



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

Probability of Detection and
Probability of Sizing Analysis

University of Wales Trinity Saint David

M.Sc. Motorcycle Engineering

Structural Integrity and Material Evaluation

Krishnadas Pachery Chandradas

Student Number:2105959

Coursework: ACEA7013 (04-01-2022)

Contents

1. Abstract.....	4
2. Introduction	4
2.1 Phased Array Ultrasonics Testing.....	4
2.2 NDT Services and Global Market	5
3 Probability of Detection	6
3.1 History and Evolution of POD	6
3.2 POD curves.....	8
4. Calculations.....	9
4.1 Mean Error	9
4.2 Standard Deviation and 95% POD.....	10
4.3 Normal Distribution Data.....	11
4.4 True Line, Linear Trendline and Correlation coefficient	11
5. Results and Discussion	13
5.1 Operator 1 Performance Analysis.....	13
5.2 Operator 2 Performance Analysis.....	14
5.3 Operator 3 Performance Analysis.....	15
5.4 Operator 4 Performance Analysis.....	16
5.5 Operator 5 Performance Analysis.....	17
5.6 Operator 6 Performance Analysis.....	19
5.7 Operator 7 Performance Analysis.....	20
5.8 Operator 8 Performance Analysis.....	21
5.9 Operator 9 Performance Analysis.....	23
5.10 Operator 10 Performance Analysis.....	24
6. Conclusion	26
7. References	27

EQUATION 1 : DETERMINATION OF ERROR.....	9
EQUATION 2: DETERMINATION OF MEAN ERROR	10
EQUATION 3: DETERMINATION OF STANDARD DEVIATION.....	10
EQUATION 4: DETERMINATION OF 95 % CONFIDENCE INTERVAL	10
EQUATION 5: DETERMINATION OF NORMAL DISTRIBUTION RANGE.....	11

TABLE 1 : TRUE SIZE OF THE FLAW	9
---------------------------------------	---

FIGURE 1: PHASED ARRAY ULTRASONIC TESTING (PAUT) ((ARISEGOBAL.COM, N.D.)	4
FIGURE 2: CONTRIBUTION OF INDUSTRIES AND NDT APPLICATIONS (GRAND VIEW RESEARCH, 2021)	5
FIGURE 3 : NUMBER OF PUBLICATIONS OF POD (JOCHEN H. KURZ, 2012).....	6
FIGURE 4 : SCHEMATIC OF POD CURVE AND CORRESPONDING 95% CONFIDENCE BOUND (ANON., 2009).....	8

GRAPH 1: NORMAL DISTRIBUTION OF AN OPERATOR	11
GRAPH 2: TRUE LINE AND TRENDLINE OF AN OPERATOR DATA.....	12
GRAPH 3 :NORMAL DISTRIBUTION OF OPERATOR 1	13
GRAPH 4: OPERATOR 1 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	13
GRAPH 5: NORMAL DISTRIBUTION OF OPERATOR 2	14
GRAPH 6: OPERATOR 2 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	14
GRAPH 7: NORMAL DISTRIBUTION OF OPERATOR 3	15
GRAPH 8: OPERATOR 3 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	16
GRAPH 9: NORMAL DISTRIBUTION OF OPERATOR 4.....	16
GRAPH 10: OPERATOR 4 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	17
GRAPH 11: NORMAL DISTRIBUTION OF OPERATOR 5	18
GRAPH 12: OPERATOR 5 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	18
GRAPH 13: NORMAL DISTRIBUTION OF OPERATOR 6.....	19
GRAPH 14 : OPERATOR 6 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	19
GRAPH 15 : NORMAL DISTRIBUTION OF OPERATOR 7	20
GRAPH 16: OPERATOR 7 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	21
GRAPH 17 : NORMAL DISTRIBUTION OF OPERATOR 8	22
GRAPH 18: OPERATOR 8 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	22
GRAPH 19 : NORMAL DISTRIBUTION OF OPERATOR 9	23
GRAPH 20 : OPERATOR 10 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	23
GRAPH 21 : NORMAL DISTRIBUTION OF OPERATOR 10	24
GRAPH 22: OPERATOR 10 PERFORMANCE DATA (TRUE SIZE VS DETECTED FLAW SIZE).	24

1. Abstract

Non-destructive testing (NDT) is the process of testing and evaluating materials without destroying them. The reliability of the Non-Destructive Testing & Evaluation (NDT&E) technique is one of the most fundamental aspects of the industrial inspection process.

In NDT&E, the Probability of Detection curve (POD) is an accepted measure of reliability, which is usually expressed as a function of flaw size. Probability of detection and sizing tolerances are now commonly required in several industries for non-destructive evaluation, distinctly sizing of wall in metallic components is still a complex and prime matter in question with respect to structural integrity. In diverse industries, NDT is an essential methodology which have wide perspective on understanding the reliability of newly manufactured components and those under operation.

This report will investigate the inspection performance of ten operators analysing the size of defects detected using the ultrasonic thickness gauging procedure. The outcome of the data measuring 21 pre-analysed defects aid to understand the efficiency of the operators and equipment using the probability of detection method (POD). The continuative objective of this report is to generate quantitative POD results and to integrate these results in a probabilistically and statistical analysis.

2. Introduction

2.1 Phased Array Ultrasonics Testing

With respect to high penetrating power and sensitivity, Ultrasonic Testing (UT) has been extensively used in steel structure NDT to date. In addition to its ability to detect both deep and internal flaws, it is sensitive enough to detect and identify extremely small defects.

Typically, pulse waves of ultrasounds ranging from 0.1-15- 50 MHz are transmitted and imperfections are revealed by reducing the amount of sound detected (Wolter, et al., 2019)



Figure 1: Phased Array Ultrasonic Testing (PAUT) ((arisegobal.com, n.d.)

2.2 NDT Services and Global Market

Automation is a hallmark of today's industrial production, which is why quality control inspection is also essential. In addition to material laboratories, the quality characteristics of raw materials, semi-finished products, and finished products should be evaluated concurrently or alongside production. A continuous and comprehensive record of critical material characteristics is essential in order to monitor or to control the manufacturing processes in order to optimize them in terms of quality, efficiency, and cost. These objectives can only be achieved by the application of appropriate non-destructive testing (NDT) methods. In the beginning, NDT techniques have been almost exclusively used for defect detection in order to prevent the catastrophic failure of safety-relevant structures after they have been in service (Ono, 2020). However, nowadays the role of NDT is extended and it includes all phases of the product's life-cycle, from raw material over production to operation, disposal, and recycling (Hanke, 2016)

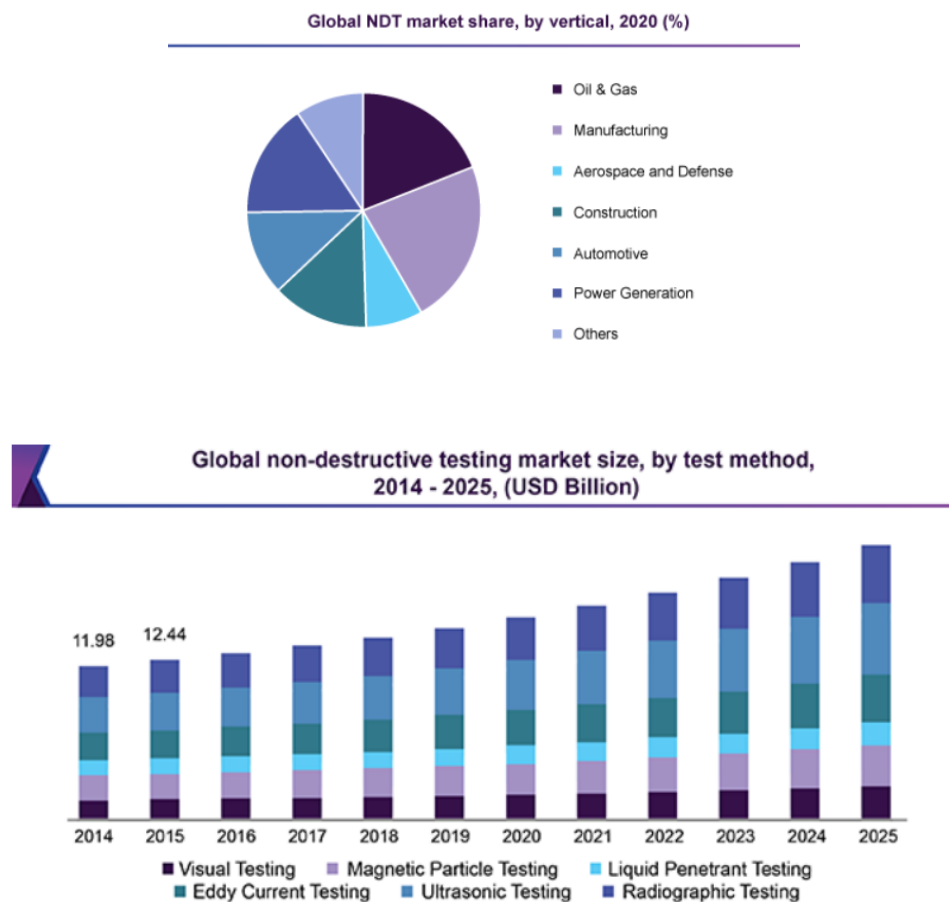


Figure 2: Contribution of Industries and NDT applications (Grand View Research, 2021)

3 Probability of Detection

3.1 History and Evolution of POD

In either of the fracture mechanics approaches, the ability to detect defects by NDT techniques is quantified through a parameter with statistical significance is termed as “Probability of Detection (POD)”. POD is defined as the fraction of targets of nominal size, expected to be found, given their existence (Annis C, 2009). POD data generation is usually carried out either using HIT (detected)/MISS (undetected) approach or a (signal response) vs. a (flaw size) approach as per (Annis C, 2009) standard which describes the procedure to be adopted along with the various parameters such as sample shape, size and defect classification to be considered.

