1.0 INTRODUCTION

Aluminum cylinders are widely used in different gas and beverage services, under water diving (SCUBA), Fireman breathing (SCBA), medical (O2), beverage (CO2), and industrial gases (toxic and nontoxic).

Cylinders made of aluminum alloy 6351-T6 are known to be susceptible to sustained load cracking (SLC) in the neck and shoulder area of the cylinders. Research and Special Programs Administration (RSPA) has been notified of nine SLC related ruptures of DOT-3AL cylinders made out of aluminum alloy 6351-T6 in the past 14 years. Four of the nine ruptures resulted in serious injuries. Investigations conducted by both industry and the Department of Transportation (DOT) have indicated that although as a whole these cylinders are safe, as is the case with any such product some safety issues remain. Research and Special Programs Administration (RSPA) issued several safety advisory notices for high pressure aluminum seamless and aluminum composite hoop-wrapped cylinders made of the 6351-T6 aluminum alloy:

> “RSPA estimates that approximately seven million cylinders have been manufactured using aluminum alloy 6351-T6. RSPA presently does not know which cylinders among this population have the potential for similar failure. Cylinders made of aluminum alloy 6351-T6 are known to be susceptible to sustained load cracking (SLC) in the neck and shoulder area of the cylinder.”

Federal Register: July 26, 1994
DEPARTMENT OF TRANSPORTATION
Research and Special Programs Administration
Notice No. 94-7

Since issuing these advisory notices further work into the nature of sustained load cracking (SLC) has offered new insight into its causes and effects. However, one key factor in preventing further damage from this type of failure has to date been missing: a reliable method of early detection. Several nondestructive testing (NDT) methods of inspection have been developed, but their efficacy has not been established.

In response the RSPA contracted the Nondestructive Testing Information Analysis Center (NTIAC) to quantitatively evaluate three common methods of cylinder inspection: visual testing (VT), eddy current testing (ET), and ultrasonic testing (UT). Conducting an extensive round-robin testing procedure, NTIAC evaluated the performance of each technique in terms of its accuracy and reliability in detecting flaws in the neck and shoulder region of the cylinder. This report is a summary of our findings and concludes with a recommendation to the RSPA on a testing technique and procedure.

2.0 BACKGROUND

2.1 Significance of the Problem

Those SCUBA and SCBA compressed gas cylinders made of the 6351 alloy and heat treated to the T-6 temper condition have shown a tendency to develop sustained-load cracking (SLC). While the exact cause and mechanism of SLC is not yet fully understood, it is believed that the cracks primarily originate from the bottom of the neck in the crown of the cylinder. The crack propagates outward both radially and axially, as shown in Figure 2-1. In those cylinders that are found to contain multiple cracks, the most common configuration is for two cracks to develop approximately 180° apart from one another. It is thought that the cracks propagate slowly as a form of creep cracking mechanism: originally it was held that the cracks occurred only in those cylinders with high levels of lead and bismuth, although SLC has been found in cylinders with much lower lead levels.

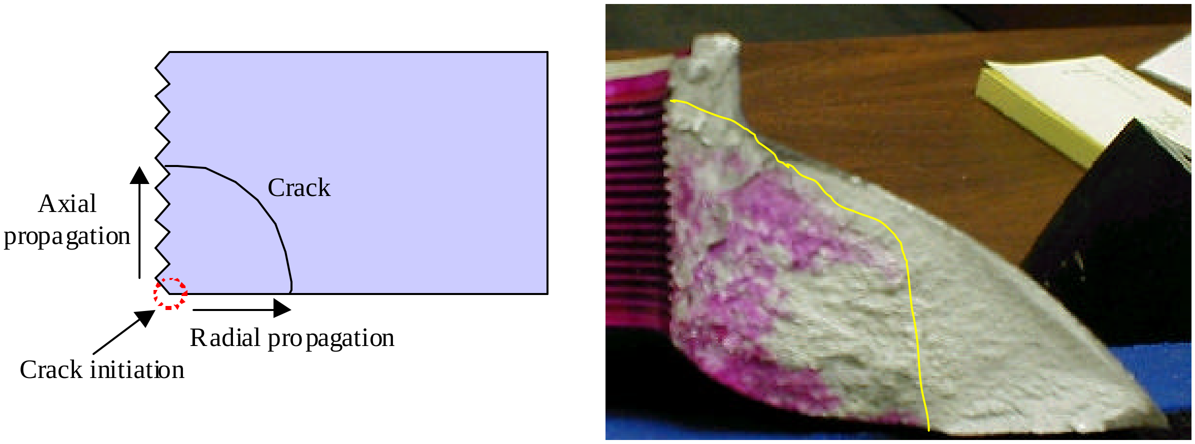


Figure 2-1. Illustration of the Site of Crack Initiation and its Propagation through the Cylinder

By far the most common failure mode is in bottle leaking, which occurs when the crack breaks the outer surface of the cylinder, allowing the compressed gas to escape. In a few isolated incidents, however, the failure mode was rupture rather than leak.

2.2 Methods of Inspection

2.2.1 Visual Testing (VT)

The most commonly-practiced form of nondestructive inspection in use today is visual inspection, due primarily to its simplicity and low cost. Visual testing (VT) typically uses a source of light and a dental mirror to inspect the neck and shoulders of the cylinder. A crack is seen as a thin, "feathery" line that propagates axially up the neck of the bottle, which is distinguishable from a scratch in that a scratch is typically straighter and brighter. Common practice across all the methods of NDT is that any crack that is seen to cross more than two threads is a severe flaw that warrants scrapping the bottle. Figure 2-2 shows a crack detected during a visual inspection (the contrast of the picture has been artificially heightened to illustrate the position of the crack).

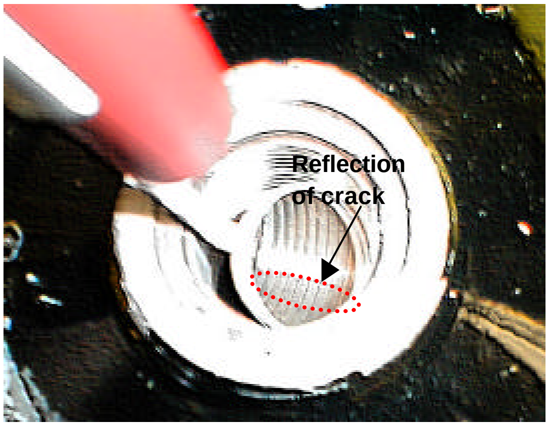


Figure 2-2. VT Equipment used to Visually Inspect the Compressed Gas Cylinders for this Project

Another definition common across the inspection techniques is the definition of a fold in the bottle. Formed during the initial manufacture of the bottle, a fold can appear as a flaw during an inspection, but are not considered as flaws. As such, the successful inspection technique must take into consideration the existence of folds in its execution.

2.2.2 Eddy Current Testing (ET)

The second most commonly-practiced method of NDE for SCUBA and SCBA compressed gas cylinders is eddy current inspection. Several commercial systems are available on the market today. These systems are becoming more popular with dive shops and other groups that routinely conduct cylinder inspections, as it is held that ET is a more sensitive method of inspection than visual inspection. Eddy current inspections of compressed gas cylinders are conducted by threading a plastic probe into the neck of the cylinder: the probe contains an eddy current sensor radially oriented in order to provide the highest level of sensitivity to the axial propagation of the crack (Figure 2-3).

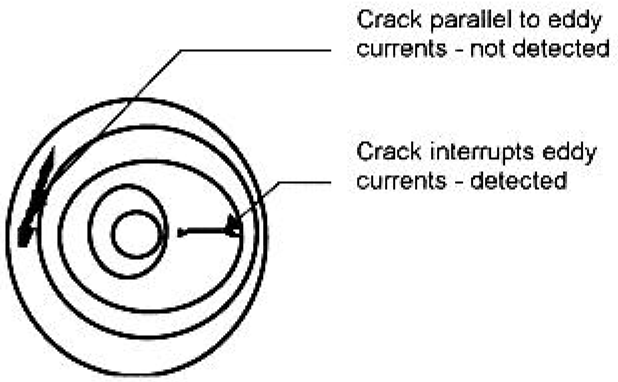


Figure 2-3. Orientation of the Flaw with Respect to the Eddy Currents has a Direct Bearing on the Detection of the Flaw

A typical eddy current system measures the impedance of a sample: in the presence of a flaw the eddy currents in the sample are interrupted, causing a change in the complex impedance of the sample at this position. In the SCUBA / SCBA inspection systems, this is displayed as a “blip” on a B-scan LED display. In turning the plastic probe through a complete revolution (i.e., threading the probe in or out), a given flaw’s signal will re-occur at the same position: a count of the total number of signals for this flaw gives the number of threads the flaw crosses. The ET equipment used for this investigation is shown in Figures 2-4 and 2-5.

2.2.3 Ultrasonic Testing (UT)

This is another inspection method that can be used for SLC inspection of SCUBA / SCBA cylinders. Such systems were prototyped in the past, and are based upon similar designs used for inspecting industrial gas cylinders.

The system NTIAC used for this project sends the ultrtasonic shear wves energy around the shoulders of the cylinder, which has the effect that the system is actually “looking” forward (i.e., clockwise) from its transducer position by approximately 90°. By looking at the shoulder, the ultrasonic system has the potential to detect a crack at the earliest stage, since it is looking primarily at the position most frequently observed as the site of crack initiation. In contrast, both VT and ET primarily inspect the threads in the neck of the cylinder, although VT inspectors are trained to inspect deeper into the bottle as well. As a side-effect of inspecting from the shoulders rather than the neck, UT inspection is not examining threaded neck of the cylinder for the size of a crack as do VT and ET, but rather detect SLC in the shoulder of the cylinder. The UT inspectors must have higher level of training to produce accurate flaw profile. A picture of the UT inspection “arm” is shown in Figure 2-6, and its placement and use on a SCBA bottle is shown in Figure 2-7.

Figure 2-4. Eddy Current Equipment used in the Project. This is one of several systems that are commercially available for eddy current inspection of thread regions. The authors, the RSPA and the DOT do not endorse any specific system manufacturer.

Figure 2-5. Closeup of an Eddy Current Probe Threaded onto the Calibration Block

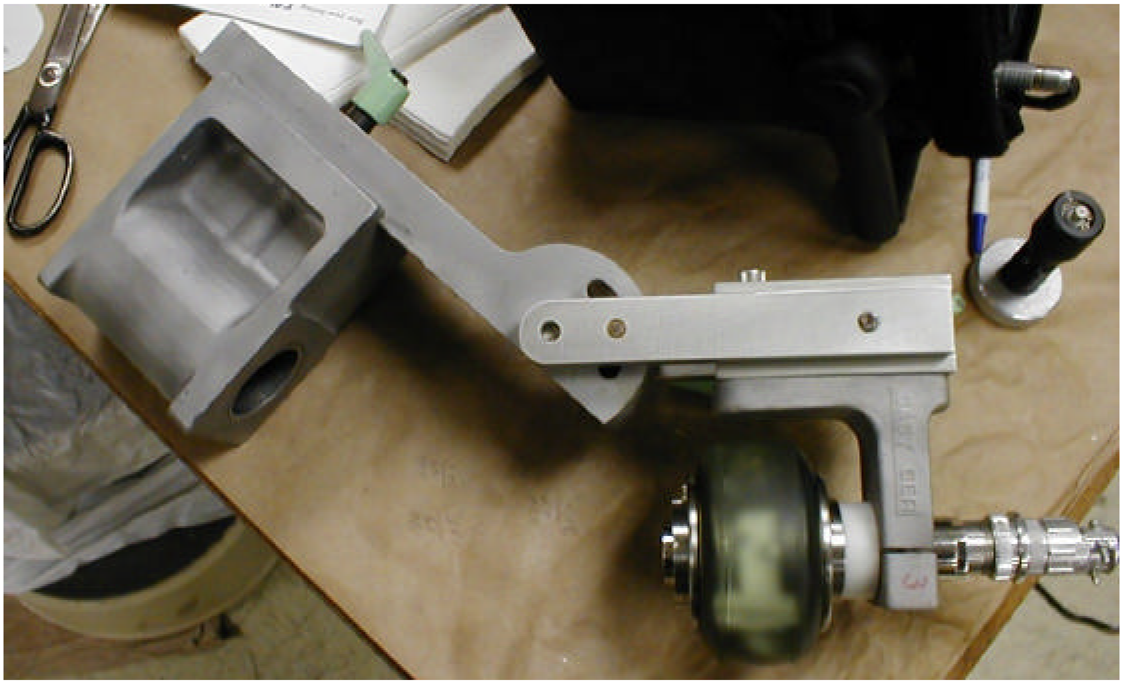


Figure 2-6. UT Inspection System, Consisting of One Pitch and One Catch Probe Inside a Fluid-Filled Wheel

Figure 2-7. Example of a UT Inspection In Progress

3.0 TEST PROCEDURE

3.1 Nondestructive Testing

The original outline of the testing procedure is illustrated in Figure 3-1. To initiate the project NTIAC arranged to have 51 cylinders[1] shipped from Luxfer's California and North Carolina facilities for the test population. These cylinders varied in their age and condition, having been used in the field for some time prior to being "condemned" by an inspector and sent back to Luxfer. The cylinders were to be first inspected via the three methods of NDT, after which destructive testing would be used to gauge their results.

Since the ultimate goal of this research is to evaluate three NDT techniques and recommend a testing technique that can be used by people in the field, it is of utmost importance to go beyond the traditional scope of NDT. In many other NDT applications, the equipment and testing is conducted by someone with years of experience and training. In this case, however, the NDT method is to be employed by people unfamiliar with NDT as a discipline, such as a part-time high school student working afternoons in a local dive shop. As such, NTIAC and RSPA both recognized the need to explore the dependence of skill on each inspection technique: an inspector can receive a minimum amount of training before he or she is deemed fit to practice the technique.

Each inspection technique was thus undertaken by an inspector of one of three skill levels:

- Expert (designated as "X"): the inspector is either the system vendor, a representative of the manufacturing company, or someone with years of experience in the specific technique.

- Skilled ("S"): the inspector has no experience with the specific inspection technique as applied to SCUBA/SCBA cylinders, but is familiar with NDT and other types of inspection.

- Technician ("T"): although familiar with general laboratory practices, the inspector is not familiar with NDT or its associated technologies.

Table 3-1 indicates each of the NDT inspections performed during the project, along with their designation.

Table 3-1. Nine Different NDT Inspections were Conducted for each of the 51 Cylinders

NDT Technique	Skill Level		
	Expert (X)	Skilled (S)	Technician (T)
Visual Testing (VT)	**VTX**	**VTS**	**VTT**
Eddy Current Testing (ET)	**ETX**	**ETS**	**ETT**
Ultrasonic	**UTX**	**UTS**	**UTT**

---

[1] The original contract with RSPA stipulated that at least 50 cylinders were to be acquired. Luxfer had one extra bottle available for this project, and as such while the work plan in Figure 3-1 indicates that 50 cylinders were to be inspected the actual total was 51 cylinders.

Testing (UT)			

Each of these inspections followed the same procedure. Each inspector was briefed on the requirements for their inspection: specifically, to identify all the flaws in each bottle and determine a size for each, whether expressed in thread counts (VT and ET) or as small, medium, or large (UT). This departs from the procedure used in the field, in the sense that a single flaw over two threads in size is sufficient cause to condemn a bottle. The inspector was walked through a sample inspection by the test coordinator, for the purpose of illustrating the data recording procedure. Once the inspector indicated they understood the requirements and the procedure they began the inspection, with periodic check-ins from the test coordinator.

