean		19.47	81.7	37.51	41.0	55.67	25.5
Direction 2 (length)	11	76.87	84.0	166.80	31.0	241.20	28.0
	12	76.81	82.0	165.60	47.0	241.20	30.0
	13	76.84	79.0	165.60	30.0	241.20	21.0
	14	76.84	90.0	166.80	24.0	242.40	17.0
	15	76.81	70.0	165.60	42.0	241.20	28.0
Mean		76.83	81.0	166.08	34.8	241.44	24.8

Directional labels for the mild steel block sample are shown in figure 5 whilst figure 6 depicts the raw data for this sample. Raw measurements of results, seen through table 1, indicate good correlation of distance measurements leading to confidence in the performing engineer. The variance in amplitude between measurements may have numerous contributors, one likely factor comes from transducer alignment where small variations in position and angle can affect wave energy (Polonivhko, 2007). However, for both directional measurements, the amplitude regressions fall within expected tolerances (Mix., 2005).

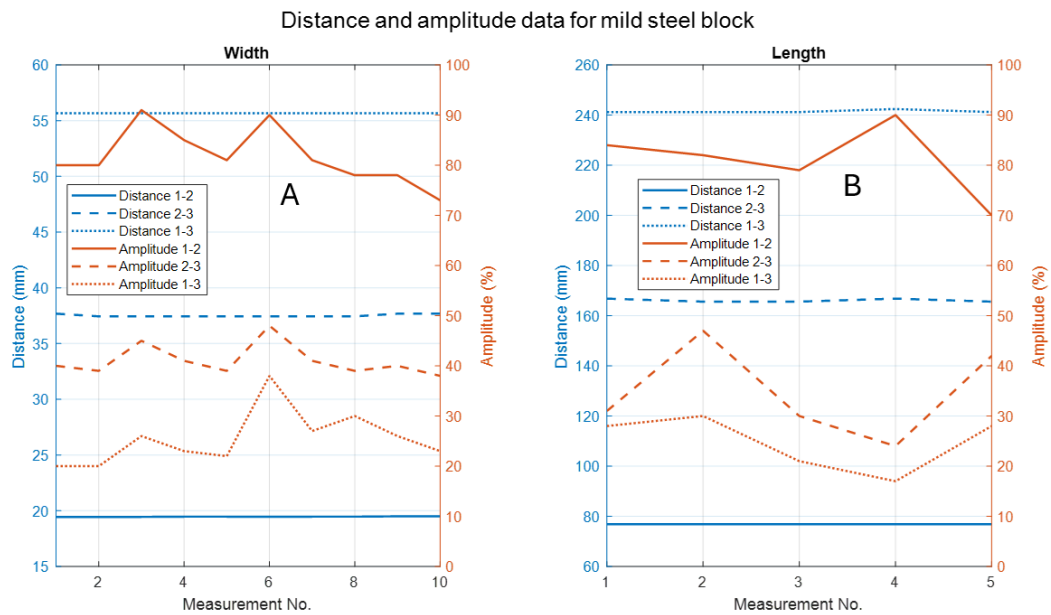


Figure 6 – A) Raw data for the mild steel block width B) Raw data for the mild steel block length

In analysing a component that is considerably longer in one direction, measurements taken across the longer dimension may yield several wave signals of varying amplitude on the ultrasonic scanner. This is because of the angle of beam spread causing the interactions with the sidewalls of the component rather than the final wall directly (Iowa State University, n.d.), thereby complicating data analysis as in Figure 6. To overcome this, the echo with the highest amplitude is selected, as it is the strongest wave reflection and most direct, shortest path, which minimizes interference from sidewalls and other factors.

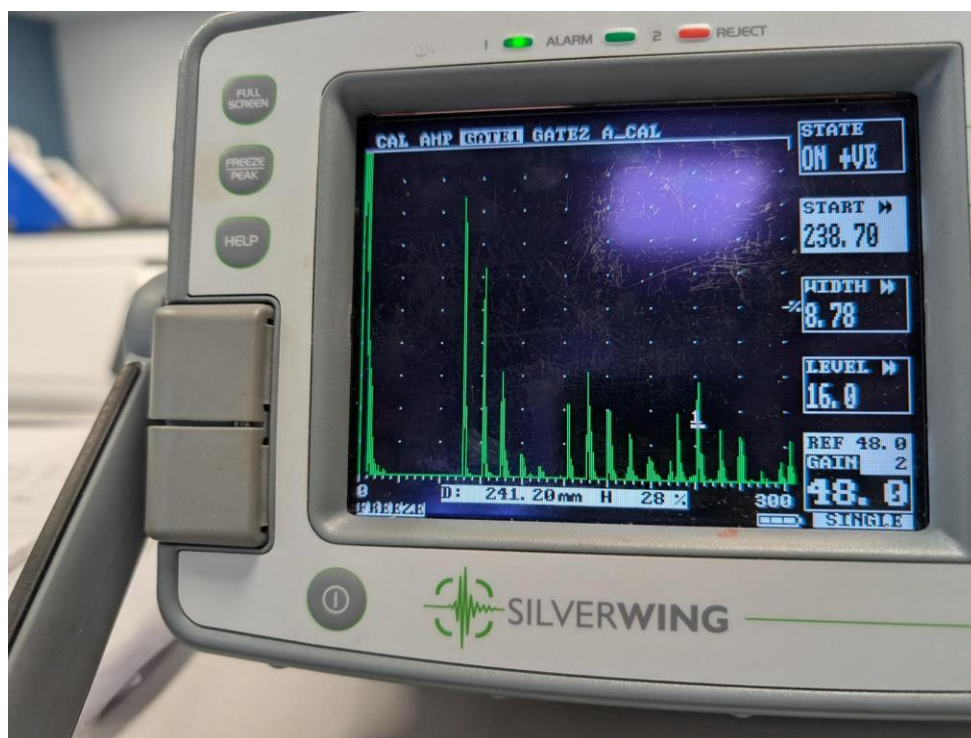


Figure 7 - Sidewall effects for testing on mild steel block length

Measurements taken for the brake disc show larger deviance with both distance and amplitude of the wave signal, see table 2 and figure 7. This is due to surface undulations from wear on the hub surface, preventing an optimal contact point for all measurements.

Table 2 - Raw data for the brake disc sample

Brake disc hub						
Measurement	Echo 1		Echo 2		Echo 3	
	Distance (mm)	Amplitude (%)	Distance (mm)	Amplitude (%)	Distance (mm)	Amplitude (%)
1	6.87	80.0	12.84	37.0	18.96	15.0
2	7.22	75.0	12.96	29.0	19.20	11.0
3	7.22	76.0	12.96	35.0	19.08	16.0
4	6.82	74.0	12.84	29.0	18.84	10.0
5	6.92	75.0	12.96	26.0	19.20	8.0
6	6.92	71.0	12.96	32.0	19.08	14.0
7	7.22	74.0	12.84	31.0	18.96	12.0
8	6.92	72.0	12.96	30.0	19.08	11.0
9	6.87	75.0	12.84	34.0	18.72	15.0
10	6.85	74.0	12.72	36.0	18.72	16.0
11	6.81	80.0	12.60	40.0	18.60	17.0
12	6.84	91.0	12.60	54.0	18.60	26.0
13	7.22	73.0	12.84	39.0	18.84	20.0
14	6.87	82.0	12.72	41.0	18.72	20.0
15	7.20	75.0	12.84	27.0	18.96	9.0
Mean	6.98	76.5	12.83	34.7	18.90	14.7

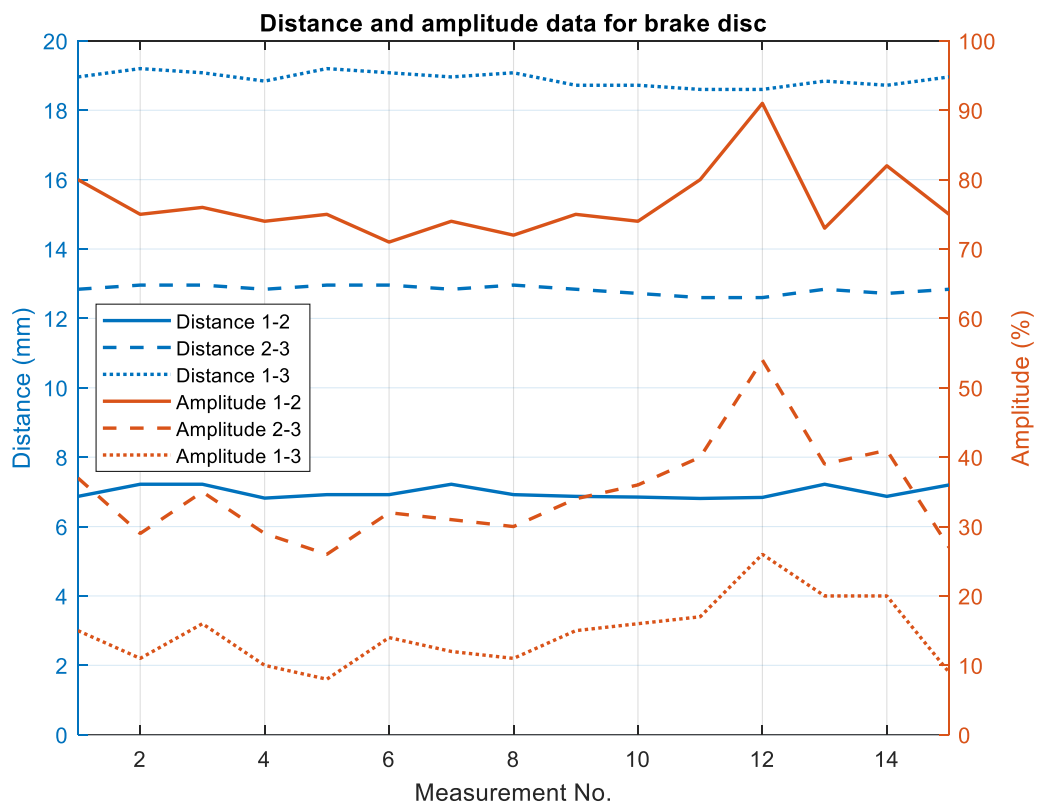


Figure 8 - raw data for brake disc sample represented graphically

3. Calculations and error estimates

3.1. Attenuation coefficient

The attenuation coefficient is calculated through equation 3, in the case of this experiment. This experiment used the measured distance as opposed to the calculated distance from the scanner, this is done to minimize reliance on the scanner due to issues with calibration dating, as seen in figure 9, meaning confidence in results is low.