In the field of NDT, the number of publications dealing with the topic POD has increased significantly during the last few decades (Figure 3)

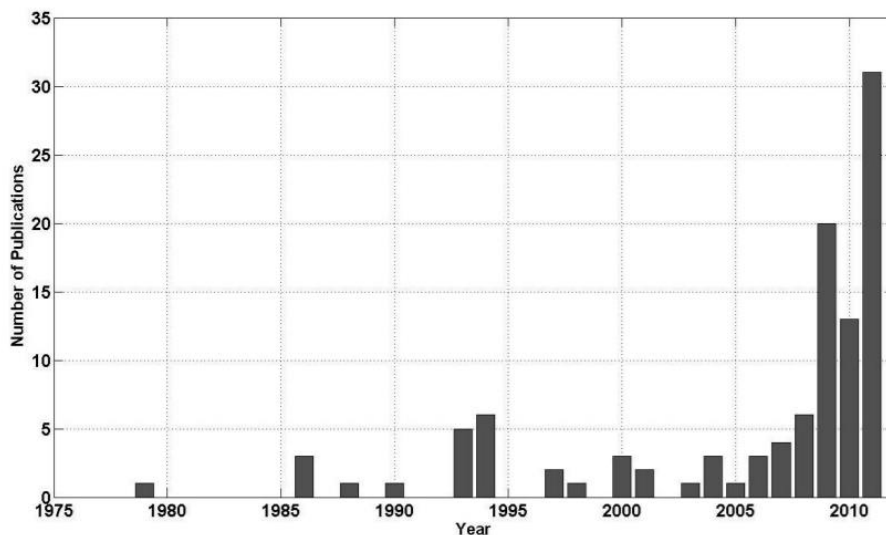


Figure 3 : Number of publications of POD (Jochen H. KURZ, 2012)

The use of the POD functions as a metric for quantifying the NDE capability have undergone considerable development since late 1960's and early 1970's (Health and Safety Executive, 2006) (Health and Safety Executive, 2006). In 1969, the National Aeronautics and Space Administration (NASA) recognized the need for NDE quantification for design and production of the space shuttle. In order to meet the damage tolerance requirement, NASA appraised NDE reliability based on the largest flaw, which is possible to be missed rather than the smallest flaw that is possible to be detected (National Aeronautics and Space Administration, 2009) . The need for this transition was felt because detecting the smallest flaw by NDE method was more unmanageable and less reliable over detecting the largest flaw that could be missed by a NDE method.

The NASA methodology of quantifying NDE capability was incorporated in the standard NASA-5009 and was soon adopted by the US Air Force (USAF) by issuing the document MIL A-83444 (National Aeronautics and Space Administration, 2009). Thereafter, NASA incorporated the POD concept into its Space Shuttle design requirements in 1973 (W.D.Rummel, 1997). The NASA POD development project was pioneering work in quantifying the NDE capabilities and was soon accepted as a standard method. In order to construct the initial plot of the POD curve, moving averages were used, and fitting a curve through these points was done manually (R.T.Anderson, 1973). Thereafter, large number of methods were formulated to draw the POD curves. The equation $POD(a) = N_d/n$ served as the main equation of these methods and binomial statistics was used to calculate the confidence bounds (B.G.W.Yee, 1976). Nevertheless, this non-parametric method was inadequate as it produced coherent results for a single flaw size, but often led to the requirement of extremely large sample sizes to obtain reasonable lower confidence bounds on the POD (A.P.Berens, 1987).

Thereafter, in 1978 a final report (W.H.Lewis, 1978) of a large NDE reliability program carried out by the USAF and statisticians based on the logistic regression to analyse hit/miss data was issued (P.W.Hovey, 1981). The program was called “Have-Cracks-Will-Travel” and the experimental data from the program included numerous examples of two cracks of approximately equal length having significantly different POD. These results clearly demonstrated that besides crack length, the POD was influenced by many other factors (P.W.Hovey, 1983). It is to be noted that the “Have-Cracks-Will-Travel” data was recorded in terms of ‘hit/miss ‘only. Based on the analysis of the above data, Berens and Hovey (P.W.Hovey, 1981) proposed a probabilistic description of POD, “as a curve through the averages of detection probabilities for all cracks of the same length”. They called such a curve as regression curve. The aforementioned approach was documented in an American Society for Testing of Materials (ASTM) Special Technical Publication (P.W.Hovey, 1983). It is mentioned in the aforementioned document that log odds models of the POD provide better estimates of the POD with less scatter, as compared to the non-parametric method which used binomial based estimates.

3.2 POD curves

Pod curves are the generally accepted method for characterizing and quantifying non-destructive inspection capability and reliability (Berens, 1989) (Anon., 2009). It is usually expressed as a function of defect size (or other variables). Defect size at 90% POD with a 95% degree of confidence which is abbreviated as $a_{90/95}$, is a particular value of interest, as shown in figure 4 (Anon., 2009). NDT officials often say, “90% POD with 95% confidence.”. The ‘90/95’ flaw size information is used as a reference. It basically gives insight on how a method/system is performing most of the time. The confidence (interval) is a statistical tool to approximate the POD for a testing method. (Ginzel, 2006)

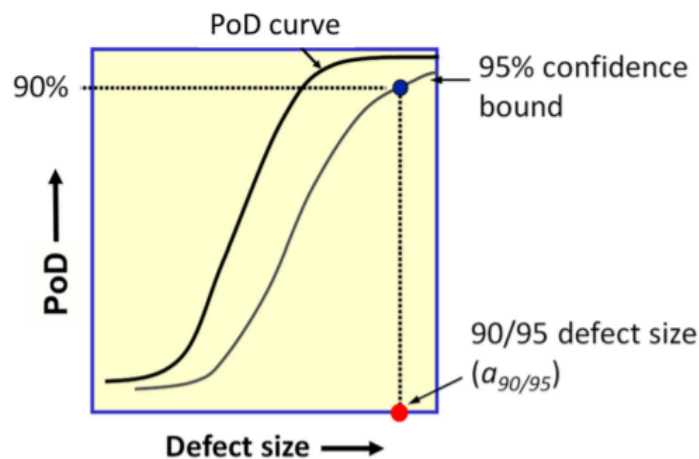


Figure 4 : Schematic of POD curve and corresponding 95% confidence bound (Anon., 2009)

According to Airbus, the aircraft structure has to be designed in such a way that, missing of defects $< a_{90/95}$ is tolerable (KÜCK, 2009). POD curves are usually estimated from data taken from inspections. Normally the quantitative stimulus responses $\hat{a}$ are recorded as inspection results. Sometimes, inspection results are recorded only in terms of whether or not a flaw was found. Commonly, 1 indicates a flaw was found; 0 indicates a flaw was not found. Data of 30 flaw detection of this nature are called hit/miss data (also called binary data). The POD function can be formulated either by the quantitative response data ($\hat{a}$ vs. a data) or the hit/miss data. The flaw size a (or other variables) should be uniformly spaced on a Cartesian scale over an appropriate range, which is the recommended practice (Anon., 2009). It should be noted that if a flaw is so large that it is always found (or saturates the recording device) or so small that it is always missed (or produces a signal which is obscured by the system noise), it provides only limited information about the POD function.

4. Calculations

4.1 Mean Error

Mean is the value that produces the lowest amount of error from all other values in the data set. In order to evaluate the operator efficiency, the process of detection of 21 flaw sizes has to be taken into account. Efficiency of the operator can fluctuate drastically if the sizing of one or two flaw indicates higher or lower with respect to its true size. True size of the flaw is given in Table 1 to compare the measurements taken by the operator.

Table 1 : True Size of the flaw

Defect	True Size(mm)
A	0.25
B	0.27
C	0.28
D	0.4
E	0.4
F	0.5
G	0.52
H	0.55
I	0.7
J	0.98
K	1.12
L	1.15
M	1.1
N	1.3
O	1.6
P	1.8
Q	2.1
R	2.7
S	4.9
T	6
U	7.5

Error of the operator measurement is calculated by using equation 1

$$Error(mm) = True\ size(mm) - Operator\ Measurement(mm)$$

Equation 1 : Determination of Error

All the measurements are calculated in millimetres. If the Error is negative, operator oversized the flaw size. When the error is positive, operator undersized the defects. Error is strictly not dependent on operator experience and performance, but can change according to the quality and reliability of the device used. These can be recorded by analysing the error produced from the real data. The mean error of the operator can be calculated using equation 2

$$Mean\ Error(mm) = Sum\ of\ the\ errors(mm) \div Total\ number\ of\ Defects$$

or

$$\bar{x} = \frac{1}{n} \left(\sum_{i=1}^n x_i \right) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Equation 2: Determination of Mean Error

If the operator has consistence in accuracy of measurements by obtaining true size of the defect as measured value, mean error of the operator evaluates finite efficiency. This can improve performance of operator if size of flaw is less, which is hard to detect.

4.2 Standard Deviation and 95% POD

A standard deviation (or σ) is a measure of how dispersed the data is in relation to the mean value. Low standard deviation means data are clustered around the mean where scores are close together providing a narrow range. High standard deviation indicates data are more spread out or widely scattered. Mathematically standard deviation is calculated using the equation 3

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Where $\mu = \bar{x}$

Equation 3: Determination of Standard Deviation

Since the calculations are done on the excel spreadsheet standard deviation of the error sample can be calculated using the command STDEV.S(Value)

A 95% confidence interval is a range of values that the operator can be 95% certain containing the true mean of the population. This is not the same as a range that contains 95% of the values. The 95% Probability of Detection is calculated mathematically by using the equation 4

$$95\% \text{ POD} = 1.95 \times \text{Standard Deviation}$$

Or

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}}$$

Where $\bar{x}$ =Mean z =95% s = Standard deviation n =Total number of Defects

Equation 4: Determination of 95 % Confidence Interval

The 95% confidence interval can be either positive or negative with relation to oversizing and under sizing measurements obtained by the operator that contains true mean of the error.