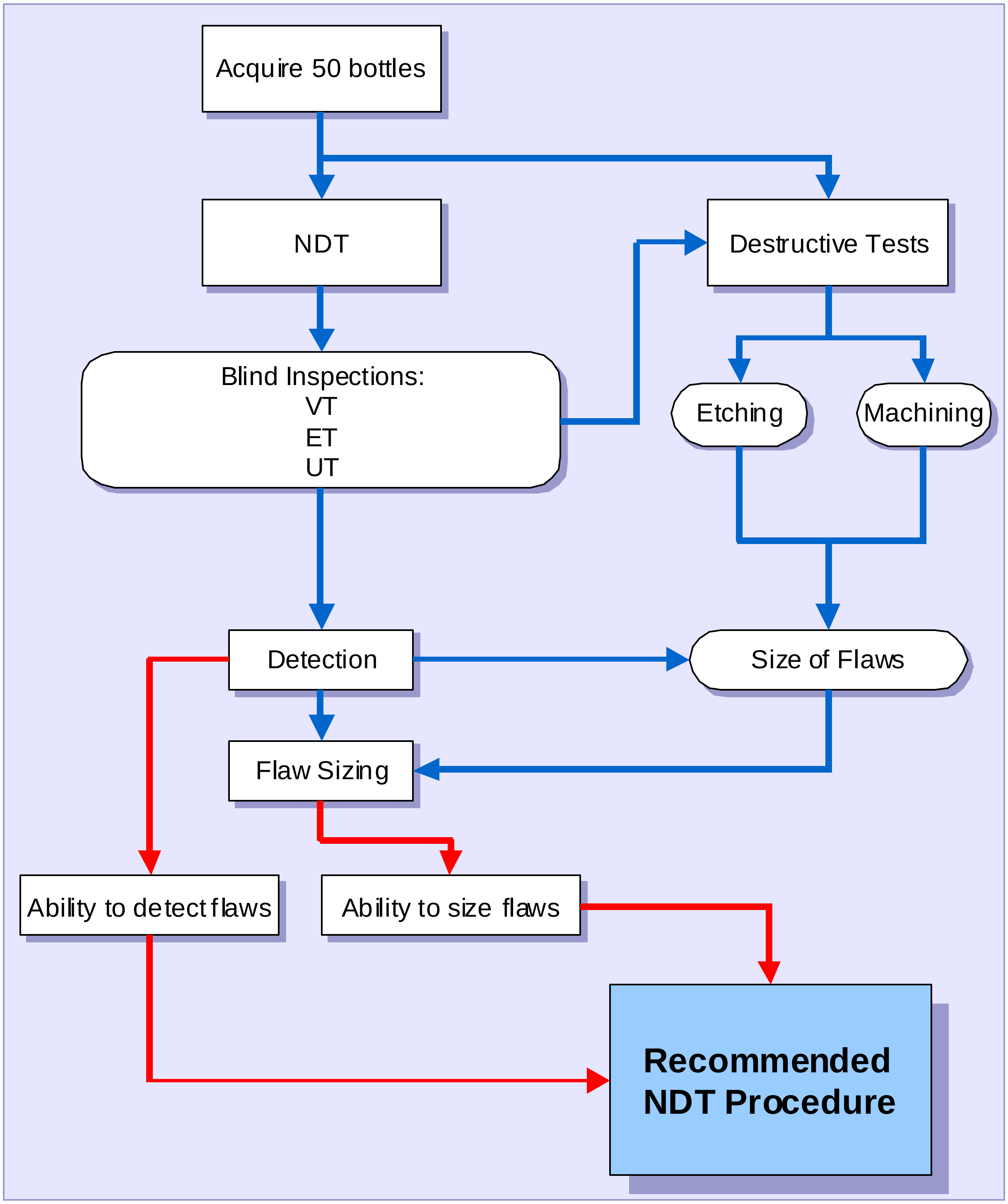


Figure 3-1. Illustration of the Test Progression Used in the Project

The data recording procedure was as follows. When an inspector found a flaw, the first step was to use a piece of masking tape to mark on the outside of the bottle the approximate angular position of the flaw.[1] The inspector was then to make a determination of the size of the flaw, which was done according to the technique:

- A VT inspector visually counted the number of threads the flaw was seen to cross.
- ET inspectors threaded the probe in and out of the neck of the bottle: each time the system found the signal represents one crossed thread, and so counting the number of "blips" gives the number of threads the flaw crosses.
- UT inspectors either called the flaw as small, medium, or large (UTX), or recorded the signal's duration as the wheel rotates past the flaw (UTS, UTT).

Once this was completed for each flaw found on the bottle, the inspector filled in a data sheet for the inspection. The inspector would choose to simply tabulate their findings on the left side of the page, including any comments, and / or to visually indicate the flaw positions using the representative overhead view of the cylinder on the right. Each data sheet records the technique, date, and bottle number, and was signed by the inspector. Once the data sheet had been completed, the inspector took the bottle (with the tape markings still on it) and the data sheet and moved it to the side for later recording.

The compilation of the test data was the responsibility of the test coordinator. Each data sheet recorded by the inspector was compared against the markings left upon the bottle by the inspector. This was done for two reasons: first, to ensure that the angles recorded on the data sheet matched the angles indicated by the tape markings; and second, to act as a method of data backup. In the event of the loss or destruction of the data sheet, the markings on the bottle would be used to fill out a replacement. This redundancy of data ensured that no inspection results were lost. It is also noteworthy that the test coordinator did not participate in any of the inspections, to avoid biasing the results. The test coordinator did participate in the instructional sessions conducted by system vendors in preparation for writing this report, but otherwise did not conduct any testing nor shared any preliminary results with the inspectors.

The compiled NDT inspection spreadsheet was used for subsequent analysis. By far the most important factor for inspections of this type is "detectability," the ability to detect flaws reliably. Since the destructive analysis was to occur after the analysis of the NDT results, and only on a subset of cylinders, certain assumptions were made as described in the Results section. Also important for a method of NDT is its ability to accurately size a flaw, although in this case flaw sizing is secondary to detectability in that, by current field practices, any flaw over 2 threads in size rejects the bottle: field inspectors will typically only size a flaw as far as to determine whether it is 2 threads or larger and will accept / reject a bottle on this basis. Gauging an inspection's flaw sizing ability relies on input from the destructive tests, in that the results of the destructive analysis for a given bottle are compared with the flaw size as reported by the NDT inspector. The procedure for destructive analysis is described in the following section.

3.2 Destructive Analysis

---

[1] Each bottle was marked with a 0° reference point prior to the inspections.

The purpose of the destructive analysis is to provide some quantifiable benchmark results for quantitatively determining the performance of each inspector. Two types of analysis were conducted, each of which is discussed in turn.

Etching

Several cylinders were selected from our test population for destructive testing. Criteria for selecting this subset include

- Representing a range of crack sizes, from the very small to the very large;

- Using no more than two cracks per bottle, of which the cracks must lie approximately 180° apart

- Including both SCUBA and SCBA compressed gas cylinders in the study; and

- Using cylinders for which there was widespread agreement on the position of a crack (i.e., that there was in fact a crack at the position indicated).

Once a visual inspection verified the existence of a crack in the marked position(s) the cylinder head was removed for easier preparation. Each head was immersed separately in a solution of 10% NaOH for approximately 20 minutes, or until upon visual inspection the heads were seen to have been cleaned of debris.

The cleaned and etched cylinder head was then cut in half along a line perpendicular to any existing cracks (Figure 3-2). Each half of the cylinder head thus included a single crack, which was broken open by applying a force directly over the crack as shown in Figures 3-3 and 3-4. Prior to this, however, each half was placed in a vacuum chamber for 15-20 minutes, during which time a dye penetrant was drawn inside the crack. The dye thus indicates the extent of the crack after the crack faces are broken open.