$$A = A_0 e^{-az} \quad \text{Equation 1}$$

Rearrange for attenuation:

$$-a \left(\frac{\text{Napers}}{\text{mm}} \right) = \frac{\log\left(\frac{A}{A_0}\right)}{z} \quad \text{Equation 2}$$

Adapting for pressure:

$$-a \left(\frac{\text{dB}}{\text{mm}} \right) = \frac{20 \log_{10}\left(\frac{A}{A_0}\right)}{z} \quad \text{Equation 3}$$

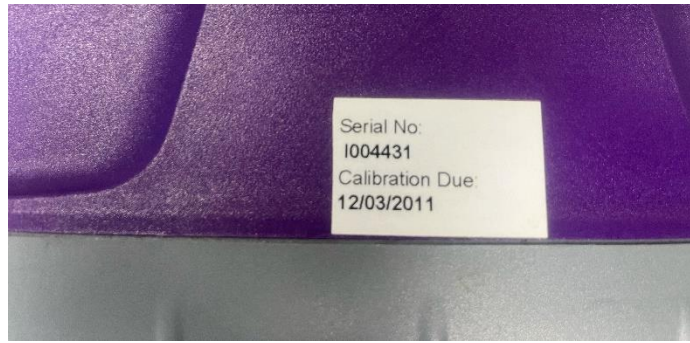


Figure 9 - Calibration date for ultrasonic scanner

Once calculated and plotted, it is apparent that the variance in results seen previously have affected the attenuation coefficients recorded for the brake disc vastly, the range of results being 0.41dB/mm, compared to 0.11dB/mm and 0.07dB/mm for the steel block width and length, as displayed through figure 10. This is largely due to reasons specified previously, which due to it being an irregular shape and an unprepared surface causes the large array of results.

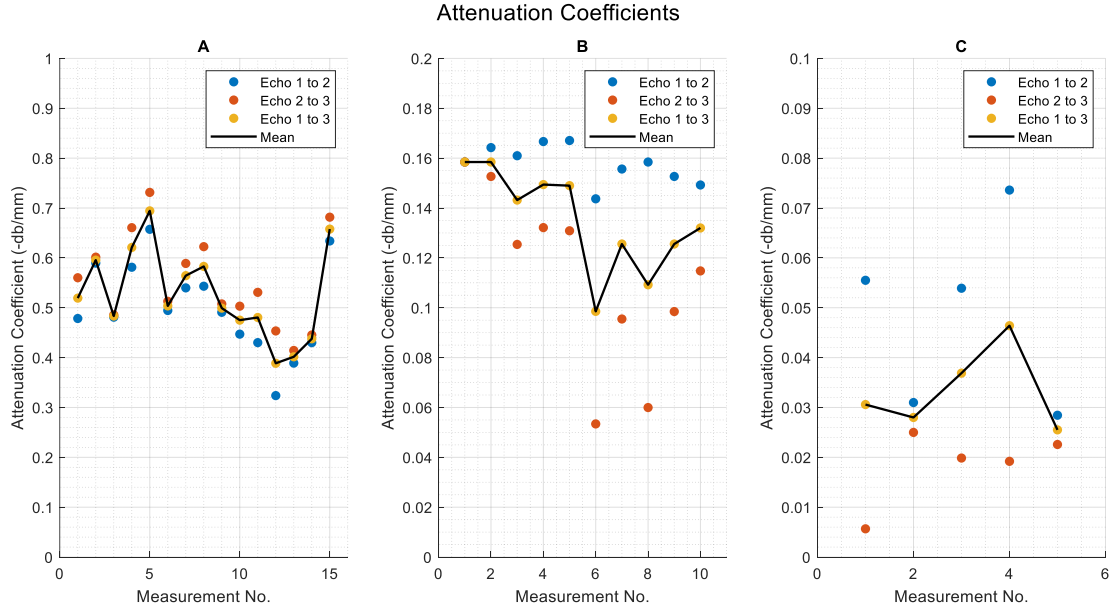


Figure 10 – A) Attenuation coefficients for the brake disc sample B) Attenuation coefficients for the mild steel block sample for the width C) Attenuation coefficients for the mild steel block sample for the length

3.2. Standard deviation & confidence limits

Statistical analysis of the data requires calculations of various values that allow interrogation of data accuracy and precision. Equations used to calculate the standard deviation and confidence limits can be seen through equations 4 and 5.

$$\sigma = \sqrt{\frac{\sum(x_i - u)^2}{N}} \quad \text{Equation 4}$$

To compute the average of a data set, equation 5 is applied, allowing computation of the confidence limits used.

$$\text{Mean} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{Equation 5}$$

$$CL = \text{Mean} \pm 1.96 * STDev \quad \text{Equation 6}$$

Within equation 6, the use of the multiplier 1.96 is a calculated value, this stems from the inverse cumulative probability which shows how many standard deviations a data point is from the mean. This calculation is shown through equation 7. Alternatively, these values can be searched through z tables shown in table 3 (Hazra, 2017).

$$Z = \text{Inverse cumulative probability} = 1 - \frac{(1-CL)}{2} \quad \text{Equation 7}$$

$$\text{For } CL = 95\%, \quad Z = 1.96$$

Table 3 - Z table values for various confidence levels (Hazra, 2017)

Confidence level	Critical (z) value to be used in confidence interval calculation
50%	0.67449
75%	1.15035
90%	1.64485
95%	1.95996
97%	2.17009
99%	2.57583
99.9%	3.29053

3.3. Normal distribution

To calculate the normal distribution used to plot a bell curve of the data, the use of the MATLAB function 'normpdf' is required. The formula used by this function is shown in equation 8.

$$f(x | \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{Equation 8}$$

3.4. Near field distance

The initial fluctuations in wave propagation are a product of destructive interference near the crystal, this area is known as the near field distance and can be calculated as a product of the wavelength and crystal diameter, shown through equation 9 are the calculations for both components.

$$\lambda = \frac{\text{Velocity}}{\text{Frequency}} \quad \text{Equation 9}$$

For Mild steel block:

$$\lambda = \frac{5564 \times 1000}{5 \times 10^6}$$

$$\lambda \approx 1.11\text{mm}$$

For Brake disc:

$$\lambda = \frac{3714 \times 1000}{5 \times 10^6}$$

$$\lambda \approx 0.74\text{mm}$$

Now with known wavelengths, it is possible to calculate near field distances for both components through equation 10.

$$NF = \frac{D^2}{4\lambda} \quad \text{Equation 10}$$

For Mild steel block:

$$NF = \frac{8.9^2}{4 \times 1.11}$$

$$NF \approx 17.80mm$$

For Brake disc:

$$NF = \frac{8.9^2}{4 \times 0.74}$$

$$NF \approx 26.66mm$$

Within each of these distances, the results obtained would be unreliable due to constructive and destructive interference meaning flaw detection within this region would be unsuitable. In this case the near field distance for the mild steel block does not interfere with the location of the testing, however for the brake disc, the testing region lies within the calculated near field, leading to inaccuracies within results. Similarly, when testing deep into the far field, there can be significant losses in resolution (A. Kumar, 2022).

4. Results and discussion

4.1. Mild steel block

Attenuation coefficients for the mild steel block are shown in table 4, confidence in the engineer can be determined through figure 11, where the accuracy is assessed through statistical analysis techniques. By investigating the confidence limits we can assess the reliability of the estimated results, the resolution provided by the ultrasonic scanner gives confidence to $\pm 0.1\%$ and $\pm 0.01\text{mm}$, meaning any result to a resolution higher than $\pm 0.1\text{dB/mm}$ is an estimate.

Table 4 - Attenuation coefficients for mild steel block sample

Mild steel rectangular block				
Measurement		Attenuation coefficient		
		1 to 2	2 to 3	1 to 3
Direction 1 (width)	1	-0.158	-0.158	-0.158
	2	-0.164	-0.153	-0.158
	3	-0.161	-0.125	-0.143
	4	-0.167	-0.132	-0.149
	5	-0.167	-0.131	-0.149
	6	-0.144	-0.053	-0.099
	7	-0.156	-0.095	-0.126
	8	-0.158	-0.060	-0.109
	9	-0.153	-0.098	-0.126
	10	-0.149	-0.115	-0.132
Mean		-0.158	-0.109	-0.133
Direction 2 (length)	11	-0.056	-0.006	-0.031
	12	-0.031	-0.025	-0.028
	13	-0.054	-0.020	-0.037
	14	-0.074	-0.019	-0.046
	15	-0.028	-0.023	-0.026
Mean		-0.047	-0.019	-0.033

Figure 11 illustrates the absolute normal distribution and frequency of all each echo range, the best distribution of results and accuracy is for the first two echo's, in comparison to the others the variability amongst data is reduced significantly. This point is further reinforced through figure 12, which depicts a regression in both accuracy and attenuation coefficient as the wave progresses, suggesting signal decay with distance (Tu, 2006).