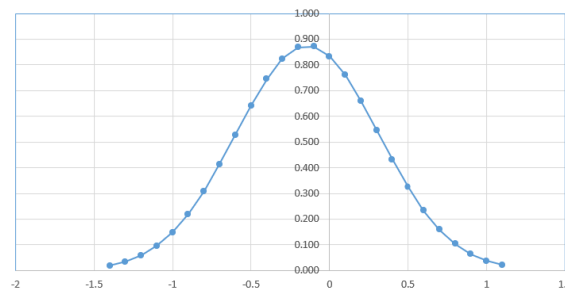
4.3 Normal Distribution Data

For a given value of x the standard normal distribution can be calculated if the mean and standard deviation is calculated for the sample of defects. Normal distribution is represented as z and can be calculated mathematically by using equation 5.

$$Z = \frac{X - \text{Mean}}{\text{Standard Deviation}}$$

Equation 5: Determination of Normal Distribution range.

Since the calculations are done on the excel spreadsheet Normal distribution of the error sample can be calculated using the command NORM.DIST(X, Mean, Standard Deviation, False) for probability mass function. The spread of normal distribution for a given range of values for an operator is mentioned in Graph 1.



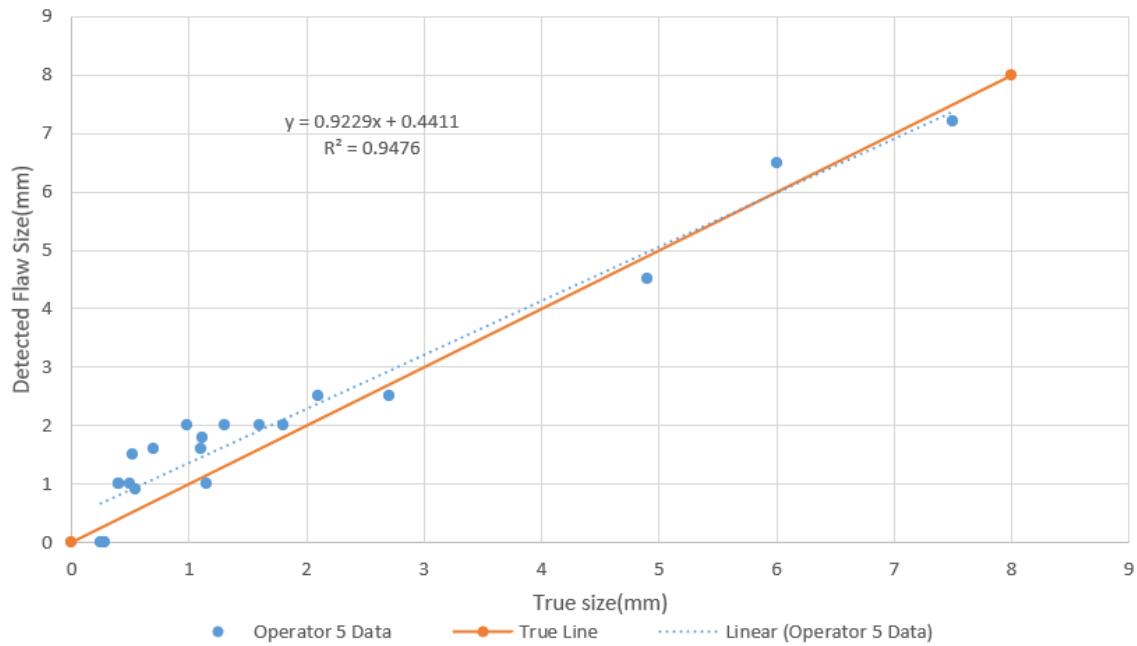
Graph 1: Normal distribution of an Operator

4.4 True Line, Linear Trendline and Correlation coefficient

Operator performance and probe efficiency can be determined by using the scatter plot of true size of defects to operator measurements. True line represents the methodology where error is zero. Operator measurements close to true line gives you an understanding of performance, where true size of the flaw is equal to the measured size by operator.

Linear trendline is also known as regression line with gradient and offset generated according to the measurement taken by the operator. A trendline is typically a line or curve that connects or passes through two or more points in the series, showing a trend. The trendline is data's best fit plotted with the formula $y=mx +c$ where m and c are constants. Value c is the offset that intersect the y axis at starting point and m is the slope of the curve. If $m<1$, that means the operator tends to undersize defects. The resulting line gives a better idea where the measurements are likely to be extended. The relationship between a true line and a trendline of an operator is on graph 2.

The squared value of correlation coefficient or coefficient of determination gives the goodness of fit of data point on regression line. An R^2 of 1 indicates that the regression predictions perfectly fit the data. Scatter points is related to operator measurement as a function of true size.



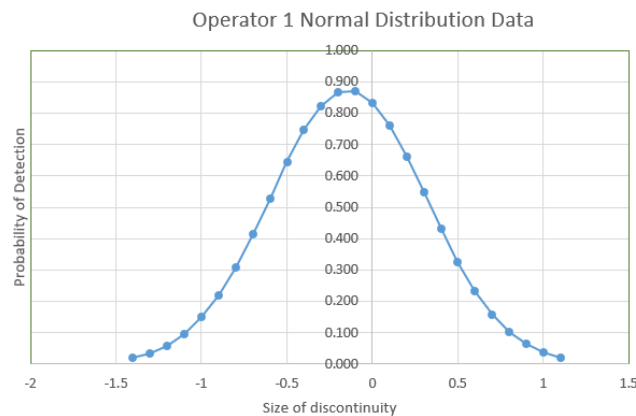
Graph 2: True line and Trendline of an operator data

From Graph 2 the linear trendline intersects the true line at a specific value. The scatter data below True line gives you total number of oversized values to its true size whereas above true line gives the oversized measurement data. From graph, The operator tends to undersize defect for a value of 5.6 mm of true size and more. This represents a trend at which the operator is likely to oversize until point of intersection for flaws with minimalistic size and undersize for flaws with larger size. If the slope m of the trendline is greater than or equal to one, the true line and linear trendline never intersects for the range of flaw size.

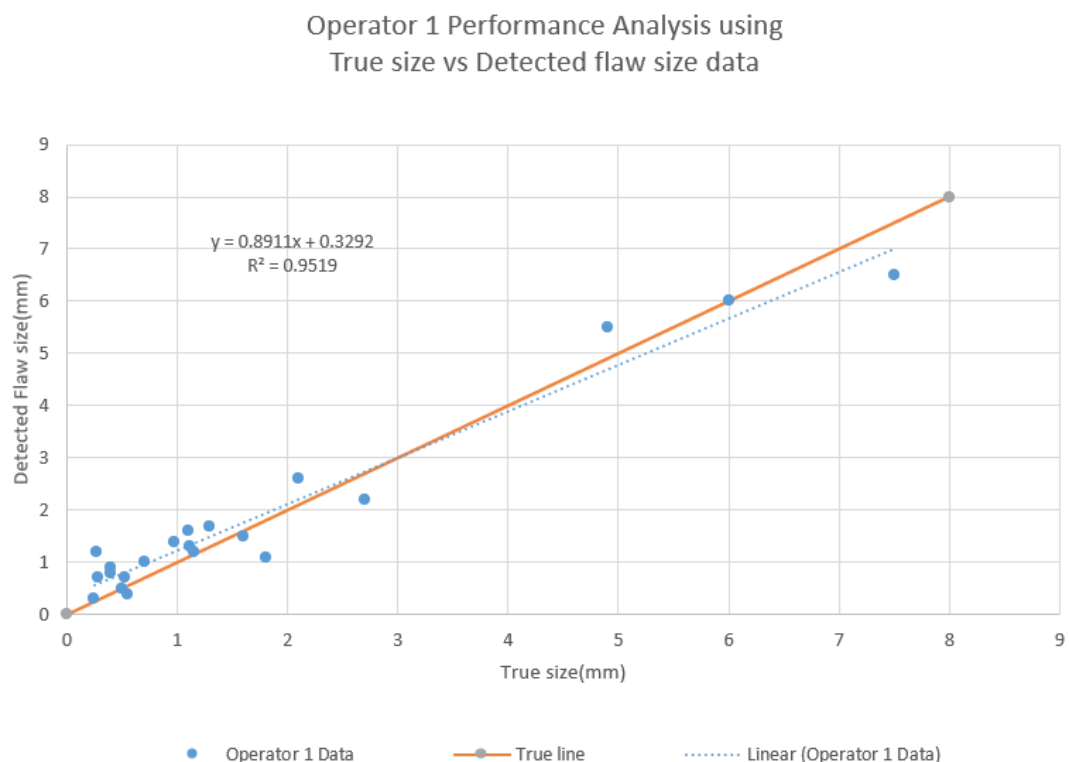
5. Results and Discussion

5.1 Operator 1 Performance Analysis

From measurements obtained by operator 1 with an ultrasonic gauge, the lowest error was recorded as -.14 mm, the mean of 21 detected flaws. A standard deviation of .45 with a 95 % confidence level of .889 represents the spread of data between -1.4 mm and +1.1mm. Normal distribution histogram of the operator is shown in graph 3 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 4.