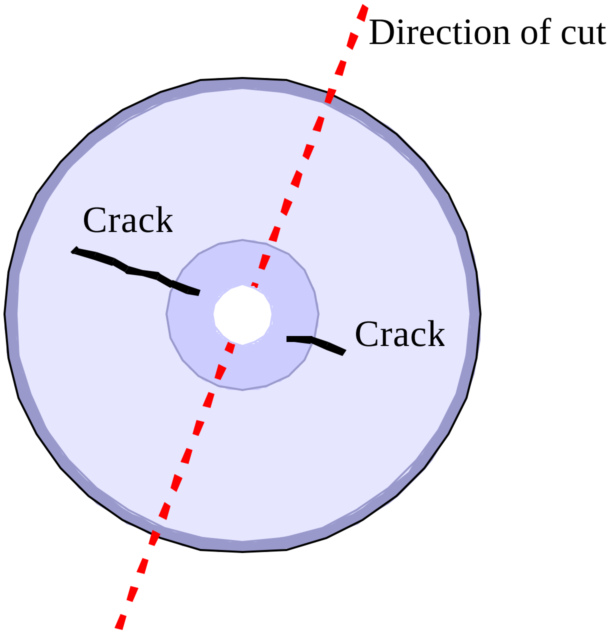


Figure 3-2. Each Cylinder Head was Cut to Isolate the Cracks of Interest

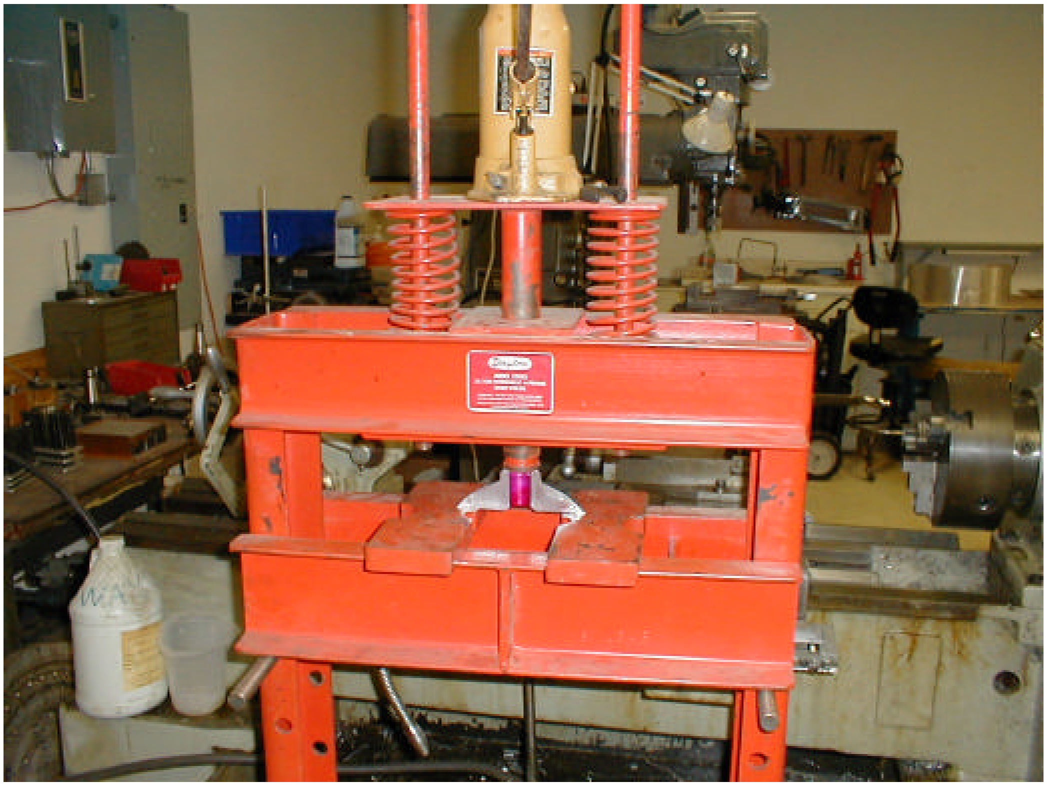


Figure 3-3. The Apparatus used to Break Open the Cracks in a Cylinder Head

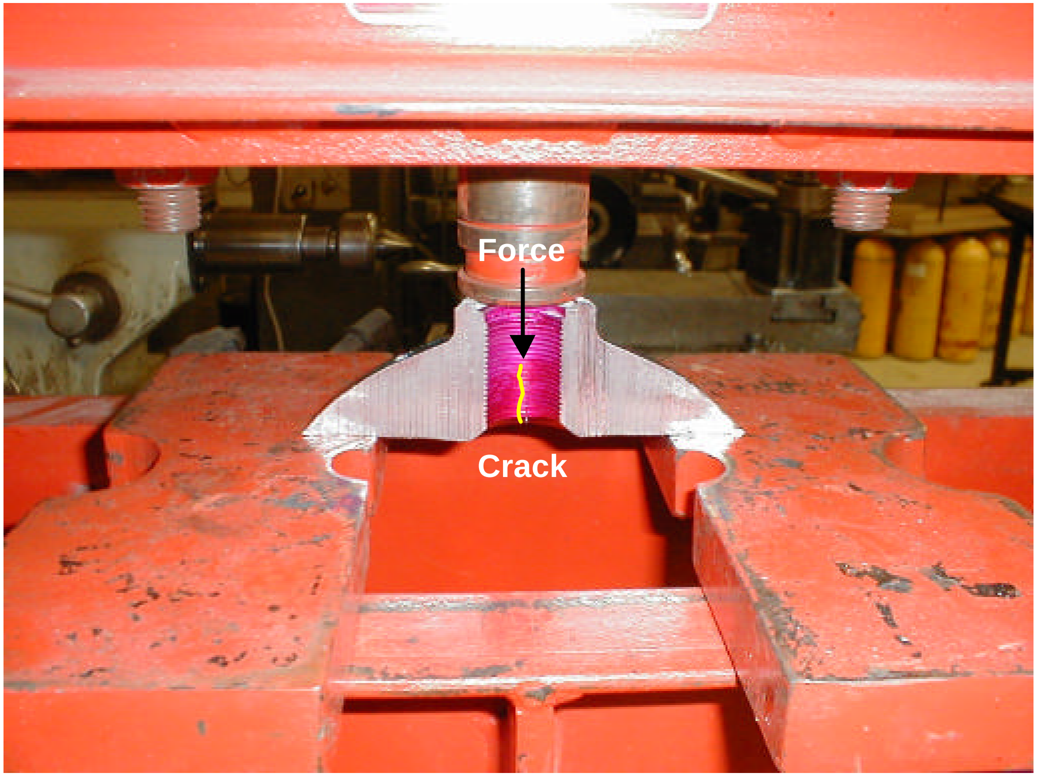


Figure 3-4. Close up View of the Crack Opening Procedure

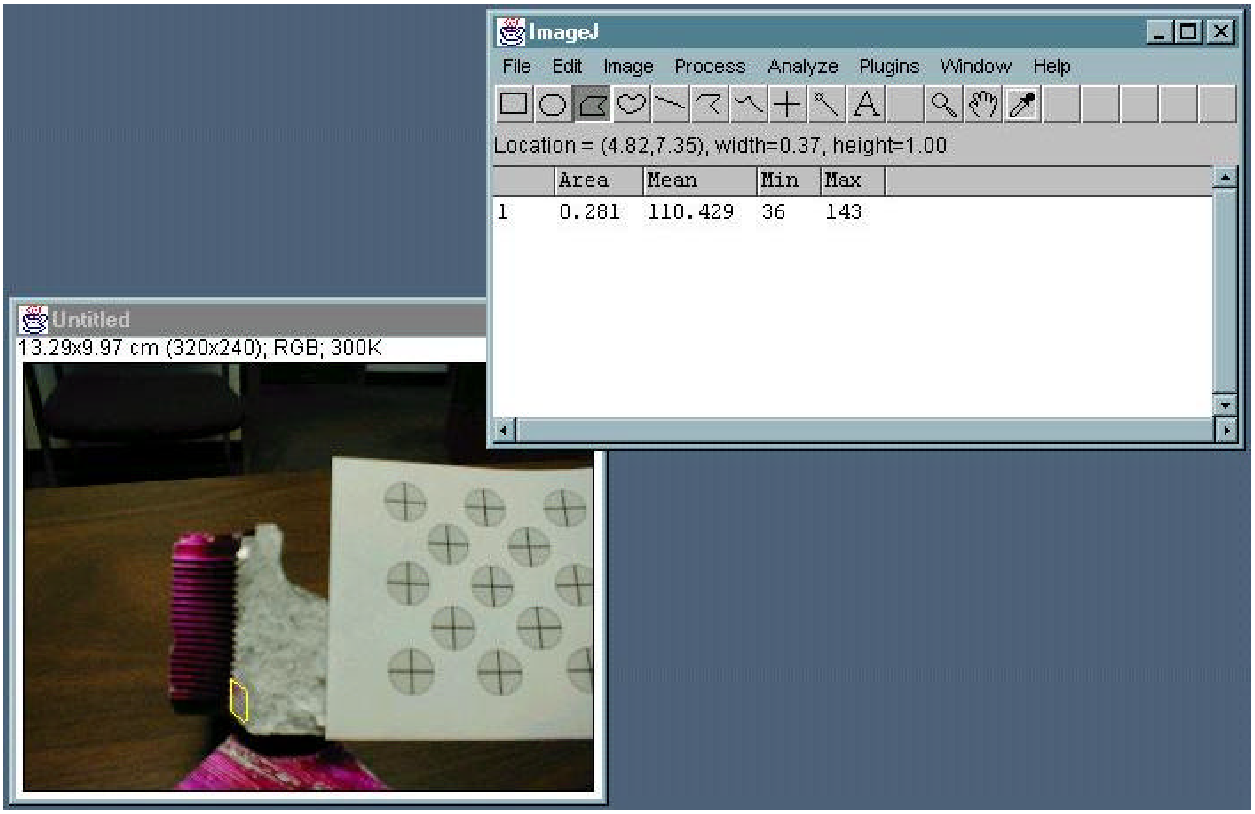
The resulting crack faces were subsequently analyzed with a Java application called ImageJ, produced and freely distributed by the National Institute of Health (NIH). ImageJ is based upon NIH Image, used for over 10 years in research and industry. In choosing ImageJ for the analysis, NTIAC had several goals in mind:

1. **To use a technique that enjoyed widespread acceptance in not only the NDE community but in the larger field of image analysis.** ImageJ's algorithms are used at NIH and Argonne National Labs, and its analysis routines are the basis of a package distributed by Scion Image with their frame grabber boards.

2. **To produce quantifiable and repeatable results that were independent of the data analysis operator.** Each of the analysis routines is based on fundamental math routines rather than direct human measurement, and as the Java source code is freely available these routines can be verified for accuracy.

3. **To produce results that could easily be shared and manipulated electronically.** Interested parties need only obtain a copy of ImageJ for themselves and the digital pictures to examine its results.

When a picture of the crack face is calibrated to a known physical dimension, ImageJ can be used to analyze regions of the image in various ways. As an example, Figure 3-5 shows a screen shot of ImageJ being used to find the cross-sectional area of a crack face.



Location = (4.82,7.35), width=0.37, height=1.00

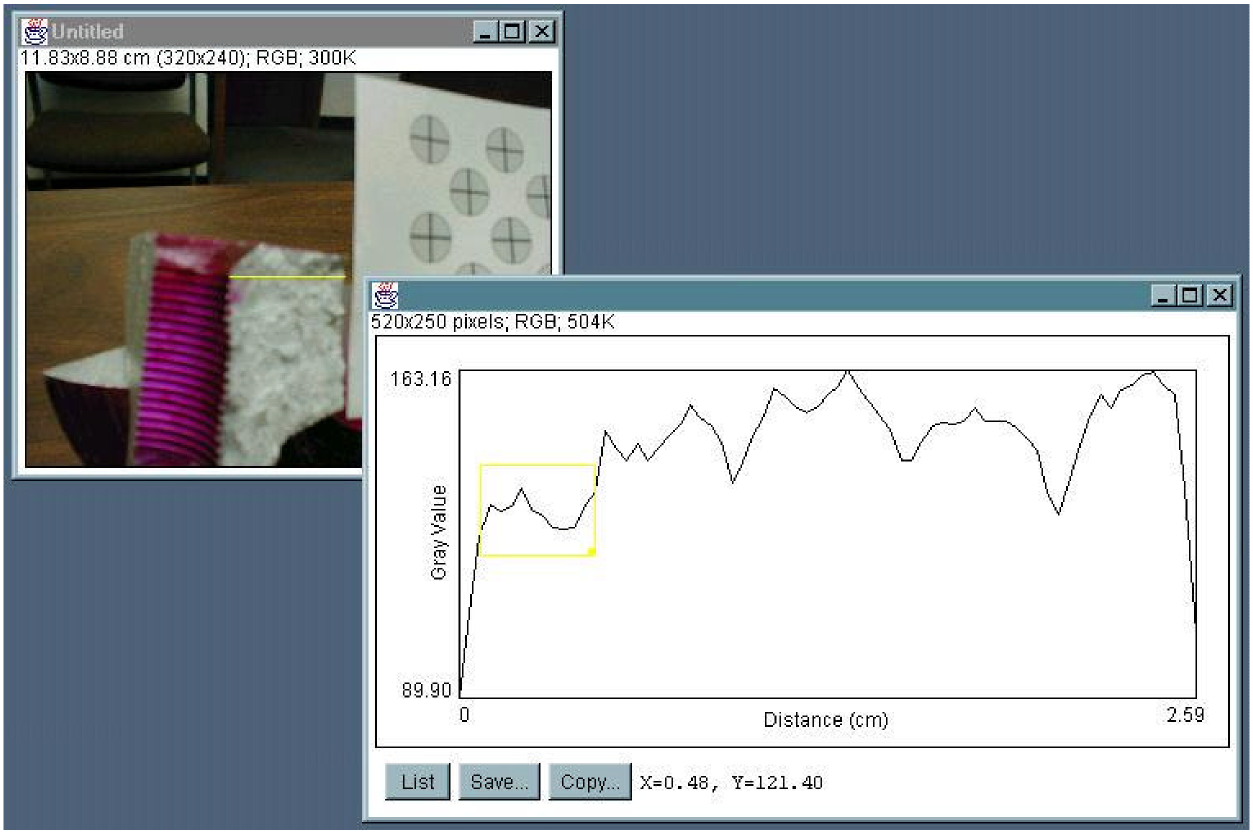
	Area	Mean	Min	Max
1	0.281	110.429	36	143

13.29x9.97 cm (320x240); RGB; 300K

Figure 3-5. Screen Shot of the Java Application ImageJ Produced by NIH, being used to Find the Cross-sectional Area of a Crack in a SCUBA Tank

In each of the pictures, a grid pattern was included to assist in calibration within ImageJ. Each of the circles seen in the crack image in Figure 3-5 is of a known diameter of 1.00 centimeters. The calibration procedure is then to measure the diameter of a circle in the picture's horizontal and the vertical directions in order to correct for the picture's aspect ratio. The extent of the crack was determined visually, but for purposes of verifying repeatability, the extent of the crack face shown in Figure 3-6 was determined by additionally plotting a grayscale profile of the crack. The crack's deepest point of penetration into the aluminum was measured visually in ImageJ at approximately 0.50 centimeters: the profile plot of Figure 3-6 estimates this extent at 0.48 centimeters. Thus, visually determining the extent of the crack was seen to agree quite well with the alternative procedure.

The pictures of each crack face were taken with a standard digital camera. Each face was arranged so as to provide maximum contrast between the crack face and the virgin aluminum. In the case of a wide crack, this task was made simpler because of the greater depth of penetration of the dye. For a tight crack, however, the dye was not able to penetrate as deeply into the crack, and as a result it becomes necessary to pay more attention to the ambient lighting while photographing the crack face. In all cases, it was noted that a certain degree of clarity was lost in going from the actual object to a digital photograph, and again in printing a hard copy of the photo, and as such NTIAC will ship the actual cracks with its final report to the RSPA.



```
11.83x8.88 cm (320x240); RGB; 300K
520x250 pixels; RGB; 504K
X=0.48, Y=121.40
```

Figure 3-6. Measuring the Extent of a Crack. The highlighted region shows the approximate extent of the crack along the line in the upper left image.

Machining

As part of the destructive analysis undertaken during this effort to quantitatively evaluate the various methods of nondestructive testing, NTIAC contracted Bryant-Lee Associates of San Antonio, Texas to conduct electrical discharge machining (EDM) tests of two cylinders, P4205 (SCUBA) and T94863 (SCBA). Their full report is presented as an appendix, but is summarized both here and under the Results section of this report for convenience.

The following procedure was used to section and measure the samples:

1. Before the samples were EDM-machined, the samples were red-dye penetrant inspected to determine the approximate crack length along the bore and the depth into the cylinder. Figure 3-7 shows the cracks identified by inspection. Sample A had a crack visually observed about 0.31-inch from the reference surface (shown in Figure 3-8) and about 0.94-inch radial crack from the ID bore. Sample B had a visible crack that was about 1.1-inch from the reference surface and 0.63-inch radial crack from the ID bore.

2. The yellow paint on the cylinders, shown in Figure 3-9, was stripped from the OD surface.

3. The two cylinder samples, T94863 (Sample A) and P4205 (Sample B), were EDM-machined. The EDM setup is shown in Figure 3-10.

4. The EDM-machined wafers were a nominal 0.100-inch thick face-to-face (about 0.85-inch wafer thickness and 0.015-inch EDM cut width).

5. EDM wafers were parallel to reference surface, shown in Figure 3-8.

6. Each wafer was vibroengraved with a wafer number; for example, A1, A2, B1, B2, etc. Cylinder T94863 had 14 EDM wafers and Cylinder P4205 had 20 wafers.

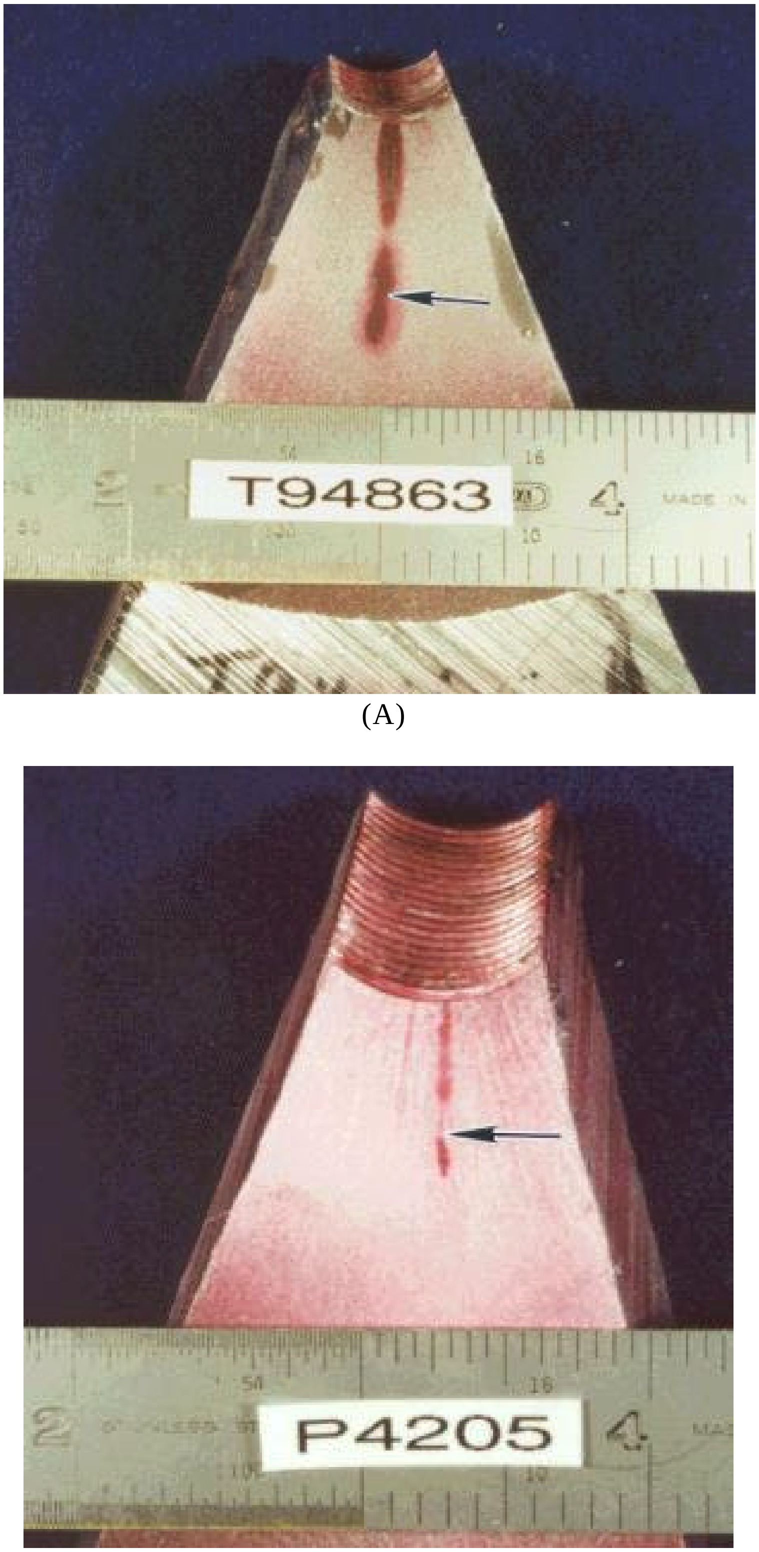
7. BLA measured and recorded the distance from reference surface for each EDM wafer.

8. To examine the crack tip, if present, each wafer was manually sanded and polished on diamond polishing wheels to a 1-micron finish.

9. The as-polished position was estimated by adding the amount of metal removed to the EDM wafer position.

10. Each wafer was examined for a crack using a stereomicroscope. The crack tip was determined by examining the wafer using a metallograph at 50x.

11. BLA determined the crack depth by measuring the distance from the ID bore (thread crest) to the crack tip. For wafers that exceeded the borehole, the crack depth was measured from the ID surface, for example; wafers A11 and B16.



(A)

(B)

Figure 3-7. Photographs of the Red-Dye Penetrant Indication for (A) Cylinder Samples T94863 and (B) P4205. Arrows indicate cracks.

