U.S. Department of Transportation

Federal Railroad Administration

Probability of Detection Evaluation Results for Railroad Tank Cars

Office of Research, Development, and Technology
Washington, DC 20590

DOT/FRA/ORD-16/13

Final Report
May 2016

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof. Any opinions, findings and conclusions, or recommendations expressed in this material do not necessarily reflect the views or policies of the United States Government, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government. The United States Government assumes no liability for the content or use of the material contained in this document.

NOTICE

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2016	Technical Report

4. TITLE AND SUBTITLE Probability of Detection Evaluation Results for Railroad Tank Cars	5. FUNDING NUMBERS DTFR53-C-00012 Task Order 213
6. AUTHOR(S) and FRA COTR Greg Garcia and Lucas Welander (TTCI), Ward Rummel (D&W Enterprises), and Francisco Gonzalez III (FRA)	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Transportation Technology Center, Inc. 55500 DOT Road Pueblo, CO 81001	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Department of Transportation Federal Railroad Administration Office of Research, Development, and Technology 1120 Vermont Avenue, NW Washington, DC 20590	10. SPONSORING/MONITORING AGENCY REPORT NUMBER DOT/FRA/ORD-16/13

11. SUPPLEMENTARY NOTES
COTR: Francisco González, III

12a. DISTRIBUTION/AVAILABILITY STATEMENT This document is available to the public through the FRA Web site at http://www.fra.dot.gov.	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Transportation Technology Center, Inc. (TTCI) used the approach developed for the National Aeronautics and Space Association to determine the probability of detection (POD) for various nondestructive test (NDT) methods used during inspection of railroad tank cars. The Federal Railroad Administration and TTCI, along with industry participation, have performed NDT PODs on tank car circumferential butt welds, longitudinal fillet welds and leak test samples requiring inspection under the *Code of Federal Regulations*. This report provides the quantitative results obtained during this research effort, which addresses system safety and risk analysis during handling and transportation of railroad tank cars that carry hazardous materials.

14. SUBJECT TERMS Bubble leak, damage tolerance analysis, direct visual testing, flaw size, hit or miss, liquid penetrant, magnetic particle, minimum crack size, nondestructive evaluation, nondestructive inspection, nondestructive testing, probability of detection, radiography, remote visual, ultrasonic	15. NUMBER OF PAGES 82
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

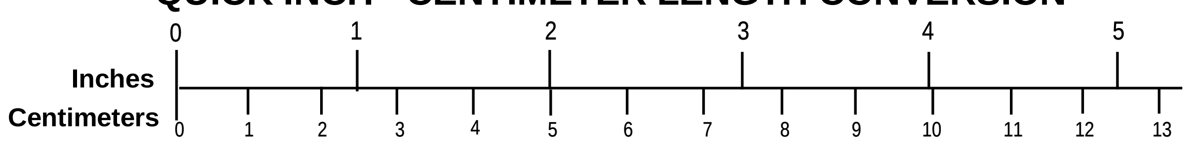
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO

°F	-40°	-22°	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Contents

Illustrations

Tables

Executive Summary

Transportation Technology Center, Inc. (TTCI), under the sponsorship of the Federal Railroad Administration (FRA), is working with the tank car industry to increase the reliability of railroad tank car structural integrity inspections. In support of the Tank Car Nondestructive Test (NDT) Program, TTCI researchers and industry participants have evaluated a variety of NDT methods used to inspect tank cars. Accomplishments to date include:

- Baseline inspections of railroad tank cars
- Validation of NDT methods
- Development of baseline probability of detection (POD) curves
  - NDT methods per Code of Federal Regulations (CFR): liquid penetrant, magnetic particle, radiography (see FRA/ORD/DOT-9/10), ultrasonic, and visual
  - Other NDT methods: bubble leak, eddy current
- Establishment of a tank car defect library
- Development of master gages
- Qualification of a bubble leak test inspection procedure

A rulemaking issued by the U.S. Department of Transportation (DOT) revises the Hazardous Materials Regulations (HMR) to replace the hydrostatic pressure test for regulation of tank cars with appropriate NDT methods. HM-201 Requalification is a federal regulation governing the qualification of DOT and Association of American Railroads' tank cars. It eliminates the hydrostatic tank test previously used and replaces it with nondestructive testing, which provides a better method of detecting defects and ensuring tank car safety. The rulemaking also requires that the test methods should be quantified to demonstrate the sensitivity and reliability of the inspection and test techniques. The rule changes are located in the Federal Register (Title 49: Transportation, Chapter I-Research) and Special Programs Administration, Department of Transportation (Part 179-Specifications for Tank Cars and Part 180-Continuing Qualification and Maintenance of Packagings) [1].

The basis for assurance of the structural integrity and for life-cycle management of engineering structures based on material, loads, and nondestructive inspection (NDI) has been established and is the primary basis for fleet management of aircraft structures. The well-established principles and tools developed for aircraft applications have been adopted and applied to a wide range of engineering structures, components, and materials in the public domain and are the basis for the methodology that TTCI has applied to railroad tank car structures.

Addressing the Revised HMRs

The CFR requirement under Section 179.7(b) (10) includes: "Procedures for evaluating the inspection and test technique employed, including the accessibility of the area and the sensitivity and reliability of the inspection and test technique and minimum detectable crack length." Section 180.509 of the CFR identifies requirements for inspection and test of specification tank cars, paragraph (e) structural integrity inspection tests. The CFR authorizes liquid penetrant, magnetic particle, radiography, ultrasonic, and visual testing as allowable NDT methods for structural integrity inspections. Alternative NDT methods may be allowed for railroad tank car inspections under special exemptions that are issued by FRA Office of Railroad Safety."

Baseline NDT of Railroad Tank Cars

NDT technicians from the railroad tank car industry assisted TTCI in identifying and documenting current industry practices. Industry representatives also performed a baseline inspection of four tank cars using the CFR authorized NDT methods along with the acoustic emissions NDT method. The technicians who assisted in this effort perform tank car inspections regularly as assigned by their respective companies. The areas of focus were the required inspection areas as identified in 49 CFR 180.509, including the circumferential butt welds and longitudinal fillet welds.

The tank cars used during the baseline inspections have been stored at the FRA's Transportation Technology Center (TTC), Pueblo, CO, as part of the Tank Car Defect Library. The tank cars are available for future evaluations to provide capability comparisons as NDT technology is developed and implemented for tank car inspections.

Validation of NDT Methods

TTCI used information generated by the aerospace and nuclear industries to create a methodology that validates railroad tank car NDT processes. A NDT process includes the NDT systems and procedures used for the inspection, as well as the NDT equipment, operator, inspection environment, and the object being inspected.

Researchers used the NDT method validation to assess the reliability and implementation costs associated with an NDT process. The use of a validation methodology to assess the applications, advantages, and limitations of NDT methods is a valuable tool to ensure inspection reliability.

Development of Baseline POD Curves

If a damage tolerance approach to determine inspection intervals is used for an engineered structure—in this case a railroad tank car—the quantification of the detectable flaw size for the NDT methods used during inspection is required. Traditionally, NDT methods have not been quantified and assumed capabilities have often been found to be in error. However, damage tolerance techniques have led to an evolution in NDT understanding, methods, and requirements. National Transportation Safety Board safety recommendations R-92-21 through R-92-24 address the suggested process of performing reliable inspection of railroad tank cars based on a damage tolerance approach. Damage tolerance design and maintenance are expected to improve the reliability and confidence level of tank car acceptance and maintenance. NDT that is quantified using the POD approach, a key measure of NDT effectiveness, is integral to damage tolerance requirements [2].

TTCI has worked with the FRA and the tank car industry to develop baseline POD curves for the allowed NDT methods. Initial evaluations were performed on the inspection of tank car circumferential butt welds. Subsequent efforts focused on both the butt welds and longitudinal fillet welds requiring inspection under the CFR.

Tank Car Defect Library

Under the sponsorship of FRA, TTCI has created a defect library containing sample artifacts such as railroad tank cars and sections of railroad tank cars. The library also contains tank cars donated by the industry and manufactured artifacts developed at TTC. Manufactured artifacts include test panels used for POD evaluations and master gages developed for inspection sensitivity verification. These specimens contain discontinuities developed in service as well as manufactured flaws simulating location and type of discontinuities expected in service.

The defect library was created to provide the tank car industry with resources similar to the aerospace and nuclear industries. Establishing a defect library and validation center offers the industry a facility to perform comprehensive, independent, and quantitative evaluations of new and enhanced inspection, maintenance, and repair techniques.

Master Gages

Baseline PODs have been developed by TTCI using standard industry NDT procedures. This data can provide a basis for design and life-cycle maintenance assumptions in general NDT inspections. The data is to be anchored by application and response to tank car master gages. The PODs have been established to provide a capability that can be used for qualification of equivalent NDT procedures and for personnel skill demonstrations.

The primary measure of reliability in NDT is repeatability and reproducibility. Master gages developed from the test tank cars are tools that can be used to perform a response comparison to calibration artifacts used in the field. The master gages are stored at TTC to preserve and periodically revalidate response linearity of the calibration artifacts.

1. Introduction

The POD approach has been used to quantify the capabilities of NDT methods allowed for inspection of railroad tank cars. This report provides the results of industry POD evaluations conducted by researchers from TTCI under the sponsorship of the FRA.

1.1 Background

A rulemaking issued by the United States DOT revises the HMRs to replace the hydrostatic pressure test with appropriate NDT methods. The rule change is contained in the *Federal Register*, Title 49, CFR Part 180.509, "Requirements for inspection and test of specification tank cars," paragraph (e) "Structural integrity inspection tests [1]." The CFR authorizes liquid penetrant (PT), magnetic particle (MT), radiography (RT), ultrasonic (UT), and optically aided visual testing (VT) as allowable NDT methods for structural integrity inspections and tests. Other NDT methods may be allowed under special exemption issued by the FRA's Office of Railroad Safety. Also, included under the requirements of 49 CFR Part 179.7 is the need to qualify not only NDT personnel but the procedures used to perform NDT reliably [1].

In order to be effective, Federal regulations require that the NDT methods have a demonstrated sensitivity and reliability for finding the type and size of flaws likely to cause a tank car failure. In the early 1970s, an internationally accepted quantitative approach that determines the POD was developed for the National Aeronautics and Space Association (NASA) and was published in NASA CR-2369, February 1974 [3]. TTCI, under contract with the FRA and with industry participation, uses the NASA approach to determine the POD for various NDT methods used in inspection of railroad tank cars.

PODs have been performed on tank car circumferential butt welds, longitudinal fillet welds, and leak test samples requiring inspection under the CFR. This report provides the quantitative results obtained during this research effort and addresses system safety and risk analysis during handling and transportation of railroad tank cars carrying hazardous materials.

1.2 Objectives

This research effort evaluates NDT methods authorized under the CFR for use in qualification or requalification of railroad tank cars. To provide direction and insight into the current capabilities of the industry when using the allowed NDT methods, the test methods are quantified of through POD.

1.3 Overall approach

NDT ensures the structural integrity of engineering components, systems, and structures during initial fabrication and acceptance, in-service maintenance inspections, and life extension. The reliability of the chosen NDT method is a key consideration in the safety and economic operation of most of the systems in public service.

The rulemaking issued by the DOT revised the HMRs, which now requires that the hydrostatic pressure test is replaced with appropriate NDT methods. The rulemaking also requires that the test methods used have been quantified to demonstrate the sensitivity and reliability of the inspection and test technique.

Per the CFR, acceptable NDT methods include PT, MT, RT, UT, and VT to provide assurance of structural integrity and other methods, such as bubble leak testing, for system functional safety.

POD evaluations performed under this research effort were conducted using railroad tank car panels containing flaws (fatigue cracks) that propagated at the toe of the girth seam (butt) welds and at the longitudinal termination of fillet welds. Leak paths were also manufactured in seals and gasket locations around the tank car manway assembly, safety valve assemblies, connection flanges, and bottom outlet nozzles. Manufacture of the test panels along with the creation of a defect library has been performed to provide the tank car industry with the means for third party assessment of NDT systems used for tank car inspection.

1.4 Scope

TTCI, under the sponsorship of the FRA, is working with the tank car industry to increase the reliability of railroad tank car structural integrity inspections. In support of the Tank Car NDT Program, TTCI researchers and industry participants have evaluated a variety of NDT methods used to inspect tank cars. Previous work has been documented in FRA report numbers DOT/FRA/ORD-01/04 and DOT/FRA/ORD-09/10, and it involved the production of physical test specimens which represent industry-inspected components [4, 5]. These test specimens were used to baseline industry detection capabilities. This report assesses specific industry procedures, specific inspector performance capabilities, and the validation of the test specimens used.

2. Probability of Detection

When fatigue and fracture mechanics analyses were developed and integrated into design, maintenance, and the life extension of engineering systems, they provided the ability to quantify confidence in safety and structural integrity. That is, the method can quantify the crack size that can be sustained as a function of load. Damage tolerance and slow crack growth properties are used to quantify safety margins, required maintenance intervals, and life extension. The crack size that can be reliably detected is a critical factor in such analyses, thus quantitative NDT is required. *Probability of detection* is the metric that is generally applied to quantifying the detection capability (of small cracks) of an NDT procedure. Reliable detection is required which, in turn, requires quantification of the flaw size that can be detected and confidence in the detection of that flaw size.

2.1 Nondestructive Testing Effects on Probability of Detection

NDT is an indirect method for detecting a flaw (crack or other anomaly), and detection is a function of multiple variables associated with the detection and measurement processes. Known variables in NDT detection include properties of the:

- Flaw (anomaly)
- Test Object
- NDT Method
- NDT Materials
- NDT Equipment
- NDT Procedure
- NDT Process
- Calibration
- Acceptance Criteria and Decision
- Human Factors

The nature of NDT processes has inherent lower (and upper) limits of detectability; i.e. the smallest flaw that can be reliably detected and the largest flaw that has ever been missed.

2.1.1 Signal and Noise Effects

When we perform NDT, we are challenged to separate the signal produced by the flaw from the background response (noise) that is inherent to the measurement. Note that this is not electronic noise but rather the response due to material condition, surface finish, and geometry. For a “well behaved” inspection process, the signal and noise may be visualized as Figure 1 shows.

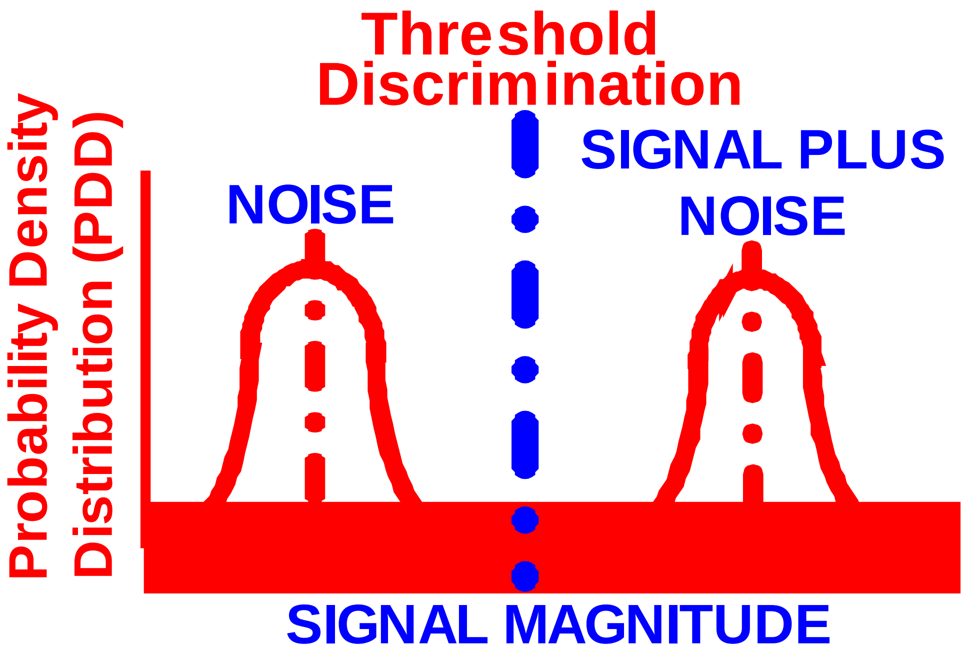


Figure 1. Signal and Noise Distributions for a NDT Procedure

(Note that detection requires discrimination of signal from noise.)

For purposes of visualization, the distributions are shown as Gaussian. In some applications, the distributions may vary with flaw size and with flaw type. Figure 1 shows a clear discrimination between signal and noise, and thus there is a clear detection opportunity. As the flaw size decreases, the signal distribution moves closer to the noise distribution and detection is more difficult. When the distributions overlap, we are unable to discriminate, and the NDT procedure is unable to detect small flaws. As the lower limit of detection is approached, false calls (noise interpreted as signal) will result in those cases where the signal and noise distributions overlap. Figure 2 is a visualization of the effects of signal and noise distributions on detection capability. Note that flaw signal is assumed to increase with increasing flaw size for the NDT process applied. This assumption must be validated for each application.

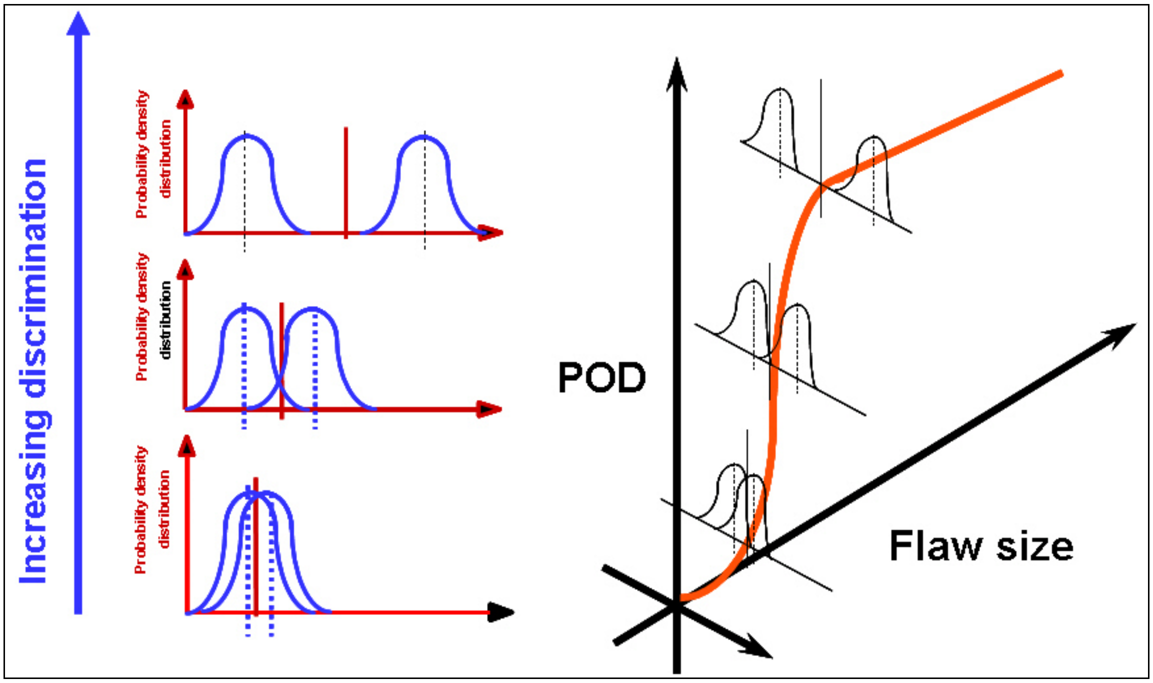


Figure 2. Signal and Noise Effects of Flaw Size

(A constant threshold discrimination level is assumed.)

For nonautomated NDT processes and systems, the threshold discrimination level may vary. This may be due to a slight variation in the NDT procedure applied or due to the variance in discrimination by the human operator. A good inspection process may result in poor discrimination as a result of variance in the applied discrimination threshold. Figure 3 shows a schematic of this condition.