Graph 3 :Normal Distribution of Operator 1



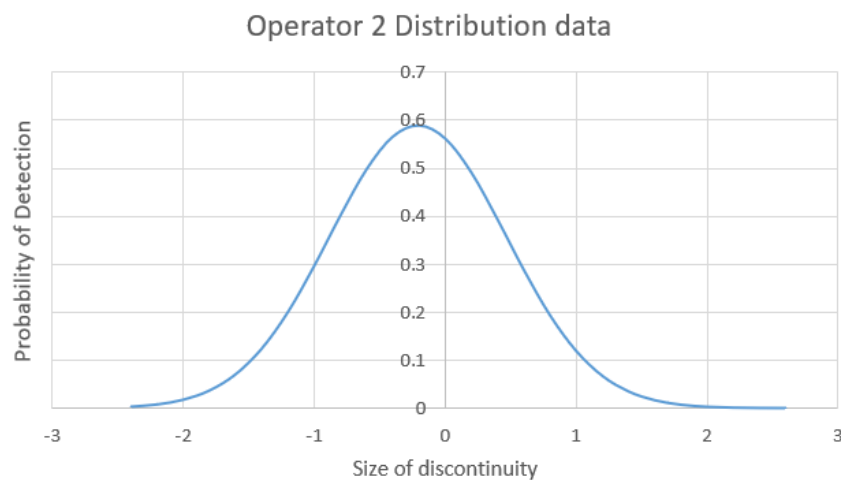
Graph 4: Operator 1 Performance data (True size vs Detected Flaw size).

A linear trendline of slope .8911 outlines that the operator tends to undersize defects for flaw size more than 3 mm. With a coefficient of determination of .95, an offset of .3292 mm is considered to give a linear trendline that is close to the true line. 6 flaw measurement of higher inaccuracy is noted.

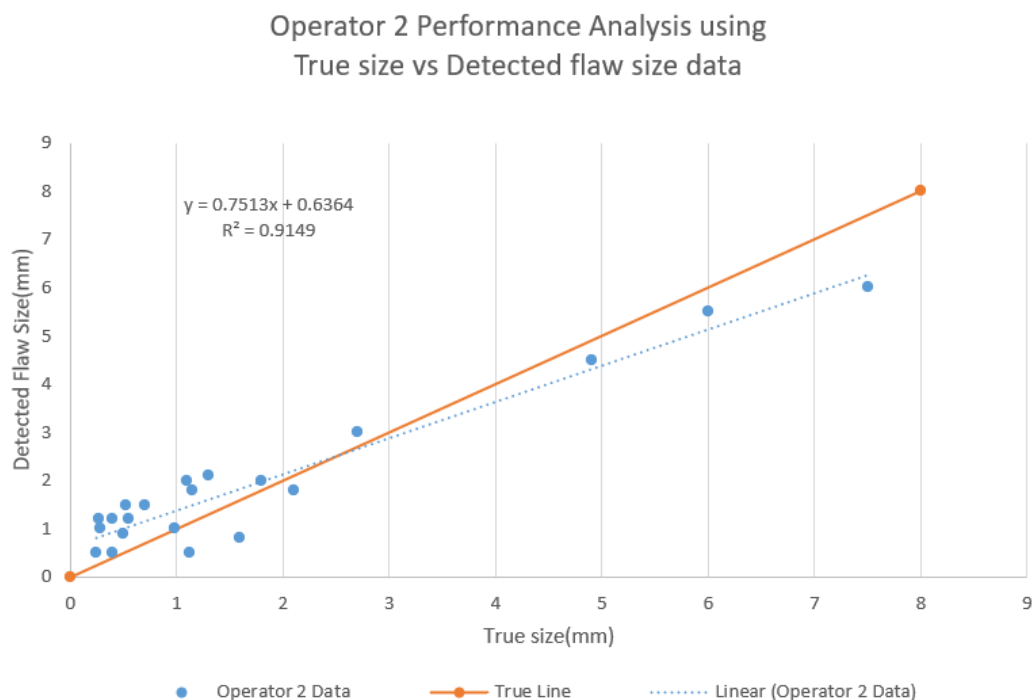
According to the operator measurement detected flaw size lies close to the trendline which outlines an efficient operation. From graph 4 it is clear that a total number of 5 undersized and 14 oversized measurements are being measured by the operator with 2 accurate flaw measurement with zero error.

5.2 Operator 2 Performance Analysis

From measurements obtained by operator 2 with an ultrasonic gauge, the lowest error was recorded as -.20 mm, the mean of 21 detected flaws. A standard deviation of .67 with a 95 % confidence level of 1.318 represents the spread of data between -2.2 mm and +1.8mm. Normal distribution histogram of the operator is shown in graph 5 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 6.



Graph 5: Normal Distribution of Operator 2



Graph 6: Operator 2 Performance data (True size vs Detected Flaw size).

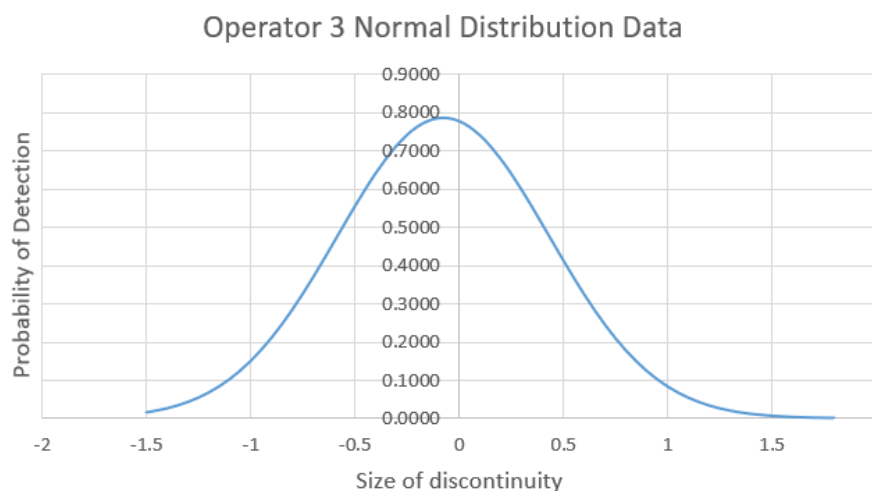
A linear trendline of slope .7513 outlines that the operator tends to oversize defects to a great extent for the flaw measurement less than 2.3mm. With a coefficient of determination of .9149, an offset of .3292 mm is considered to give a linear trend line that is broadly away to the true line.

According to the operator measurement detected flaw size lies broadly away to the true line for majority of defects over economic evaluation error outlines an inefficient operation. From graph 6 it is clear that a total number of 6 undersized and 15 oversized measurements are being measured by the operator with zero accurate flaw measurement.

13 flaw measurement of higher inaccuracy is noted with oversized readings, which affects unconditionally on the overall performance of the operator with correlation coefficient. 9149. This data concludes that operator is less experienced for the operation while providing more opportunities and training can develop the performance of operator.

5.3 Operator 3 Performance Analysis

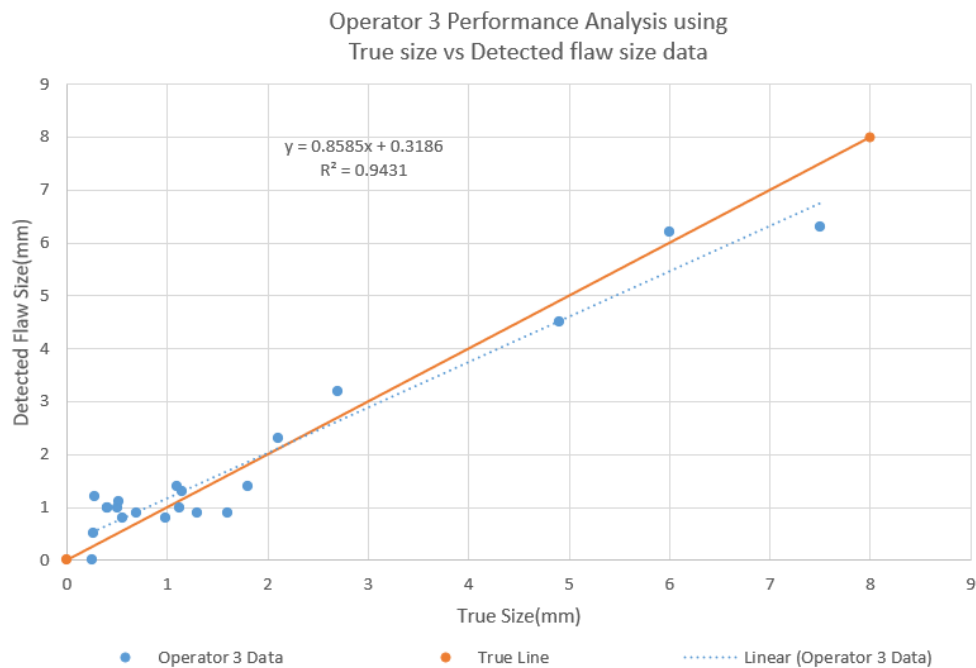
From measurements obtained by operator 3 with an ultrasonic gauge, the lowest error was recorded as -.08 mm, the mean of 21 detected flaws. Mean error of less than .1 mm makes operator valuable for an operation. A standard deviation of .51 with a 95 % confidence level of .989 represents the spread of data between -1.4mm and +1.5mm. Normal distribution histogram of the operator is shown in graph 7 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 8.



Graph 7: Normal Distribution of Operator 3

A linear trendline of slope .8585 outlines that the operator tends to undersize defects for flaw sizes more than 2.2mm. With a coefficient of determination of .94, an offset of .3186 mm is considered to give a linear trendline that is close to the true line for the range of defects.

According to the operator measurement detected flaw size lies close to the trendline which outlines an efficient operation but inefficient than operator 1 due to following reasons. From graph 8 it is clear that a total number of 8 undersized and 13 oversized measurements are being measured by the operator with zero accurate flaw measurement with zero error.