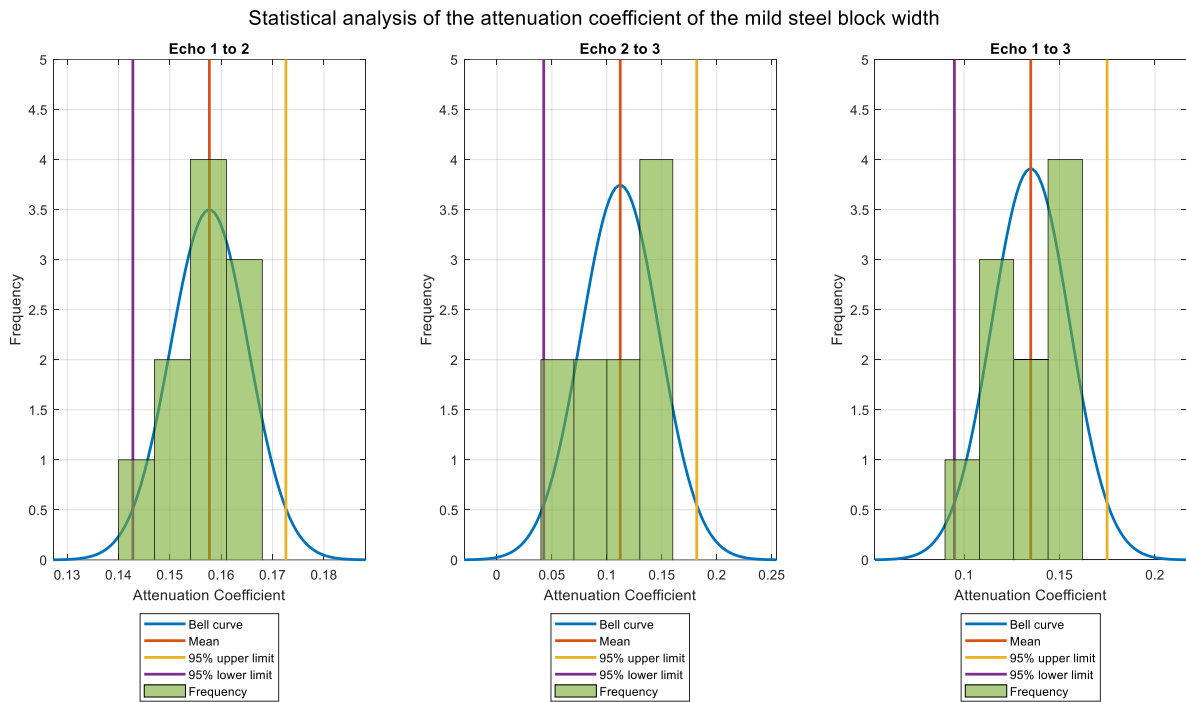


Figure 11 - Statistical analysis of multiple echoes for the mild steel sample width

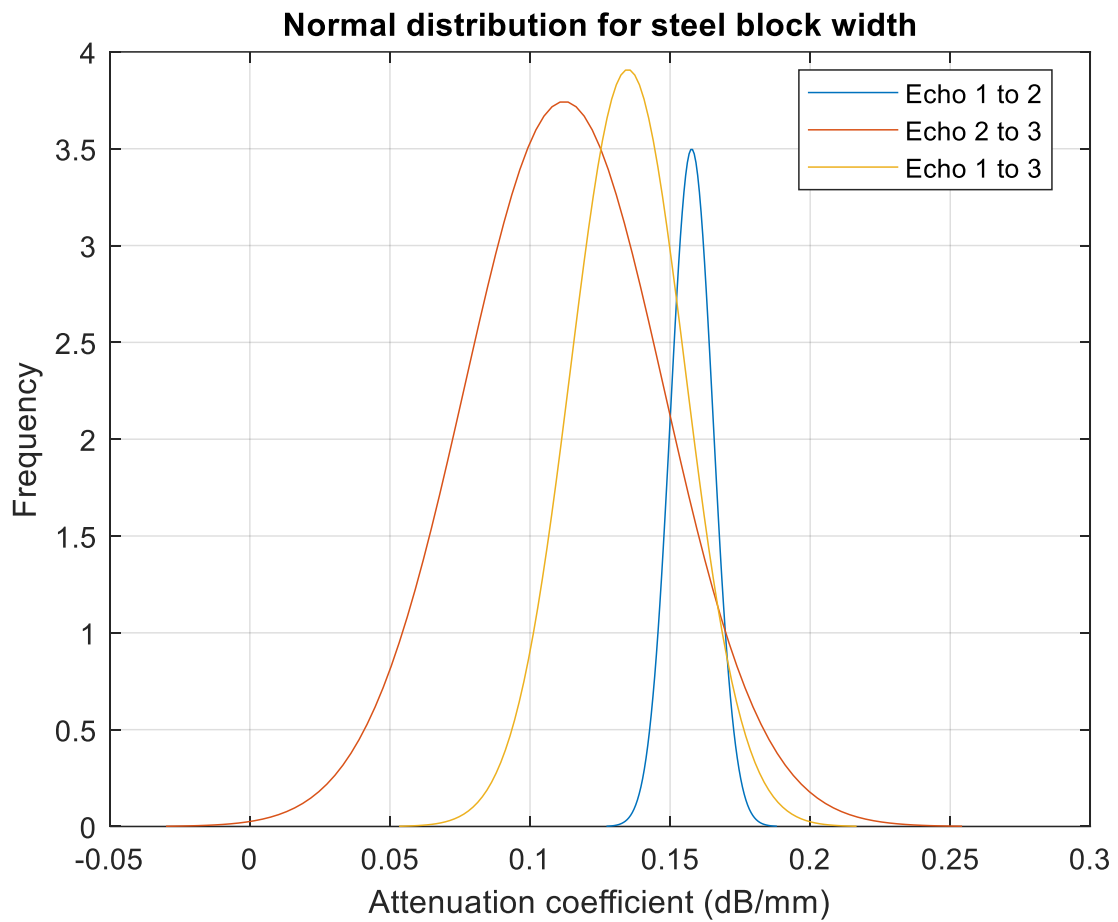


Figure 12 - Comparison of normal distributions for the mild steel block sample

The increased variability in data for echoes 2 to 3 and 1 to 3 highlights growing inconsistencies in wave amplitude measurements. This would tend to indicate greater interference, scattering, or noise as the wave travels through the material (Bai, 2019). Figure 12 shows this trend-the data for these echoes has a greater spread, which is attributed to material inhomogeneity caused by variations in grain sizing (Mekhlouf, 2009). Figure 13 further shows that there is noise in the data for echo 3, which contributes to poor consistency. The noise stems from material boundaries that interfere with the primary signal because the secondary reflections may combine with the third echo to reduce wave amplitude clarity. The trend in the attenuation coefficients indicates increased energy absorption by the medium with distance, translating to smaller returned signals, a fact that is also observed in wave propagation (Ono, 2020).