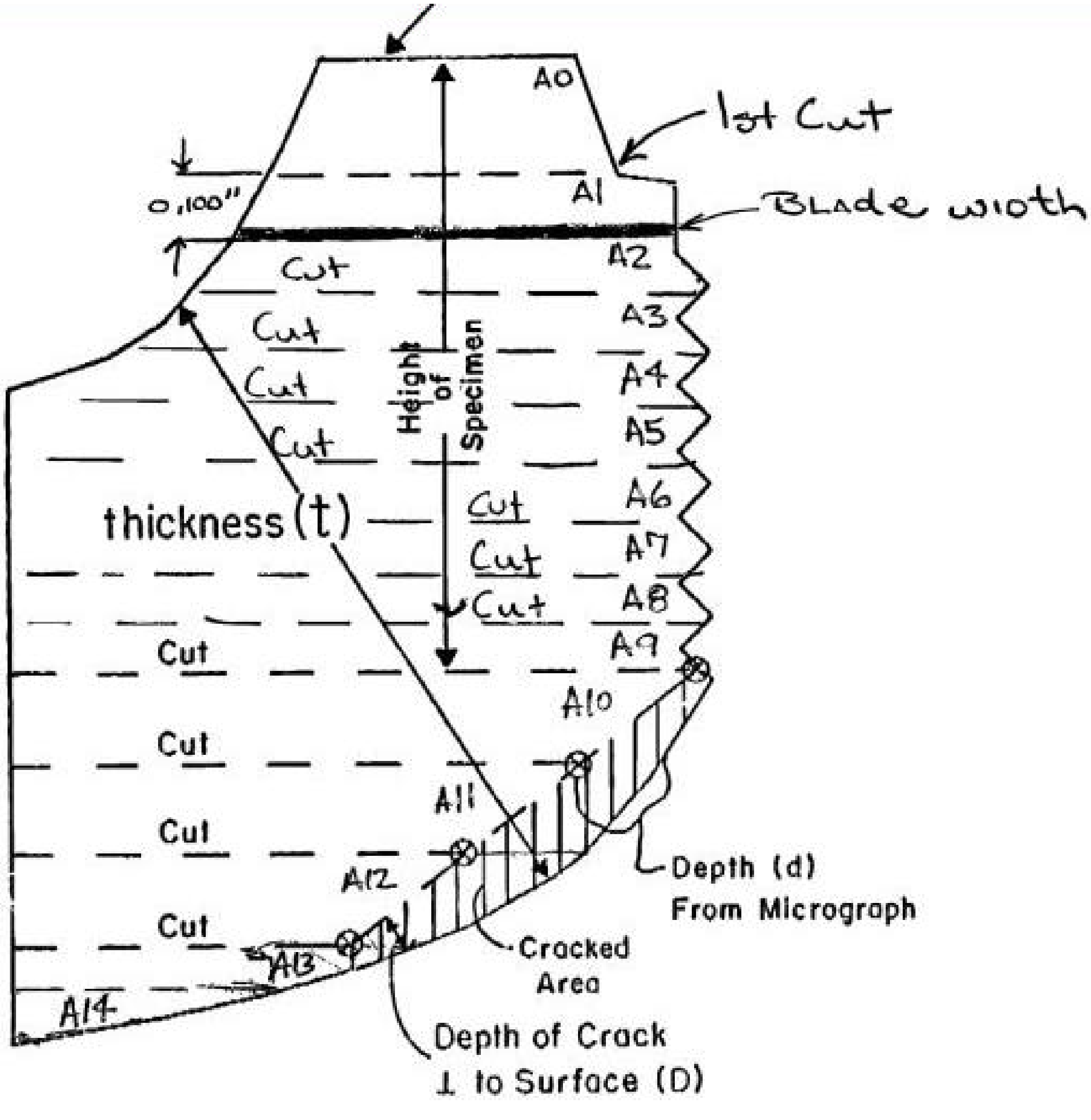
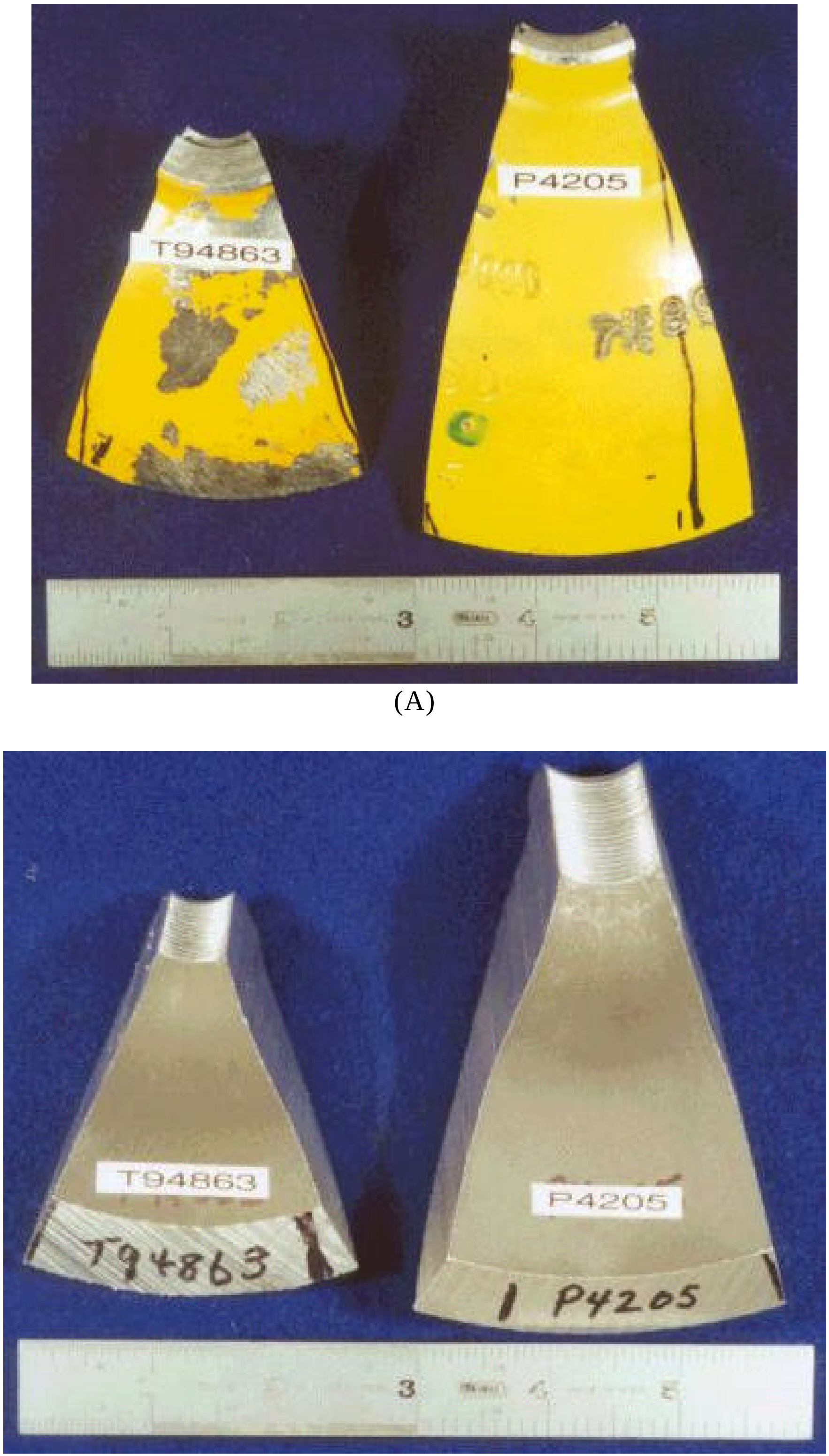


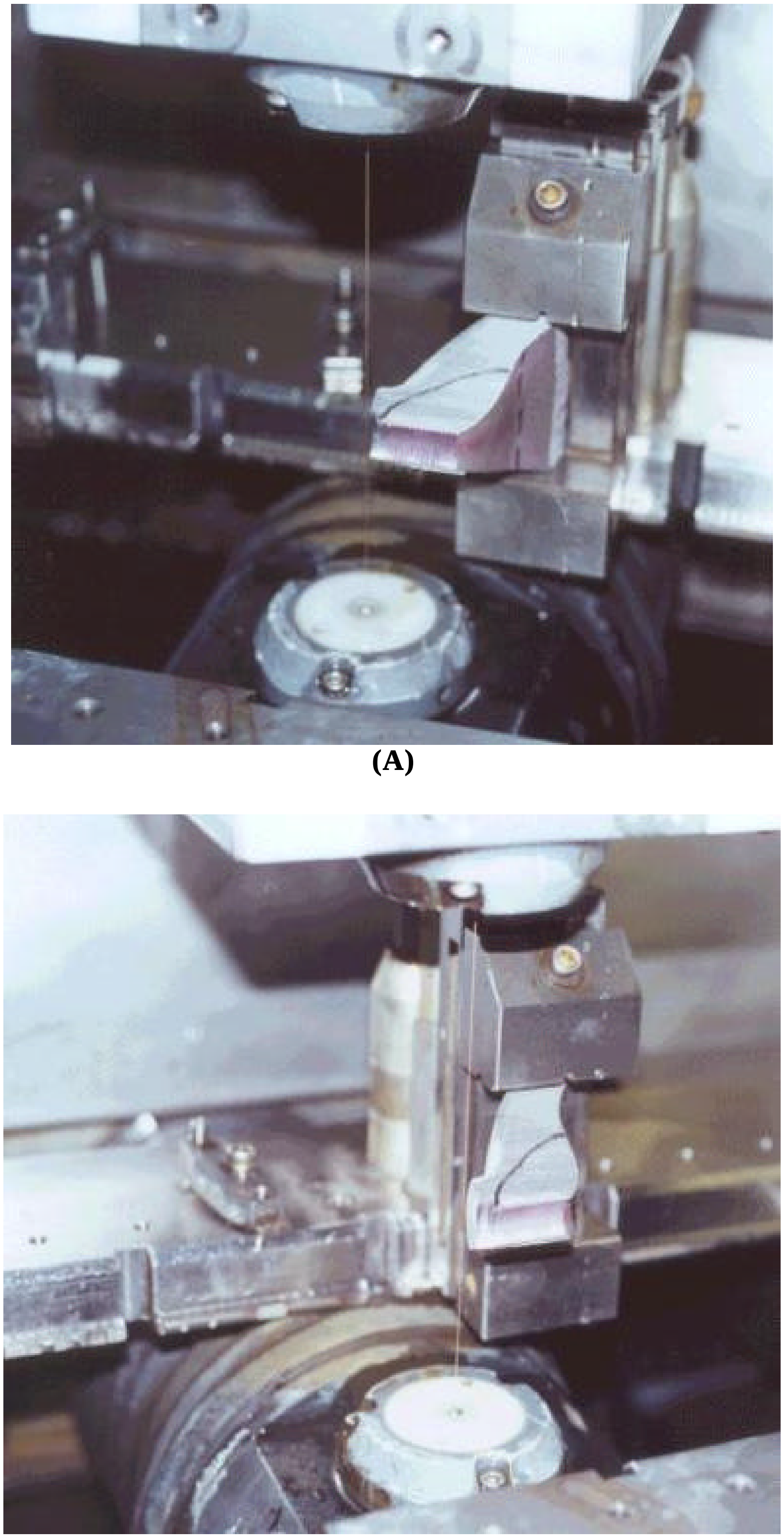
Figure 3-8. Sketch Showing how the EDM Wafers were Taken



(A)

(B)

Figure 3-9. Photographs of Cylinder Sample T94863 and P4205 Showing a) the OD Surface and b) the ID Surface



(A)

(B)

Figure 3-10. EDM Set-Up for Cylinder Sample

4.0 RESULTS

4.1 Detectability

The NDE data were analyzed by making two key assumptions:

1. Any flaw that was detected by 7 or more inspectors was real; and

2. Any flaw that was detected by 3 or fewer inspectors was not real.

An inspector was said to have detected the flaw if the angular position he or she recorded for the flaw was within a pre-determined angular range of the mean angular position of the flaw.

Results are summarized in Tables 4-1 and 4-2. To gauge the performance of each inspector it is easiest to look to the far right column, the Normalized or Reliability Index[2]. This number represents the number of false calls (whether a false negative or a false positive) that an inspector would make on 100 inspections. The results of the VE testing, a combination of visual and eddy current, are obtained by taking the best results of the VT and ET for each column, which reflects the proposed system of conducting one test and confirming with the other. As can be seen, VET ranked highest amongst all the inspections.

To compare techniques across skill levels and vice versa, two extra sets of data were produced, representing averages for the particular skill level or technique in question. Results were within expected parameters.

Table 4-1. Summary of Results Using an Angular Width of ±20°

	Real Flaws: Total 76					False Calls: Total 114					Normalized Index
	Called	%	Missed	%	Rank	False Positives	%	Rank	False Call Index	Rank	
VTX	74	97.37%	2	2.63%	2	8	7.02%	3	10	3	**5.26**
ETX	74	97.37%	2	2.63%	2	27	23.68%	7	29	7	**15.26**
UTX	73	96.05%	3	3.95%	3	14	12.28%	4	17	5	**8.95**
VTS	64	84.21%	12	15.79%	5	4	3.51%	1	16	4	**8.42**
ETS	73	96.05%	3	3.95%	3	5	4.39%	2	8	2	**4.21**
UTS	58	76.32%	18	23.68%	7	22	19.30%	5	40	8	**21.05**
VTT	75	98.68%	1	1.32%	1	24	21.05%	6	25	6	**13.16**
ETT	71	93.42%	5	6.58%	4	43	37.72%	9	48	10	**25.26**
UTT	62	81.58%	14	18.42%	6	29	25.44%	8	43	9	**22.63**
VETX	74	97.37%	2	2.63%	2	8	7.02%	3	10	3	**5.26**
VETS	73	96.05%	3	3.95%	3	4	3.51%	1	7	1	**3.68**
VETT	75	98.68%	1	1.32%	1	24	21.05%	6	25	6	**13.16**
Averages:											
By Inspection Technique											
VT	71.00	93.42%	5.00	6.58%	3.00	12.00	10.53%	1.00	17.00	2.00	**8.95**
ET	72.67	95.61%	3.33	4.39%	2.00	25.00	21.93%	3.00	28.33	3.00	**14.91**
UT	64.33	84.65%	11.67	15.35%	4.00	21.67	19.01%	2.00	33.33	4.00	**17.54**
VET	74.00	97.37%	2.00	2.63%	1.00	12.00	10.53%	1.00	14.00	1.00	**7.37**
By Skill Level											
X	73.67	96.93%	2.33	3.07%	1.00	16.33	14.33%	2.00	18.67	1.00	**9.82**
S	65.00	85.53%	11.00	14.47%	3.00	10.33	9.06%	1.00	21.33	2.00	**11.23**
T	69.33	91.23%	6.67	8.77%	2.00	32.00	28.07%	3.00	38.67	3.00	**20.35**

---

[2] A glossary of terms is presented in Appendix A.