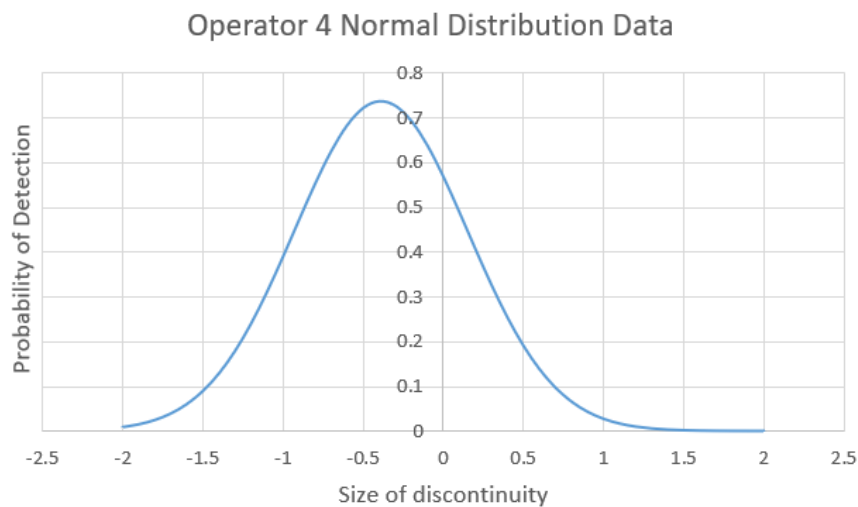


Graph 8: Operator 3 Performance data (True size vs Detected Flaw size).

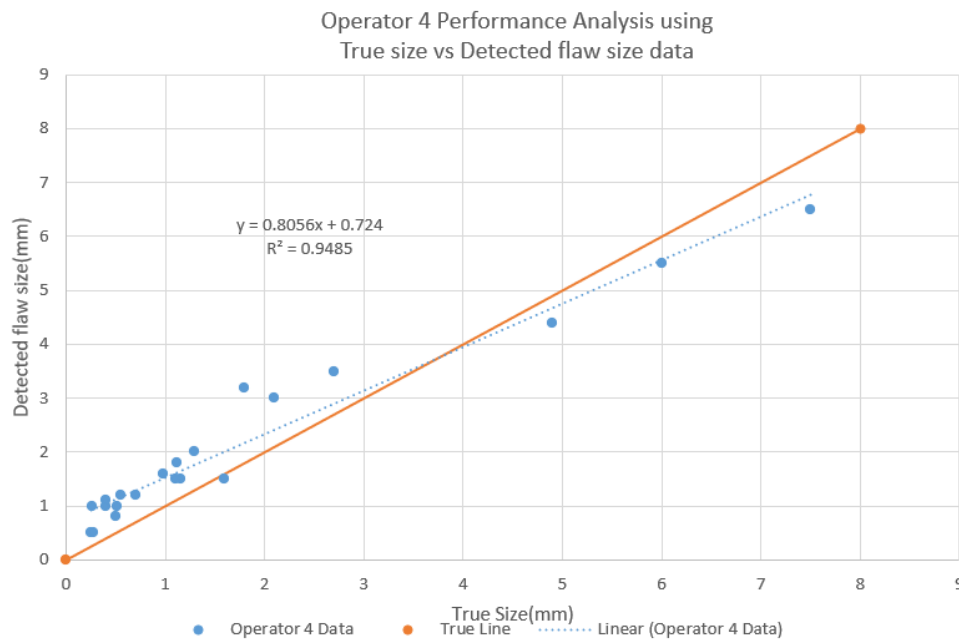
5 flaw measurement of higher inaccuracy is noted which affects unconditionally on the overall performance of the operator with correlation coefficient .9431.

5.4 Operator 4 Performance Analysis

From measurements obtained by operator 4 with an ultrasonic gauge, the lowest error was recorded as -.38 mm, the mean of 21 detected flaws. A standard deviation of .54 with a 95 % confidence level of 1.05 represents the spread of data between -2mm and +1.5mm. Normal distribution histogram of the operator is shown in graph 9 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 10.



Graph 9: Normal Distribution of Operator 4



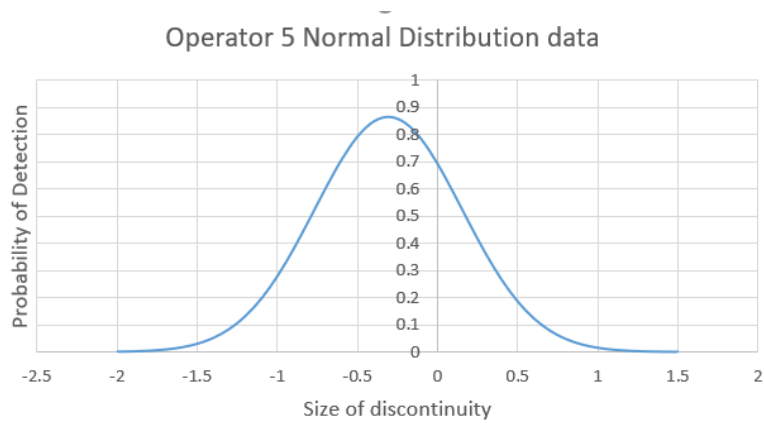
Graph 10: Operator 4 Performance data (True size vs Detected Flaw size).

A linear trendline of slope .8056 outlines that the operator tends to undersize defects for flaw sizes more than 3.7mm. With a coefficient of determination of .9485, an offset of .724 mm is considered where higher inaccuracy in performance is being noted to give a linear trendline that is far away to the true line. According to the operator measurement detected flaw size lies far away to the true line which outlines an inefficient operation. From graph 10 it is clear that a total number of 4 undersized and 17 oversized measurements are being measured by the operator with zero accurate flaw measurement.

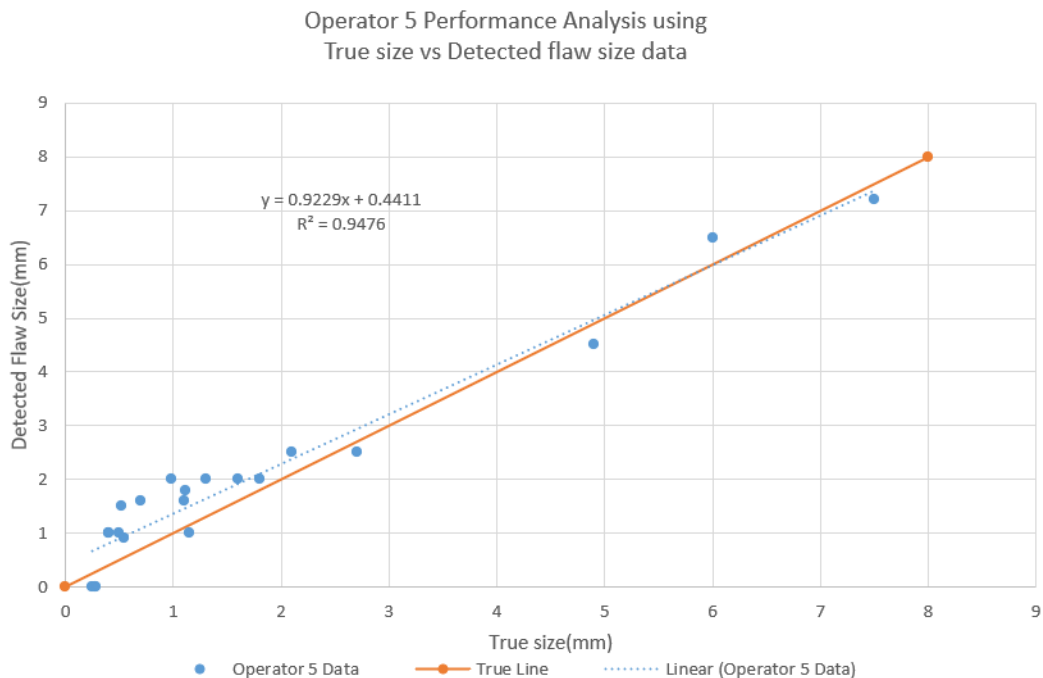
20 flaw measurement of higher inaccuracy is noted which affects unconditionally on the overall performance of the operator with correlation coefficient .9485. This data concludes that operator is less experienced for the operation while providing more opportunities and training can develop the performance of operator.

5.5 Operator 5 Performance Analysis

From measurements obtained by operator 5 with an ultrasonic gauge, the lowest error was recorded as -.31 mm, the mean of 21 detected flaws. A standard deviation of .46 with a 95 % confidence level of .898 represents the spread of data between -1.5 mm and +1mm. Normal distribution histogram of the operator is shown in graph 11 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 12.



Graph 11: Normal Distribution of Operator 5



Graph 12: Operator 5 Performance data (True size vs Detected Flaw size).

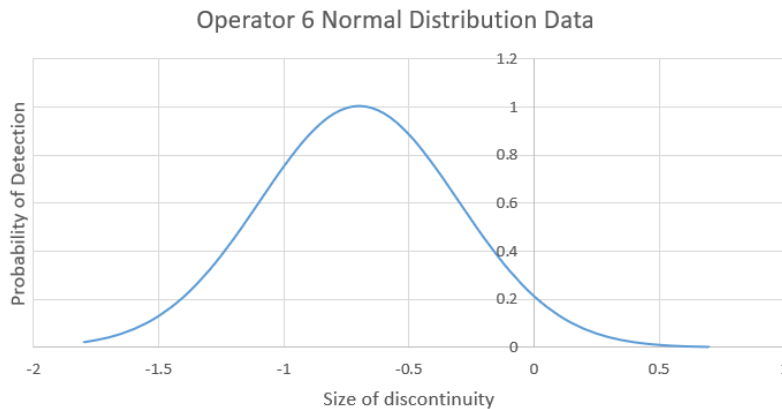
A linear trendline of slope .9229 outlines that the operator tends to undersize for flaw size more than 5.6mm. With a coefficient of determination of .9476, an offset of .4411 mm is considered to give a linear trendline that is close to the true line.