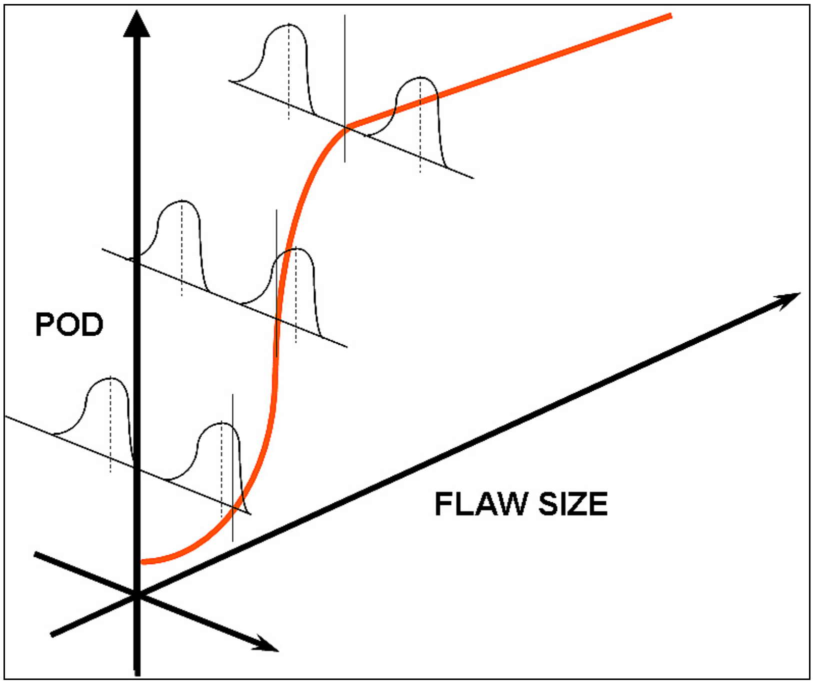


Figure 3. Effect of Varying Threshold Discrimination on Detection

2.1.2 NDT Process Effects

Quantification of the detection capability for a multiple parameter-multivariate detection process may be approached using statistical tools; however, the applicability and validity must always be approached by considerations of the variables inherent to the NDT process.

Figure 4 shows the possible results of a decision (threshold discrimination) process for each detection opportunity. The desired output is a majority of true positive (T.P.) results with few misses or false negatives (F.N.), and few false positives (F.P.). The interdependence of the matrix quantities is denoted by:

T.P. + F.N. = Total opportunities for positive calls (detection plus misses)

F.P. + T.N. = Total opportunities for negative calls (false calls plus true negatives)

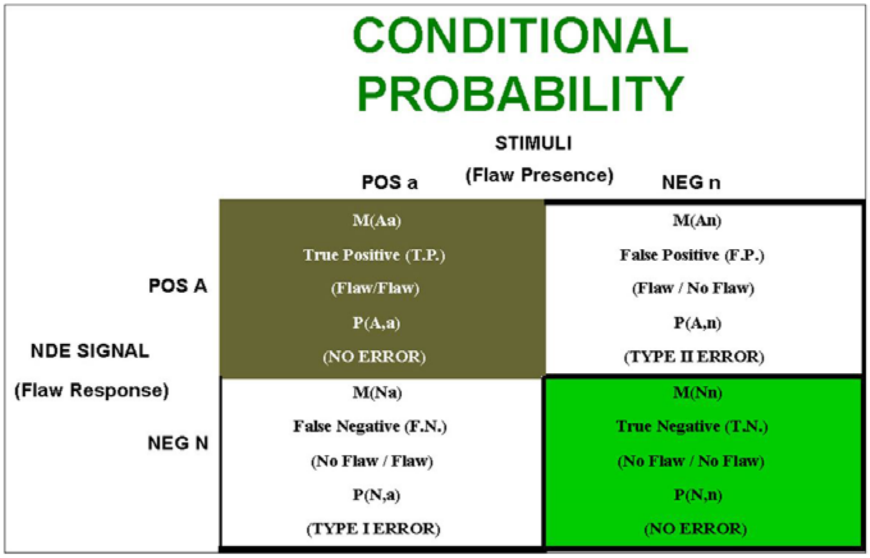


Figure 4. Conditional Probability of the Results of Multiple Inspection Opportunities

The specificity (discrimination capability) of the procedure or the POD may be expressed as:

$$\text{POD} = \frac{\text{T.P.}}{\text{T.P.} + \text{F.N.}} \quad \text{or} \quad \frac{\text{total positive calls (REJECTS)}}{\text{opportunities for positive calls (REJECTS)}}$$

The nonspecificity of the procedure or the probability of false alarms (POFA) may be expressed as:

$$\text{POFA} = \frac{\text{F.P.}}{\text{T.N.} + \text{F.P.}} \quad \text{or} \quad \frac{\text{total false positives (FALSE CALLS / ALARMS)}}{\text{opportunities for false positives (ACCEPTANCE)}}$$

The POD is the metric that is generally used to assess the detection capabilities of an inspection procedure. This involves submitting a statistically significant number of flaws to an inspection process and analyzing the outcome in terms of POD as a function of flaw size. Note that flaw signal is assumed to increase with increasing flaw size, as Figure 5 shows.

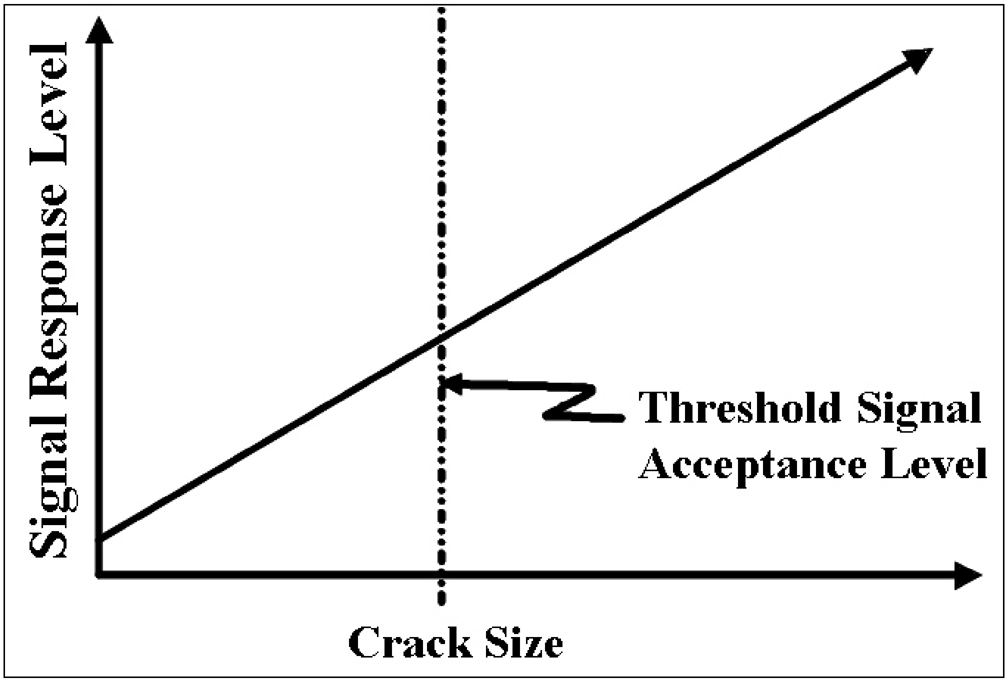


Figure 5. An Increase in Signal with Increasing Flaw Size is Assumed and Must be Validated to Ensure that the NDT Procedure is Producing the Intended Result

The output of a POD capabilities assessment is a POD curve, as Figure 6 shows. This is the result of analyzing inspection data produced from a representative number of flaws (size and size range) and includes the influence of all the variables that are inherent to the detection process. Thus, it is a probabilistic estimate of the detection capability of a process at a snapshot in time. If any of the multiple variables in the application of an NDT procedure are changed, a new assessment must be completed and a new POD curve generated.

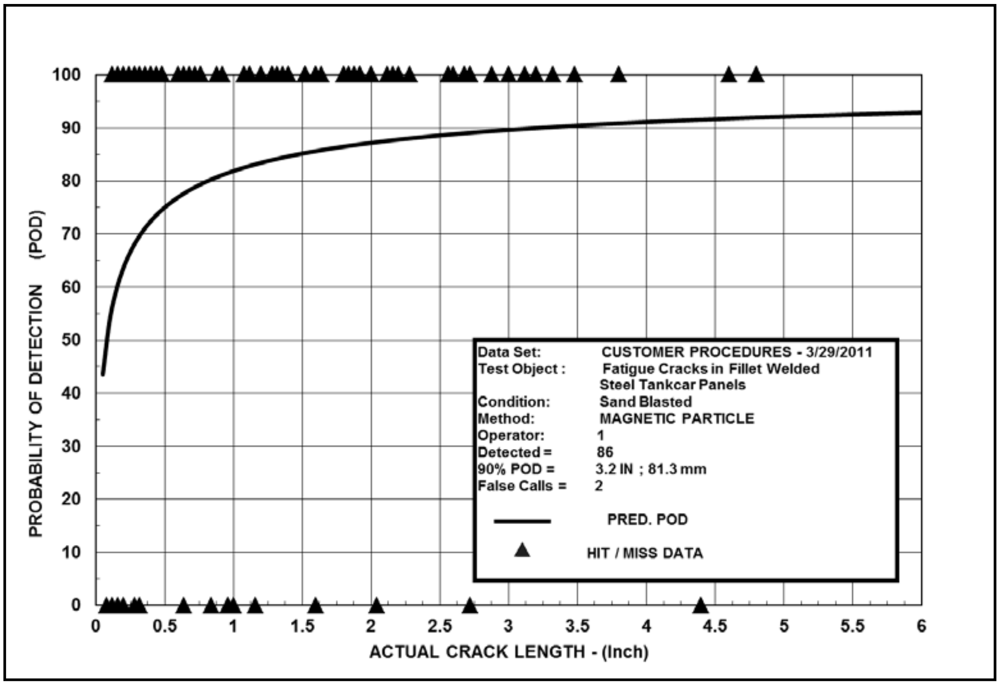


Figure 6. Typical POD Curve

(Note that detections are not single valued at a given flaw (crack) size and flaws of the same size may produce a detection or a miss.)

The discrimination point that is generally reported and used as a basis for design and maintenance acceptance decisions is the point at which the POD curve passes through the 90-percent POD level. From sampling theory, the number of cracks detected at that point will define the confidence level in the crack size used. Typically included in the POD chart are icons showing hits across the horizontal line at 100 percent and misses across the horizontal line at 0. In Figure 6, the icon used is a triangle, but others, such as Xs, squares, and circles, are also used. As Figure 6 shows, flaws of the same size can be detected or not.

2.2 Probability of Detection Methodology

In this report, NDT capabilities assessment is by POD methodology. Because crack to crack variances and NDT process variances must be addressed, the POD method was developed as a probabilistic method of analysis. In short, the method assumes that the result of any NDT method is that of discriminating between distributions of signal and noise analyses and that the system process is consistent with the POD model.

By taking signal response data as a function of flaw size and fitting it to a log linear plot, a slope and intercept can be derived and input into the POD model. It is assumed that the log linear relationship can be reproduced by rigid NDT system "calibration" and thus the POD or discrimination capability for an NDT system or procedure can be quantified. By convention, the accepted discrimination level is at the point where the POD curve passes through the 90-percentile point. This single valued output is then input to structural analyses as the basic capability of NDT discrimination and acceptance. The POD methodology is an accepted metric for validation of the capability of an NDT procedure for comparison of NDT procedures and for assessment of skill levels of NDT operators.

2.2.1 POD Approach

The original POD curves, under the NASA program, were produced using a moving average method [2]. That method required very large data sets, thus other methods were pursued to reduce data quantity requirements.

The most popular and most often used POD approach is the Berens and Hovey model [6]. This model assumes an increasing signal with increasing flaw size and uniform signal variance. It is a model description of the merger of signal and noise as previously shown in Figure 2 and is the method that has been applied to the analysis of railroad tank car inspection capabilities data.

The Berens and Hovey model uses a cumulative lognormal distribution function and is approximated by the log-odds model. The data may be described by:

$$\mathbf{POD(a)} = \frac{\exp\left[\alpha + \beta \ln(\alpha)\right]}{1 + \exp\left[\alpha + \beta \ln(\alpha)\right]}$$

where a = crack length

The maximum likelihood method is used to estimate the α and β parameters of this statistical model. The LOGIT model may be described as:

$$\mathrm{POD(a)} = F(\alpha + \beta(\mathrm{Log}\ (a))),$$

where α and β are parameters to be fit to the data and F is an increasing function of (a).

The Berens and Hovey model is the basis of Military Standard 1823 and the analysis tools that are provided [7].

The POD maximum likelihood analysis method is useful when using smaller quantities of data to provide a predictive estimate for:

- detecting the capability of an NDT procedure,
- validating the use of a procedure, and
- assessing the effects of variances in NDT procedure application by comparing the detection capabilities of various NDT procedures and operators.

It is important to note that POD is not reliability, but is a component of reliability. Inspection reliability is:

- Reproducibility, characterized by rigor in calibration
- Repeatability, characterized by rigor in process control
- Capability, characterized by the first principle’s physics of the applied NDT process

Unless reproducibility and repeatability are in control, NDT capabilities data is not in control and data is not representative of the inspection process. For NDT methods, such as liquid penetrant and magnetic particle inspections, both the consistency of the inspection materials used and the sequence of application are critical to process repeatability. For inspection methods, such as eddy current or ultrasound, which involve human pattern recognition and/or signal observation, a consistent threshold level in detection (NDT process acceptance criteria) is required.

2.2.2 False Call Effects on Probability of Detection

A false call is when the operator either records or identifies a flaw that does not exist. False calls do not influence the POD curve (when based solely on a hit or miss approach). An operator could theoretically have a high POD and a correspondingly high false call rate. Optimal results should manifest a high POD with a low false call rate. Because false calls may lead to further inspection by using additional NDT methods and techniques, fleet owners may experience costs associated with unnecessary maintenance, downtime, and repairs. The decision to use the NDT method and technique should therefore be balanced against the POD results and the number of false calls.

A variation in the threshold level results in both misses and false calls. This is evident when an operator is inadequately trained. An inexperienced operator may call all indications or signals as detections (hits) with a resulting high false call level. Data produced with a high false call rate is not valid and POD analysis results from such data are not valid. Generazio (AGARD-LS-190) has shown that the true probability of a valid detection falls off rapidly with increasing false calls, as Figure 7 shows [8]. A 90 percent detection level is not achieved when the false call rate (FCR) or false positive rate (FPR) exceeds 5%.

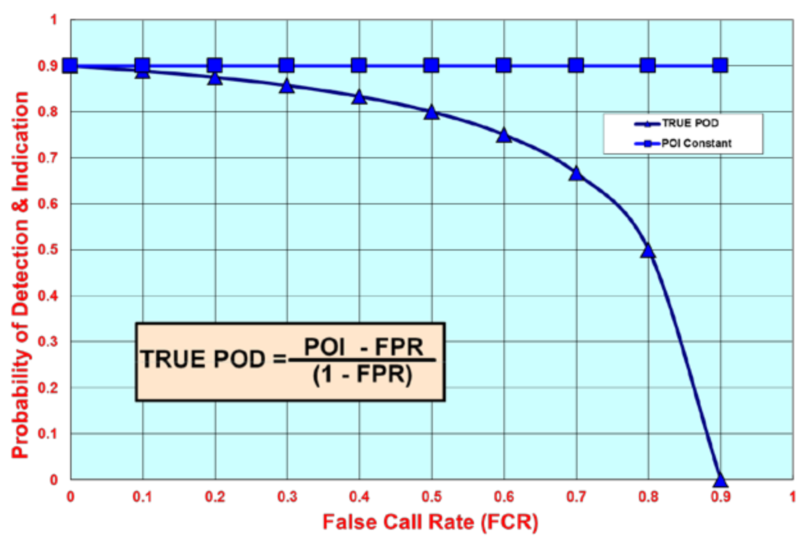


Figure 7. Probability of Indication (POI) and POD as a Function of False Call Rate [8]

(Note the POD decreases below the 90% level at a FPR >5%.)

An inexperienced operator could potentially call every inspection site as a flawed area, which would invalidate the inspection sequence and the POD for that data set, as Figures 8 and 9 show. Both POD curves show similar results but closer analysis of the inspection results shows that the curve in Figure 8, which represents that of an experienced certified operator, has six false calls. The curve in Figure 9 shows 41 false calls for an inexperienced noncertified operator.

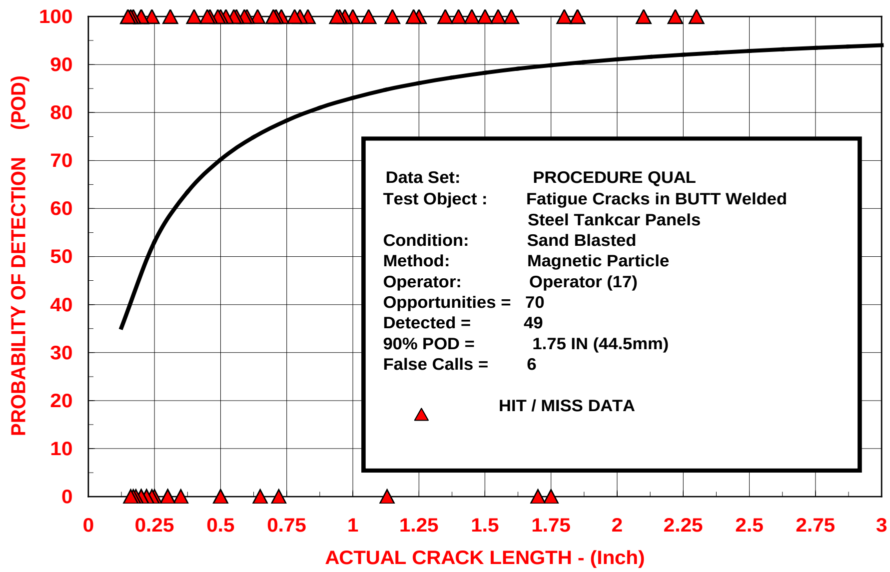


Figure 8. Experienced or Certified Operator POD Results

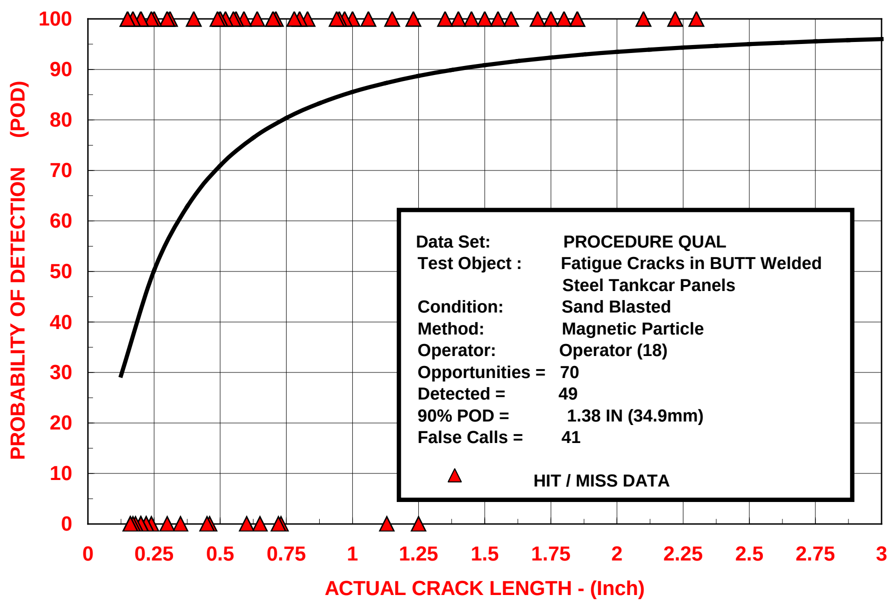


Figure 9. Inexperienced or Noncertified Operator POD Results

To aid the operators involved in POD evaluations, master gages were provided with both notches and cracks to reinforce their familiarity with flaw responses from the test panels. The master gages consisted of electro-discharge machined notches and fatigue cracks of various sizes manufactured into samples removed from tank cars retired from service.

The process used during tank car NDT POD evaluations requires each operator to inspect and size the cracks and slots in the master gage test panels before, at intervals during, and after completing the inspection of the larger tank car section panels. Results are then recorded, and the data is used as an indicator of potential variation in the applied operator discrimination (threshold acceptance) level during completion of the inspection sequences. When a large variation in discrimination and sizing is found, the FPR for that operator is usually high, and validity of the inspection sequence is therefore in question.

2.3 Probability of Detection Results

The following summarizes the results of the analysis of NDT data documented by TTCI during POD evaluations performed at TTC from 2009 to 2011. The report constitutes an analysis of data collected during quantification of CFR-accepted NDT methods as applied to inspection of railroad tank car butt weld and fillet weld required inspection locations as identified in the CFR. The documented POD approach uses a “hit or miss” protocol.

The data that was collected for this analysis constituted the inspection results of different industry participants using inspection procedures and equipment provided by the companies they represent. Evaluations were performed on test panels containing girth (butt) welds and fillet welds taken from retired railroad tank cars. Fatigue cracks were previously induced in both panel sets in order to provide a POD evaluation of NDT methods. Cracks range in size from 0.15 in. to 3.25 in. for the butt weld samples and 0.080 in. to 6.00 in. for the fillet welds. The variety of cracks from smallest to largest provides a range of inspection opportunities that are representative of cracked components from field service.