Table 4-2. Summary of Results Using an Angular Width of ±30°

	Real Flaws: Total 77					False Calls: Total 95					Normalized Index
	Called	%	Missed	%	Rank	False Positives	%	Rank	False Call Index	Rank	
VTX	75	97.40%	2	2.60%	2	8	8.42%	3	10	4	**5.81**
ETX	76	98.70%	1	1.30%	1	23	24.21%	6	24	7	**13.95**
UTX	75	97.40%	2	2.60%	2	11	11.58%	4	13	5	**7.56**
VTS	65	84.42%	12	15.58%	5	3	3.16%	1	15	6	**8.72**
ETS	74	96.10%	3	3.90%	3	5	5.26%	2	8	2	**4.65**
UTS	63	81.82%	14	18.18%	6	17	17.89%	5	31	9	**18.02**
VTT	76	98.70%	1	1.30%	1	24	25.26%	7	25	8	**14.53**
ETT	73	94.81%	4	5.19%	4	42	44.21%	9	46	11	**26.74**
UTT	64	83.12%	13	16.88%	7	25	26.32%	8	38	10	**22.09**
VETX	76	98.70%	1	1.30%	1	8	8.42%	3	9	3	**5.23**
VETS	74	96.10%	3	3.90%	3	3	3.16%	1	6	1	**3.49**
VETT	76	98.70%	1	1.30%	1	24	25.26%	7	25	8	**14.53**
Averages:											
By Inspection Technique											
VT	72.00	93.51%	4.00	5.26%	3.00	11.67	12.28%	1.00	15.67	2.00	**9.11**
ET	74.33	96.54%	1.67	2.19%	2.00	23.33	24.56%	3.00	25.00	3.00	**14.53**
UT	67.33	87.45%	8.67	11.40%	4.00	17.67	18.60%	2.00	26.33	4.00	**15.31**
VET	75.33	97.84%	0.67	0.88%	1.00	11.67	12.28%	1.00	12.33	1.00	**7.17**
By Skill Level											
X	75.33	97.84%	0.67	0.87%	1.00	14.00	14.74%	2.00	14.67	1.00	**8.53**
S	67.33	87.45%	8.67	11.26%	3.00	8.33	8.77%	1.00	17.00	2.00	**9.88**
T	71.00	92.21%	5.00	6.49%	2.00	30.33	31.93%	3.00	35.33	3.00	**20.54**

Discussion

In general terms, the eddy current system is designed to find any and all anomalies in a cylinder: this is borne out by the data, which show that while ET has the lowest incidence of missed real flaws it also had the second highest incidence of false positives. In contrast, VT had the second lowest incidence of missed flaws and the lowest incidence of false positives, while UT missed the most real flaws and had the highest rate of false positives. This is likely due to the fact that the UT system requires trained inspector to understand the entry angle of the shear waves and usage of the UT system in order to determine the size and position of a flaw. The actual position of a flaw must be calculated in accordance skip distances of shear waves for the curved surface (shoulder of the cylinder).

The detectability study suggests that the best option in terms of an approach to inspecting these cylinders is in the “VE” technique, in which eddy current testing and visual testing are combined in a single procedure. It was observed that ET followed by VT produces the desirability effect of high detection rates with both few false positives and few false negatives.

4.2 Flaw Sizing (Comparison with Etching Procedure)

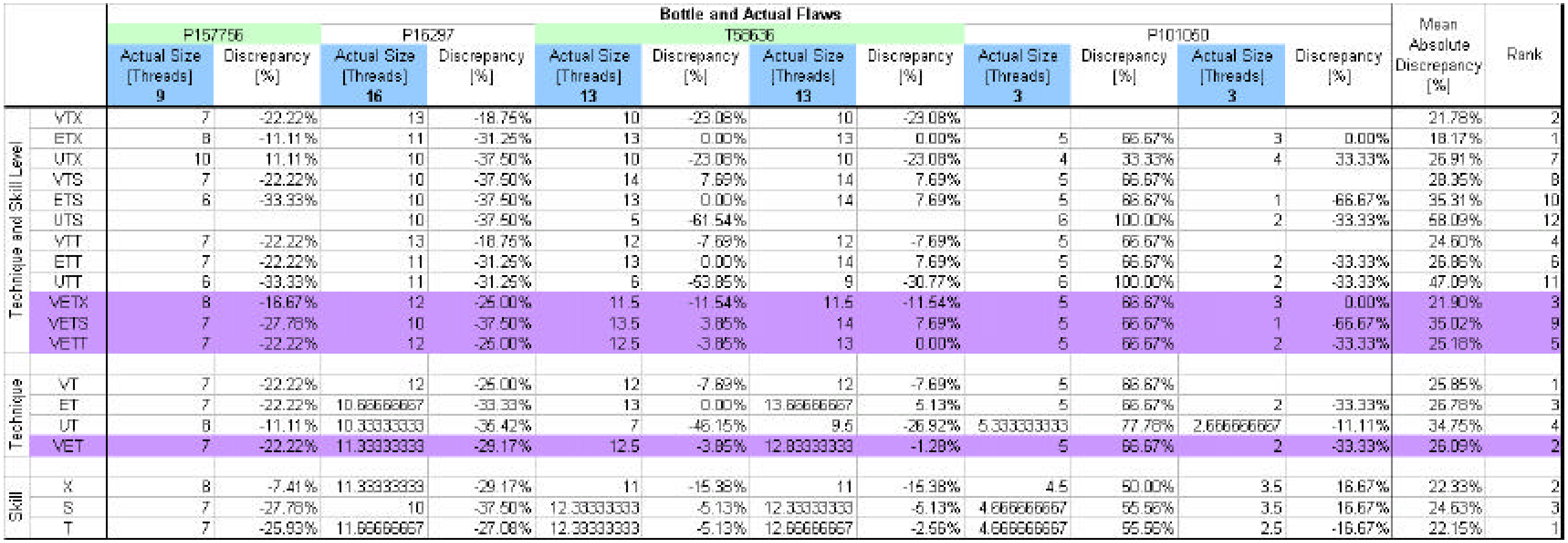
The cylinders chosen for analysis are summarized in Table 4-3. Crack positions and sizes are reported as the mean values indicated for those NDE techniques that detected the crack. It is difficult to compare flaw positions reported by NDE with those found during the destructive analysis since the destructive analysis effectively assumes the same flaw positions. It can be said, however, that in neither the etching procedure nor in the machining procedure conducted by Bryant-lee was a flaw not found during the destructive analysis. These results are to be compared with the results of the etching analysis conducted by NTIAC to provide quantifiable information about the size of the flaws in this subset of cylinders.

Table 4-3. Cylinders Chosen for the Flaw Sizing Analysis

Bottle Number	Bottle Type	Mean Crack Positions [Degrees from Reference Point]	Mean Crack Sizes [Number of Crossed Threads]
P15756	SCUBA	315	7
P16297	SCUBA	135	12
P101050	SCUBA	15 180	5 2
T58636	SCBA	44 209	10 11

Figures 4-1 to 4-12 graphically display the results of this analysis as summarized in Table 4-4, both in terms of the actual inspections and in terms of skill level and technique. Each inspection technique is a “spoke” on the wheel: a red data point indicates a negative size difference (i.e., the inspection reported a smaller flaw than what was measured destructively), a green data point is a positive size difference (the inspection reported a larger flaw than the measured). Each circle out from the center of the figure represents a 10% difference from the actual measured flaw size: a red data point on the third circle from the center on the VTX spoke thus indicates that for this particular bottle the expert visual inspection returned a flaw size that was 30% smaller than the actual measured flaw size.

Table 4-4 introduces the Mean Absolute Discrepancy (MAD) as a measure of the technique’s success in sizing a crack. The MAD score is the absolute value of the average discrepancy between the reported flaw size and the actual flaw size: a MAD of 25% means that on average the technique will report a flaw that is 25% off (larger or smaller than) of the actual size as determined via destructive analysis.



		Bottle and Actual Flaws												Mean Absolute Discrepancy [%]	Rank
		P157756		P16297		T58636				P101050					
		Actual Size [Threads] 9	Discrepancy [%]	Actual Size [Threads] 16	Discrepancy [%]	Actual Size [Threads] 13	Discrepancy [%]	Actual Size [Threads] 13	Discrepancy [%]	Actual Size [Threads] 3	Discrepancy [%]	Actual Size [Threads] 3	Discrepancy [%]		
Technique and Skill Level	VTX	7	-22.22%	13	-18.75%	10	-23.08%	10	-23.08%					21.78%	2
	ETX	8	-11.11%	11	-31.25%	13	0.00%	13	0.00%	5	66.67%	3	0.00%	18.17%	1
	UTX	10	11.11%	10	-37.50%	10	-23.08%	10	-23.08%	4	33.33%	4	33.33%	26.91%	7
	VTS	7	-22.22%	10	-37.50%	14	7.69%	14	7.69%	5	66.67%			28.35%	8
	ETS	6	-33.33%	10	-37.50%	13	0.00%	14	7.69%	5	66.67%	1	-66.67%	35.31%	10
	UTS			10	-37.50%	5	-61.54%			6	100.00%	2	-33.33%	58.09%	12
	VTT	7	-22.22%	13	-18.75%	12	-7.69%	12	-7.69%	5	66.67%			24.60%	4
	ETT	7	-22.22%	11	-31.25%	13	0.00%	14	7.69%	5	66.67%	2	-33.33%	26.85%	6
	UTT	6	-33.33%	11	-31.25%	6	-53.85%	9	-30.77%	6	100.00%	2	-33.33%	47.09%	11
	VETX	8	-16.67%	12	-25.00%	11.5	-11.54%	11.5	-11.54%	5	66.67%	3	0.00%	21.90%	3
	VETS	7	-27.78%	10	-37.50%	13.5	3.85%	14	7.69%	5	66.67%	1	-66.67%	35.02%	9
	VETT	7	-22.22%	12	-25.00%	12.5	-3.85%	13	0.00%	5	66.67%	2	-33.33%	25.18%	5
Technique	VT	7	-22.22%	12	-25.00%	12	-7.69%	12	-7.69%	5	66.67%			25.85%	1
	ET	7	-22.22%	10.66666667	-33.33%	13	0.00%	13.66666667	5.13%	5	66.67%	2	-33.33%	26.78%	3
	UT	8	-11.11%	10.33333333	-35.42%	7	-46.15%	9.5	-26.92%	5.333333333	77.78%	2.666666667	-11.11%	34.75%	4
	VET	7	-22.22%	11.33333333	-29.17%	12.5	-3.85%	12.83333333	-1.28%	5	66.67%	2	-33.33%	26.09%	2
Skill	X	8	-7.41%	11.33333333	-29.17%	11	-15.38%	11	-15.38%	4.5	50.00%	3.5	16.67%	22.33%	2
	S	7	-27.78%	10	-37.50%	12.33333333	-5.13%	12.33333333	-5.13%	4.666666667	55.56%	3.5	16.67%	24.63%	3
	T	7	-25.93%	11.66666667	-27.08%	12.33333333	-5.13%	12.66666667	-2.56%	4.666666667	55.56%	2.5	-16.67%	22.15%	1

Table 4-4. Summary of the Flaw Sizing Study

In 83% of the samples, both the average and median absolute error in thread count was one or two threads regardless of the total crack length. Similarly, 83% of average flaw lengths were undersized.

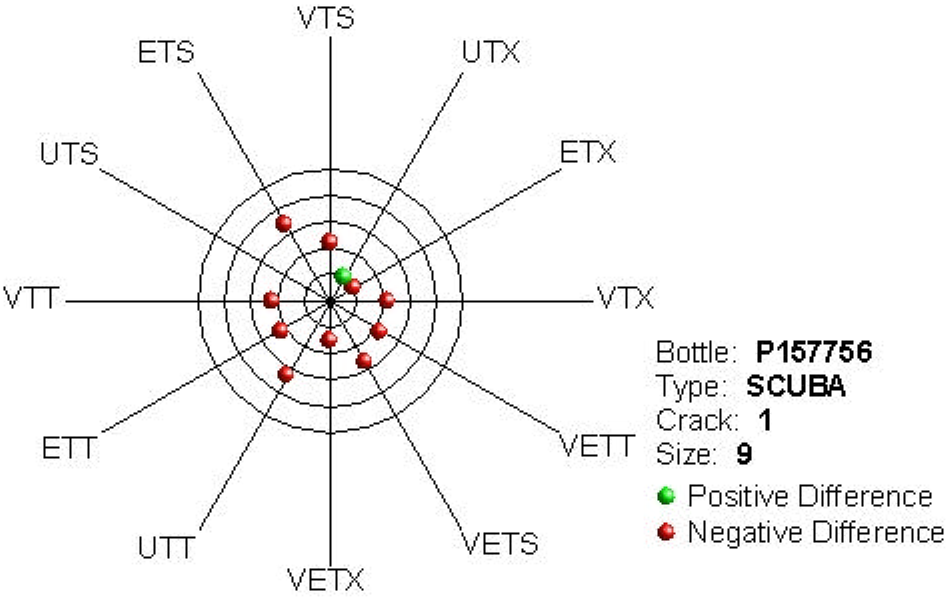


Figure 4-1. Reported Flaw Size for the Crack in SCUBA Bottle P157756 for Each Inspection

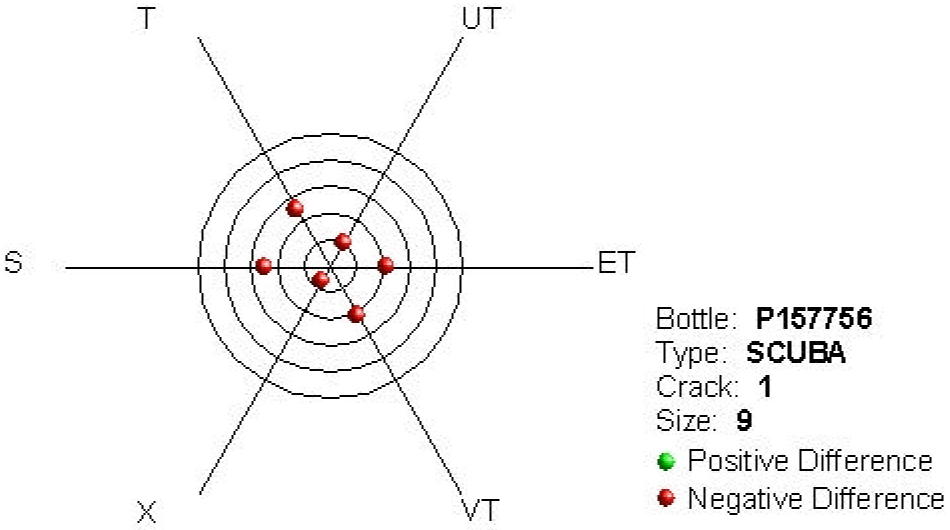


Figure 4-2. Average Flaw Sizes for the Crack in P157756 for Technique and Skill Level

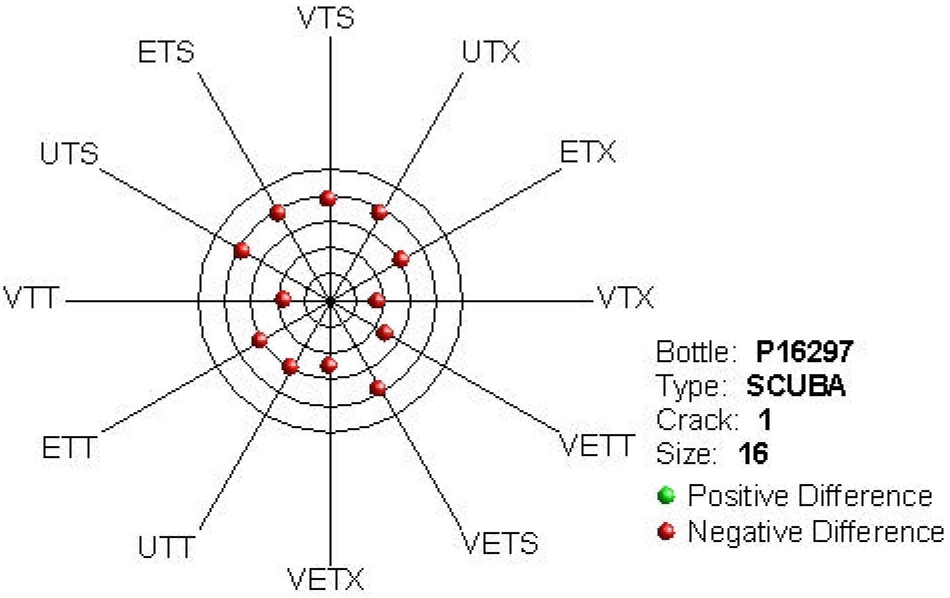


Figure 4-3. Reported Flaw Size for the Crack in SCUBA Bottle P16297 for Each Inspection