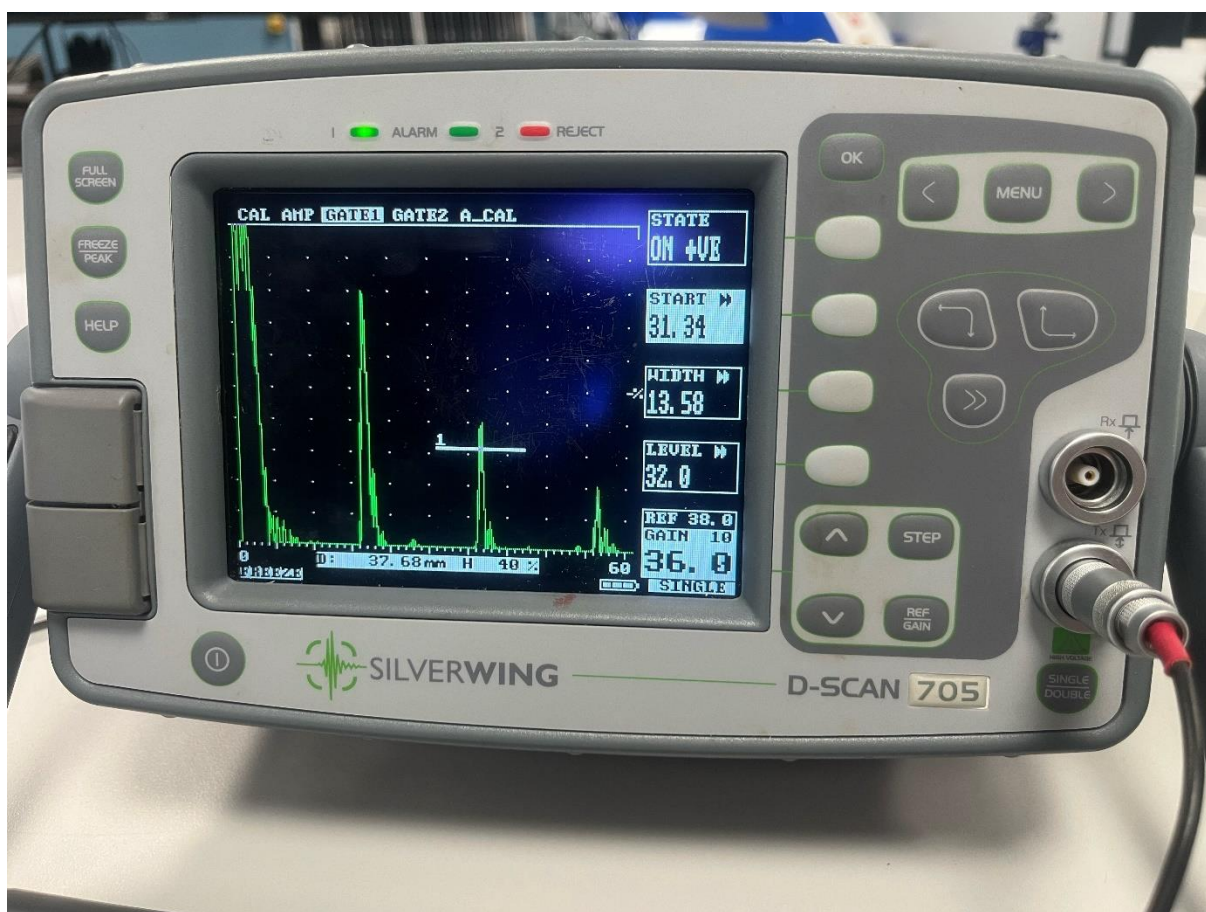


Figure 13 - Illustrating 'noise' in data for third echo

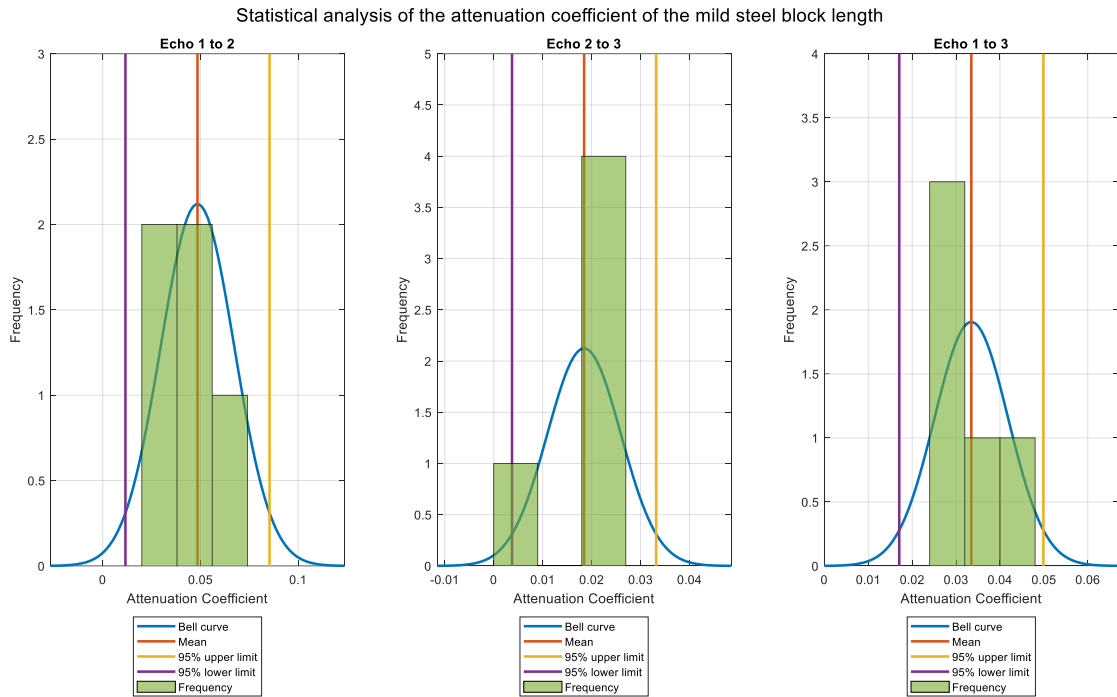


Figure 14 - Statistical analysis of multiple echoes for the mild steel sample length

Due to the limited surface area on the top of the steel block, only five measurements were taken, reducing resultant resolution and limiting analysis, yet conclusions can still be drawn from Figure 14. Compared to previous data, the recorded attenuation coefficients are noticeably lower, potentially due to the grain structure. The grain direction may hinder wave propagation and reduce returned wave amplitudes, which is relevant in this case (NDT-Inspect, 2021). Another contributor is the reflections from sidewalls as mentioned previously, these reflections absorb wave energy thus reducing the received wave amplitude and hence reducing attenuation coefficient.

4.2. Brake disc

Table 5 shows attenuation coefficients for the brake disc, there is more variation in data compared to the mild steel block. This could be due to the poor surface finish of the unprepared component that gives rise to unexpected variation. Furthermore, measurements within the near field, as identified from previous calculations, give poor representation of attenuation coefficients. In the near field, the beam is not fully developed and has fluctuating intensities. A larger transducer might reduce the near field distance, hence providing more stable results (Iowa State University, n.d.).

Table 5 - Attenuation coefficients for the brake disc sample

Brake disc hub			
Measurement	Attenuation coefficient		
	1 to 2	2 to 3	1 to 3
1	-0.478	-0.560	-0.519
2	-0.590	-0.601	-0.595
3	-0.481	-0.486	-0.483
4	-0.581	-0.661	-0.621
5	-0.657	-0.731	-0.694
6	-0.494	-0.513	-0.504
7	-0.540	-0.589	-0.564
8	-0.543	-0.622	-0.583
9	-0.491	-0.508	-0.499
10	-0.447	-0.503	-0.475
11	-0.430	-0.531	-0.480
12	-0.324	-0.453	-0.389
13	-0.389	-0.414	-0.402
14	-0.430	-0.445	-0.438
15	-0.634	-0.682	-0.658
Mean	-0.491	-0.534	-0.512

Figure 15 shows the normal distributions of attenuation coefficients, where the range for the brake disc is wider compared to the mild steel sample. Brake discs are manufactured from heterogeneous materials, such as cast iron with graphite flakes or composites, which have a non-uniform structure and hence more scattering and absorption. (Shin, 2013). This effect is more pronounced with increased wave travel, which explains the higher variation observed. Moreover, the attenuations are higher in the second and third wave propagations since ultrasonic waves with high frequencies are absorbed and scattered greatly, leading to the greater attenuation coefficients in later echoes.

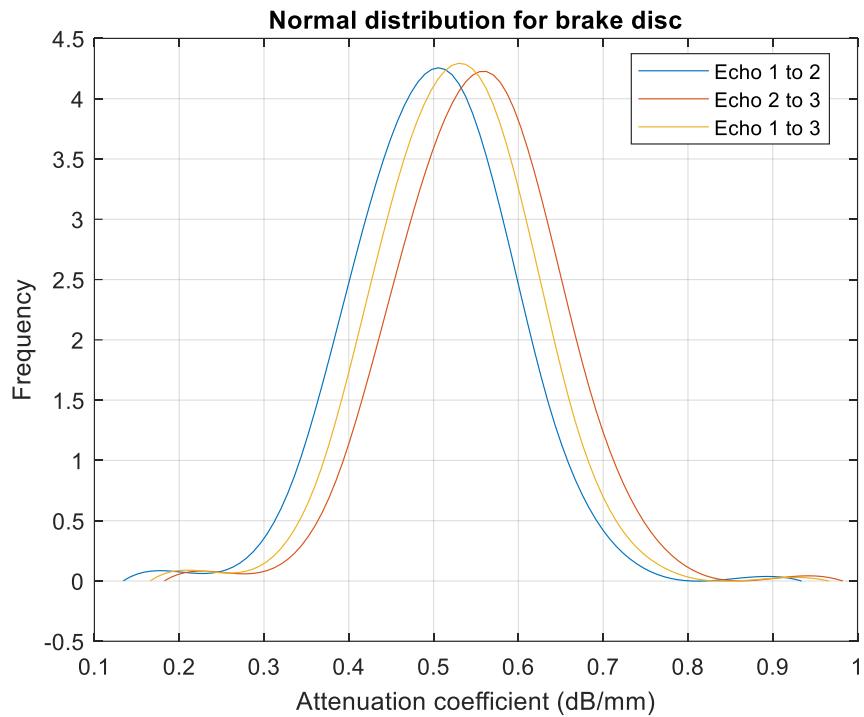


Figure 15 - Comparison of normal distributions for multiple echoes on brake disc

The histogram data, see figure 16, shows little deviance above or below the 95% confidence limits for all echo ranges, this suggests that the testing was performed with a consistent accuracy, meaning the variance amongst results is caused by external factors as discussed previously. The amalgamation of these factors results in the large deviation, if each were singular, the deviance in results would be less significant.

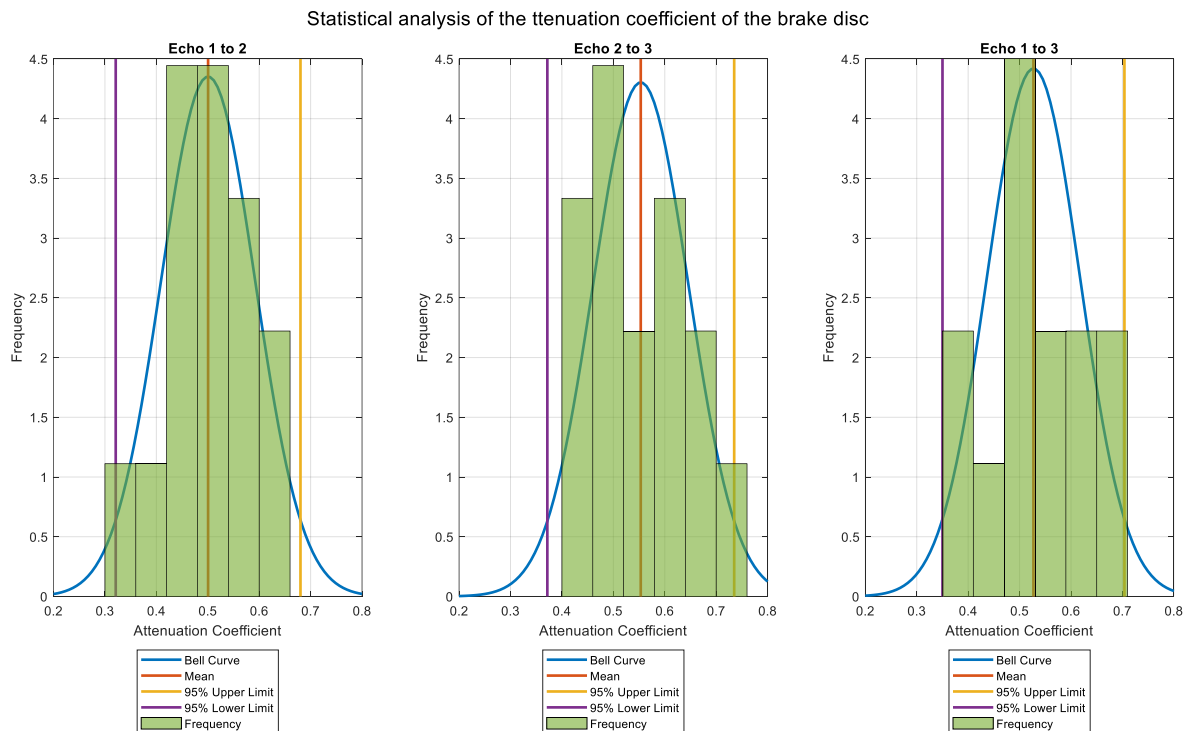


Figure 16 - Statistical analysis of multiple echoes for the brake disc sample

5. Conclusion

This experiment determined the attenuation coefficients for the mild steel block and brake disc using ultrasonic testing. The mild steel block shows consistent values because of the homogeneous structure and uniform alignment of grains. On the other hand, the brake disc showed huge variation due to its heterogeneous composition, surface irregularities, and scattering effects. Other factors contributing to variability involve unprepared surfaces and interference in the near field. Material inhomogeneity, interference, and boundary effects bring about higher attenuation for brake disc during the second and third wave propagations.