Results indicated both the capabilities of the NDT methods and the difficulties in applying some of those methods to railroad tank car butt and fillet welds. The data reflects operator variability in the application of the various inspection methods and the effect from false calls on detection capability.

Any given NDT method’s capability is specific to variables related to flaw characteristics such as size, orientation, and state of stress (compression or tension). The test object, inspection equipment, calibration, written procedure, and its related processes, acceptance criteria, human factors, and environmental conditions are also variables affecting NDT capability.

Related to human factors, the operator’s ability to inspect an item within a given time period under a particular job quota and maintain production levels introduces an inherent need to inspect at a given rate. Consequently, the POD curve is influenced by an operator’s ability to discriminate flaws at a standard inspection rate (i.e., a time study, which is the rate that is generally based on the speed to which other operators in a facility inspect the same item). For example, if two operators evaluate a test sample, one operator may spend 15 minutes while another operator may spend 30 minutes, depending on their comfort level for decision processing during flaw discrimination. During the POD evaluations, operators were asked to inspect the tank car specimens based on the average inspection rate for a typical shop environment.

The POD curves represented in this report provide a quantitative measure of the effectiveness of an NDT method; this gives us an opportunity to evaluate the need to use one method over another given the nature (criticality) of the area under observation and the desired sensitivity. The curves serve as a baseline, which allows changes to NDT variables to become measurable if another study of the capabilities of the method is performed and the resulting change is observed.

2.3.1 Direct Visual

Visual testing employs electromagnetic radiation at frequencies visible to the human eye. Direct visual testing (DVT), also referred to as optically aided visual testing, can be enhanced by using tools, such as magnifiers, flashlights, and dental picks, to mechanically enhance or aid the human sensory system. Interrogation using DVT incorporates the senses (looking, feeling, and smelling) to assist in interpreting the condition of the item being inspected.

The DVT method's advantages are that it is economical, expedient, and requires relatively little training or equipment for many applications. However, the DVT method is limited because it is only applicable to external or surface conditions and it is dependent on the inspector's visual acuity. Visual acuity may be influenced by operator variables such as physiological processes, psychological state, experience, health, and fatigue. The DVT method provides a fast, economical method to perform tank car inspection and provide effective determination of many surface discontinuities, but it is dependent upon lighting and operator influences. Using DVT alone may not be suitable without the aid of supplemental equipment and/or test methods to adequately determine small discontinuities and tightly closed cracks.

DVT results show variability among operators and the influence of test articles. The girth (butt) weld inspection showed higher operator variability than the fillet weld inspections. False calls were recorded during the evaluations and POD curves show results from a pure hit or miss assessment and results with the influence of false calls taken into account.

DVT Butt Weld POD Results

Butt weld POD evaluation results for industry participants A and B are provided in this section. Results show individual results of each company's participants and the combined average for the company. Before performing inspections of the test panels, the participants performed setup, which included calibration in accordance with the applicable company procedure. After calibration, the participants began their inspections of the butt weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

DVT Butt Weld POD Results for Company A

Company A had three technicians participate in the DVT PODs. Table 1 lists the POD results at different crack lengths, while Table 2 lists the results with the false call adjustment included in the results. Figures 10 through 12 show the POD curves, with and without the false call adjustment, for each operator. Figure 13 shows a comparison of the results of each operator, and Figure 14 shows the combined average of all three operators.

Table 1. Company A Butt Weld DVT POD Percentages (%)

Flaw Size	Operator	Operator	Operator	Combined Results

(inch)	6	17	19	
0.25	52	31	18	34
0.50	60	35	20	38
0.75	65	37	21	41
1.00	68	39	21	43
1.25	70	40	22	44
1.50	72	42	22	45
1.75	74	43	23	46
2.00	75	43	23	47
2.25	76	44	24	48
2.50	77	45	24	49
2.75	78	45	24	49
3.00	78	46	24	50
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	1	0	2	3

Table 2. Company A Butt Weld DVT POD Percentages (%) with False Call Adjustment

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	51	31	15	32
0.50	59	35	17	37
0.75	64	37	18	40
1.00	67	39	18	41
1.25	69	40	19	43
1.50	71	42	19	44
1.75	72	43	20	45
2.00	73	43	20	46
2.25	74	44	20	46
2.50	75	45	21	47
2.75	76	45	21	47
3.00	77	46	21	48
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	1	0	2	3

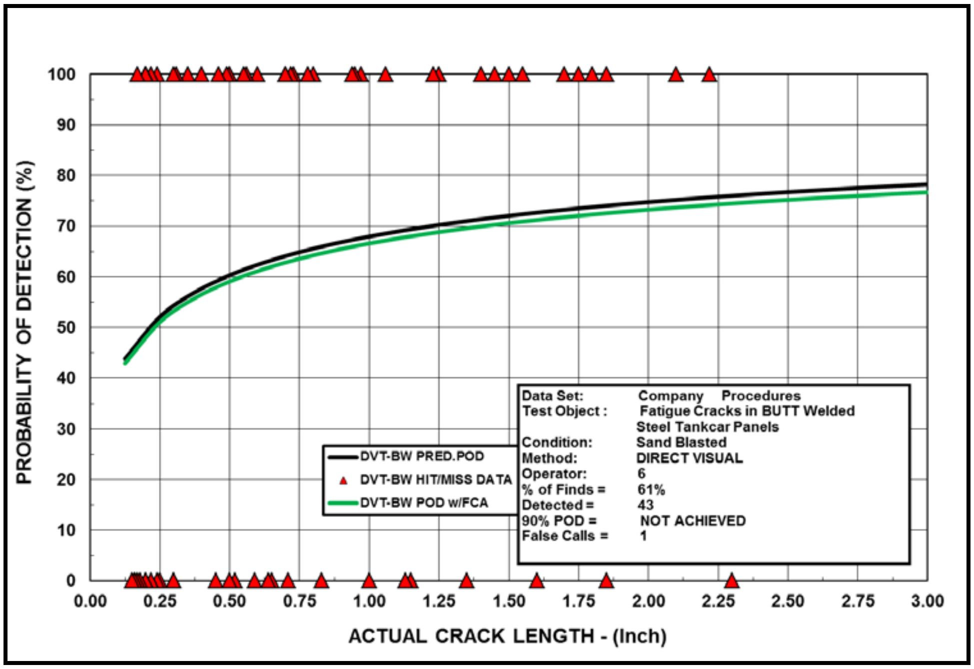


Figure 10. DVT Butt Weld POD Results for Operator 6

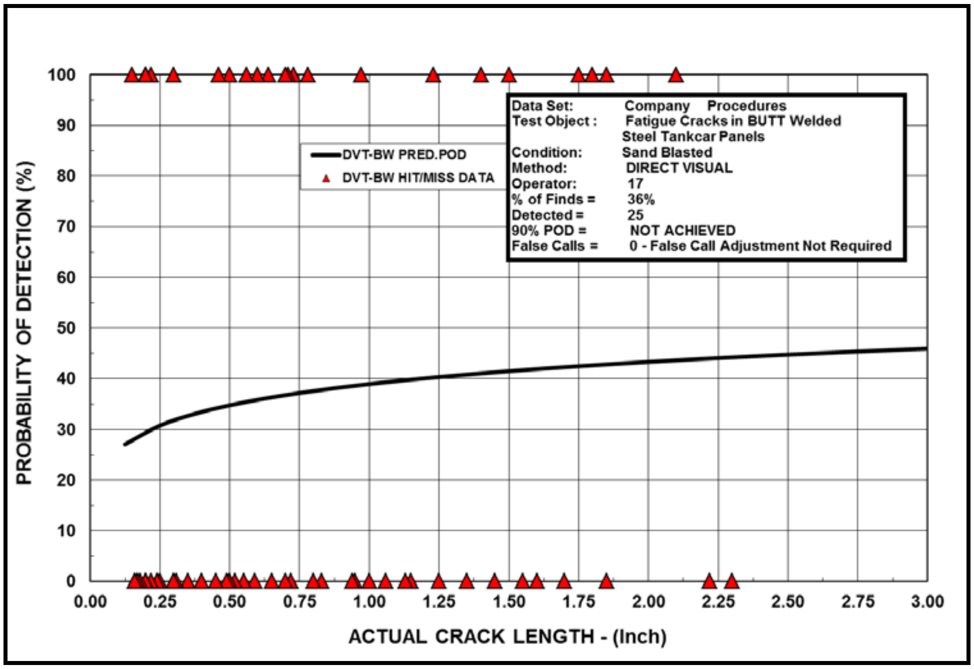


Figure 11. DVT Butt Weld POD Results for Operator 17

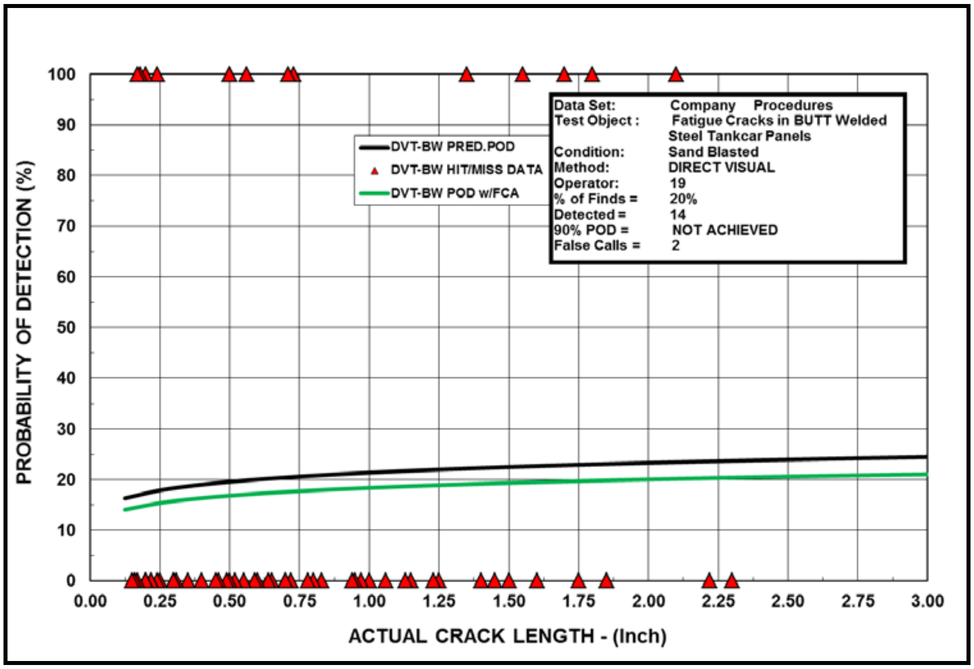


Figure 12. DVT Butt Weld POD Results for Operator 19

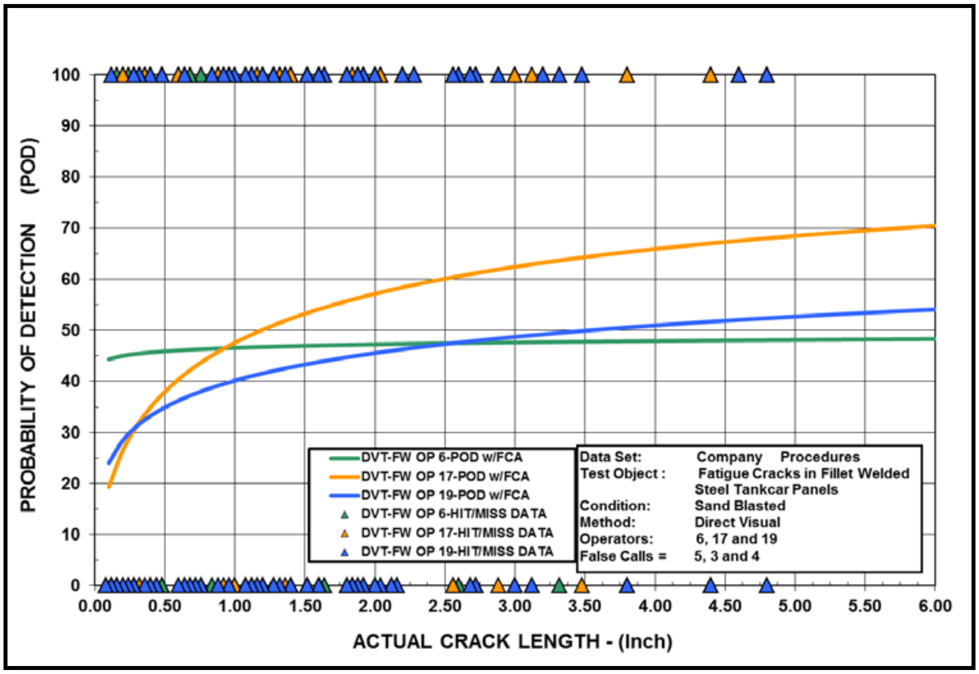


Figure 13. DVT Butt Weld POD Company A Operator Comparisons

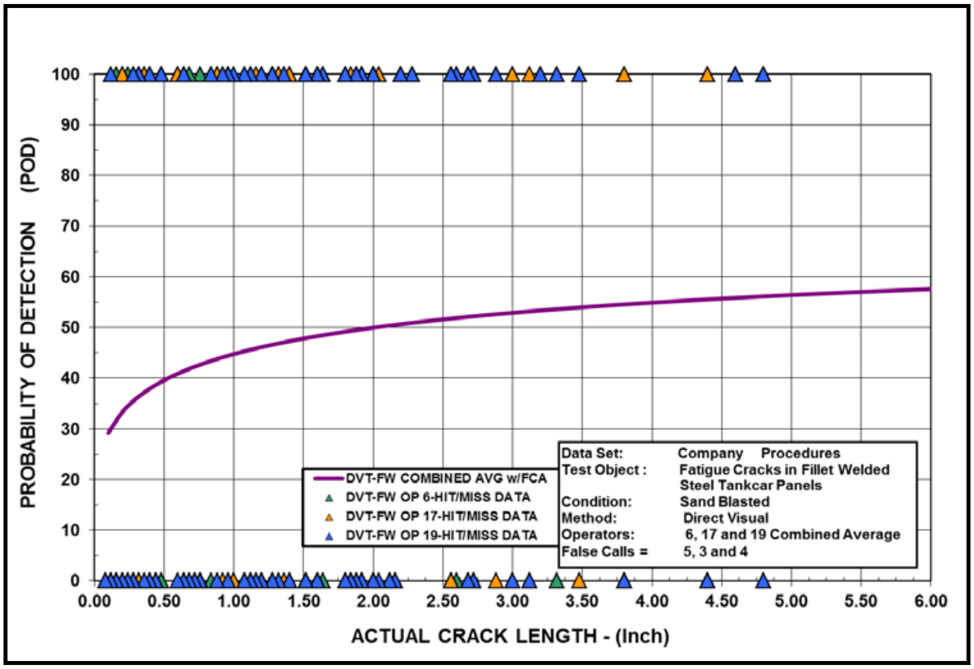


Figure 14. DVT Butt Weld POD Combined Average for Company A

DVT Butt Weld POD Results for Company B

Company B had three technicians participate in the DVT PODs. Table 3 lists the POD results at different crack lengths. Figures 15 through 17 show the POD curves for each operator. Figure 18 compares the results of each operator, and Figure 19 shows the combined average of all three techncians.

Table 3. Company B Butt Weld DVT POD Percentages (%)

Flaw Size (inch)	Operator 15	Operator 20	Operator 27	Combined Results
0.25	37	20	58	38
0.50	44	33	66	48
0.75	49	42	69	54
1.00	52	49	72	58
1.25	55	54	74	61
1.50	57	58	75	64
1.75	59	62	77	66
2.00	60	65	78	68
2.25	62	67	79	69
2.50	63	69	79	70
2.75	64	71	80	72
3.00	65	73	81	73
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	68	4	97	169

Note: Grey areas depict a high number of false calls that skew POD results.