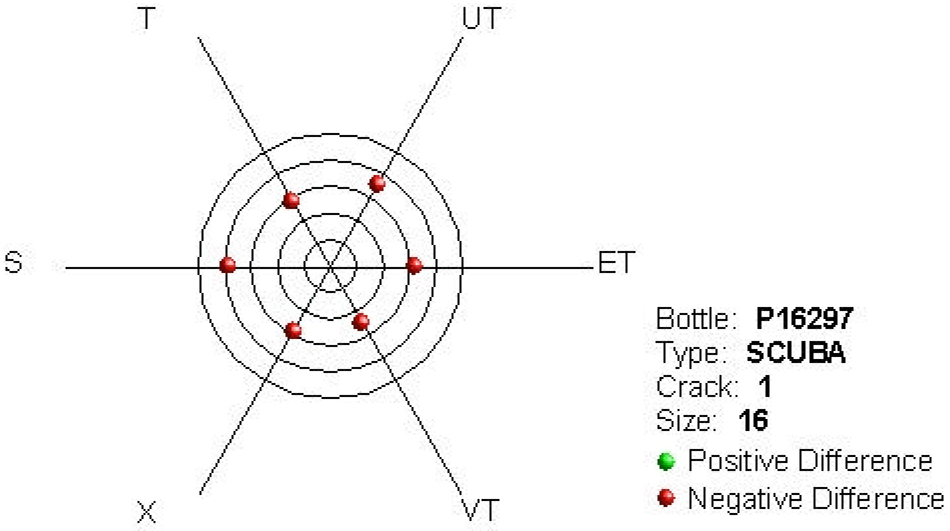


Figure 4-4. Average Flaw Sizes for the Crack in P16297 for Technique and Skill Level

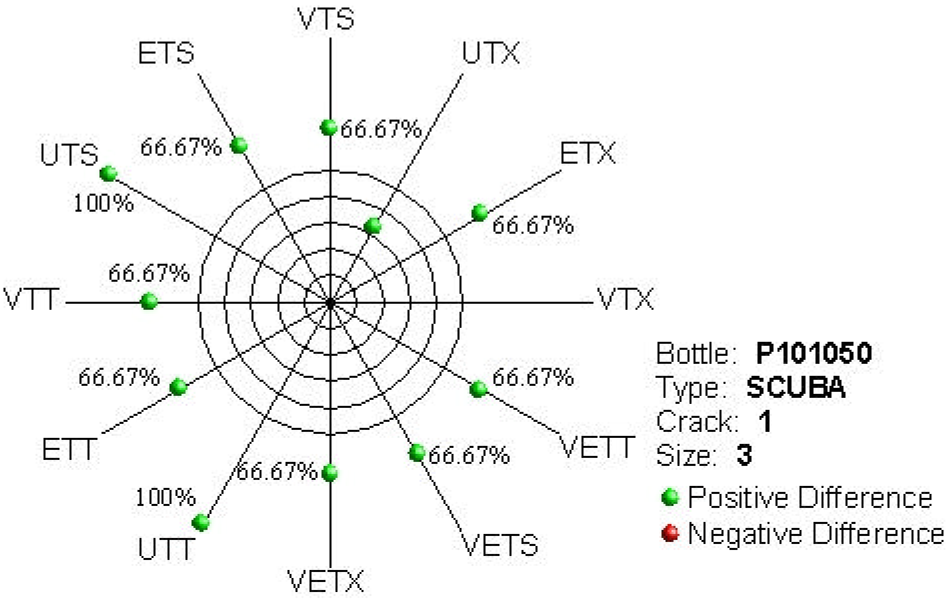


Figure 4-5. Reported Flaw Size for the First Crack in SCUBA Bottle P101050 for Each Inspection

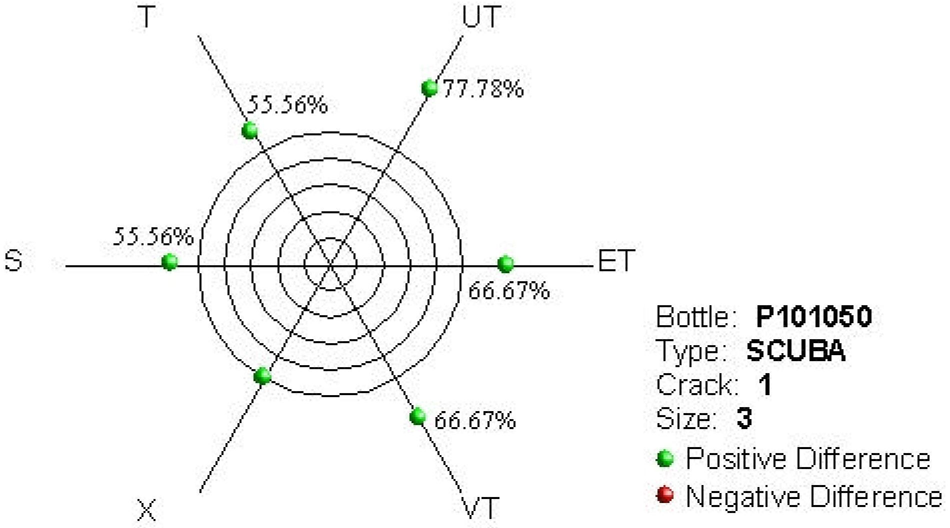


Figure 4-6. Average Flaw Sizes for the First Crack in P101050 for Technique and Skill Level

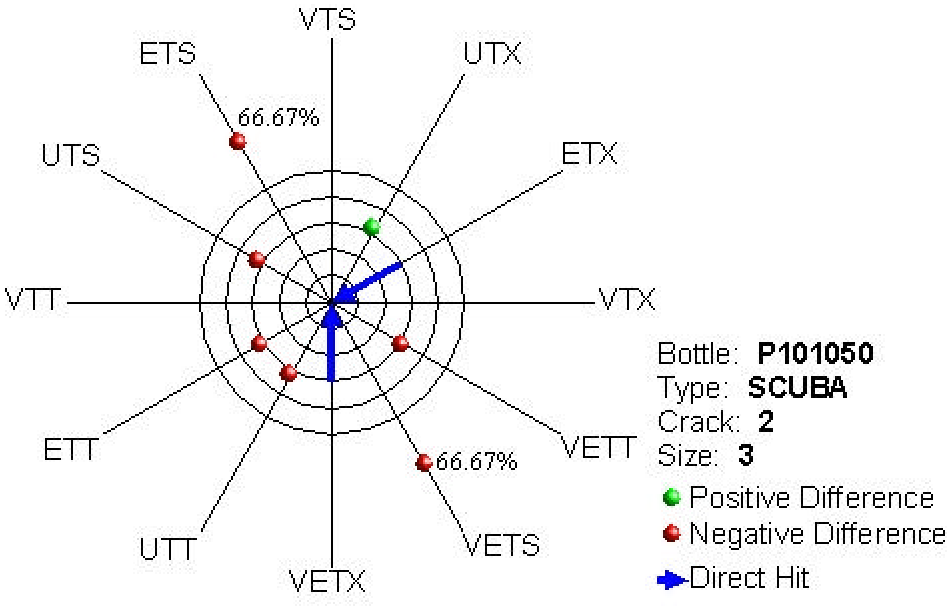


Figure 4-7. Reported Flaw Size for the Second Crack in SCUBA Bottle P101050 For Each Inspection

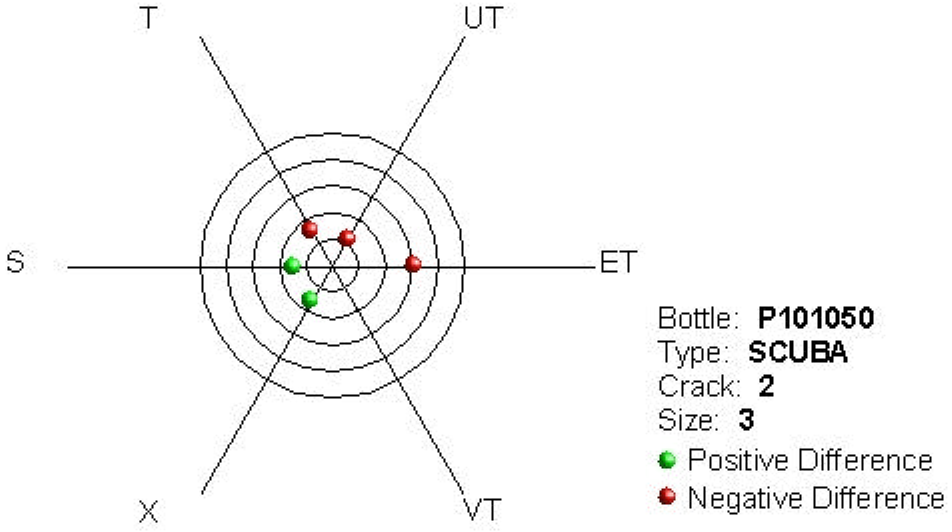


Figure 4-8. Average Flaw Sizes for the Second Crack in P101050 for Technique and Skill Level

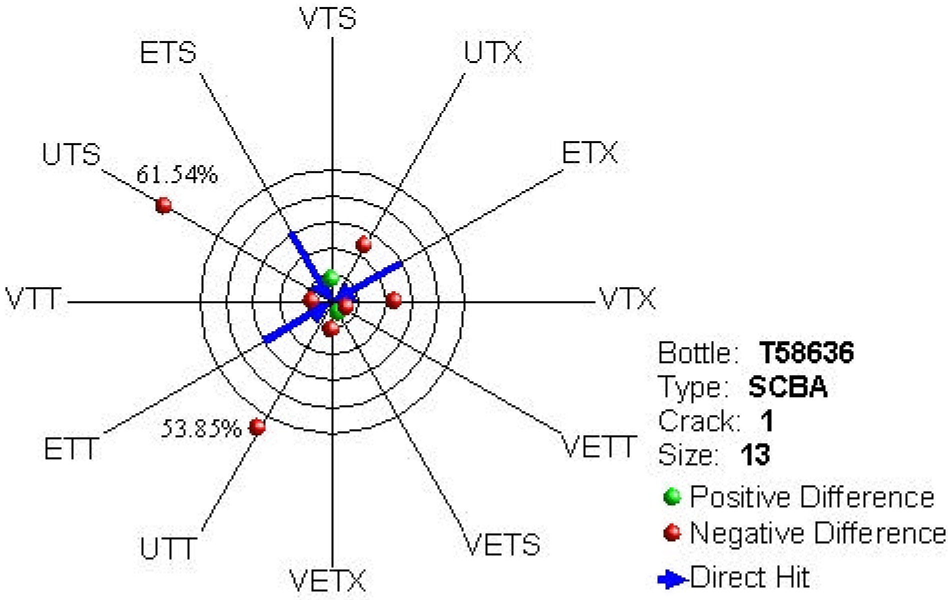


Figure 4-9. Reported Flaw Size for the First Crack in SCBA Bottle T58636 for Each Inspection

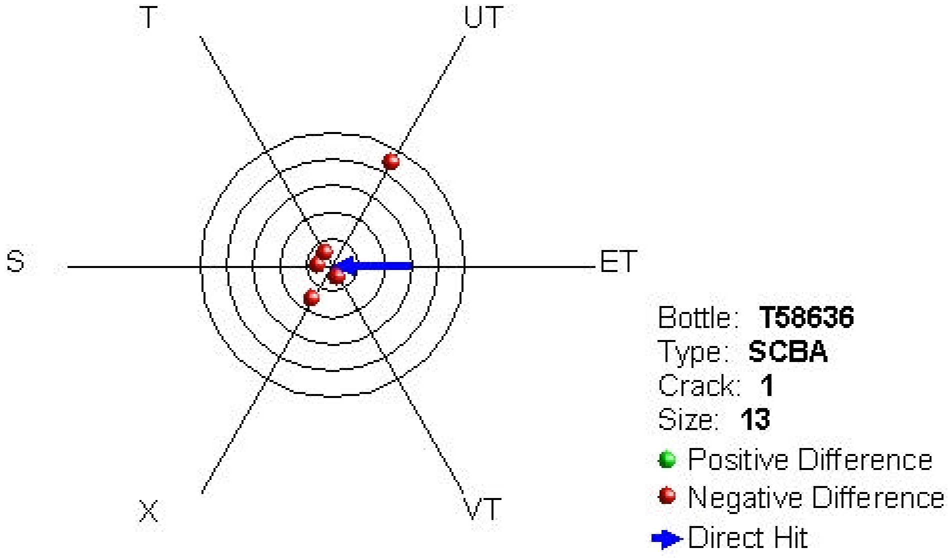


Figure 4-10. Average Flaw Sizes for the First Crack in T58636 for Technique and Skill Level

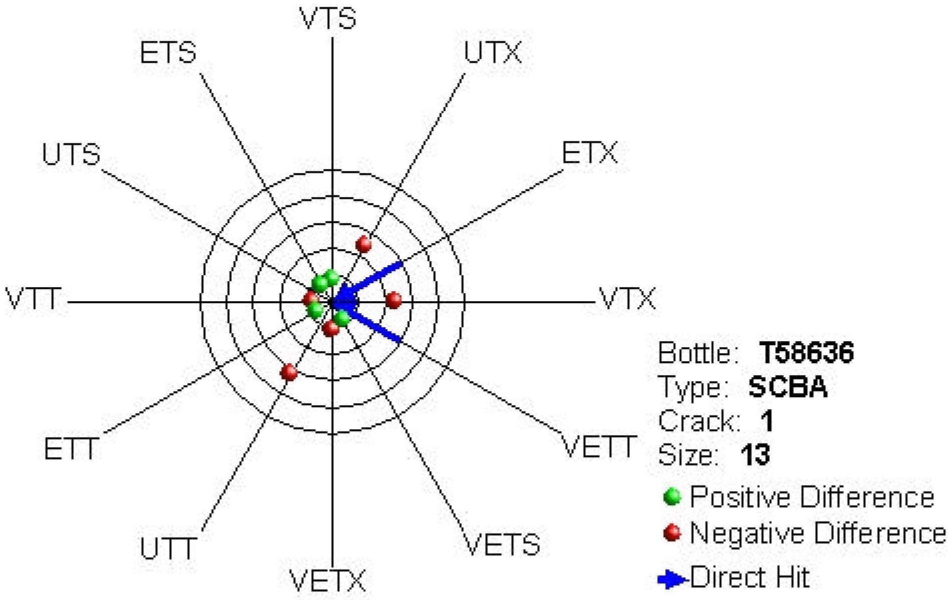


Figure 4-11. Reported Flaw Size for the Second Crack in SCBA Bottle T58636 for Each Inspection