Histogram data was consistent within 95% confidence limits; however, these deviations were generally due to surface irregularities, transducer alignment, and wave scattering. The ultrasonic test is reliable for the calculation of attenuation coefficients but less effective in the case of materials with irregular geometries or unprepared surfaces. Future testing can be enhanced by better surface preparation, transducer alignment, and considering near-field effects. Despite its limitations, the results confirm ultrasonic testing to be one of the major non-destructive techniques that assess material properties and understand technique constraints.

References

- A. Kumar, W. A., 2022. High resolution in non-destructive testing: A review. *Journal of applied physics*, Volume 132.
- Bai, L. V. A. D. B. W., 2019. *Detection and charecterisation of defects in highly scattering materials using ultrasonic arrays*, Bristol: s.n.
- Hazra, A., 2017. Using the confidence interval confidently. *Journal of thoracic disease*, 9(10), pp. 4124-4129.
- Iowa State University, n.d. *Nondestructive evaluation techniques*. [Online]
Available at: <https://www.nde-ed.org/NDETechniques/Ultrasonics/EquipmentTrans/beamspread.xhtml>
[Accessed 08 12 2024].
- Mekhlouf, S. Z. M. B. Y., 2009. Residual Stress Analysis in the Titanium by Eddy Current Method. *e-Journal of Nondestructive Testing* , 15(4).
- Mix., P. E., 2005. *Intorduction to nondestructive testing: a training guide*. 2nd ed. s.l.:Wiley-Interscience.
- NDT-Inspect, 2021. *Ultrasonic Testing (UT) Handbook*. [Online]
Available at: <https://ndtinspect.com/ultrasonic-testing-ut-handbook/>
[Accessed 08 12 2024].
- Ono, K., 2020. A Comprehensive Report on Ultrasonic Attenuation of Engineering Materials, Including Metals, Ceramics, Polymers, Fiber-Reinforced Composites, Wood, and Rocks. *applied sciences*, 10(7).
- Polonivhko, V. R. J., 2007. *Effects of transducer geometry and beam spreading on acoustic Doppler velocity measurements near boundaries*. Vancouver, BC, Canada, OCEANS 2007.
- Shin, M. W. J. G. H. K. J. K. K. H. Y. J. H., 2013. The Effect of Residual Stress on the Distortion of Gray Iron Brake Disks. *Journal of materials engineering and performance*, Volume 22, pp. 1129-1135.
- Tu, H. Z. J. C. Q., 2006. Attenuation estimations using envelope echo data: Analysis and simulations. *Ultrasound in medicine and biology*, 32(3), pp. 377-386.
- Wang, Z. C. X. M. H. K. Y. D. Z., 2018. Effect of surface roughness on ultrasonic testing of back-surface micro-cracks. *applied sciences*, 8(8).

Xu, R. B.-G. R. K. M. L. A. M. S. a. O. J., 2024. Frequency dependant amplitude response if different couplant materials for mounting piezoelectric sensors. *NDT & E International* , Volume 141.

University of Wales Trinity Saint David

Wales institute of Science and Arts



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

Probability of sizing analysis

Structural integrity & material evaluation ACEA7013

Joseph Britton

2104278

December/2024

Table of Contents

Table of figures	iii
Table of tables.....	iii
Nomenclature	iv
Abstract.....	iv
1. Introduction	1
1.1. Method	1
2. Measurements and graphs of raw measurements	2
3. Calculations and error estimates	3
3.1. Mean.....	3
3.2. Standard deviation	3
3.3. CV.....	3
3.4. RMSE.....	3
3.5. R^2	3
3.6. Slope and Y intercept.....	3
3.7. Confidence levels.....	4
4. Results and discussion.....	5
5. Conclusion	11
References	12

Table of figures

Figure 1 - Raw data for operator sizing shown graphically.....	2
Figure 2 - Coefficient of variation for all operators	5
Figure 3 - Root mean square error for remaining operators	6
Figure 4 - Raw data for operator 2 with linear plot	6
Figure 5 - A) Comparison of normalised and unnormalized scores B) Weighted scores for operators .	7
Figure 6 - A) Operator 1 95% confidence limits B) Operator 4 95% confidence limits C) Operator 7 95% confidence limits.....	8
Figure 7 - A) Corrected data for Operator 1 B) Corrected data for Operator 7	9
Figure 8 - Corrected data weighted scoring for all operators.....	10

Table of tables

Table 1 - Raw data of operator sizing	2
Table 2 - R^2 , Slope and Y intercept for remaining operators	7
Table 3 - Corrected data with final scoring for all operators.....	9

Nomenclature

Symbol/Term	Definition	Units
$\bar{x}$	Sample mean	mm
S	Sample standard deviation	mm
z	Z-score	Unitless
CV	Coefficient of variation	%
RMSE	Root mean square error	Unitless
R^2	Coefficient of determination	Unitless
Slope (m)	Gradient of linear regression	Unitless
Y-intercept (C)	Y intercept of linear regression	Unitless
x	Independent variable	mm
y	Dependant variable	mm
Mean (μ)	Average value of a dataset	mm
Standard Deviation (σ)	Measure of variability	mm
Confidence Interval (CI)	Statistical range indicating the reliability of results	mm
POD	Probability of detections	Unitless
N	Number of data points	Unitless
SS_{res}	Residual sum of squares	Unitless
SS_{tot}	Total sum of squares	Unitless

Abstract

The sizing accuracy of ten technicians subjected to the novel ultrasonic thickness gauging technique is evaluated in this report. All their performances were measured against the known defect sizes using different statistical analyses: linear regression, coefficient of variation, root mean square error, and R^2 . A normalizing and weighting of such metrics were used to provide a scoring system for identifying the most reliable operator. The results showed a significant variation across operators, thus pointing out the need for a comprehensive approach towards statistical evaluation. Operator 7 had the maximum accuracy and consistency, validated through normalized scoring and analysis of the confidence interval. This study has brought into light the effectiveness of multiple statistical measures when combined for robust performance evaluation in defect sizing procedures.

1. Introduction

This report will discuss the sizing accuracy of 10 individual technicians utilizing a new ultrasonic thickness gauging procedure. The task involves analysing measurements of through the wall thickness to determine the accuracy of the operators through 21 defect measurements. Linear regression and statistical analysis techniques will be applied to the data set to assess all individuals. Known true size values for each defect will be used to make conclusions on the exactness of each operator, in the attempt to determine the effect of user capabilities using this procedure.

1.1. Method

Operator performance is evaluated by linear regression and statistical analysis. Linear regression models the relationship between the true defect size and reported size, whilst mean and standard deviation are used to calculate the coefficient of variation (CV), a measure of variability as a percentage of the mean. The lower the value of CV, the greater the consistency (Hayes, 2024).

The R^2 value, or coefficient of determination, represents consistency and reliability. A value closer to 1 indicates better alignment with the true values. However, it does not include bias or systematic error (Nagelkerke, 1991).

The root mean square error (RMSE) calculates predictive accuracy as an average of squared differences between predicted and observed values; lower RMSE means higher performance but sensitive to large errors (Hodson, 2022).

Confidence levels, describe the reliability of the regression models, essentially indicating the probability that the results will fall within the given tolerance range. (Statistics how to, 2024).

A composite score is then used, weighting metrics by importance to accuracy and consistency, identifies the most successful operator, a method validated within a number of studies (Ayan, 2023) (Thokala, 2012) (Kisly, 2016).

2. Measurements and graphs of raw measurements

The data for all operators is shown through table 1, it is taken as part of a larger POD study assessing a new ultrasonic thickness gauging technique. It is clear to see large dispersion of results between operators against the known true size as further illustrated through figure 1. There is a clear case of oversizing at smaller defect sizes, along with operators 2 and 6 missing defect A completely. This trend shifts when defect size increases, for defect U, only operator 1 oversize's his measurements, whilst all other undersize the defect.

Table 1 - Raw data of operator sizing

Defect	true size	Op1	Op2	Op3	Op4	Op5	Op6	Op7	Op8	Op9	Op10
A	0.25	1.00	0.00	0.50	0.50	0.50	0.00	1.00	0.30	0.50	0.80
B	0.27	1.00	0.00	1.00	0.73	1.20	0.50	1.00	1.20	0.50	0.80
C	0.28	1.00	0.00	0.50	0.76	1.00	1.20	1.00	0.70	0.50	1.30
D	0.40	1.20	1.00	1.00	0.91	0.50	1.00	1.00	0.80	0.60	1.50
E	0.40	1.40	1.00	1.10	0.98	1.20	1.00	1.00	0.90	0.20	1.60
F	0.50	1.10	1.00	0.80	0.79	0.90	1.00	1.20	0.50	0.00	1.00
G	0.52	0.90	1.50	1.00	1.01	1.50	1.10	1.00	0.70	1.00	0.90
H	0.55	0.50	0.90	1.20	0.96	1.20	0.80	1.00	0.40	1.30	1.60
I	0.70	2.00	1.60	1.20	1.35	1.50	0.90	1.00	1.00	1.60	2.20
J	0.98	1.90	2.00	1.60	1.31	1.00	0.80	1.00	1.40	0.50	1.80
K	1.12	1.00	1.80	1.80	1.37	0.50	1.00	1.50	1.30	1.80	2.00
L	1.15	2.20	1.00	1.50	1.56	1.80	1.30	1.50	1.20	1.80	2.00
M	1.10	2.00	1.60	1.50	1.60	2.00	1.40	2.00	1.60	0.50	2.20
N	1.30	2.20	2.00	2.00	1.73	2.10	0.90	1.50	1.70	1.70	2.00
O	1.60	2.00	2.00	1.50	1.56	0.80	0.90	2.00	1.50	1.20	2.00
P	1.80	2.40	2.00	3.20	2.14	2.00	1.40	2.50	1.10	2.20	2.50
Q	2.10	3.50	2.50	3.00	2.49	1.80	2.30	2.00	2.60	2.00	3.20
R	2.70	2.80	2.50	3.50	2.97	3.00	3.20	3.50	2.20	2.50	3.50
S	4.90	5.70	4.50	4.40	4.87	4.50	4.50	5.50	5.50	4.20	5.40
T	6.00	7.00	6.50	5.50	6.17	5.50	6.20	6.50	6.00	5.30	6.70
U	7.50	8.00	7.20	6.50	6.99	6.00	6.25	8.00	6.50	6.50	6.90