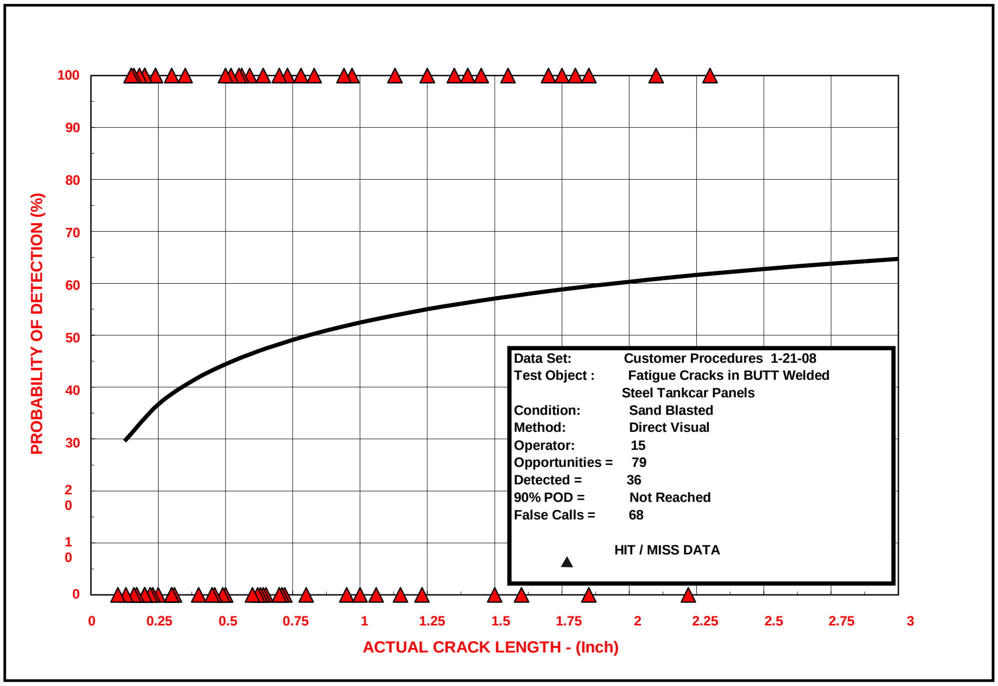


Figure 15. DVT Butt Weld POD Results for Operator 15

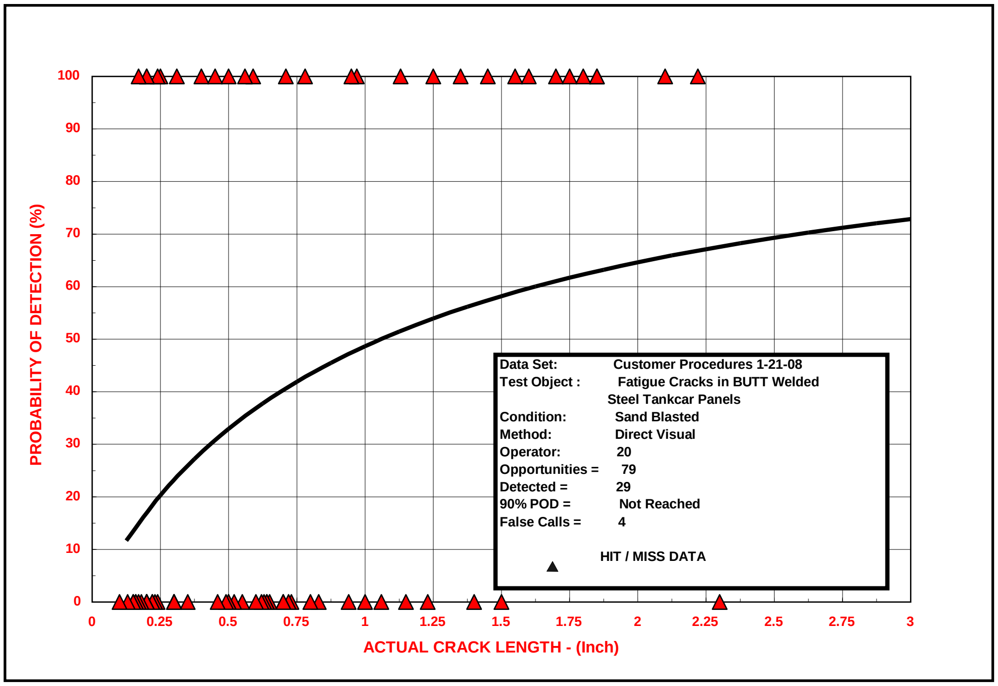


Figure 16. DVT Butt Weld POD Results for Operator 20

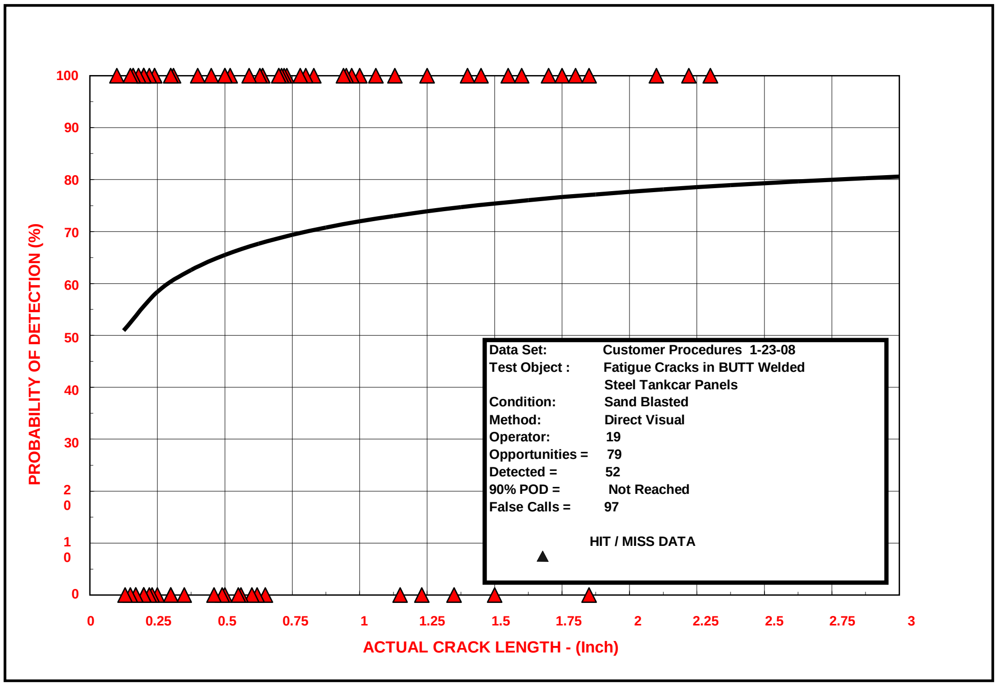


Figure 17. DVT Butt Weld POD Results for Operator 27

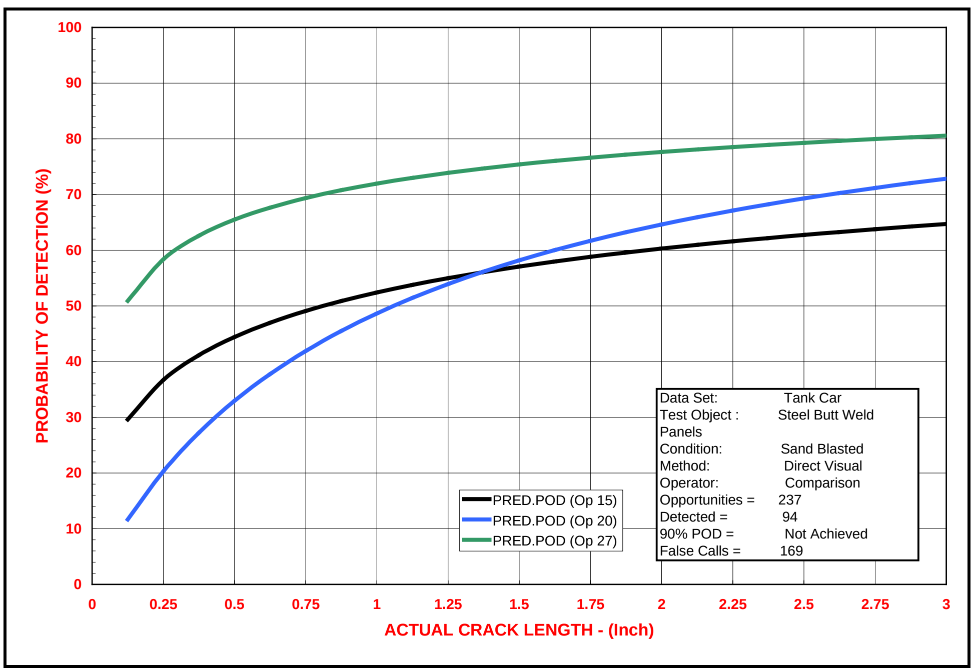


Figure 18. DVT Butt Weld POD Company B Operator Comparisons

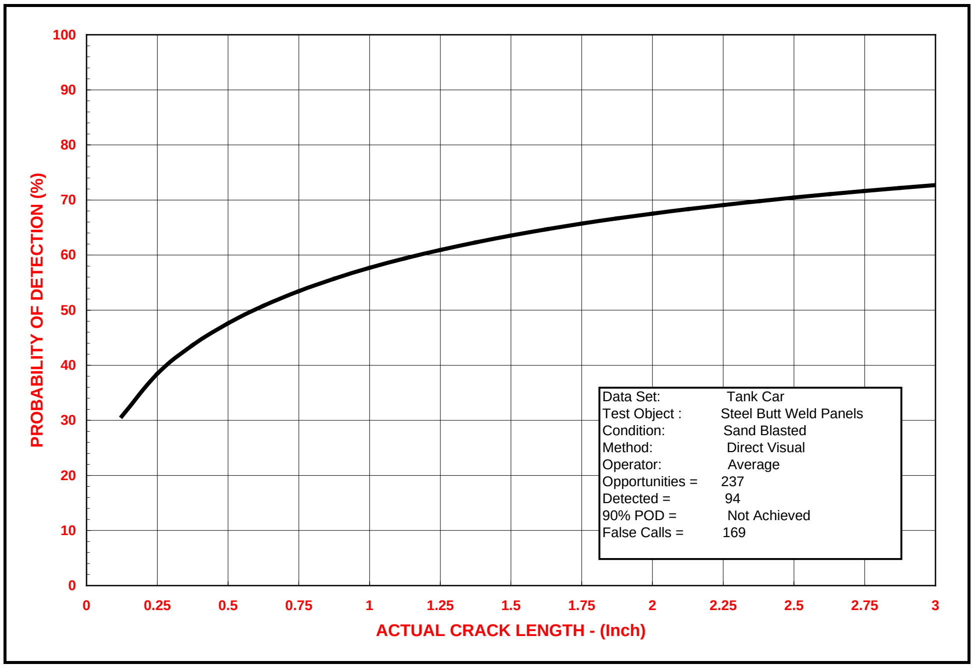


Figure 19. DVT Butt Weld POD Combined Average for Company B

DVT Fillet Weld POD Results

Fillet weld POD evaluation results for industry participants A and B are provided below. The results show the individual results of each company's participants and the combined average for the company. The participants performed setup, which included calibration in accordance with applicable company procedures. After calibration, the participants began to inspect the fillet weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections

DVT Fillet Weld POD Results for Company A

Company A had three technicians participate in the DVT PODs. Table 4 lists the POD results at different crack lengths. Figures 20 through 22 show the POD curves for each operator. Figure 23 shows a comparison of the results of each operator, and Figure 24 shows the combined average of all three operators.

Table 4. Company A Fillet Weld DVT POD Percentages (%)

Flaw Size (inch)	Techncian 6	Operator 17	Operator 19	Combined Results
0.25	50	28	31	36
0.50	50	40	38	43
0.75	51	45	41	46
1.00	51	51	44	48
1.25	51	53	45	50
1.50	52	57	47	52
1.75	52	59	48	53
2.00	52	61	49	54
2.25	52	62	50	55
2.50	52	64	51	56
2.75	52	65	52	56
3.00	52	66	53	57
3.25	52	67	54	58
3.50	53	68	54	58
3.75	53	69	55	59
4.00	53	70	55	59
4.25	53	71	56	60
4.50	53	72	56	60
4.75	53	72	57	61
5.00	53	73	57	61
5.25	53	73	58	61
5.50	53	74	58	62
5.75	53	74	58	62
6.00	53	75	59	62
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	5	3	4	12

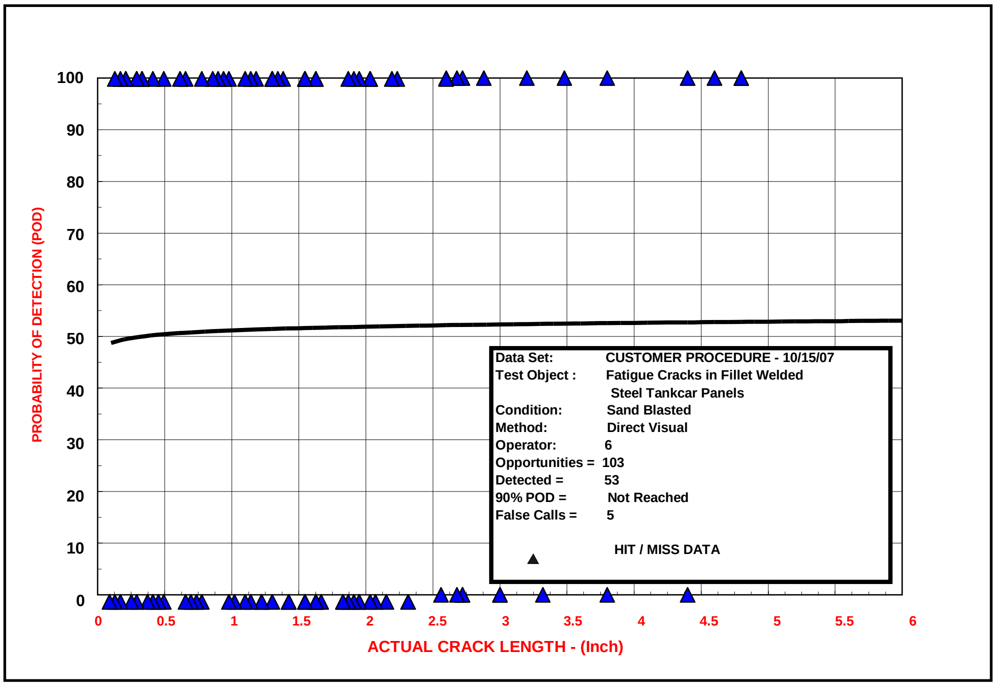


Figure 20. DVT Fillet Weld POD Results for Operator 6

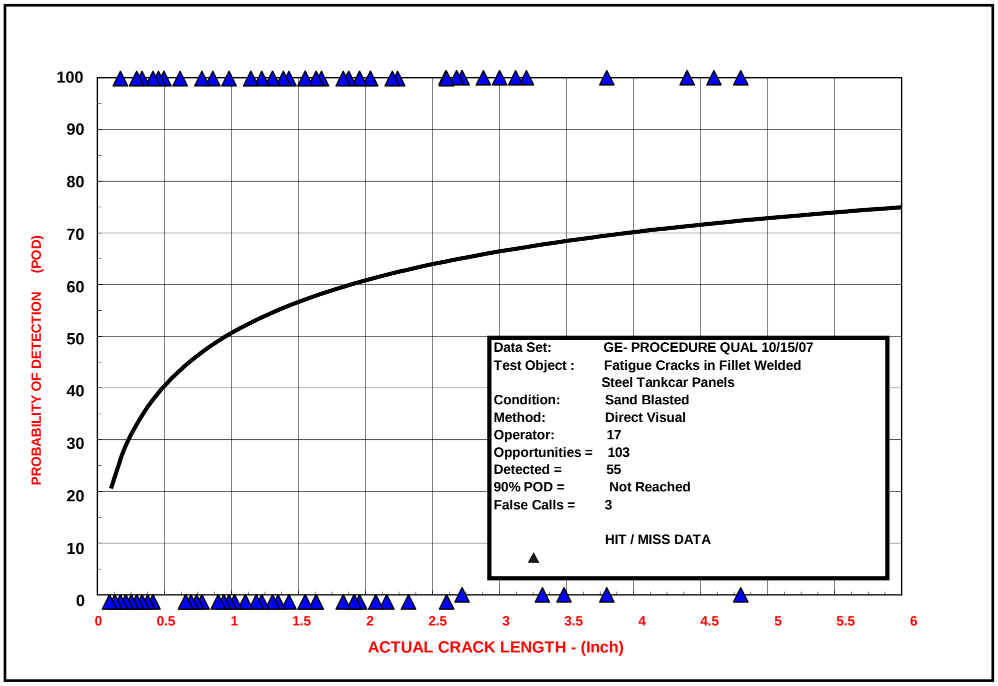


Figure 21. DVT Fillet Weld POD Results for Operator 17

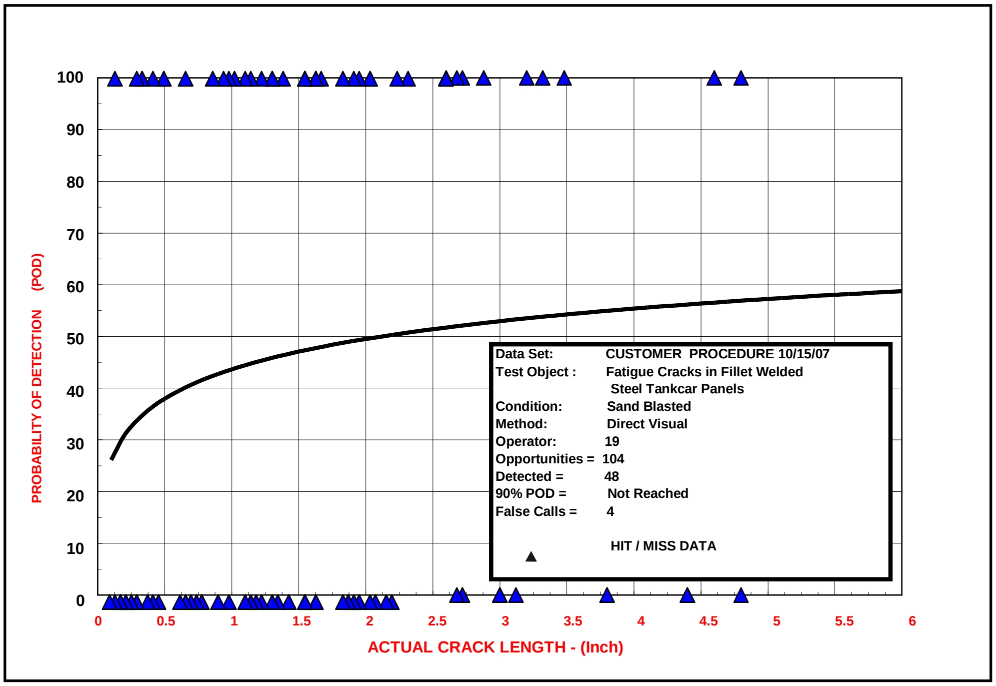


Figure 22. DVT Fillet Weld POD Results for Operator 19

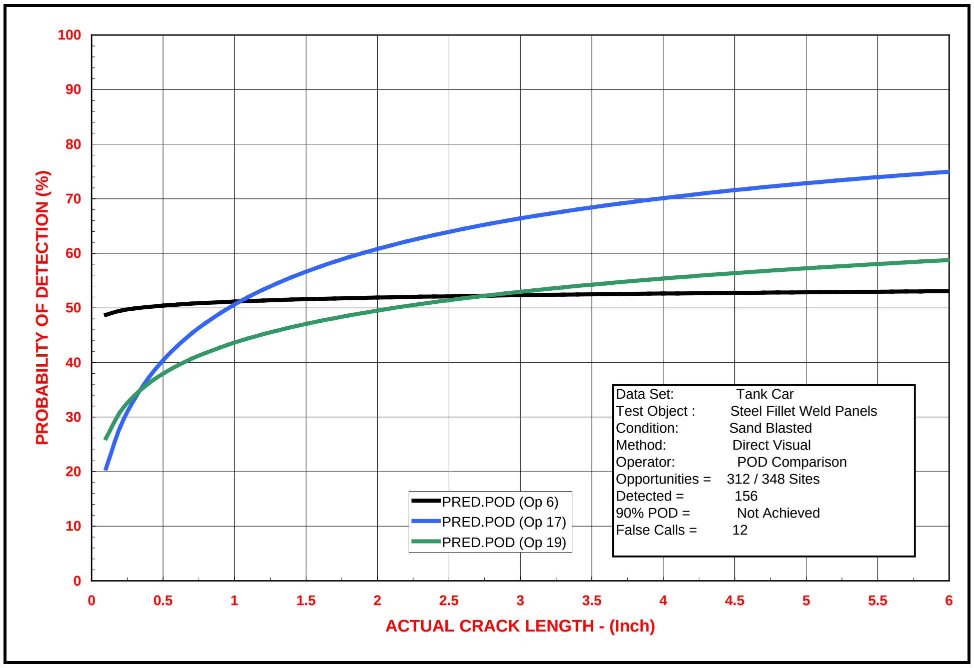


Figure 23. DVT Fillet Weld POD Company A Operator Comparisons

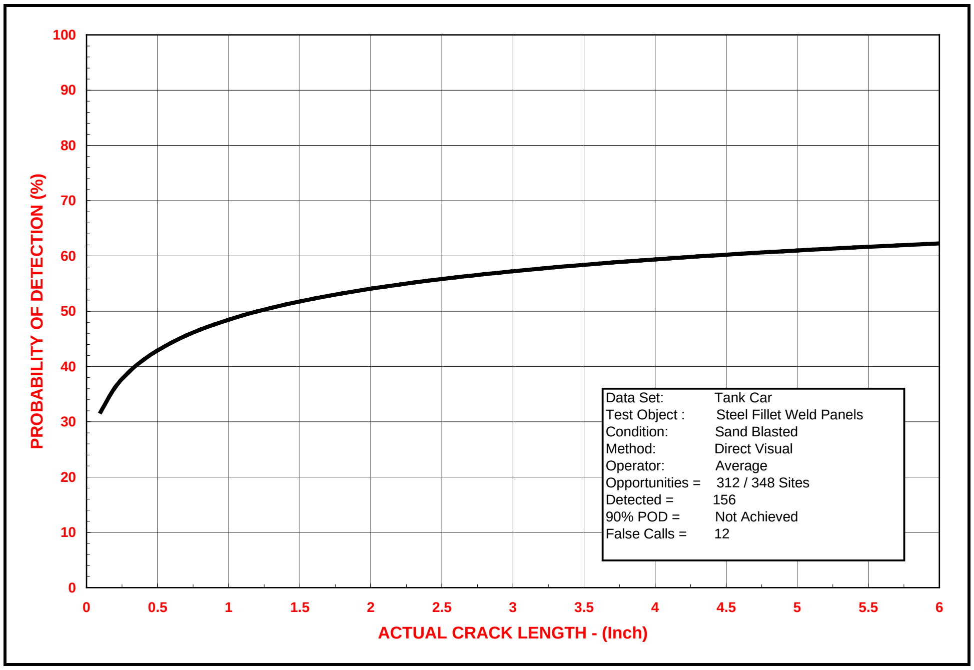


Figure 24. DVT Fillet Weld POD Combined Average for Company A

DVT Fillet Weld POD Results for Company B

Company B had three technicians participate in the DVT PODs. Table 5 lists the POD results at different crack lengths. Figures 25 through 27 show the POD curves for each of the operators. Figure 28 shows a comparison of the results of each operator, and Figure 29 shows the combined average of all three operators.