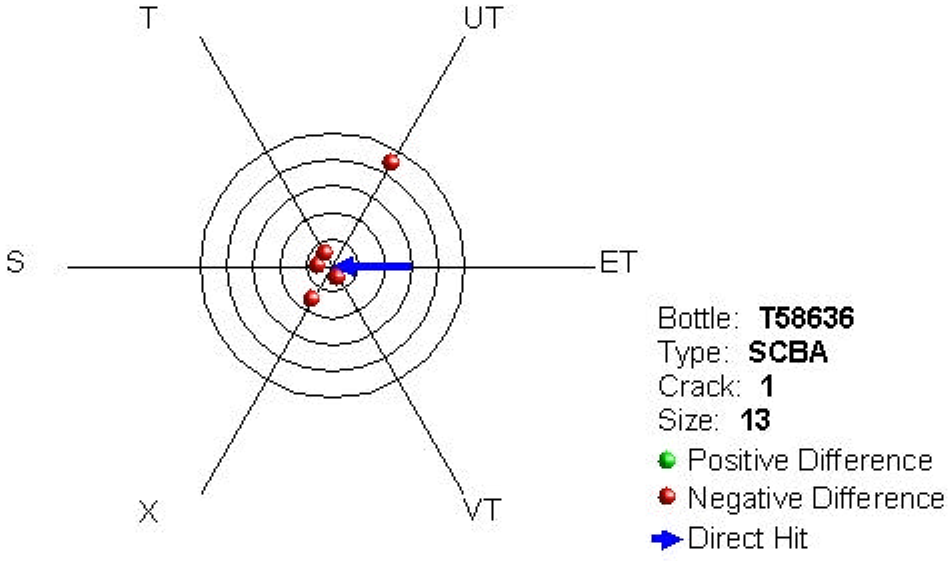


Figure 4-12. Average Flaw Sizes for the Second Crack in T58636 for Technique and Skill Level

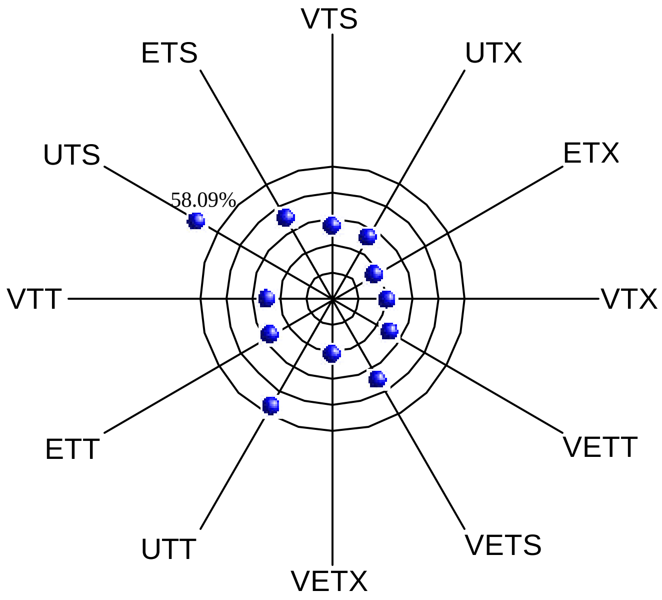


Figure 4-13. Mean Absolute Discrepancy (MAD) Results by Inspection

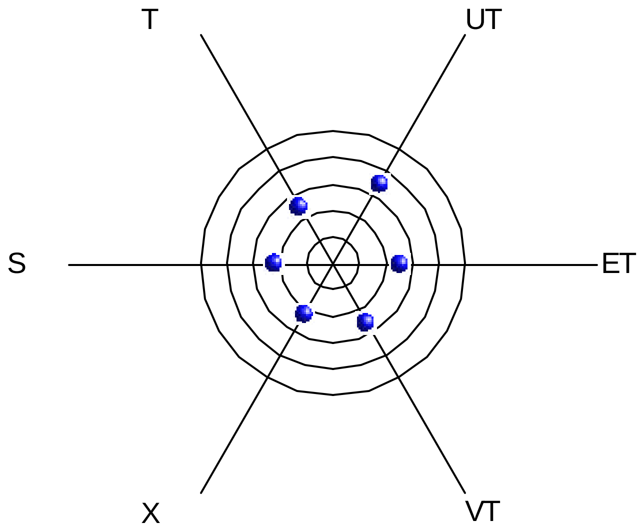


Figure 4-14. Mean Absolute Discrepancy (MAD) Results by Technique and Skill Level

Discussion

Little overall difference in terms of sizing ability can be seen between the inspection techniques: ET, VT, and VET all have a MAD of roughly 25%. Of the three, ET appears to hold the most promise in terms of accurate sizing, but again this difference is very slight: on the order of 1%.

Ultrasonic testing (UT) finished last with a MAD of approximately 35%. As UT does not use the same manner of counting threads in returning a flaw size, it is worthwhile to explore the procedure used to obtain a thread count from UT data. For each flaw call, an average thread size was determined by taking the mean of the flaw sizes reported by all the visual and eddy current testers. Similarly, for the same flaw, the mean of the ultrasonic signal durations was taken. In both cases, the standard deviation from the mean was also calculated.

Based on this data, a subset was taken that represented the highest level of correlation for the mean flaw size. Quantitatively, this was done by using only those data points for which the uncertainty in the mean flaw size was 20% or lower. Using this subset, a plot of the UT duration call versus the ET/VT thread call was produced (Figure 4-15).

Based upon this plot, a trendline was produced that most closely approximated the data. This line is then used as the translation between duration and thread size, and was found to be of the form $y = -0.0006x^3 + 0.0274x^2 - 0.0041x + 2.7704$, in which x is the flaw size in signal duration reported by the UT inspector, and y is the corresponding flaw size in threads for the same flaw. Since the other methods do not count fractions of threads, in the final stage of calibration y is taken as an integer.

The UTX inspection (ultrasonic testing conducted by the system vendor) did not employ a duration call but rather the inspector opted to record a flaw as small, medium, or large. A quick tally of the detected real flaws indicated the following size ranges:

- A small flaw was between 2 to 5 threads.

- A medium-sized flaw was between 6-9 threads.

- A large flaw was over 10 threads.

The convention used for the sizing study was to label a small flaw as 4 threads, a medium as 7, and a large as 10 threads. This introduces a certain level of discrepancy into the analysis, but these numbers were chosen both because they represent the middle of their respective size ranges and because they agreed most closely with experimental data.

It is noteworthy that the accuracy in determining a flaw size does not directly correlate with the axial length (i.e., number of crossed threads) of a flaw: in the above figures, a blue arrow signifies a “direct hit,” in which the technique correctly gauged the size of the flaw. Some of the cracks, having an axial length of only three threads, were nonetheless correctly sized by more than one inspector. This is likely a reflection of the width of the crack, i.e., how tightly the crack faces met.

In inspections of SCUBA/SCBA cylinders it is common practice to quote an aspect ratio when describing SLC features. In Table 4-5 the aspect ratios of some of the cracks are calculated, where the aspect ratio is taken as width:length. As can be seen from Table 4-5, the cracks were quite varied in terms of their geometries. In addition, one particular bottle broken open to demonstrate the etching process was later found to have an aspect ratio of 3.00:1.

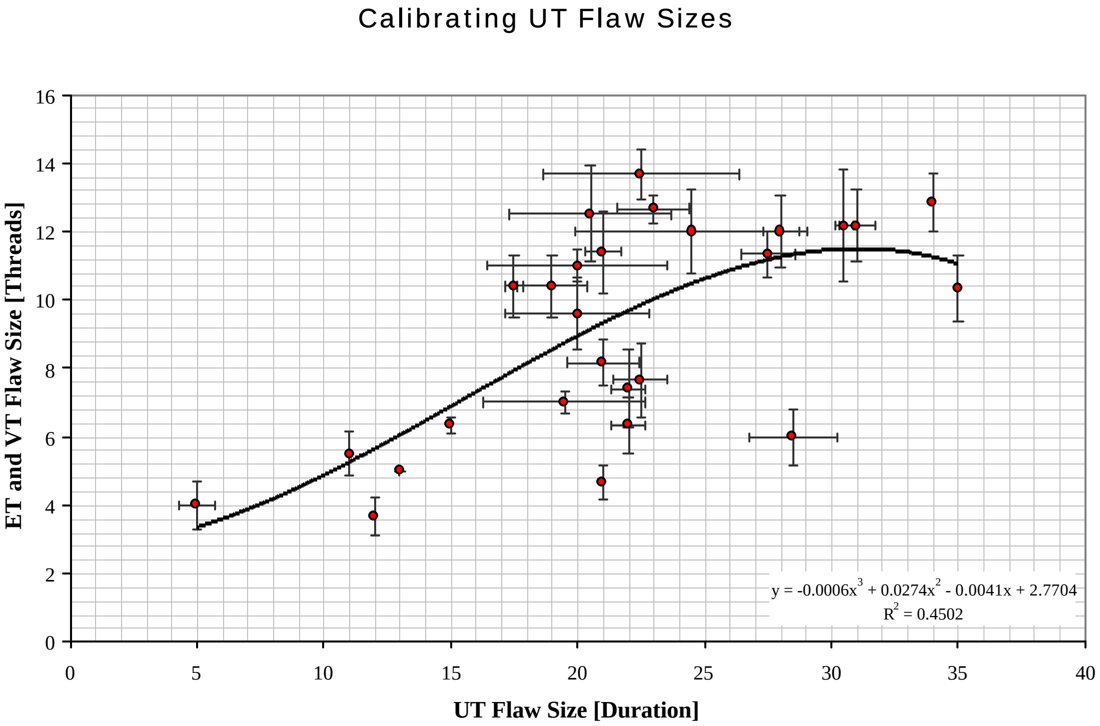


Figure 4-15. Production of the UT Calibration Curve

Table 4-5. Example of the Aspect Ratios Seen During the Flaw Sizing Study. Max 1.75:1, Min. 0.65:1

Bottle	Crack #	Face	Surface Area (cm^2)	Thread Count	Length (cm)	Width (cm)	**Aspect Ratio (W:L)**
P15756	1	A	5.567	8	1.40	2.45	**1.75:1**
		B	5.230	9	1.55	2.50	**1.61:1**
P16297	1	A	13.423	16	3.30	3.70	**1.12:1**
		B	12.377	16	3.30	3.80	**1.15:1**
T58636	1	A	7.544	13	2.30	3.20	**1.39:1**
		B	8.040	13	2.20	3.20	**1.45:1**
P101050	1	A	0.554	3	0.60	0.60	**1.00:1**
		B	0.566	3	0.60	0.70	**1.17:1**
	2	A	0.554	6	1.15	0.80	**0.70:1**
		B	0.655	6	1.15	0.75	**0.65:1**

4.3 EDM Results

The results of Bryant-Lee Associates' EDM analysis of the two cylinders is presented in full in their attached report and is summarized here for convenience.

BLA's determination of crack depths is found in Tables 4-6 and 4-7 for T94863 and P4205, respectively. Columns six and seven define the crack position and crack depth. BLA observed that the actual crack tip sometimes exceeded the observed continuous crack at low magnifications. On some wafers, such as wafer A5, the continuous crack stopped at about 0.490-inch deep. However, this was not the crack tip. The crack tip was actually at 0.711-inch deep. Discontinuous, tight cracks extended beyond the continuous crack. Figure 4-16 shows the apparent crack tip at 0.490-inch deep and discontinuous cracking at 0.665-inch deep.

Table 4-6. Wafer Location and Crack Depth for T94863

Wafer I.D.	Wafer Distance from Reference Surface(1) (inches)	A-EDM'ed Wafer Thickness (inches)	As-Polished Thickness (inches)	Metal Removed (inches)	As-Polished Distance from Reference Surface (inches)	Distance from Hole Edge(1) (inches)
A1	0.125	0.84	0.60	0.24	0.149	0.073
A2	0.225	0.85	0.68	0.17	0.242	0.368
A3	0.325	0.85	0.80	0.05	0.330	0.451
A4	0.425	0.84	0.78	0.06	0.431	0.539
A5	0.525	0.85	0.74	0.11	0.536	0.711
A6	0.625	0.84	0.82	0.02	0.627	0.868
A7	0.725	0.85	0.81	0.04	0.729	0.950
A8	0.825	0.85	0.82	0.03	0.828	0.998
A9	0.925	0.84	0.78	0.06	0.931	0.957
A10	1.025	0.85	0.76	0.09	1.034	0.869
A11*	1.125	0.85	0.77	0.08	1.133	0.965
A12*	1.225	0.85	0.75	0.10	1.235	0.803
A13*	1.325	0.85	0.74	0.11	1.336	0.465
A14*	1.425	0.86	0.76	0.10	1.435	0.330

(1) See Figure 3-8.
* Wafers Beyond the Borehole.

Crack Position

Crack Depth

Table 4-7. Wafer Location and Crack Depth for P4205

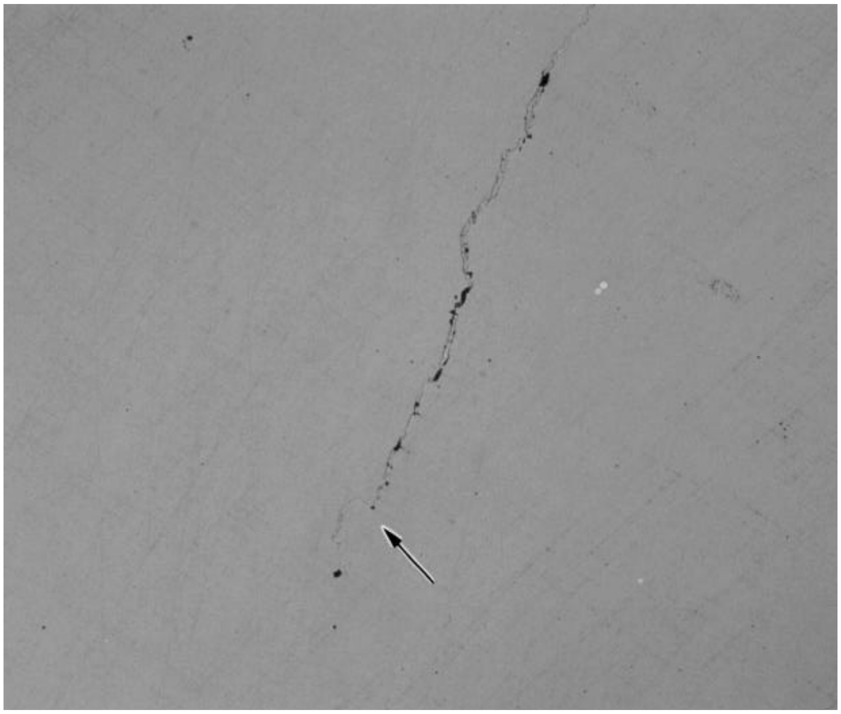
Wafer I.D.	Wafer Distance from Reference Surface(1) (inches)	A-EDM'ed Wafer Thickness (inches)	As-Polished Thickness (inches)	Metal Removed (inches)	As-Polished Distance from Reference Surface (inches)	Distance from Hole Edge(1) (inches)
B1	0.115	0.85	0.71	0.14	0.129	No Crack
B2	0.215	0.86	0.80	0.06	0.221	No Crack
B3	0.315	0.85	0.77	0.08	0.323	No Crack
B4	0.415	0.86	0.65	0.21	0.436	No Crack
B5	0.515	0.85	0.64	0.21	0.536	No Crack
B6	0.615	0.85	0.77	0.08	0.623	0.222
B7	0.715	0.85	0.79	0.06	0.721	0.308
B8	0.815	0.85	0.79	0.06	0.821	0.379
B9	0.915	0.84	0.78	0.06	0.921	0.544
B10	1.015	0.85	0.80	0.05	1.020	0.666
B11	1.115	0.85	0.79	0.06	1.121	0.742
B12	1.215	0.85	0.78	0.07	1.222	0.776
B13	1.315	0.84	0.79	0.05	1.320	0.935
B14	1.415	0.85	0.77	0.08	1.423	0.946
B15	1.515	0.85	0.77	0.08	1.523	0.994
B16*	1.615	0.85	0.77	0.08	1.623	0.928
B17*	1.715	0.85	0.76	0.09	1.724	0.736
B18*	1.815	0.85	0.80	0.05	1.820	0.548
B19*	1.915	0.85	0.79	0.06	1.921	0.308
B20*	2.015	0.86	0.79	0.07	2.022	No Crack

(1) See Figure 3-8.