According to the operator measurement detected flaw size lies away to the true line for undersized defects which outlines an inefficient operation. From graph 12 it is clear that a total number of 6 undersized and 15 oversized measurements are being measured by the operator with zero accurate flaw measurement with zero error.

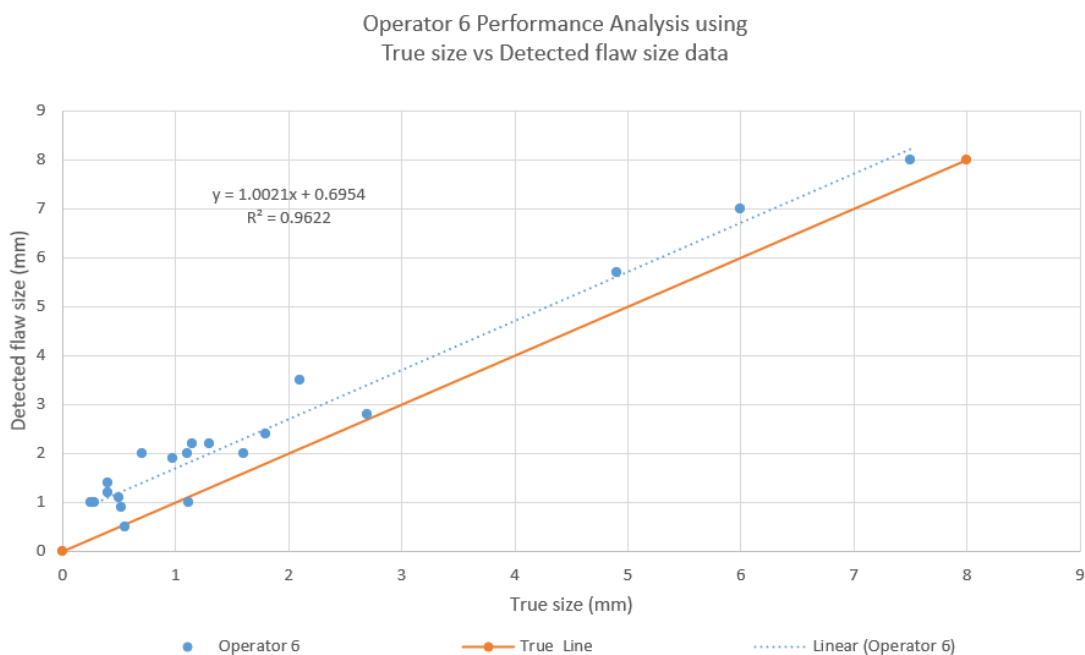
13 flaw measurement of higher inaccuracy is noted for oversized measurements which affects unconditionally on the overall performance of the operator with correlation coefficient .9496. Meanwhile the operator is considered as good due to error tolerance less than .5mm.

5.6 Operator 6 Performance Analysis

From measurements obtained by operator 1 with an ultrasonic gauge, the lowest error was recorded as -.7 mm which is inefficient for the operation, the mean of 21 detected flaws. A standard deviation of .396 with a 95 % confidence level of .774 represents the spread of data between -1.8mm and .5mm. Normal distribution histogram of the operator is shown in graph 13 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 14.



Graph 13: Normal Distribution of Operator 6



Graph 14 : Operator 6 Performance data (True size vs Detected Flaw size).

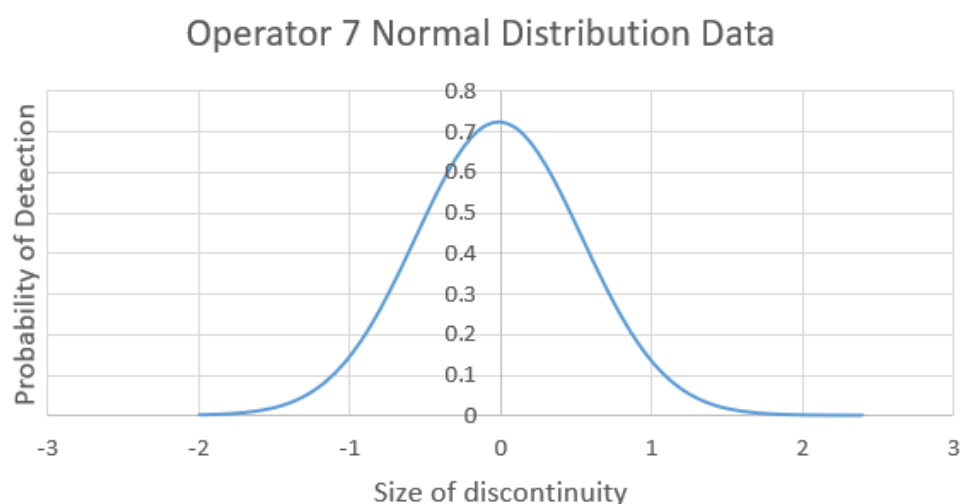
A linear trendline of slope 1.0021 outlines that the operator tends to oversize defects. With a coefficient of determination of .9622, an inefficient offset of .6954 mm is considered to give a linear trendline that is far away to the true line.

According to the operator measurement detected flaw size lies far away to the true line which outlines an inefficient operation. From graph 14 it is clear that a total number of 2 undersized and 19 oversized measurements are being measured by the operator with zero accurate flow measurement with zero error.

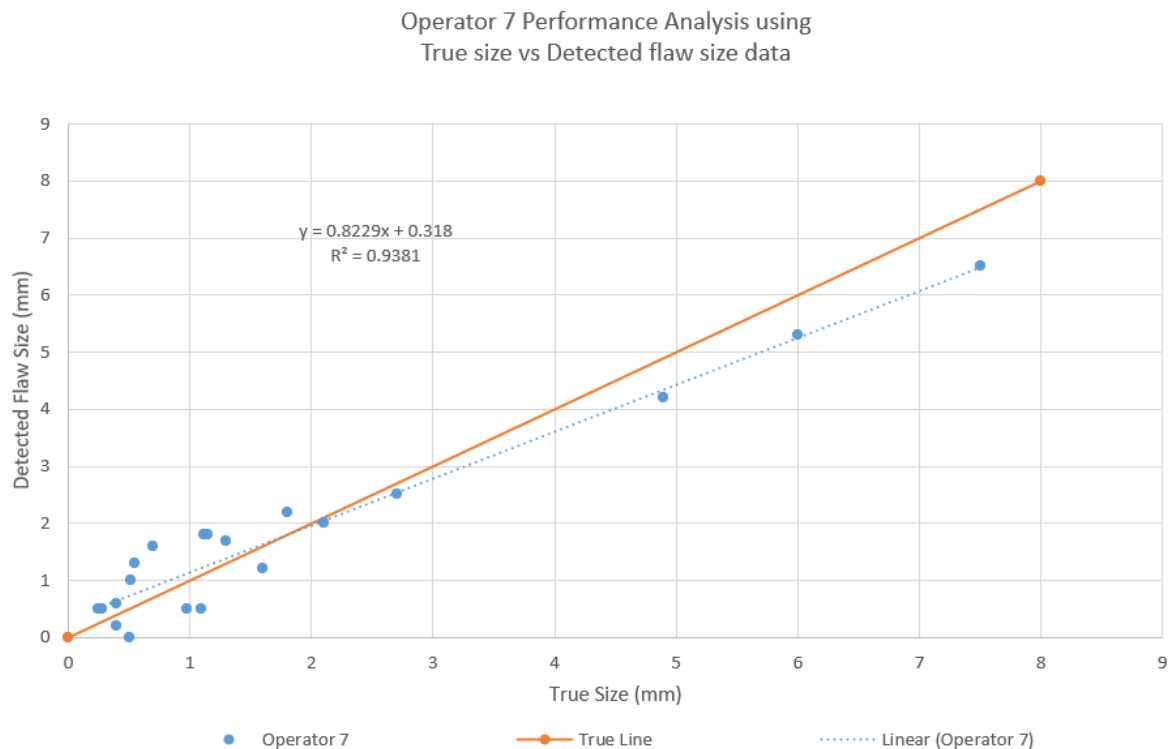
18 flaw measurement of higher inaccuracy is noted for oversized measurements which affects unconditionally on the overall performance of the operator with correlation coefficient .9622. This data concludes that operator is less experienced for the operation while providing more opportunities and training can develop the performance of operator.

5.7 Operator 7 Performance Analysis

From measurements obtained by operator 7 with an ultrasonic gauge, the lowest error was recorded as -.01 mm, the most efficient mean of 21 detected flaws compared to other operators. A standard deviation of .551 with a 95 % confidence level of 1.075 represents the spread of data between -1.5 and +1.5mm. Normal distribution histogram of the operator is shown in graph 15 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 16.



Graph 15 : Normal Distribution of Operator 7



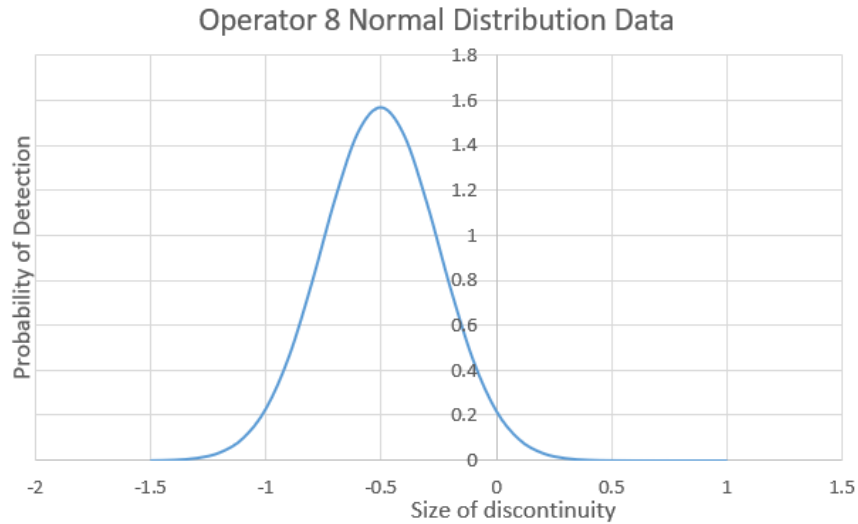
Graph 16: Operator 7 Performance data (True size vs Detected Flaw size).