Table 5. Company B Fillet Weld DVT POD Percentages (%)

Flaw Size (inch)	Operator 15	Operator 20	Operator 27	Combined Results
0.25	63	66	44	58
0.50	64	65	51	60
0.75	65	65	53	61
1.00	65	65	55	62
1.25	65	65	56	62
1.50	66	64	58	63
1.75	66	64	58	63
2.00	66	64	60	63
2.25	66	64	60	63
2.50	66	64	61	64
2.75	66	64	61	64
3.00	67	64	62	64
3.25	67	64	62	64
3.50	67	64	63	64
3.75	67	64	63	65
4.00	67	64	64	65
4.25	67	64	64	65
4.50	67	63	65	65
4.75	67	63	65	65
5.00	67	63	65	65
5.25	67	63	65	65
5.50	67	63	66	65
5.75	67	63	66	66
6.00	67	63	66	66
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	2	2	3	7

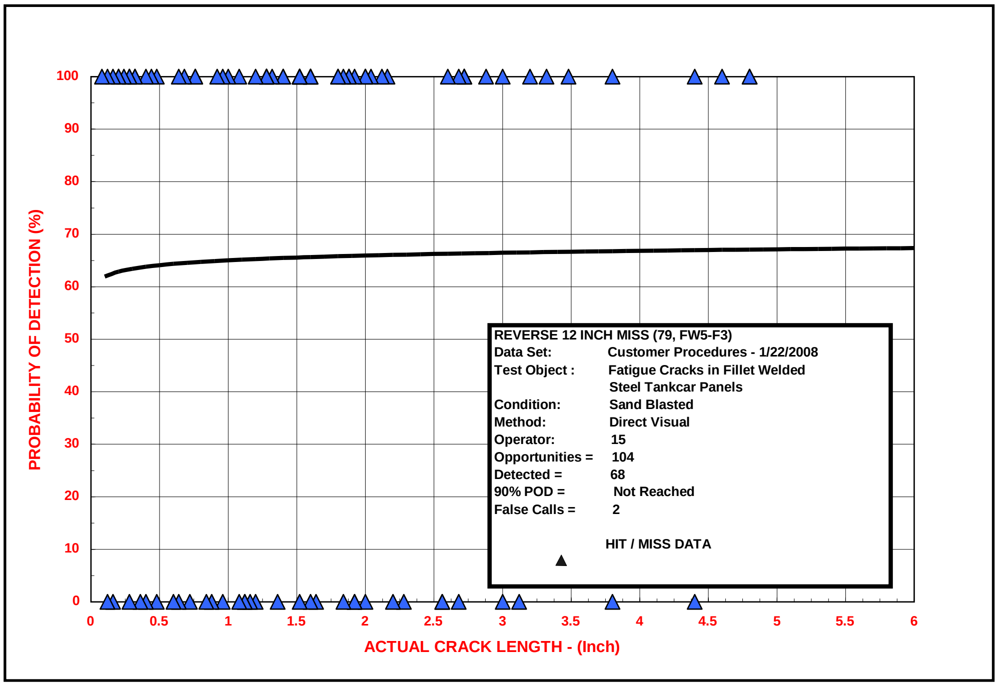


Figure 25. DVT Fillet Weld POD Results for Operator 15

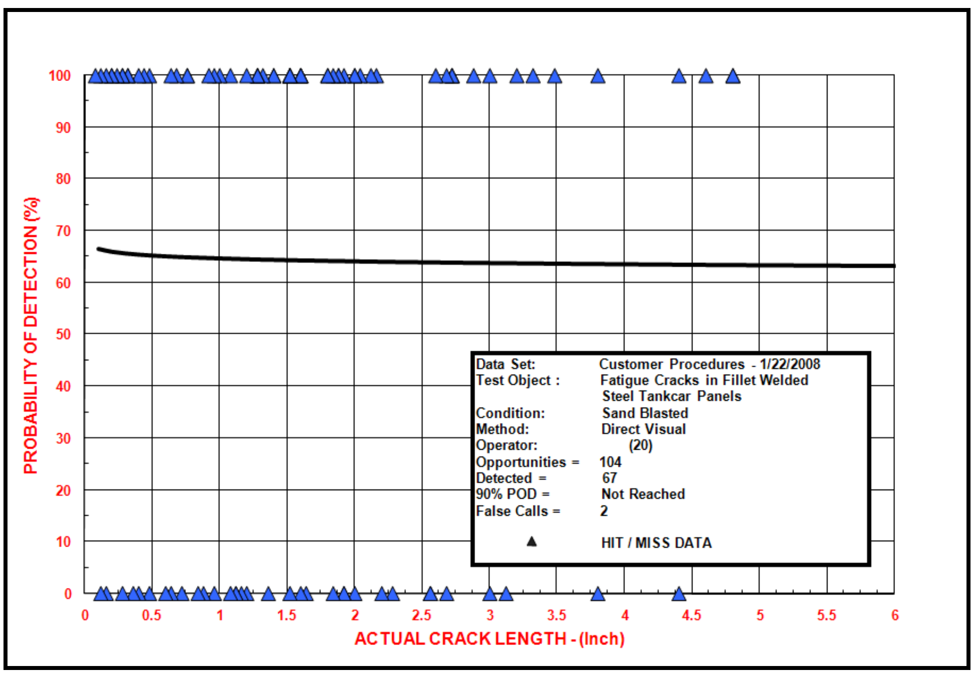


Figure 26. DVT Fillet Weld POD Results for Operator 20

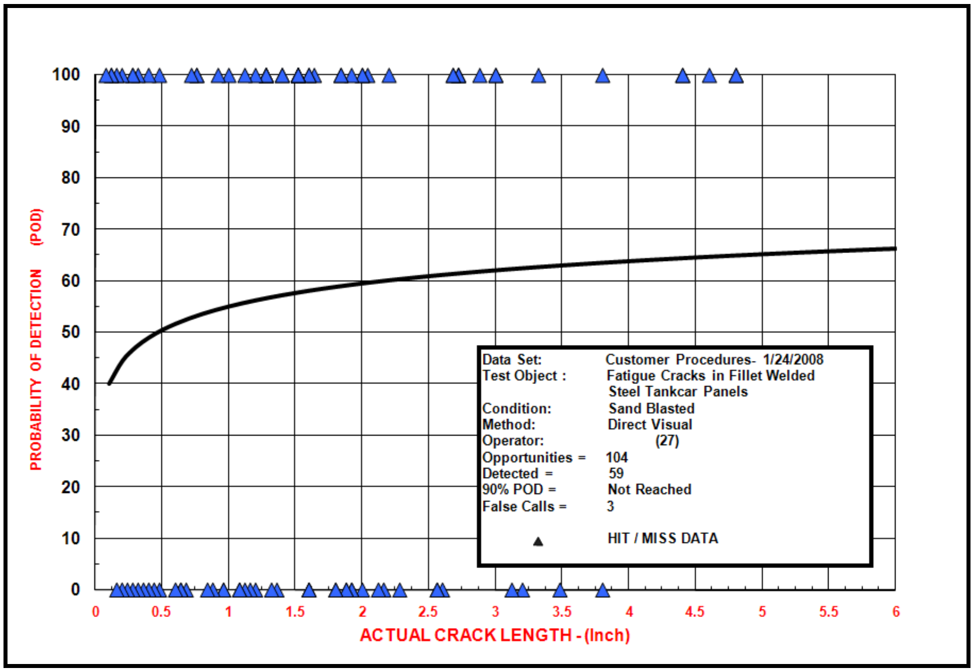


Figure 27. DVT Fillet Weld POD Results for Operator 27

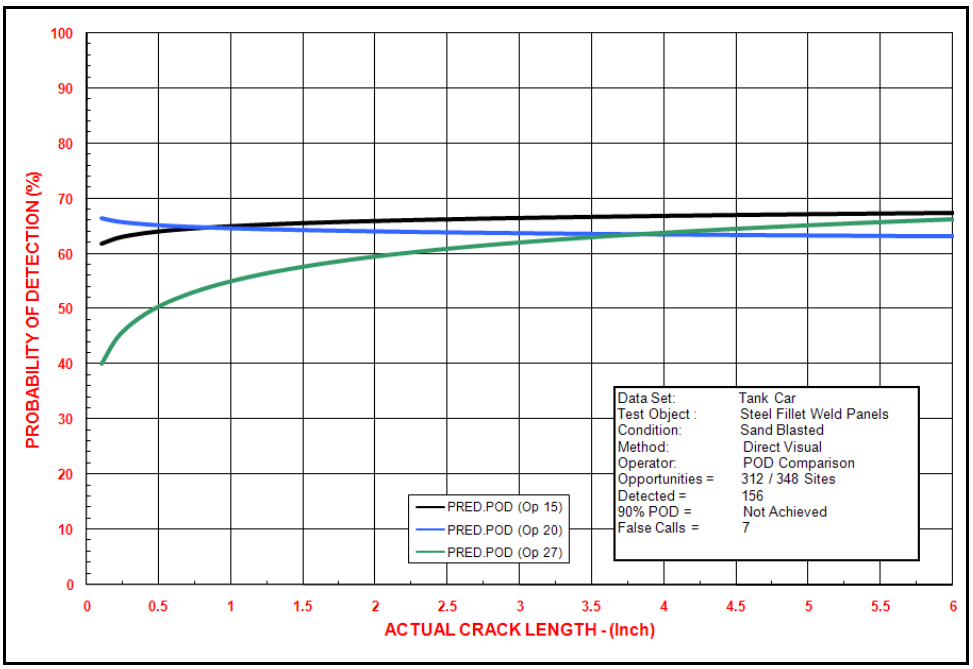


Figure 28. DVT Fillet Weld POD Company B Operator Comparisons

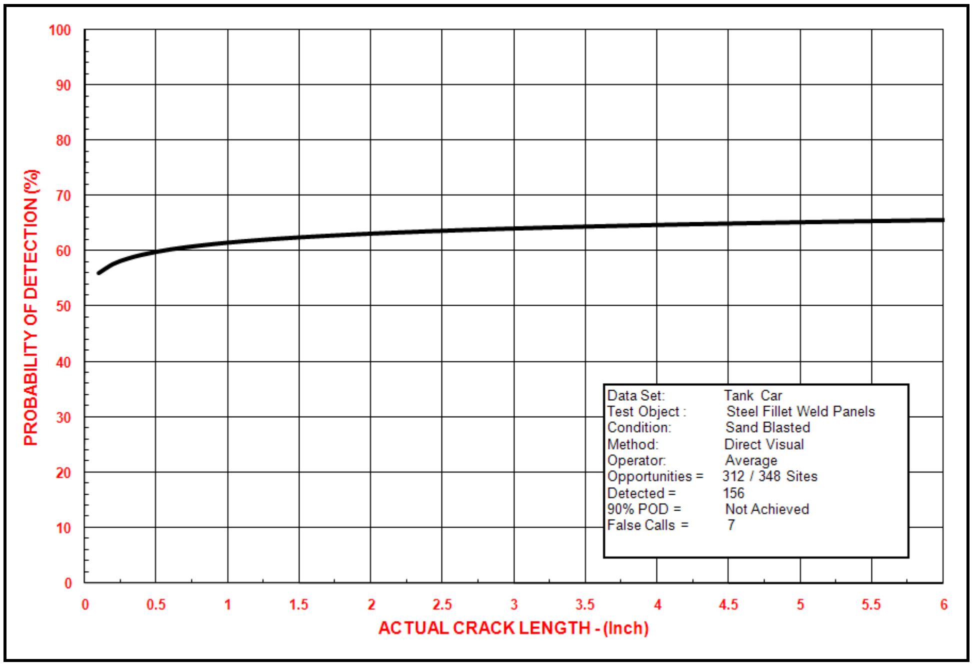


Figure 29. DVT Fillet Weld POD Combined Average for Company B

DVT Observations

Butt weld DVT results show variability among operators. The range among the operators with the lowest and highest POD percentages was 18 to 52 percent at the smallest crack size of 0.25 in. and 25 to 78 percent at the largest crack size of 3.00 in. Tables 1 through 5 list the POD percentages by flaw size for each of the participating operators. Figures 10 to 29 are POD charts for each of the operators; Figures 13, 18, 23, and 28 compares their results and Figures 14, 19, 24, and 29 showing the operator's combined averages.

The visual inspection method is economical, but it is affected by the operator's visual acuity and other sensory factors. The target flaw size for butt welds is 0.50 in, and results show a difference of 40 percent between the highest and lowest performers at that crack size with the average among all three operators being 38 percent. The combined average results for the operators showed the POD range from smallest to largest crack size to be 34 to 50 percent.

2.3.2 Remote Visual

Remote visual testing (RVT) uses electromagnetic radiation at the same visible frequencies as DVT, because it is an enhanced version of visual testing that uses optical aids. The main difference between DVT and RVT is the use of optical aids such as borescopes, fiberscopes, cameras, and other visual enhancement equipment that allows the inspector to get to locations not readily accessible without the use of these tools. Interrogation that employs RVT also incorporates looking, feeling, and smelling to assist in the interpretation of the condition of the item being inspected as does DVT.

The RVT method's advantages are that industrial borescopes and fiberscopes allow operators to inspect remote or confined areas that basic aids cannot reach. Automated equipment such as scopes or other video technology also provides real-time documentation during inspection. Limitations of RVT are that it is applicable to external or surface conditions only, and it is dependent on the operator's visual acuity and expertise with the optical equipment. The RVT method can provide effective determination of many surface discontinuities, but is again highly dependent upon lighting and operator influences.

RVT Fillet Weld POD Results

Fillet weld RVT POD evaluation results for industry participants A and B are provided below. The results show the results of each company's participants and the combined average for the company. Before the test panels were inspected, the participants performed setup, which included calibration in accordance with the applicable company procedure. After calibration the participants began their inspections of the weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

RVT Fillet Weld POD Results for Company A

Company A had three technicians participate in the RVT PODs. Table 6 lists the POD results at different crack lengths. Figures 30 through 32 show the POD curves for each operator. Figure 33 shows a comparison of the results of each operator, and Figure 34 shows the combined average of all three operators.

Table 6. Company A Fillet Weld RVT POD Percentages (%)

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	15	17	58	30
0.50	17	17	53	29
0.75	18	17	51	29
1.00	19	17	50	29
1.25	20	17	49	29
1.50	20	17	47	28
1.75	21	17	47	28
2.00	21	17	46	28
2.25	22	17	45	28
2.50	22	17	45	28
2.75	22	17	44	28
3.00	23	17	44	28
3.25	23	17	43	28
3.50	23	17	43	28
3.75	23	17	43	28
4.00	23	17	42	28
4.25	24	17	42	28
4.50	24	17	42	28
4.75	24	17	41	28
5.00	24	17	41	28
5.25	24	17	41	28
5.50	25	17	41	28
5.75	25	17	40	28
6.00	25	17	40	27
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	5	2	2	9