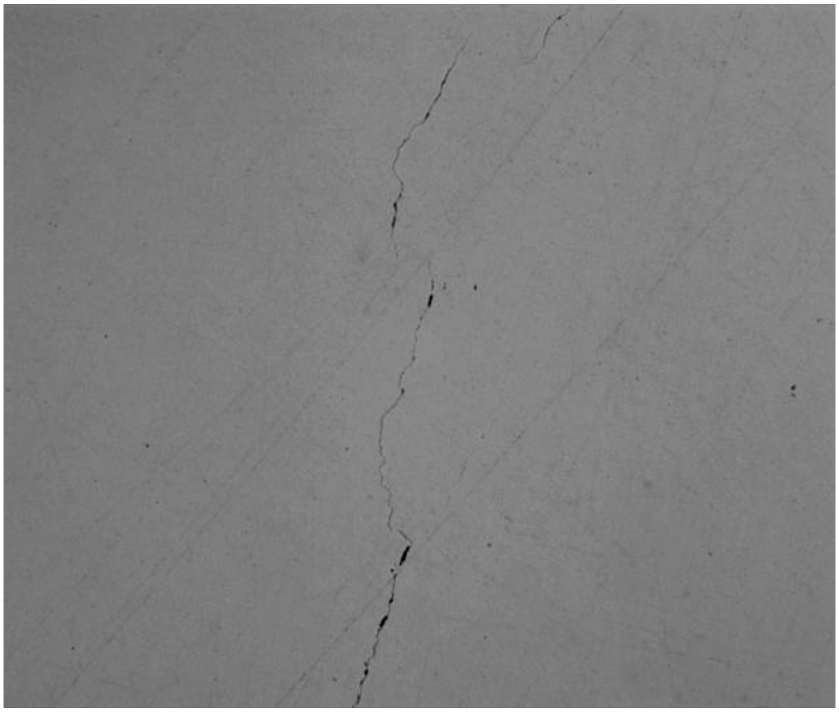
* Wafers Beyond the Borehole.

↑ Crack Position

↑ Crack Depth



a) 0.490-inch 50x



b) 0.665-inch 50x

Figure 4-16. Photomicrographs of the a) Apparent Crack Tip at 0.490-inch from Bore ID and b) Crack Tip Region at 0.665-inch from Bore ID for Wafer A5.

From Table 4-6 it can be seen that the crack in SCBA bottle T94863 begins 0.149” from the reference surface, which actually puts the crack above the first thread of the bottle. An accurate assessment of the flaw size would then put the crack as being 14 threads or larger: Table 4-8 lists the actual flaw sizes as reported by the inspectors. Similar analysis for P4205 gives a thread count of approximately 13 threads, and the actual results from the NDE inspections are presented in Table 4-9.

Table 4-8. Reported Flaw Size for the Flaw in SCBA Bottle T94863, for Which the Actual Flaw Size is Approximately 14 Threads

Technique	Reported Flaw Size [Threads]
VTX	10
ETX	2
UTX	10
VTS	10
ETS	13
UTS	6
VTT	13
ETT	13
UTT	5

Table 4-9. Reported Flaw Size for the Flaw in SCBA Bottle T94863, for Which the Actual Flaw Size is Approximately 14 Threads

Technique	Reported Flaw Size [Threads]
VTX	2
ETX	6
UTX	10
VTS	10
ETS	4
UTS	9
VTT	13
ETT	6
UTT	Not Detected

As can be seen, there exists a wide variance between the reported flaw sizes from inspector to inspector. This can perhaps be attributed to the geometry of the flaw (i.e. how “tight” the crack faces are, etc.), and also to operator error due to fatigue, and should perhaps be considered as a future research goal to study in greater detail.

5.0 CONCLUSIONS AND RECOMMENDATIONS

NTIAC’s investigation into the nondestructive inspection of aluminum gas cylinders followed this procedure:

1. Fifty-one (51) cylinders were nondestructively inspected via visual, eddy current, and ultrasonic testing for sustained load cracking (SLC). Each technique was employed by three inspectors of three different skill levels: expert, NDE specialist, and technician. The inspections were tallied and recorded by the test coordinator.
2. A subset of these fifty-one cylinders were set aside for destructive analysis. Two methods of destructive analysis were used: etching, in which the cracks are forced open mechanically after having been exposed to a dyed penetrant; and EDM analysis, in which successive layers of the cylinder are machined away to arrive at the actual depth of the crack.
3. The results of the NDE phase of the project were analyzed for two factors: the ability to detect cracks and the ability to size cracks. Each inspector was ranked according to these measurements by their Reliability Index (RI), the number of false calls an inspector would make per 100 flaws; and by their Mean Absolute Discrepancy (MAD), which measures the absolute discrepancy between an inspector's flaw size estimation and the actual flaw size as determined by destructive analysis.

The results of this investigation can be summarized as follows:

- In most cases, an inspection technique exhibited either a good RI or MAD but not both, reflective of the design and nature of the inspection technique. The best inspection technique therefore is to combine two inspection techniques into a single procedure as explained below.
- The level of skill of an inspector is a crucial factor in terms of both the RI and MAD of the inspection, but the importance of skill varied between the three techniques. The ultrasonic equipment was seen to have the highest degree of dependence upon skill level, in that only the system vendor's representative was able to perform at an adequate level.
- Of the two measures of an inspector's performance, the Reliability Index (RI) is the more important. In practice an inspector is not interested in the extent of a flaw other than determining its existence: a bottle with a flaw found to be 2 threads in size is automatically rejected without further investigation.

Based upon the results compiled by NTIAC as part of this project, our recommendation is to adopt the combined visual and eddy current testing (designated VET) as the standard for inspection of aluminum gas cylinders. Figure 5-1 provides a graphical summary of the overall results of the project; Figure 5-2 illustrates the averages across technique and skill level respectively.

Combination VET inspection is conducted by first inspecting the cylinder with the eddy current system, then by visual testing. This order is important since on average eddy current testing has the lowest rate of false negatives, about half the rate of visual testing. At the same time, however, eddy current testing also has a much higher rate of false positives (i.e., detecting a non-existent flaw), which is improved dramatically by confirming with visual testing. The procedure is then: if eddy current testing indicates a flaw, the user then removes the eddy current system from the cylinder and attempts to confirm the flaw visually. If the flaw is not visually detected, it is deemed to be a false positive on the part of the eddy current system and so the cylinder is not rejected based on this false positive.

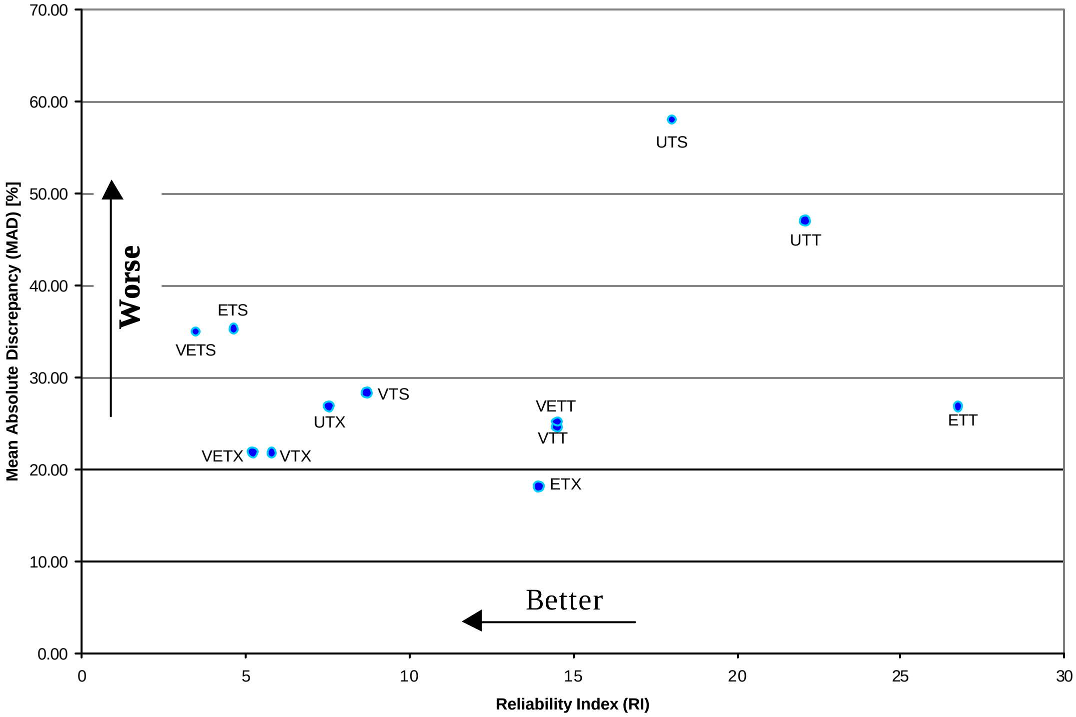


Figure 5-1. Summary of the Test Results, Plotting Each Technique's Mean Absolute Discrepancy (MAD) Against it's Reliability Index (RI). An ideal technique would be plotted at the origin (0,0), representing a technique that has no false calls and that perfectly sizes every flaw it detects.

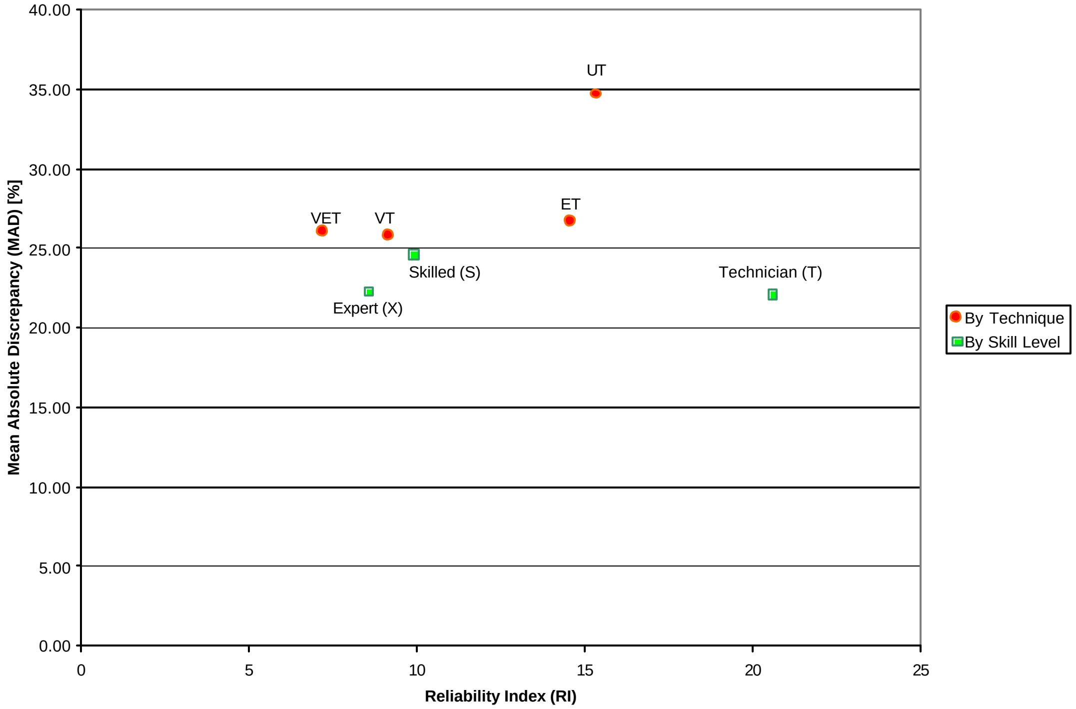


Figure 5-2. Average Performances Across Technique and Skill Level. As in Figure 5-1 the ideal case is to be at the origin (0,0).

In addition, our findings suggest that the skill level of the inspector is a key factor in the performance of the inspection. In particular, the data show that it is absolutely essential to have the inspector achieve at least the "skilled" level of expertise. While there exists a small but appreciable difference in terms of performance between an expert and a skilled inspector, the data indicate that an inspector with only a passing familiarity with the technique fares far worse in terms of his or her inspections. Here again VET testing is the best alternative, in that the difference between an expert and a skilled inspector is rather small. This is effectively a measure of the complexity of the inspection, and can perhaps best be thought of as the slope of the learning curve. An inspection such as VT for which the difference between expert and skilled levels of competence is slight but with unskilled lagging far behind represents a technique that although perhaps difficult to grasp in early stages is one that with sufficient time and practice is readily implemented. In contrast, UT inspection shows a very steep learning curve: an expert in this particular type of inspection was shown to perform as well as the other inspectors, but a skilled UT inspector with six years of experience in UT performed at virtually the same level as an unskilled inspector. It is worth mentioning that the expert inspector opted to conduct a hand scan rather than using the wheel transducer as did the other inspectors, and as such the UTX results are likely a combination of equipment and skill level. However, in further demonstration of the importance of skill a Level III UT inspector was asked to conduct the inspections with the wheel transducer, and ultimately performed so poorly that the inclusion of these secondary tests would have skewed the UT data to an unacceptable degree. In effect, the UT system showed that only the system vendor or an inspector with an equal amount of experience with the specific equipment could hope to perform at a level comparable to the other methods. The importance of proper training is best summarized by Figure 5-2.

By far the greatest spread in terms of performance across skill level is shown in eddy current testing, with a ΔRI of approximately 10 between each skill level (Figure 5-1). It is interesting to note that the skilled inspector actually performed better than the expert inspector in terms of RI, which can be attributed to the expert attempting to not miss any flaws, which resulted in a much higher level of false negatives. In addition, both the expert and the technician eddy current inspectors expressed some difficulty in locating the last thread in a cylinder: it is conceivable that some of the flaws reported by the expert and technician were actually indications that the probe had left the threaded region of the neck. In any case, the difference in skill levels serves to emphasize the importance of proper training for eddy current testing.

To summarize, the Nondestructive Testing Information Analysis Center recommends combined visual and eddy current testing (VET) for these reasons:

- In the field it is ultimately more important to find a flaw than to accurately size it. The VET technique had the best Reliability Index (RI) at 7.37, meaning that for every 100 inspection an average of 7.37 false calls will be made. In contrast, the second best technique was VT, which would make an average of 9.11 false calls per 100 inspections, which roughly corresponds to two additional cylinders per 100 inspected that are properly accepted or condemned.
- In terms of the secondary issue of flaw sizing, the VET technique ranks behind only VT with a Mean Absolute Discrepancy (MAD) of 26.09%, meaning that on average the flaw size VET reports is 26.09% larger or smaller than the actual flaw size, which is only slightly better than the MAD of VT of 25.85%.
- It does not require that the inspector be an expert in the technique, but rather that he or she be trained sufficiently to be described as “skilled.”