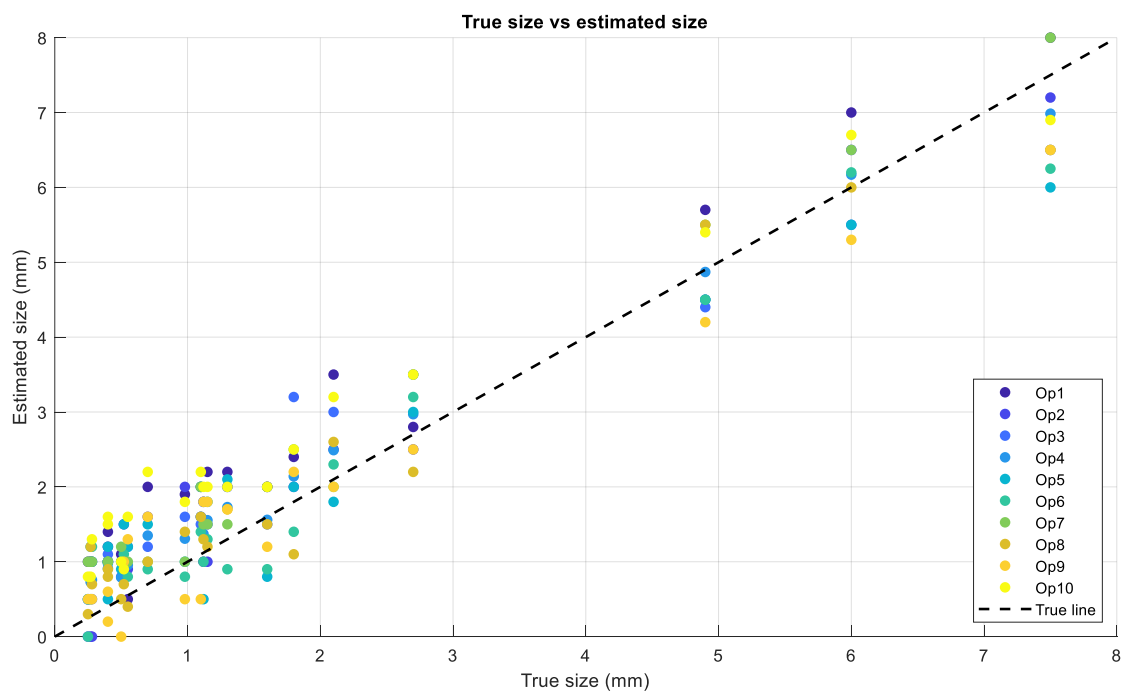


Figure 1 - Raw data for operator sizing shown graphically

3. Calculations and error estimates

3.1. Mean

The mean average, shown in equation 1, is not used for the scoring of operators directly, however it is required for the calculation of CV. This describes the sum of all data, divided by the number of data points.

$$Mean = \frac{\sum_{i=1}^n x_i}{n} \quad \text{Equation 1}$$

3.2. Standard deviation

The same is true for the calculation of standard deviation as is for mean, it is not directly used in the scoring of an operator, however it is used for the calculation of CV. The equation can be seen in equation 2 (Newcastle university, n.d.).

$$\sigma = \sqrt{\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}{N}} \quad \text{Equation 2}$$

3.3. CV

The CV describes the variability of an operator relative to the mean value and expressed as a percentage. For this task it is weighted at 20% due to its relative importance, if an operator is not consistent in their sizing, then it can be assumed that they are performing poorly. This parameter is calculated through equation 3 (Geeks for geeks, 2024).

$$CV (\%) = \left(\frac{\sigma}{Mean} \right) \times 100 \quad \text{Equation 3}$$

3.4. RMSE

RMSE, see equation 4 (Hodson, 2022), is weighted at 20% because it provides a direct measure of the operator's overall error, meaning it is critical to accuracy, although secondary to slope.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad \text{Equation 4}$$

3.5. R²

The equation to calculate R² is shown in equation 5 (Soch, 2019), as this variable is related to true values it shows its importance however as this value doesn't capture biases or absolute errors in measurements, a weight of 10% is applied for the scoring system.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad \text{Equation 5}$$

3.6. Slope and Y intercept

Shown through equation 6 is the equation for a straight line (Geeks for geeks, 2024). A slope approximately equal to 1 ensures a high accuracy of measurements, slopes above 1 imply oversizing defects while slopes below 1 depict under sizing. The intercept corresponds to a baseline offset error,

less critical than slope when judging operators' skills. In this analysis, the slope is weighted 40% because the slope is a critical component in describing proportional accuracy, which is necessary for valid defect measurement. The intercept is weighted 10% because it is less influential, since it reflects systematic bias.

$$y = mx + c \quad \text{Equation 6}$$

3.7. Confidence levels

Confidence levels are not directly used to calculate the final score of each operator. However, they serve as a visual aid on graphical representations of the data set. The equation to calculate this metric is shown in equation 7 (Penn state, 2024).

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}} \quad \text{Equation 7}$$

4. Results and discussion

To begin analysing which operator performs best, it is recommended to begin by discounting several initially. This first reduction will make it easier to analyse the remainder due to the smaller sample size. From figure 2, operators 6, 8 and 9 show high CV percentages meaning they can be discounted from future consideration. The high CV percentage suggests that these operators are sporadic in their sizing and show little trend in their sizing.

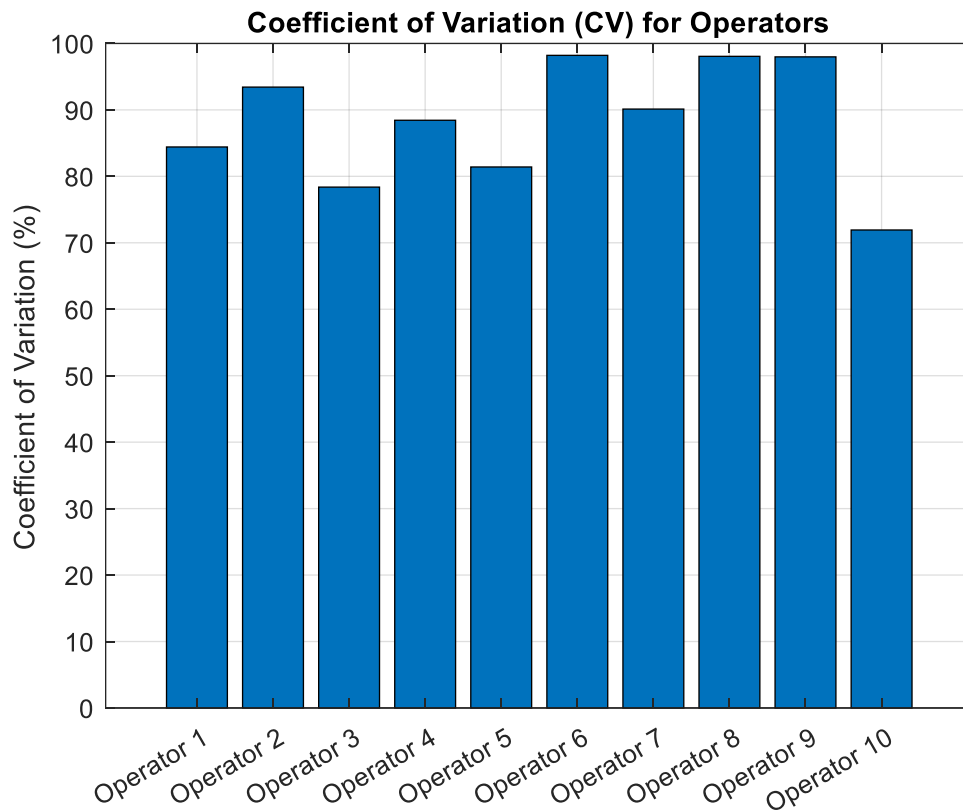


Figure 2 - Coefficient of variation for all operators

Further analysis, using RMSE values in Figure 3, underlines operators 2 and 5 with relatively big errors. Figure 4 shows the operator 2 fully missing the first three defects, which is indicative of poor accuracy and hence justifies the exclusion. Operator 5 has the highest RMSE among the remaining operators and exhibits great measurement errors, especially for the large defects, confirmed by Table 1. This inconsistency supports also the exclusion of operator 5.