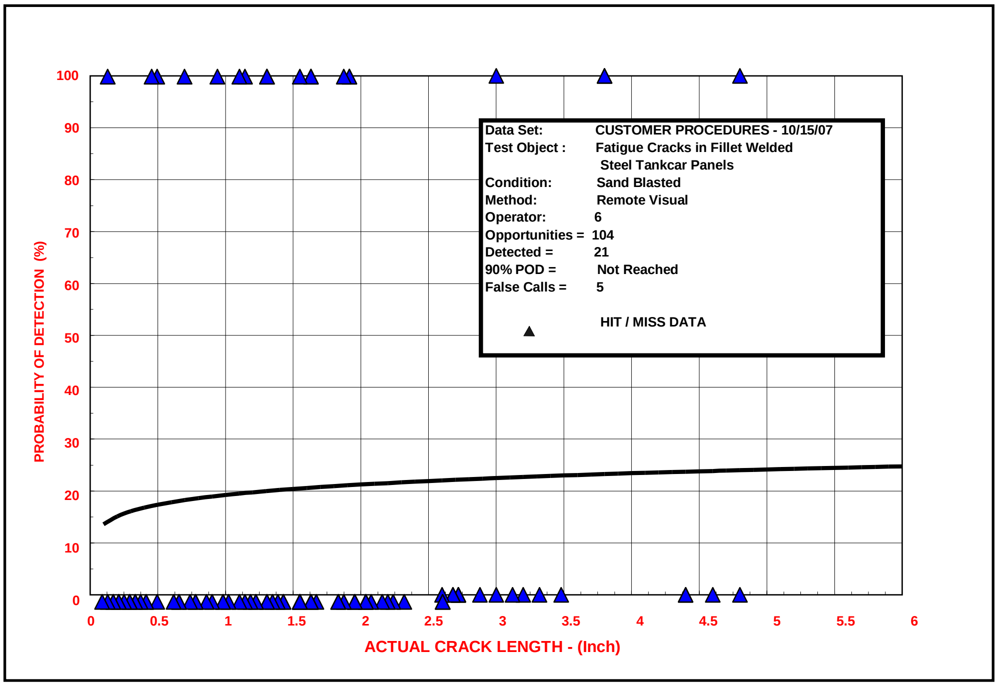


Figure 30. RVT Fillet Weld POD Results for Operator 6

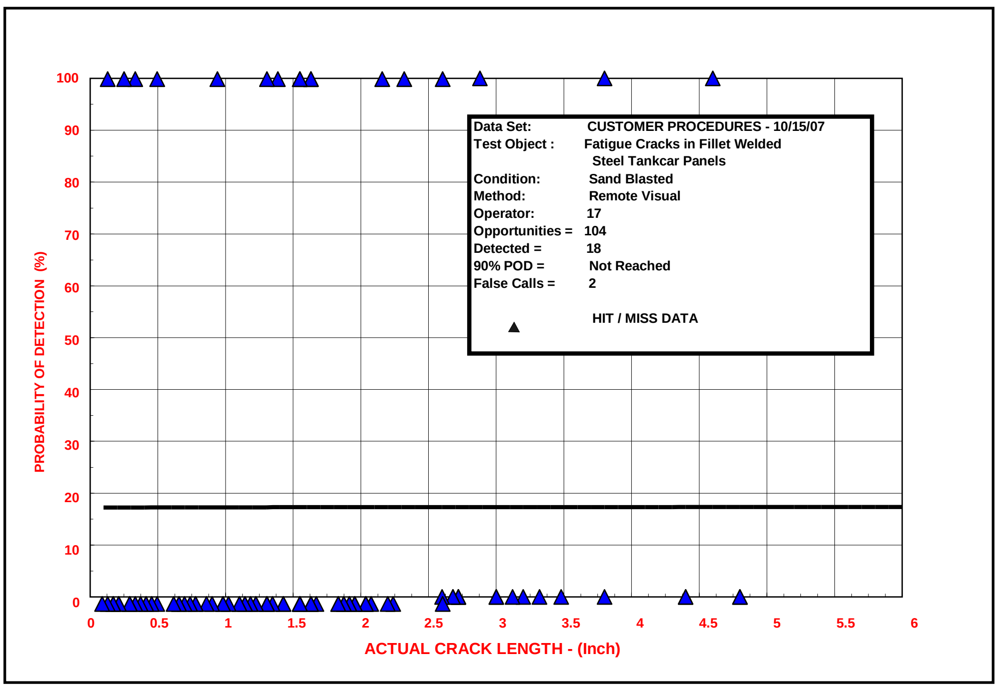


Figure 31. RVT Fillet Weld POD Results for Operator 17

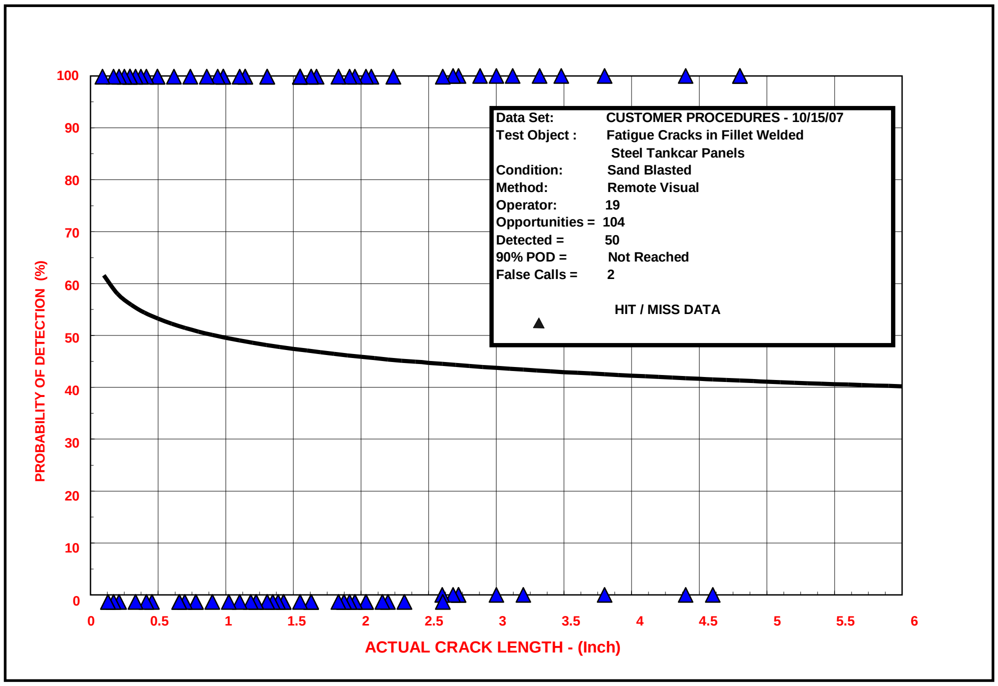


Figure 32. RVT Fillet Weld POD Results for Operator 19

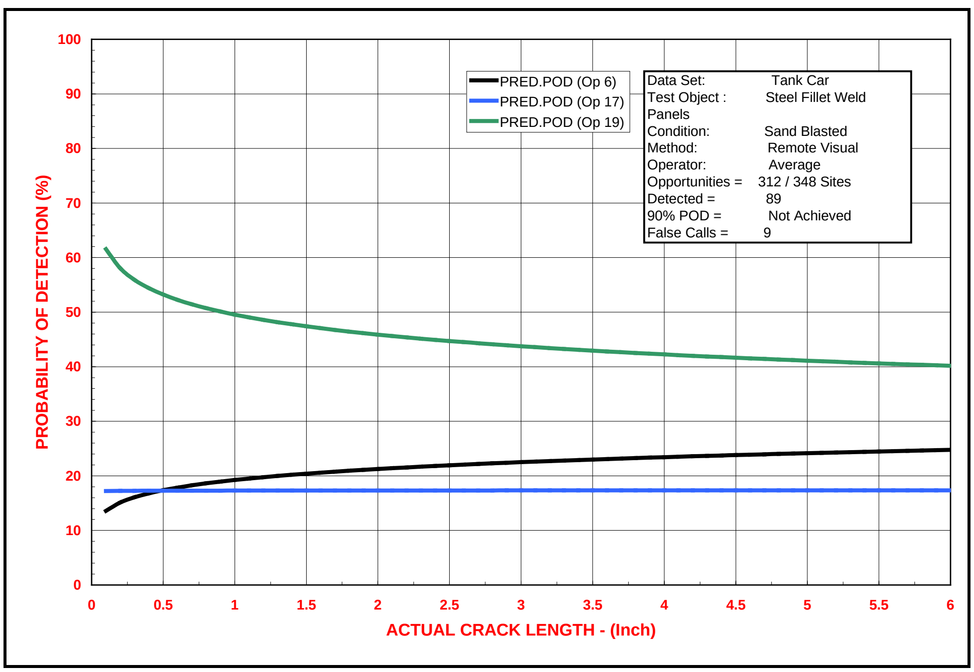


Figure 33. RVT Fillet Weld POD Operator Comparisons

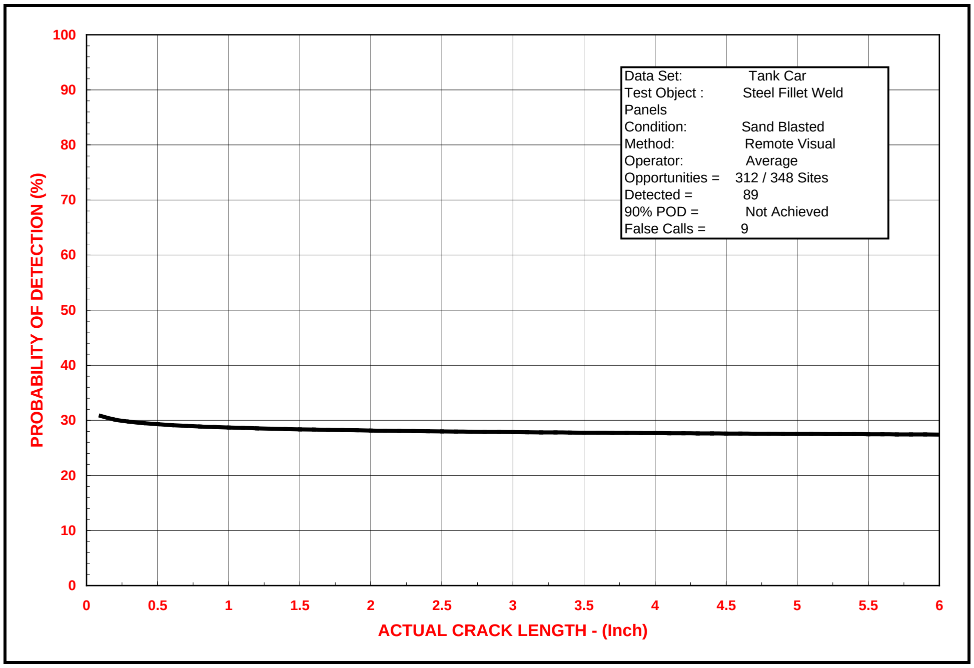


Figure 34. RVT Fillet Weld POD Combined Average

RVT Fillet Weld POD Results for Company B

Company B had four technicians participate in the RVT PODs. Table 7 lists the POD results at different crack lengths. Figures 35 through 38 show the POD curves for each operator. Figure 39 shows a comparison of the results of each operator, and Figure 40 shows the combined average of all four operators.

Table 7. Company B Fillet Weld RVT POD Percentages (%)

Flaw Size	Operator 15	Operator 16	Operator 21	Operator 27	Combined Results
0.25	48	11	98	59	54
0.50	48	14	97	68	57
0.75	49	16	97	71	58
1.00	49	18	97	74	59
1.25	49	19	96	76	60
1.50	49	20	96	77	61
1.75	49	21	96	78	61
2.00	49	22	96	79	62
2.25	49	23	97	79	62
2.50	50	24	96	80	62
2.75	50	24	95	81	62
3.00	50	25	95	81	63
3.25	50	25	95	82	63
3.50	50	26	95	82	63
3.75	50	27	95	82	63
4.00	50	27	95	83	64
4.25	50	28	95	83	64
4.50	50	28	95	84	64
4.75	50	28	95	84	64
5.00	5	29	95	84	64
5.25	50	29	94	84	65
5.50	50	30	94	85	65
5.75	50	30	94	85	65
6.00	50	30	94	85	65
90% POD	Not Achieved	Not Achieved	Not Valid	Not Achieved	Not Achieved
False Calls	8	2	9	5	24

Note: Results for Operator 21 are skewed as every weld termination was called a defect. Operator 15 results are also skewed because of a higher number of false calls.

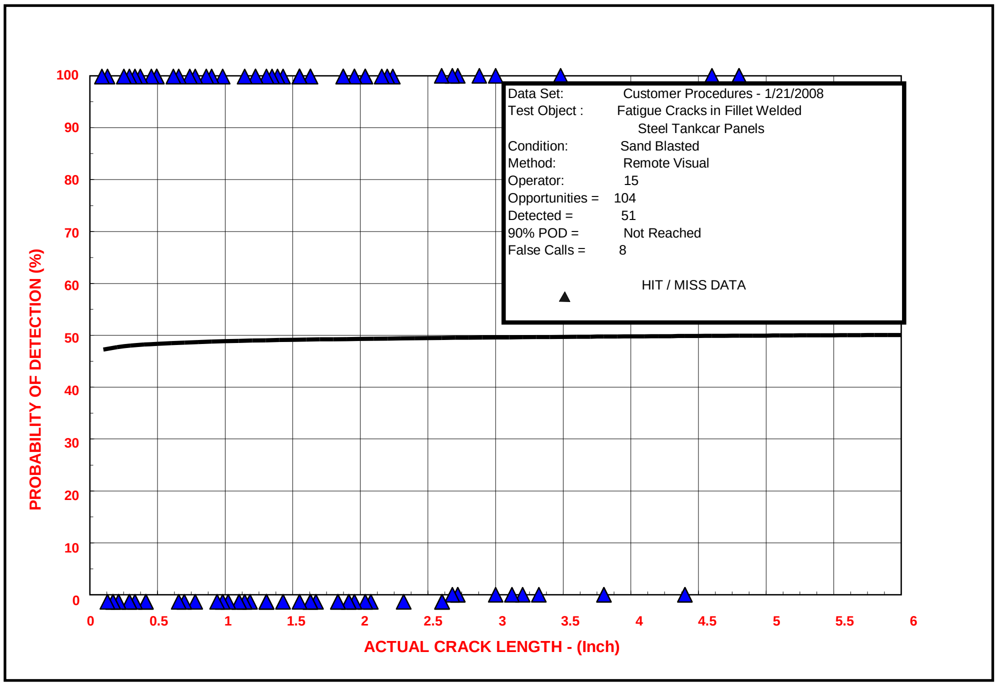


Figure 35. RVT Fillet Weld POD Results for Operator 15

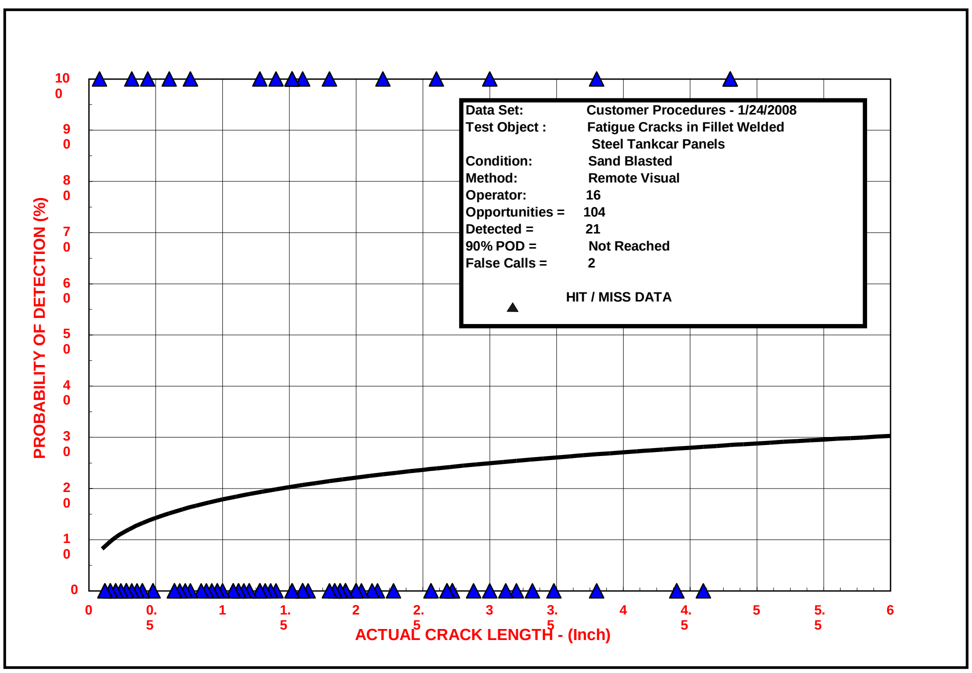


Figure 36. RVT Fillet Weld POD Results for Operator 16

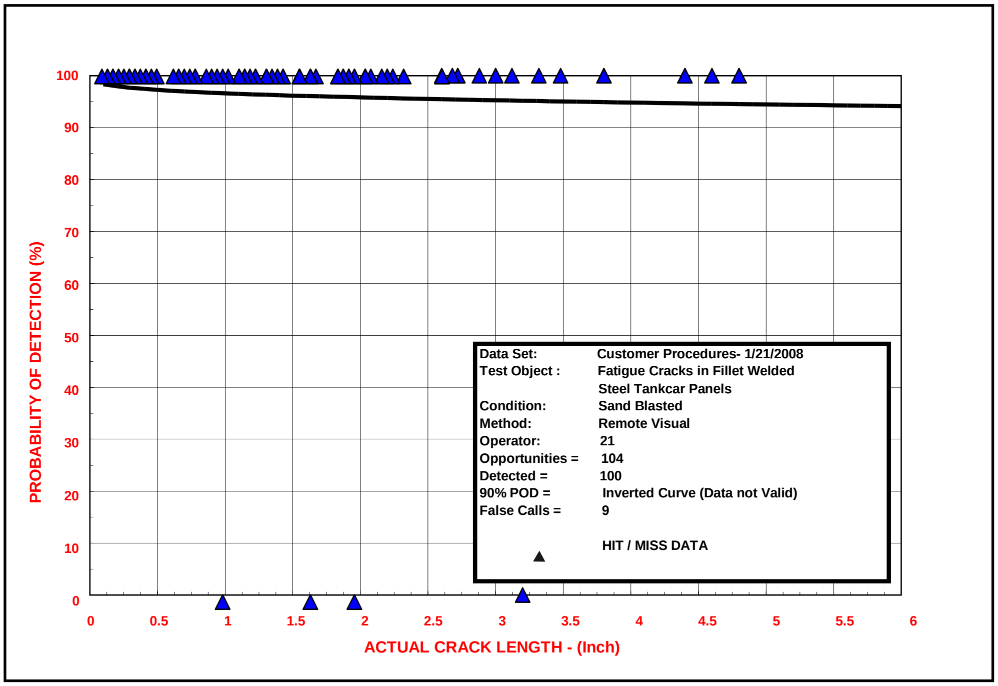


Figure 37. RVT Fillet Weld POD Results for Operator 21

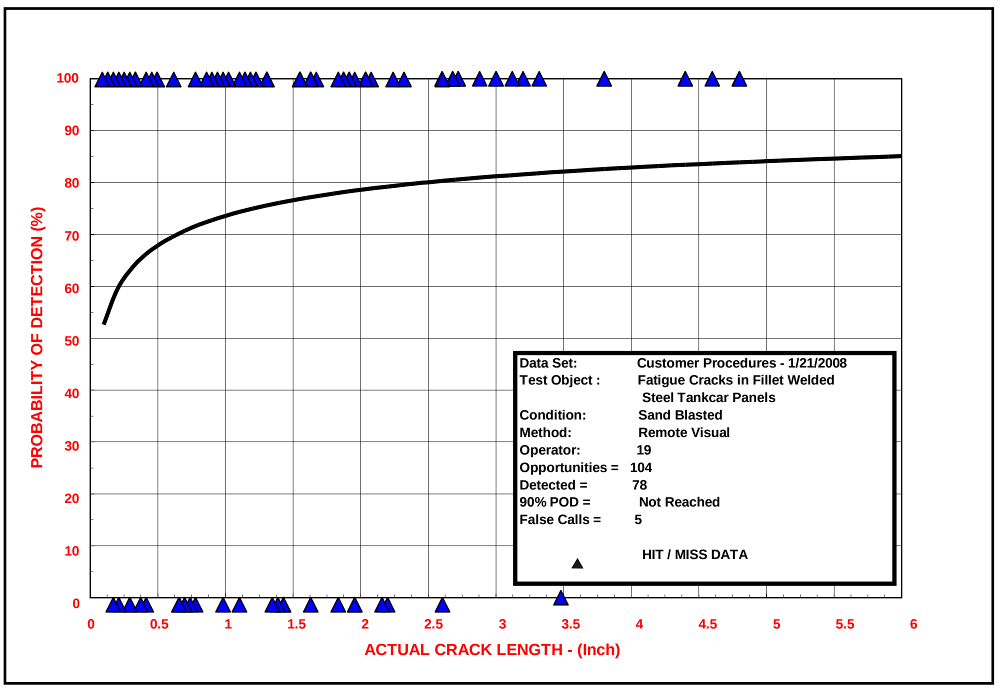


Figure 38. RVT Fillet Weld POD Results for Operator 27

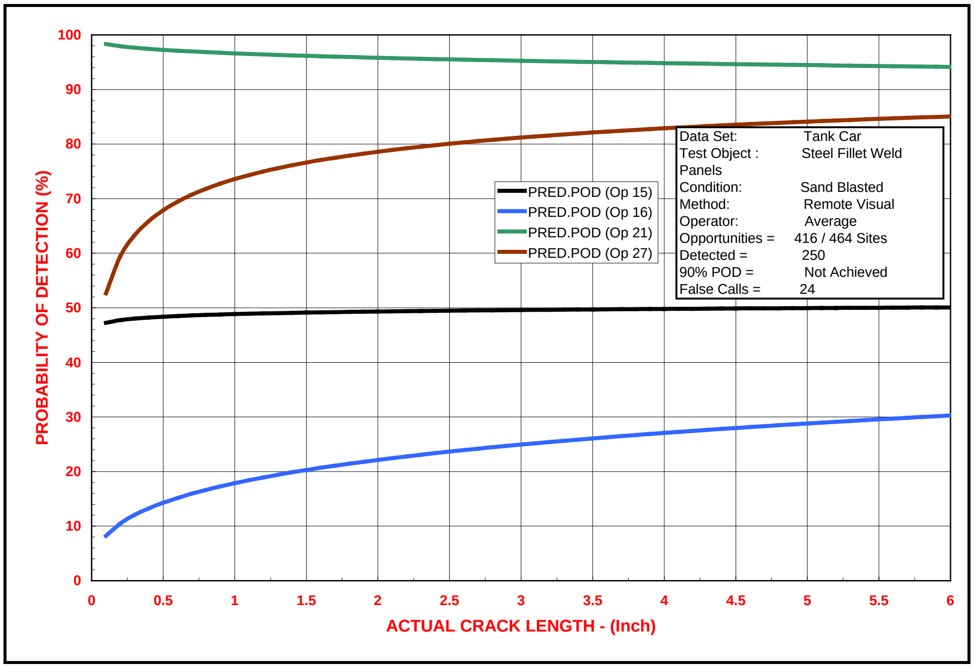


Figure 39. RVT Fillet Weld POD Company B Operator Comparisons

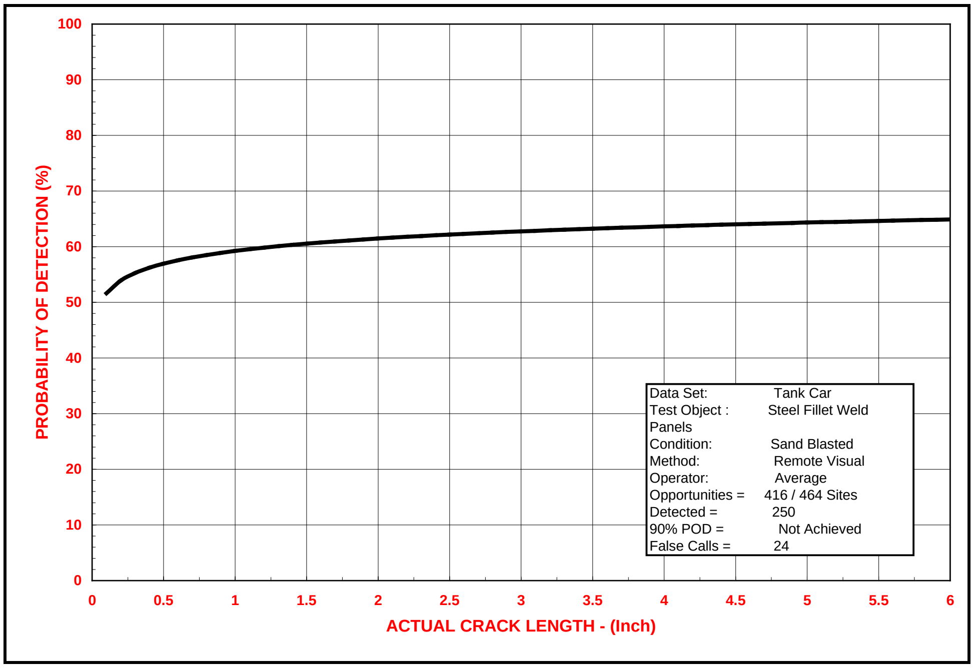


Figure 40. RVT Fillet Weld POD Combined Average for Company B

RVT Observations

RVT POD results for the fillet weld inspections show variability among the operators. Results for Operators 6 and 17 correlate with each other fairly consistently but do not vary much in POD from the smallest to largest crack size. Operator 19 shows a higher POD at smaller flaw sizes with a drop in POD as crack size becomes larger. The combined average of the operators shows the influence from both sets of results as the POD does not vary much from the smallest crack size of 0.25 in. (POD 30 percent) to the largest crack size of 6 in. (POD 27 percent), with POD highest at the smallest crack size and lowest at the largest.

2.3.3 Liquid Penetrant

Liquid penetrant (PT) is a physical and chemical NDT process that exposes discontinuities that are open to the surface by taking advantage of capillary interaction between the penetrating liquid and the surface of the part being inspected. The liquid enters surface cavities and later emerges as visual evidence of discontinuities such as cracks, porosity, laps, or seams.

PT's can be performed rapidly, simply, and over a large area (i.e., the complete surface of a part can be inspected at once). It is also economical and can be used on a variety of materials and shapes with minimal capital investment. However, PT is limited to detecting discontinuities open to the surface. It will not reveal the depth of a discontinuity and is adversely affected by foreign substances that may block or seal the cavity restricting the entry of the penetrant into the discontinuity.

The PT method provides a portable and economical NDT approach to evaluate discontinuities open to the surface. Many weld discontinuities found during tank car inspections originate at the surface or slightly below the surface (eventually propagating to the surface) suggesting that PT continues to be a valuable method for tank car inspection. Capability (reliability and repeatability) of inspections can continue to be enhanced through emphasis on operator training, equipment calibrations, and inspection procedures.

PT results show variability among operators and the influence of test articles, but the variability is not as great as it was among the tests with the DVT method. The fillet weld inspection showed higher operator variability than the girth (butt) weld inspections. Detailed results for the butt and fillet weld PT POD evaluations are provided in the following two subsections.

PT Butt Weld POD Results

Butt weld PT POD evaluation results for industry participants A and B are provided below. The results show individual results of each company's participants and the combined average for the company. Before performing inspections of the test panels, the participants performed setup, which included calibration in accordance with applicable company procedures. After calibration, the participants began their inspections of the butt weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

PT Butt Weld POD Results for Company A

Company A had three technicians participate in the PT PODs. Table 8 lists the POD results at different crack lengths. Figures 41 through 43 show the POD curves for each operator.

Figure 44 shows a comparison of the results of each operator, and Figure 45 shows the combined average of all three operators.