A linear trendline of slope .8229 outlines that the operator tends to undersize defects over 1.75 mm. With a coefficient of determination of .95, an offset of .318 mm is considered to give a linear trend line that is close to the true line.

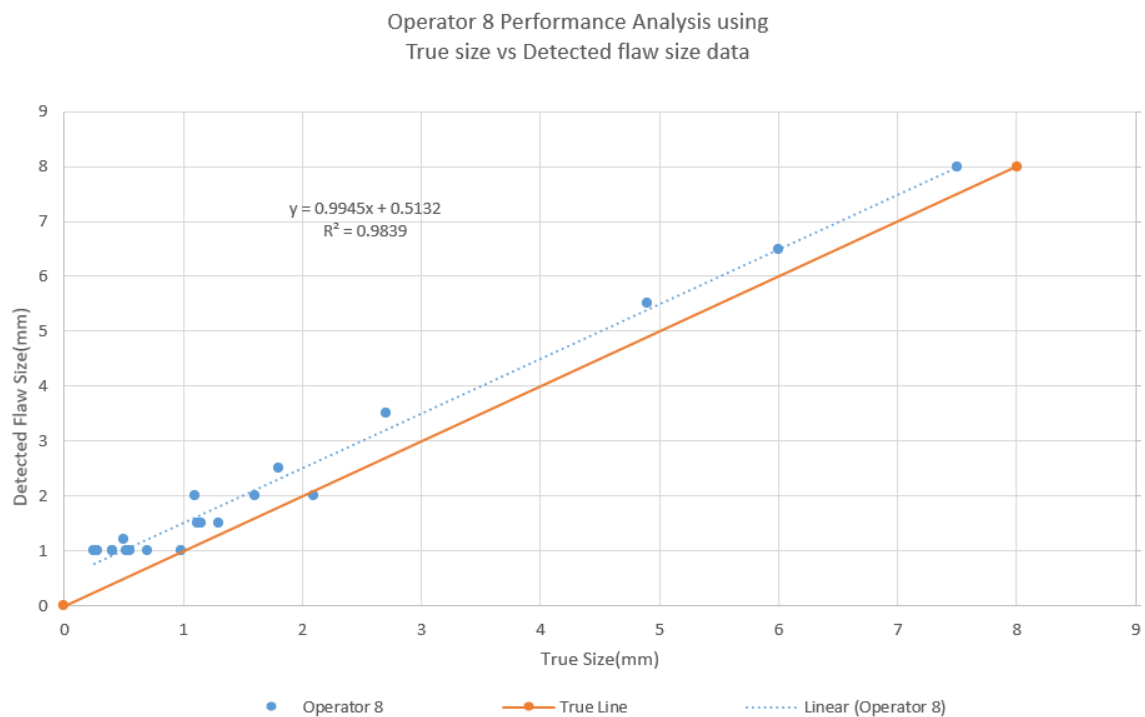
According to the operator measurement detected flaw size lies slightly scattered to the true line which outlines an efficient operation but not efficient enough due to uneven under sizing and oversizing of flaw size to its true value. From graph 16 it is clear that a total number of 10 undersized and 11 oversized measurements are being measured by the operator with zero accurate flaw measurement with zero error. This resulted in the mean error of the operator close to zero. 14 flaw measurement of higher inaccuracy is noted for oversized measurements which affects unconditionally on the overall performance of the operator with correlation coefficient .9381.

5.8 Operator 8 Performance Analysis

From measurements obtained by operator 8 with an ultrasonic gauge, the lowest error was recorded as -.5 mm, the mean of 21 detected flaws. A standard deviation of .254 with a 95 % confidence level of .495 represents the spread of data between -1.3 and +.3mm. Normal distribution histogram of the operator is shown in graph 17 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 18.



Graph 17 : Normal Distribution of Operator 8



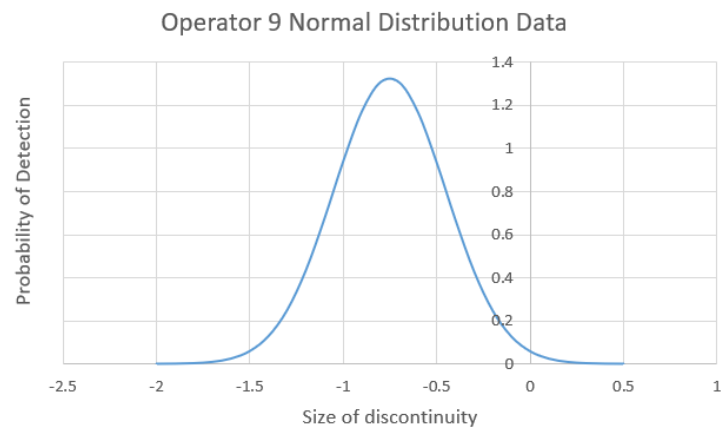
Graph 18: Operator 8 Performance data (True size vs Detected Flaw size).

A linear trendline of slope .9945 outlines that the operator tends to oversize defects. With a coefficient of determination of .9839, an offset of .5132 mm is considered to give a linear trendline parallel to the true line.

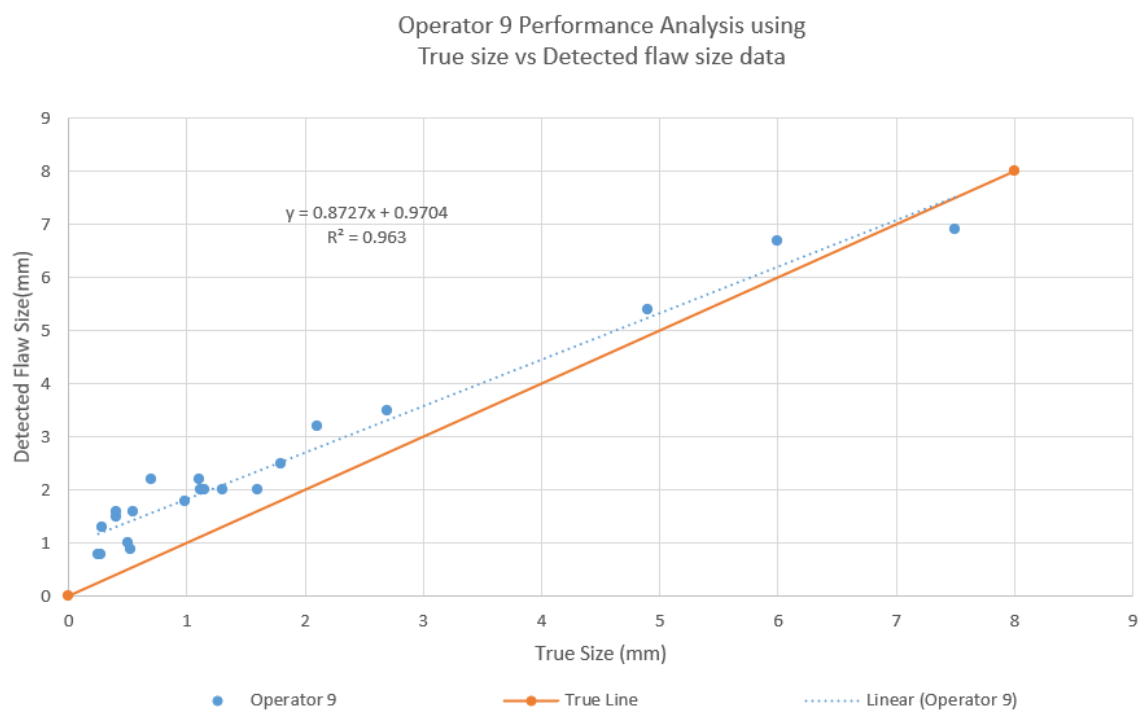
According to the operator measurement detected flaw size lies away from the true line which outlines an inefficient operation. From graph 18 it is clear that a total number of 1 undersized and 20 oversized measurements are being measured by the operator with zero accurate flaw measurement with zero error. 18 flaw measurement of higher inaccuracy is noted for oversized measurement. This data concludes that operator is less experienced for the operation while providing more opportunities and training can develop the performance of operator.

5.9 Operator 9 Performance Analysis

From measurements obtained by operator 9 with an ultrasonic gauge, the lowest error was recorded as -.75 mm, the mean of 21 detected flaws. A standard deviation of .3 with a 95 % confidence level of .585 represents the spread of data between -1.6mm and +.4 mm. Normal distribution histogram of the operator is shown in graph 19 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 20.



Graph 19 : Normal Distribution of Operator 9



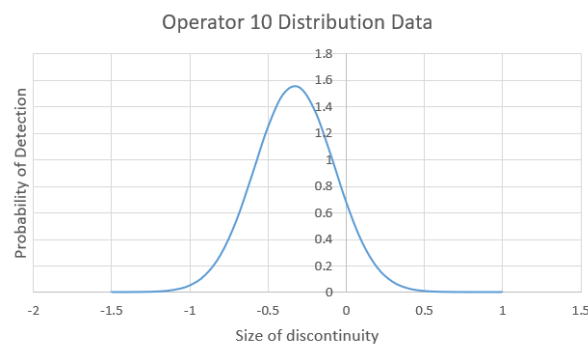
Graph 20 : Operator 10 Performance data (True size vs Detected Flaw size).

A linear trendline of slope .8727 outlines that the operator tends to oversize defects. With a coefficient of determination of .963, an offset of .9704 mm is considered to give a linear trendline that is inefficient to detect smaller flaw sizes.