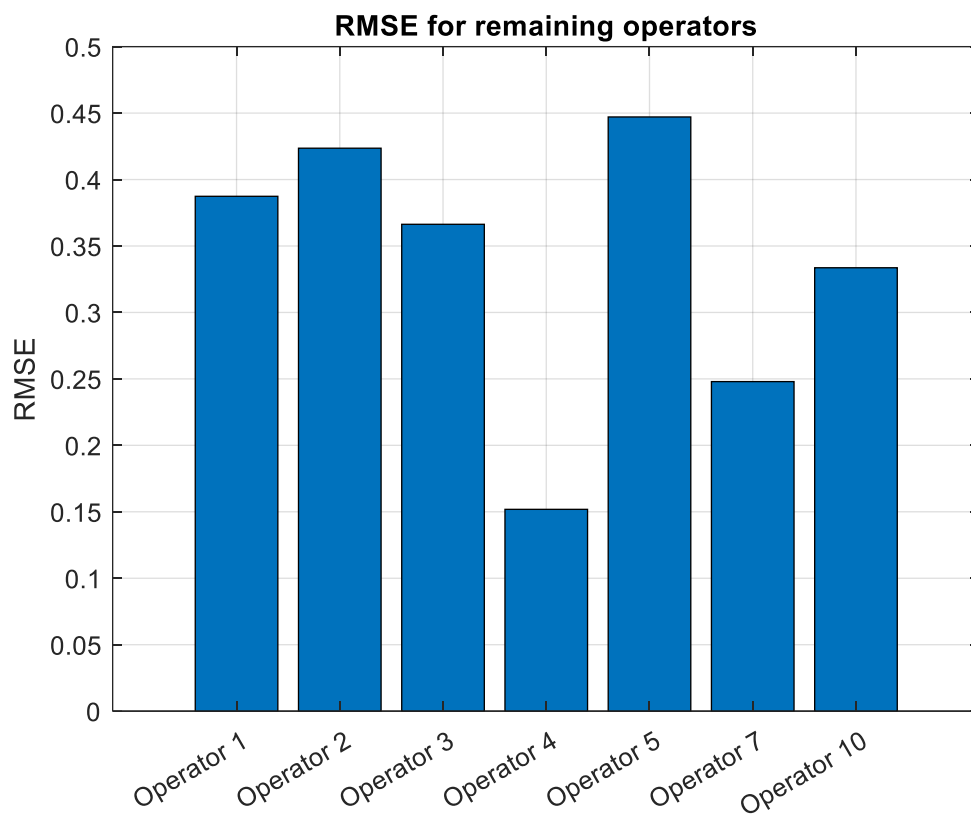


Figure 3 - Root mean square error for remaining operators

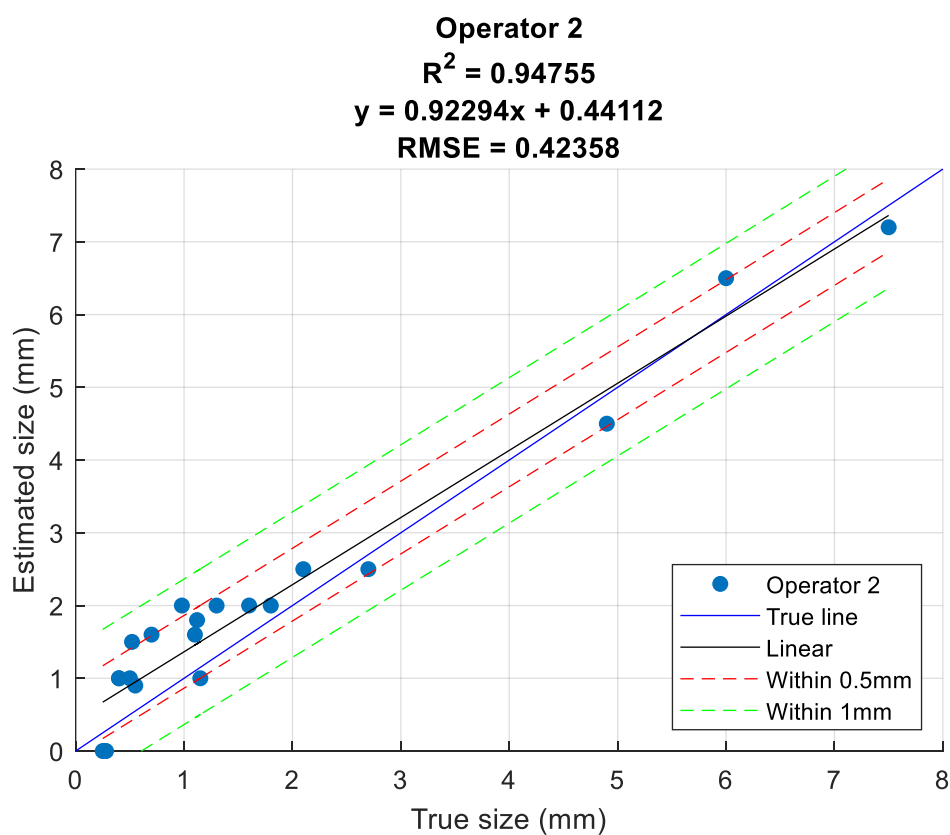


Figure 4 - Raw data for operator 2 with linear plot

The remaining data is shown through table 2, as discussed previously the largest weighted parameter for the eventual scoring system is the slope. Knowing this, the two outlying operators with the best slope values are operators 1 and 7. This is reinforced through their low values for y intercept, however, using a more wholistic view of the remaining parameters, another operator performs well, operator 4. Leaving operators 3 and 10 to be discounted.

Table 2 - R^2 , Slope and Y intercept for remaining operators

Operator	Rsquared	Slope	Intercept
{ 'Operator 1' }	0.96222	1.0021	0.69536
{ 'Operator 3' }	0.94846	0.80555	0.72397
{ 'Operator 4' }	0.99252	0.89682	0.49209
{ 'Operator 7' }	0.98393	0.99454	0.51321
{ 'Operator 10' }	0.96302	0.8727	0.97039

Now 3 operators remain, the decisions made previously can be validated through the development of the scoring system. If the current parameters were weighted and then combined into a score the scoring system would not provide any valuable data. This is due to the scaling of different parameters, meaning that they need to be normalised. Once more the parameters CV, Y intercept and RMSE must be flipped as lower values provide a better score. Normalisation of a data set can be done through equations 8 and 9.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad \text{Equation 8}$$

For CV and RMSE:

$$x' = 1 - \frac{x - \min(x)}{\max(x) - \min(x)} \quad \text{Equation 9}$$

Validation as to why it is necessary to normalise values for scoring is shown in figure 5, this displays data that is meaningless and does not provide any valuable insight into the true best operator.

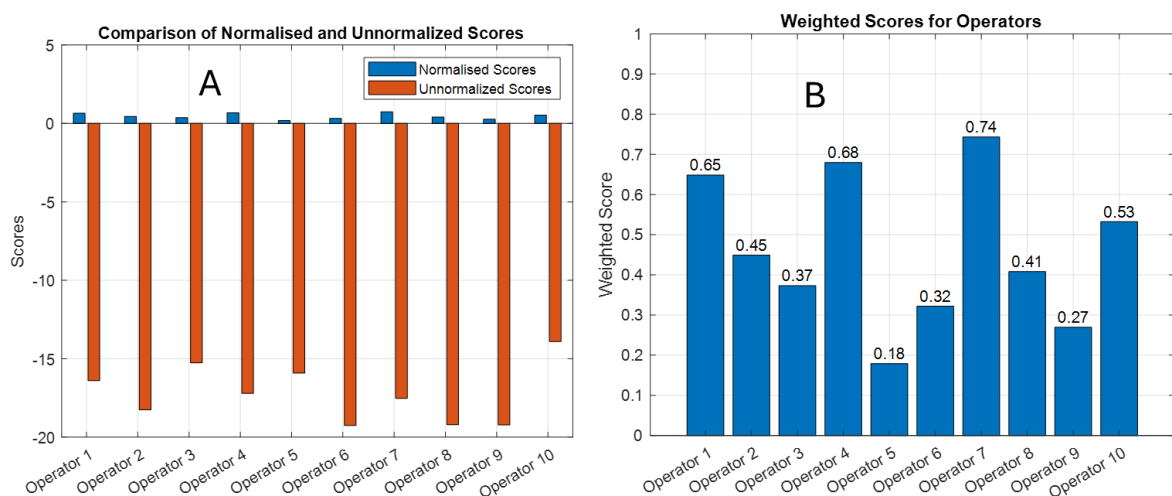


Figure 5 - A) Comparison of normalised and unnormalized scores B) Weighted scores for operators

After corrections, proper analysis and validation confirm earlier decisions. Those operators who scored below 0.5 in Figure 5 were excluded, which again validates their removal. Further evaluation of the remaining operators can be done by analysing confidence limits. Figure 6 shows that operator 4 has the narrowest confidence limits, indicating high consistency (Bobbitt, 2021), which contradicts the fact revealed by weighted scoring because higher weighting was given to the slope and operator 4 scored poorly regarding that, as shown in Table 2. The final exclusion is operator 1, who ranked third in the weighted scoring but has the widest confidence limits relative to the linear plot of the dataset. This inconsistency supports their removal from further consideration.

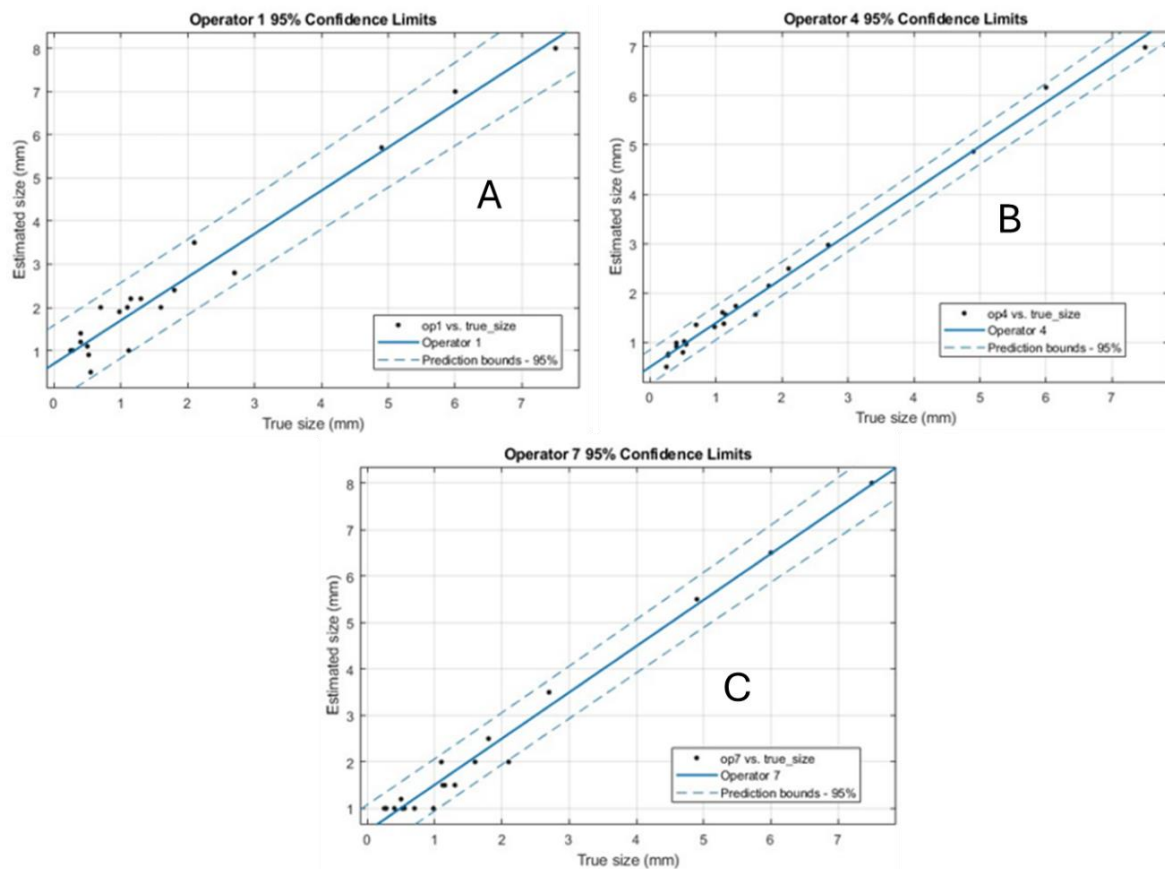


Figure 6 - A) Operator 1 95% confidence limits B) Operator 4 95% confidence limits C) Operator 7 95% confidence limits