Table 8. Company A Butt Weld PT POD Percentages (%)

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	39	36	11	28
0.50	54	52	26	44
0.75	62	61	39	54
1.00	68	68	50	62
1.25	72	72	58	67
1.50	75	75	65	72
1.75	78	78	70	75
2.00	79	80	74	78
2.25	81	82	78	80
2.50	82	83	80	82
2.75	84	85	82	84
3.00	85	86	84	85
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	5	6	3	14

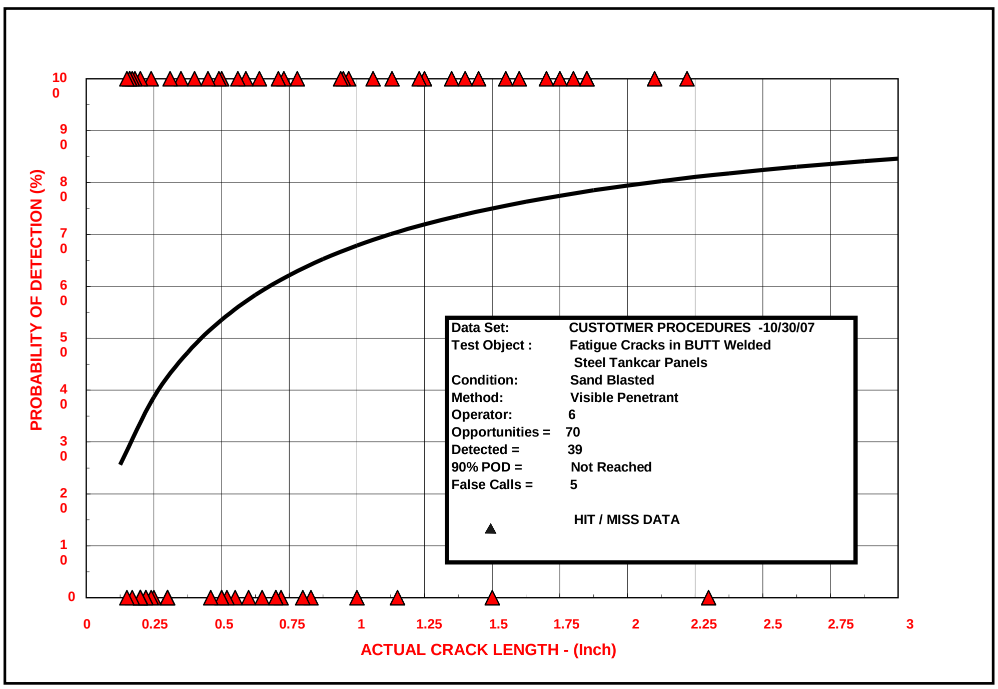


Figure 41. PT Butt Weld POD Results for Operator 6

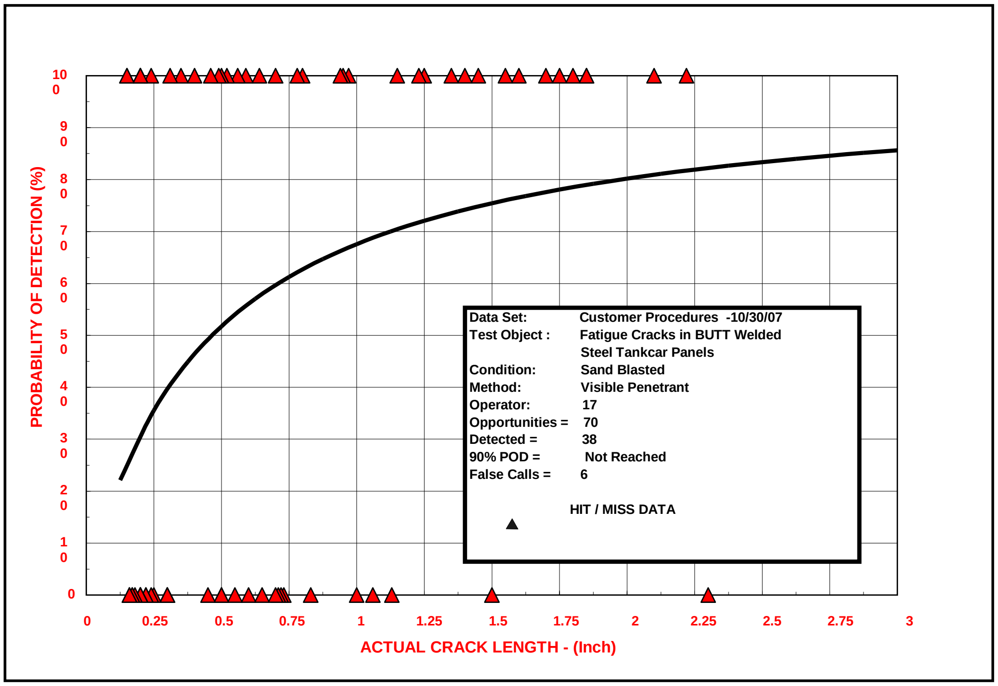


Figure 42. PT Butt Weld POD Results for Operator 17

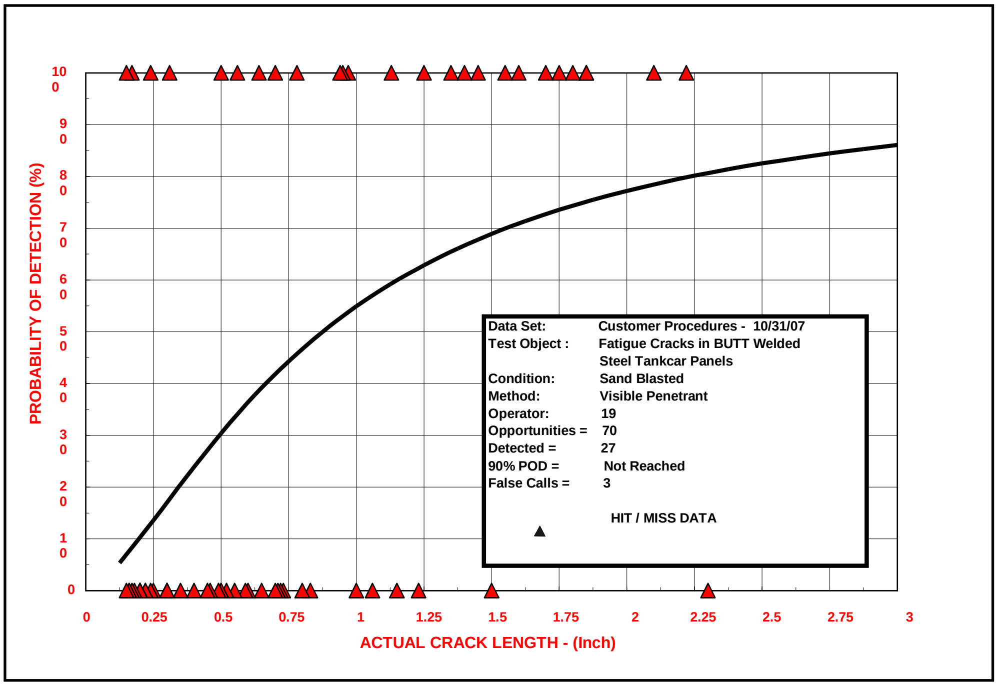


Figure 43. PT Butt Weld POD Results for Operator 19

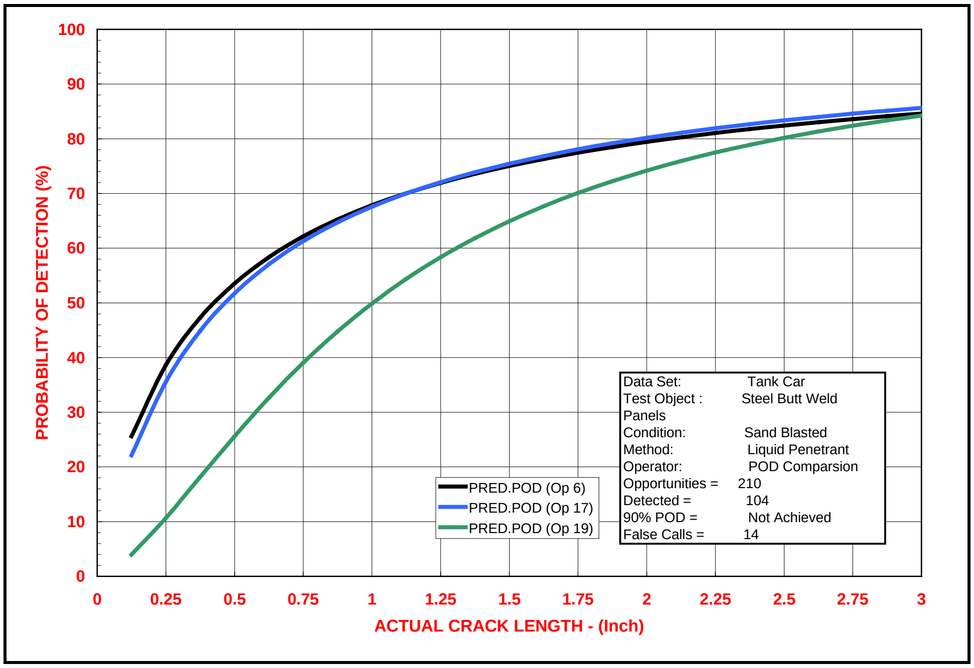


Figure 44. PT Butt Weld POD Operator Comparisons

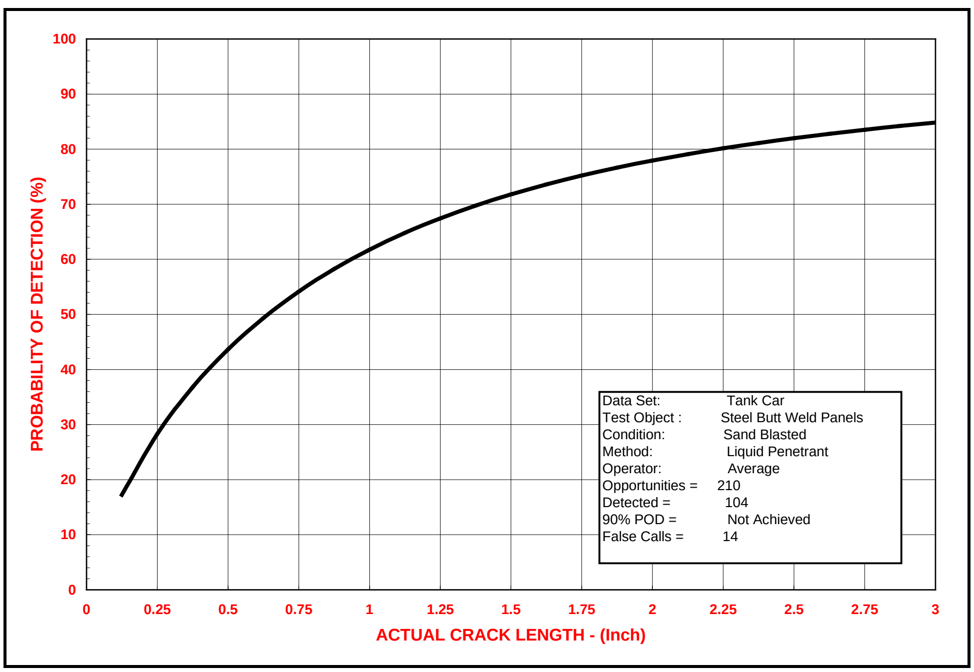


Figure 45. PT Butt Weld POD Average for Company A

PT Butt Weld POD Results for Company B

Company B had three technicians participate in the PT PODs. Table 9 lists the POD results at different crack lengths. Figures 46 through 48 show the POD curves for each operator. Figure

49 compares the operator's results, and Figure 50 shows the combined average of all three operators.

Table 9. Company B Results of Butt Weld PT POD Percentages (%)

Flaw Size (inch)	Operator 15	Operator 21	Operator 22	Combined Results	Industry Average
0.25	58	39	29	41	34
0.50	71	45	45	55	51
0.75	77	58	56	64	61
1.00	80	64	63	69	67
1.25	83	68	68	73	72
1.50	85	72	72	76	75
1.75	86	74	76	79	78
2.00	88	76	78	81	80
2.25	89	78	80	82	82
2.50	89	80	82	84	83
2.75	90	81	83	85	84
3.00	91	82	84	86	85
90% POD	2.63 inches	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	39	6	2	47	Not Recorded

Note: Grey areas depict a high number of false calls that skew the results.

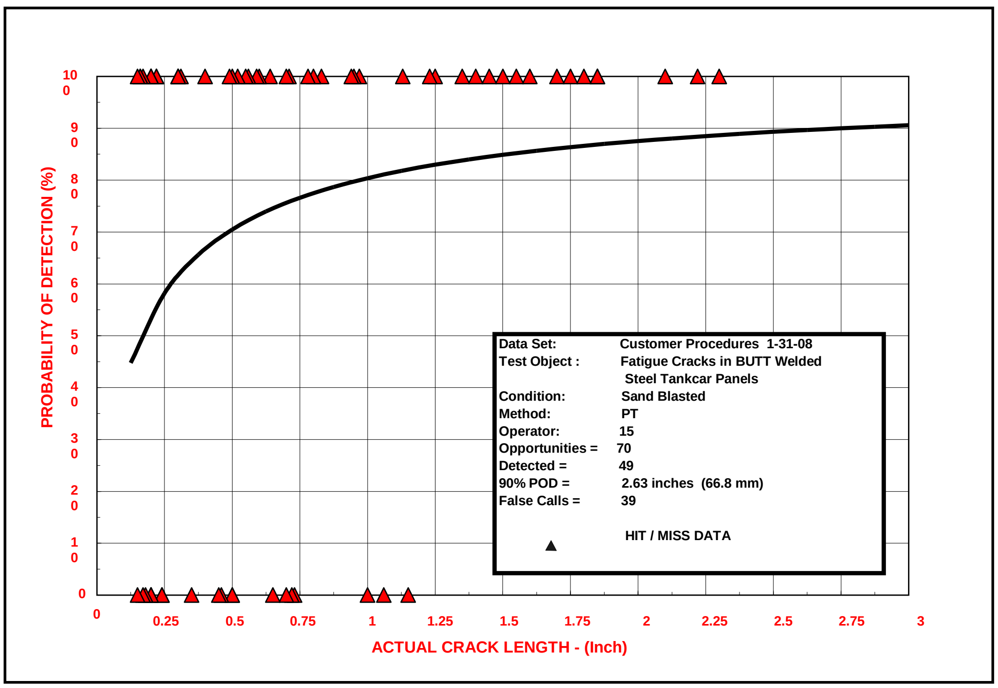


Figure 46. PT Butt Weld POD Results for Operator 15

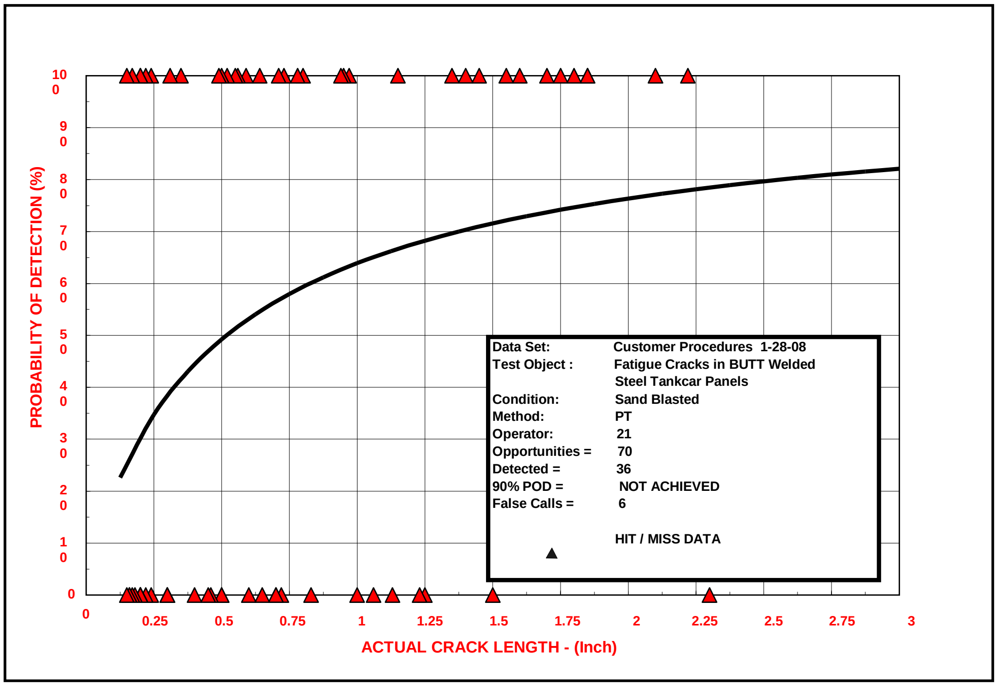


Figure 47. PT Butt Weld POD Results for Operator 21

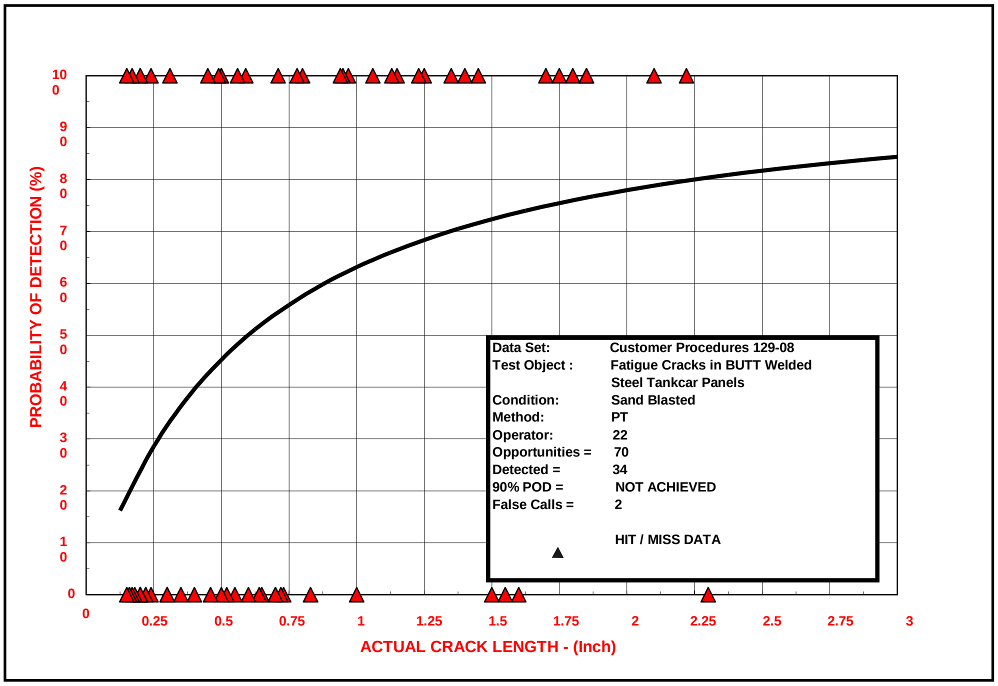


Figure 48. PT Butt Weld POD Results for Operator 22

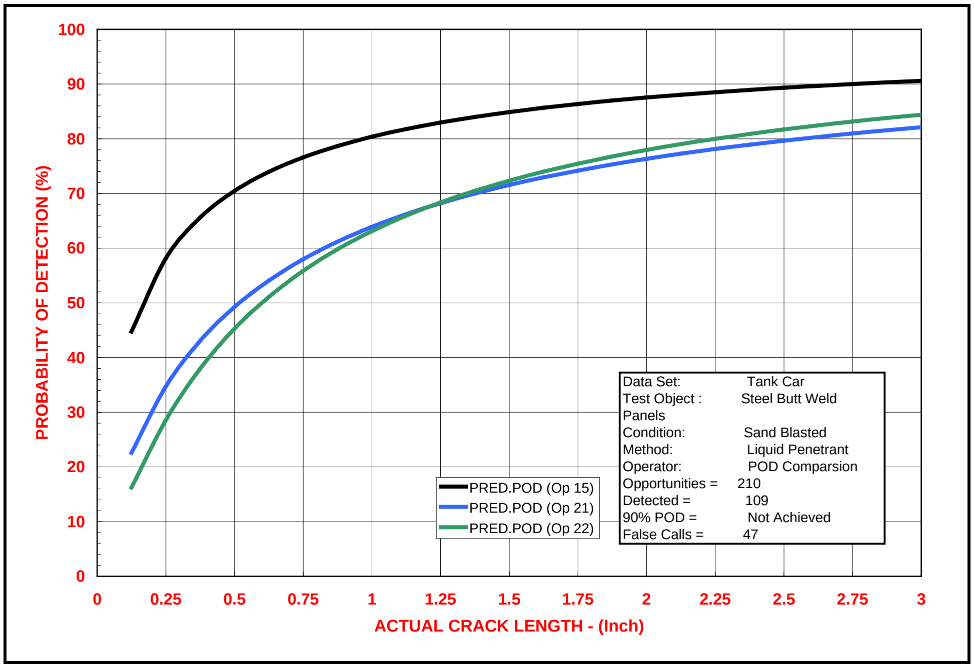


Figure 49. PT Butt Weld POD Company B Operator Comparisons

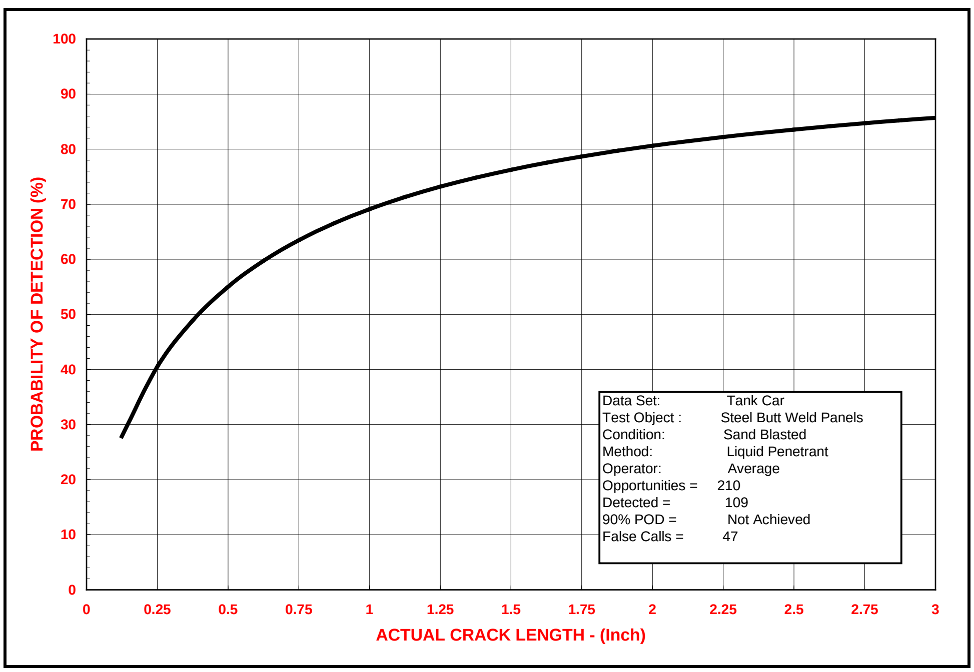


Figure 50. PT Butt Weld POD Combined Average for Company B

PT Fillet Weld POD Results

Fillet weld PT POD evaluation results for industry participants A and B are presented in the pages below. The results show individual results of each company's participants and the combined average for the company. Before the test panels were inspected, the participants performed setup, which included calibration in accordance with applicable company procedures. After calibration, the participants began their inspections of the weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

PT Fillet Weld POD Results for Company A

Company A had three technicians participate in the PT PODs. Table 10 lists the POD results at different crack lengths. Figures 51 through 53 show the POD curves for each operator. Figure 54 compares the operator's results, and Figure 55 shows the combined average of all three operators.