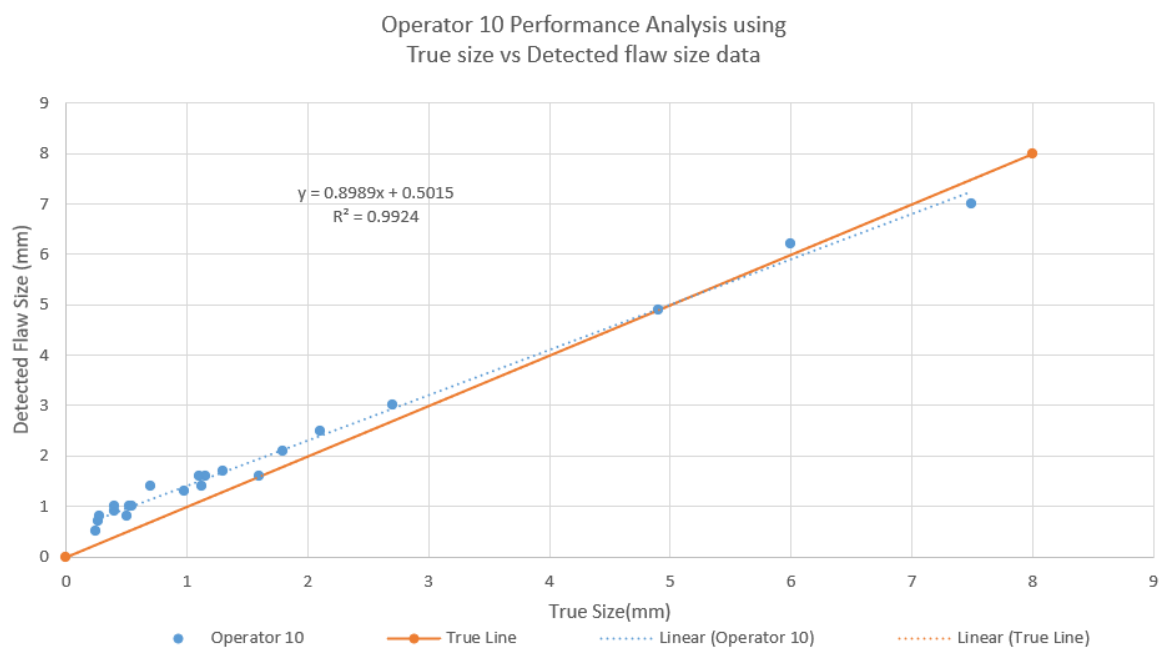
According to the operator measurement detected flaw size lies close to the trendline which outlines a satisfactory operation. From graph 20 it is clear that a total number of 1 undersized and 20 oversized measurements are being measured by the operator with zero accurate flaw measurement with zero error. 21 flaw measurement of higher inaccuracy is noted for oversized and undersized measurement. This could be due to the device inaccuracy and needs to be calibrated as the operator measurements tends to generate values close to linear trendline.

5.10 Operator 10 Performance Analysis

From measurements obtained by operator 10 with an ultrasonic gauge, the lowest error was recorded as -.33 mm, the mean of 21 detected flaws. A standard deviation of .255 with a 95 % confidence level of .499 represents the spread of data. Normal distribution histogram of the operator is shown in graph 21 whereas performance data of the operator with respect to the detected flaw size and true size is shown in graph 22.



Graph 21 : Normal Distribution of Operator 10



Graph 22: Operator 10 Performance data (True size vs Detected Flaw size).

A linear trendline of slope .8989 outlines that the operator tends to undersize defects for a flaw size over 4.9mm. With a coefficient of determination of .9924, an offset of .5015 mm is considered to give a linear trendline that is close to the true line.

According to the operator measurement detected flaw size lies close to the trendline which outlines an efficient and pristine operation. From graph 22 it is clear that a total number of 1 undersized and 18 oversized measurements are being measured by the operator with 2 accurate flaw measurement with zero error. 3 flaw measurement of inaccurate measurement is noted which makes performance of operator 10 better to that of operator 1 whose values are consistent.

6. Conclusion

According to microscopic or macroscopic approach, structural integrity of a material is solely depended upon the strength of core bond between material molecules or atoms. These bond strengths are solely relied upon the operating conditions like temperature and load acting on the material environment. Engineering parameters like stress concentration tends to generate the disassociation of molecules in microscopic approach. A flaw can be generated by disassociation inside these material over long operating hours over manufacturing defects where the reliability of engineering materials under operation is not equal to 100%. Cost of components in several industries like aerospace and automotive are relatedly high where reliability is a prior concern for the safety during operation. Hence it is essential to inspect the components to check the generation of flaw over long term material as well as the newly manufactured ones.

In this report, POD was determined and compared using a phased array ultrasonic probe and sampling phased array technique used by 10 different operators to inspect test piece with 21 flaws. A probability of detection (POD) curve can be obtained from a statistical distribution to evaluate the probability of sizing across a specific flaw using a Non-Destructive Testing method. The probability that an inspector can miss a flaw of extremely small measurements are likely to increase if there is lack of experience where hit is assured for sized flaws. In order to obtain an efficient performance 95 % confidence, bound is considered in industries as a reference to conclude efficient operator performance verifying the reliability of structural integrity of components suitable for operation. The factors that are dependent on inspection carried out obtained by comparing the performance were

- Lack of concentration of operators while operation by not following precise test methodology
- Lack of experience resulting in the under sizing and oversizing of defects with a scattered measurement for a wide range of defects under inspection.
- Device inaccuracy resulting in perfectly undersized / completely oversized data for full range of defects where detection performance of operator efficient.

Adequate training with more inspection experience can develop the inherent potential of the operators. From the data obtained from the normal distribution data and scatter plot, the report is concluded as follows

- Category A-Inspector 1,3 and 10 are considered to be the proficiently classified operators due to their excellent performance of detection. Inspector 1 is the well performed technician among the 3 with a competitive result over technician 3 and 10 with minimalistic tolerance of measurement.
- Category B-Inspector 5,7 and 8 are considered as experienced and the performance is validated as good. Performance of operator 7 was likely similar to Operator 10 and should have considered as excellent if the mean error was less than .5mm.
- Category C-Inspector 2,4 and 6 are considered as satisfactory but not good enough as their performance is categorized as average. This is due to a large-scale reading for a tolerance of above .5 mm
- Category D-Inspector 9 is the only operator mentioned as poor in performance. According to the histogram and scatter plot operator performance was affected due to device inaccuracy. Analyzing the performance with a new accurate probe should categorize this operator as good or excellent.

7. References

- A.P.Berens, 1987. NDE reliability data analysis.. In: *ASM Metals Handbook*. 9th ed. Ohio: s.n., pp. 689-701.
- Annis C, M.-H.-1., 2009. Nondestructive evaluation system reliability assessment, Department of defence handbook, Wright-Patterson AFB, USA.
- Anon., 2009. *US Air Force Aeronautical Systems Center, Military Handbook . Non Destructive Evaluation System Reliability Assessment (Department of Defence*. s.l.:s.n.
- ariseglobal.com, n.d. *ariseglobal.com*. [Online]
Available at: <https://www.ariseglobal.com/advanced-ndt-inspection-services/phased-array-ultrasonic-testing/>
[Accessed 15 December 2021].
- B.G.W.Yee, F. J. G. L. a. P., 1976. *Assessment of NDE Reliability*, Washington DC: NASA.
- Berens, A. P., 1989. NDE reliability data analysis. In: *Metals Handbook*. Ohio: ASM International, pp. 1389 -1470.
- Ginzel, E., 2006. *Introduction to the Statistics of NDT*, s.l.: NDT.net.
- Grand View Research, 2021. *Grand View Research*. [Online]
Available at: <https://www.grandviewresearch.com/industry-analysis/non-destructive-testing-equipment-services-market>
[Accessed 16 December 2021].
- Hanke, 2016. Expanding the potential of NDT across the entire product life cycle. In Proceedings of the 55th Annual Conference of the British Institute of Non-Destructive Testing, Nottingham, UK. p. 305–308.
- Health and Safety Executive, 2006. *HSE-RR-454 Probability of detection curves; derivation, applications and limitation*, Merseyside, UK: s.n.
- Jochen H. KURZ, A. J. S. D. G. D., 2012. Probability of Detection (POD) determination using ultrasound phased array for considering NDT in probabilistic damage assessments. p. 2.
- KÜCK, U. S. a. A., 2009. *Application of POD analysis at airbus*. Berlin, Germany, s.n.
- National Aeronautics and Space Administration, 2009. *NASA-STD-5009. NDE requirements for fracture critical metallic components.*, Washington DC, USA: NASA.
- Ono, K. D. G. ., E. A., 2020. Nondestructive testing, 1. general. In Ullmann's Encyclopedia of Industrial Chemistry;.
- P.W.Hovey, A. a., 1981. *Evaluation of NDE reliability characterization.* , Ohio: s.n.
- P.W.Hovey, A. B. a., 1983. *Statistical Methods for Estimating Crack Detection Probabilities. Probabilistic Fracture Mechanics and Fatigue Methods*, Ohio: s.n.
- R.T.Anderson, T. R., 1973. *Detection of Fatigue Cracks by Nondestructive Testing Methods*, Washington Dc , USA: s.n.

W.D.Rummel, G., 1997. *NDE capabilities data book*. 3rd ed. Austin, Texas: NTIAC.

W.H.Lewis, W. B. a. J. H., 1978. *Technical Report SA-ALC/MME 76-6-38-1, Reliability of Nondestructive Inspections – Final*, USA: s.n.

Wolter, B., Gabi, Y. & Conrad, C., 2019. *Nondestructive Testing with 3MA—An Overview of Principles and Applications*.