The analysis here brings to the fore that there can be disagreements between statistical criteria, necessitating a look at an integral approach across various parameters. To finalize the data, the offset from the y-intercept is taken away by subtracting this from the data value across the set. Figure 7 represents the corrected data to substantiate that, in fact, operator 7 had a better slope than operator 1, whereas in uncorrected data, the reverse was seen. Both corrected and uncorrected analyses show that operator 7 is the best performer. Most of the statistical metrics confirm this result, as operator 7 is most accurate and reliable.

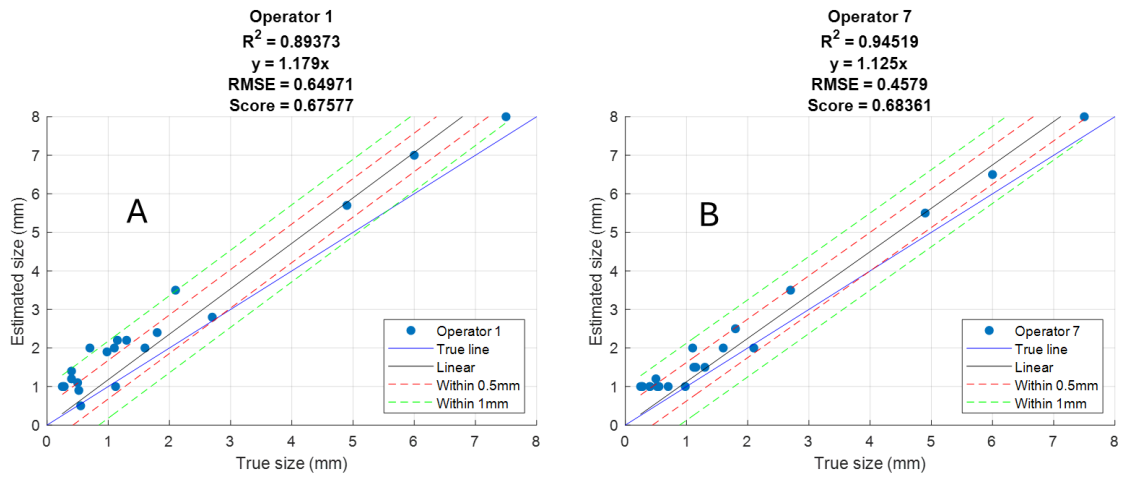


Figure 7 - A) Corrected data for Operator 1 B) Corrected data for Operator 7

To further validate this methodology of statistical analysis and the final decision made the whole of the data set was corrected as a final check for any flaws made during the analysis. The results proved that the decision of operator 7 being the best performing operator is correct. For the final data see table 3 and scoring see figure 8.

Table 3 - Corrected data with final scoring for all operators

Operator	R ²	Slope	RMSE	Score
1	0.89373	1.179	0.64971	0.67577
2	0.91555	1.0351	0.53749	0.6514
3	0.8352	0.98965	0.65505	0.63757
4	0.94834	1.022	0.39913	0.67087
5	0.8178	0.91309	0.65404	0.61758
6	0.92216	0.93687	0.47941	0.63344
7	0.94519	1.125	0.4579	0.68361
8	0.93265	0.97482	0.46239	0.64393
9	0.91738	0.90374	0.47639	0.62687
10	0.78699	1.1195	0.80069	0.65059

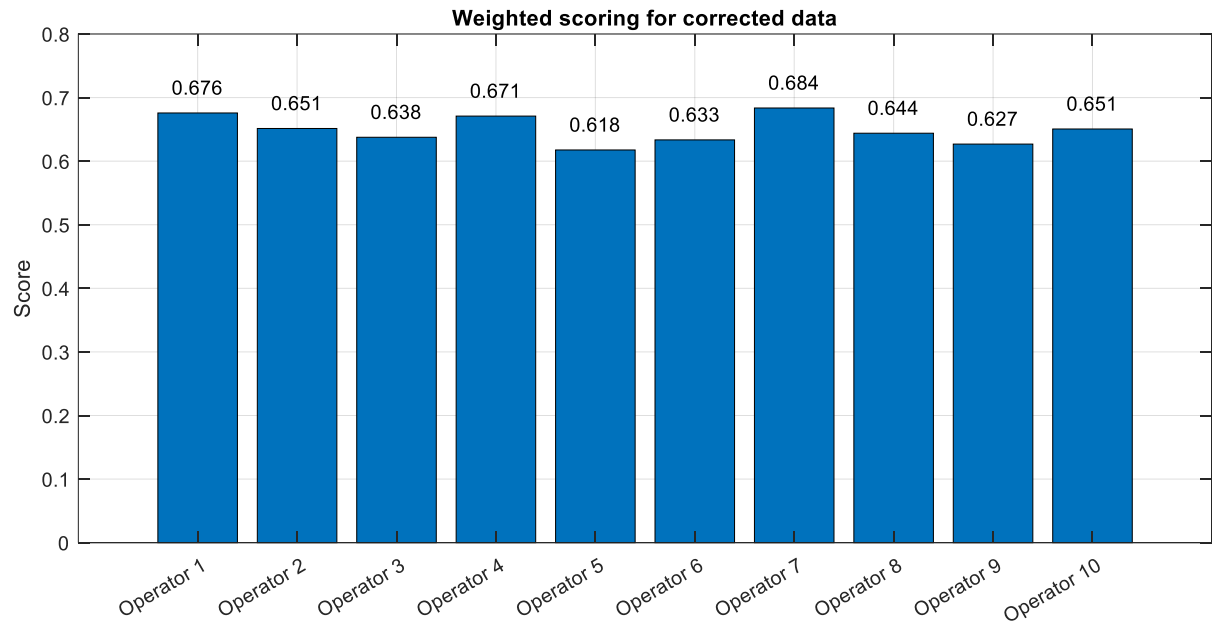


Figure 8 - Corrected data weighted scoring for all operators

5. Conclusion

In this report, various statistical analysis techniques have been used to find the sizing accuracy of the probability of detection study: namely CV, RMSE, Linear Regression, and confidence interval analyses to eventually build a confident scoring system to assess.

The results of this study have proven a scoring system as a valid analysis method if weighted correctly. This indicates a strong reliance on the slope of the plotted line relative to true defect size and estimated defect size from the operator, since this was the largest weighted parameter for this study. The final decision based on the method built throughout found operator 7 to be most reliable, demonstrating superior accuracy and consistency throughout.

It has also been established that there is a need for a holistic approach in undertaking statistical analysis, given the clear disagreements across techniques. This again emphasizes the importance of integrating multiple metrics towards the comprehending of operator performance. The processes of normalization and weighting applied in this study avoided any single parameter disproportionately driving the final scoring to provide a balanced evaluation.

References

Ayan, B. A. S. B. M. P., 2023. A Comprehensive Review of the Novel Weighting Methods for Multi-Criteria Decision-Making. *Information*, 14(5).

Bobbitt, Z., 2021. *What is Considered a Good Confidence Interval?*. [Online]
Available at: <https://www.statology.org/what-is-a-good-confidence-interval/>
[Accessed 14 12 2024].

Geeks for geeks, 2024. *Coefficient of Variation Formula*. [Online]
Available at: <https://www.geeksforgeeks.org/coefficient-of-variation-formula/>
[Accessed 14 12 2024].

Geeks for geeks, 2024. *Slope intercept form*. [Online]
Available at: <https://www.geeksforgeeks.org/slope-intercept-form/>
[Accessed 14 12 2024].

Hayes, A., 2024. *Coefficient of Variation: Meaning and How to Use It*. [Online]
Available at: <https://www.investopedia.com/terms/c/coefficientofvariation.asp>
[Accessed 10 12 2024].

Hodson, T. O., 2022. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific model development*, 15(14), pp. 5481-5487.

Kisly, D. T. A. C. M. S., 2016. *Implementation of multiple criteria decision analysis approaches in the supplier selection process: a case study*. Guimarães, Portugal, s.n.

Nagelkerke, N. J. D., 1991. A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), pp. 691-692.

Newcastle university, n.d. *Variance and standard deviation*. [Online]
Available at: <https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/descriptive-statistics/variance-and-standard-deviation.html>
[Accessed 14 12 2024].

Penn state, 2024. *4.5.2 - Derivation of the Confidence Interval*. [Online]
Available at: <https://online.stat.psu.edu/stat800/lesson/4/4.5/4.5.2>
[Accessed 14 12 2024].

Soch, J., 2019. *Proof: Derivation of R^2 and adjusted R^2* . [Online]
Available at: <https://statproofbook.github.io/P/rsq-der.html>
[Accessed 14 12 2024].

Statistics how to, 2024. *Confidence Level: What is it?*. [Online]
Available at: <https://www.statisticshowto.com/probability-and-statistics/statistics->

definitions/confidence-level/

[Accessed 10 12 2024].

Thokala, P. D. A., 2012. Multiple Criteria Decision Analysis for Health Technology Assessment. *Value in Health*, Volume 15, pp. 1172-1181.