Table 10. Company A Fillet Weld PT POD Percentages (%)

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	50	16	25	30
0.50	51	32	31	38
0.75	52	40	33	41
1.00	52	49	35	45
1.25	52	54	37	47
1.50	52	59	38	50
1.75	52	62	39	51
2.00	52	66	40	53
2.25	53	68	41	54
2.50	53	71	42	55
2.75	53	73	43	56
3.00	53	75	43	57
3.25	53	76	44	58
3.50	53	78	44	58
3.75	53	79	45	59
4.00	53	80	45	60
4.25	53	81	46	60
4.50	53	82	46	61
4.75	53	82	47	61
5.00	54	83	47	61
5.25	54	84	47	62
5.50	54	85	48	62
5.75	54	85	48	62
6.00	54	86	49	63
90% POD	Not Achieved	Not achieved	Not Achieved	Not Achieved
False Calls	2	0	1	3

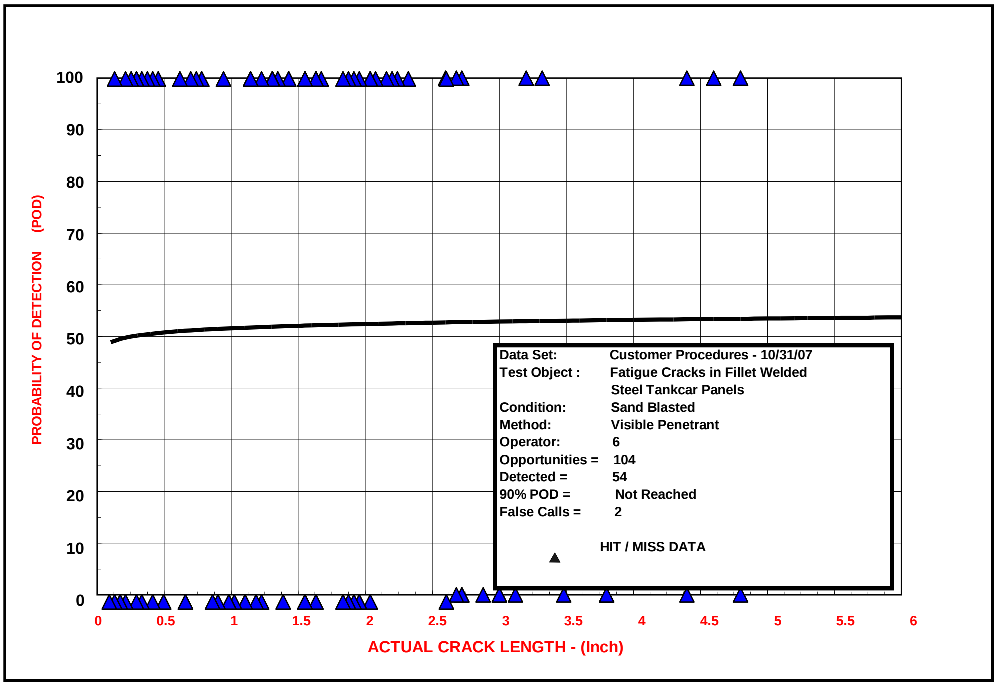


Figure 51. PT Fillet Weld POD Results for Operator 6

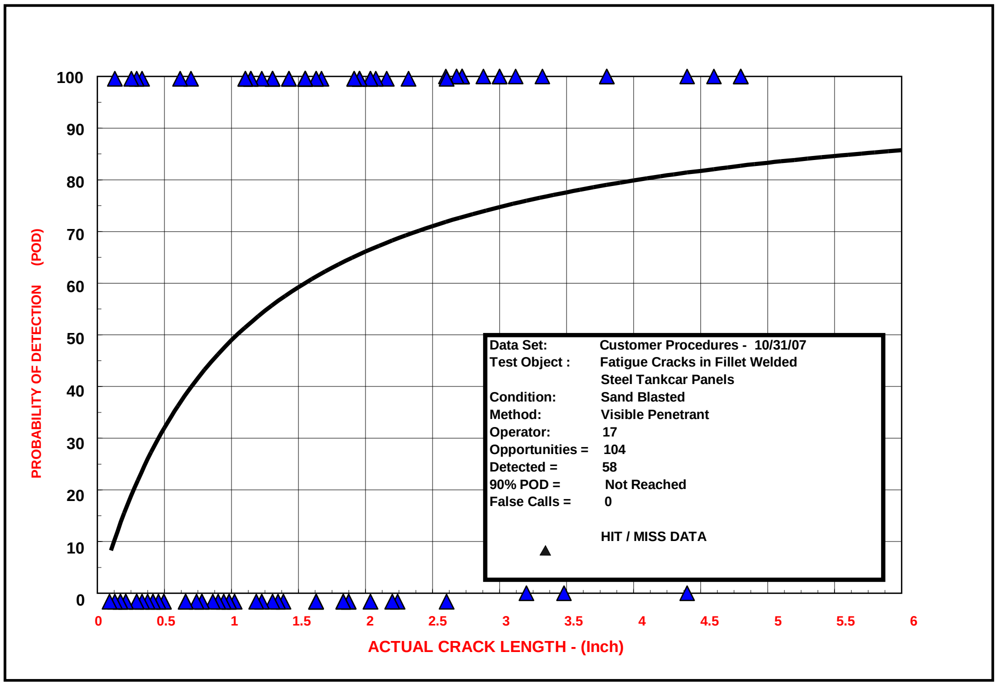


Figure 52. PT Fillet Weld POD Results for Operator 17

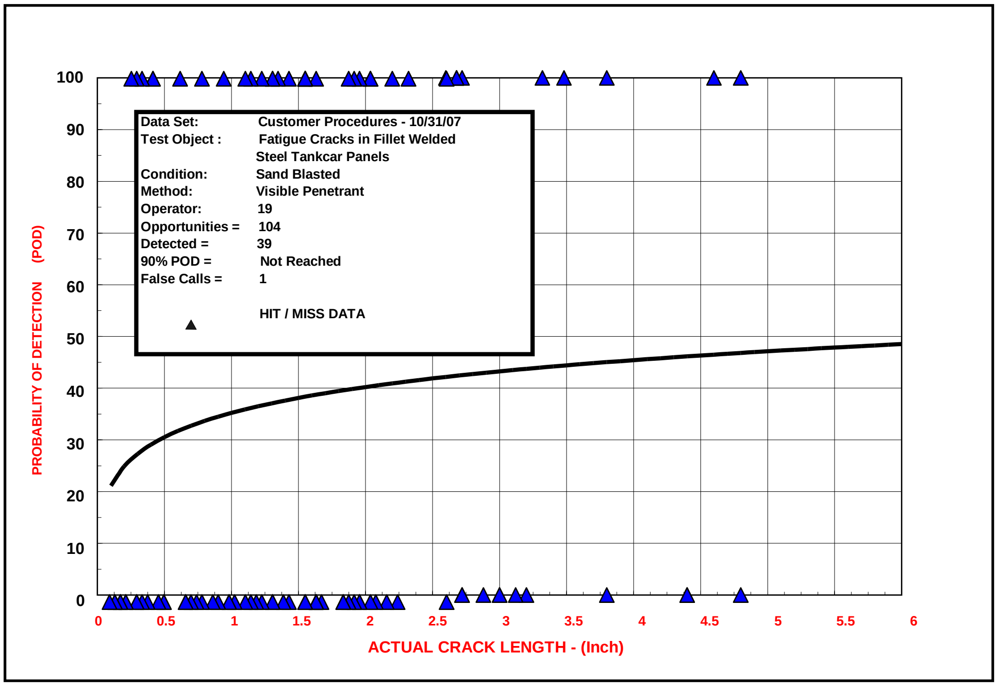


Figure 53. PT Fillet Weld POD Results for Operator 19

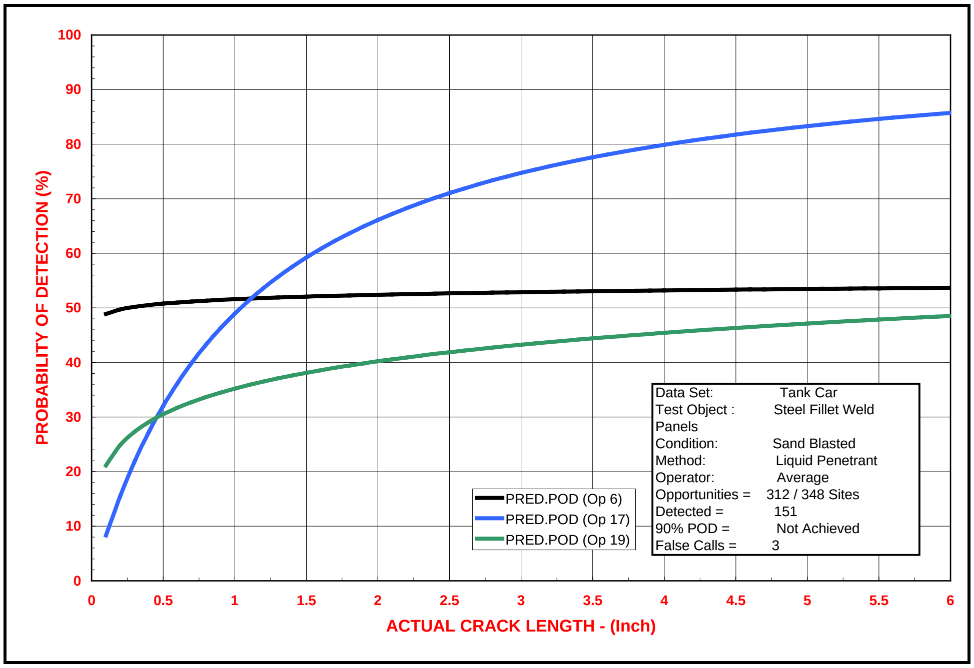


Figure 54. PT Fillet Weld POD Company A Operator Comparisons

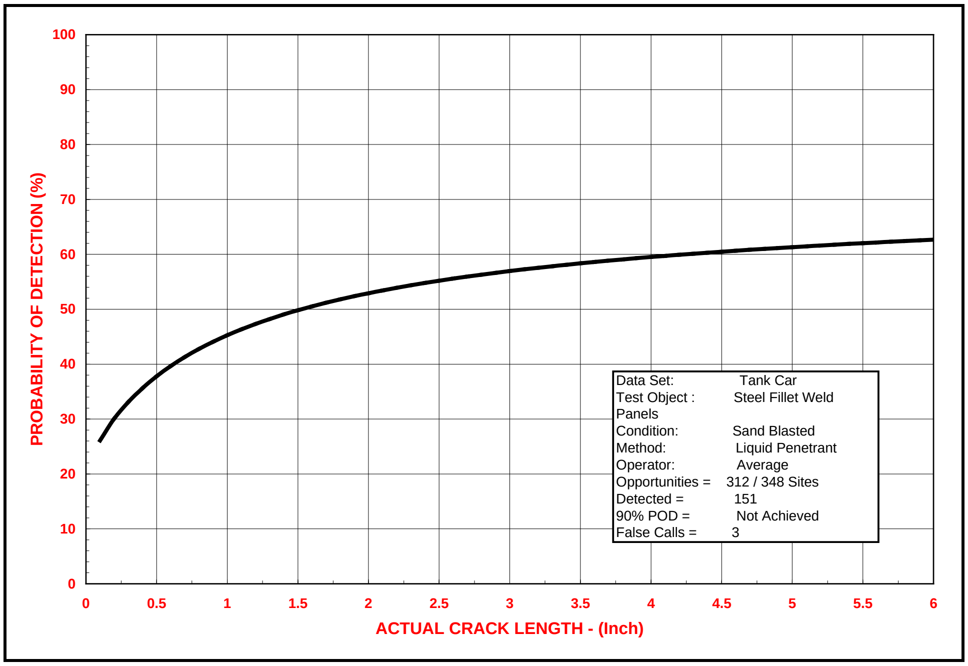


Figure 55. PT Fillet Weld POD Combined Average for Company A

PT Fillet Weld POD Results for Company B

Company B had three technicians participate in the PT PODs. Table 11 lists the POD results at different crack lengths. Figures 56 through 58 show the POD curves for each operator. Figure 59 compares each operator's results, and Figure 60 shows the combined average of all three operators.

Table 11. Company B Fillet Weld PT POD Percentages (%)

Flaw Size (inch)	Operator 15	Operator 21	Operator 22	Combined Results
0.25	26	20	11	19
0.50	38	37	23	33
0.75	43	44	30	39
1.00	49	52	37	46
1.25	51	56	41	50
1.50	55	61	47	54
1.75	57	64	50	57
2.00	59	67	53	60
2.25	61	69	57	62
2.50	62	72	59	64
2.75	64	73	61	66
3.00	65	75	63	68
3.25	66	76	64	69
3.50	67	77	66	70
3.75	68	78	67	71
4.00	69	79	69	73
4.25	70	80	70	73
4.50	71	81	71	74
4.75	71	82	72	75
5.00	72	83	73	76
5.25	72	83	74	77
5.50	73	84	75	77
5.75	73	84	76	78
6.00	75	85	77	79
90% POD	Not Achieved	Not achieved	Not Achieved	Not Achieved
False Calls	3	2	1	6

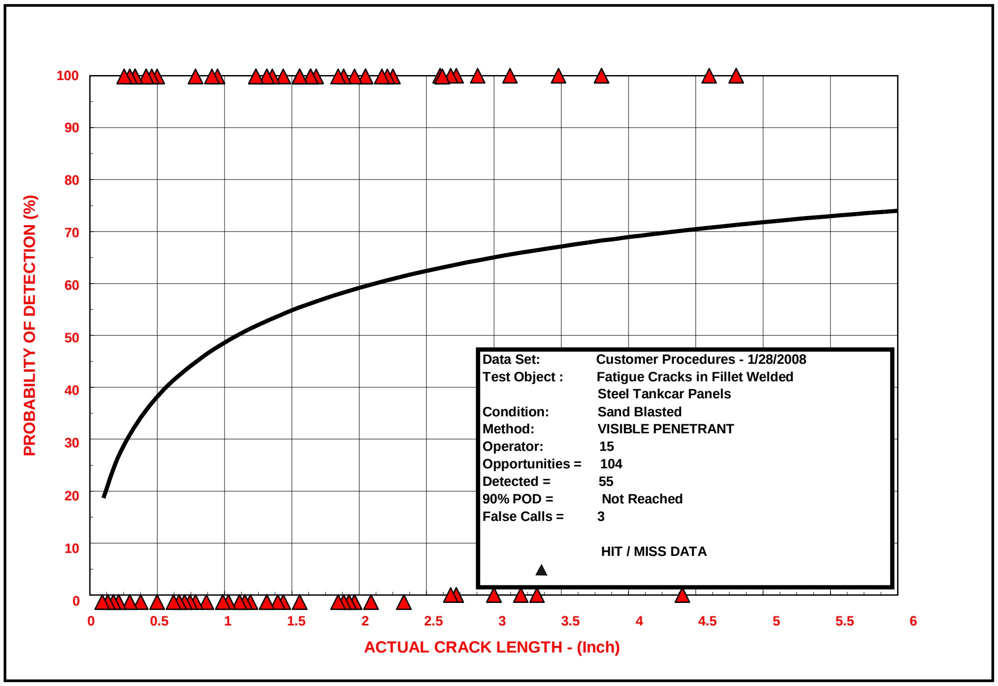


Figure 56. PT Fillet Weld POD Results for Operator 15

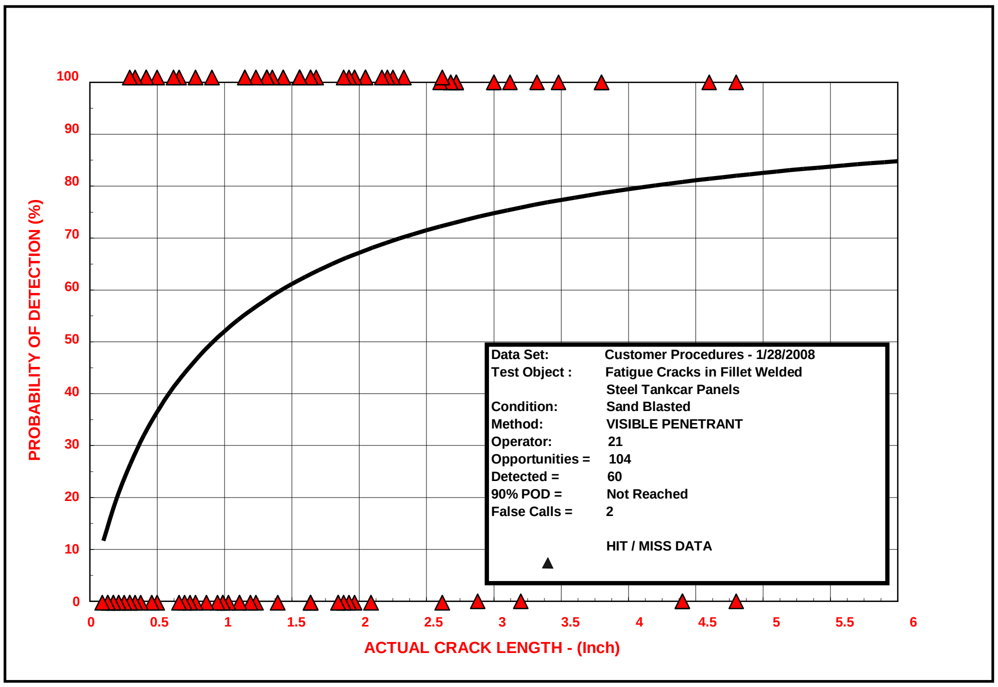


Figure 57. PT Fillet Weld POD Results for Operator 21

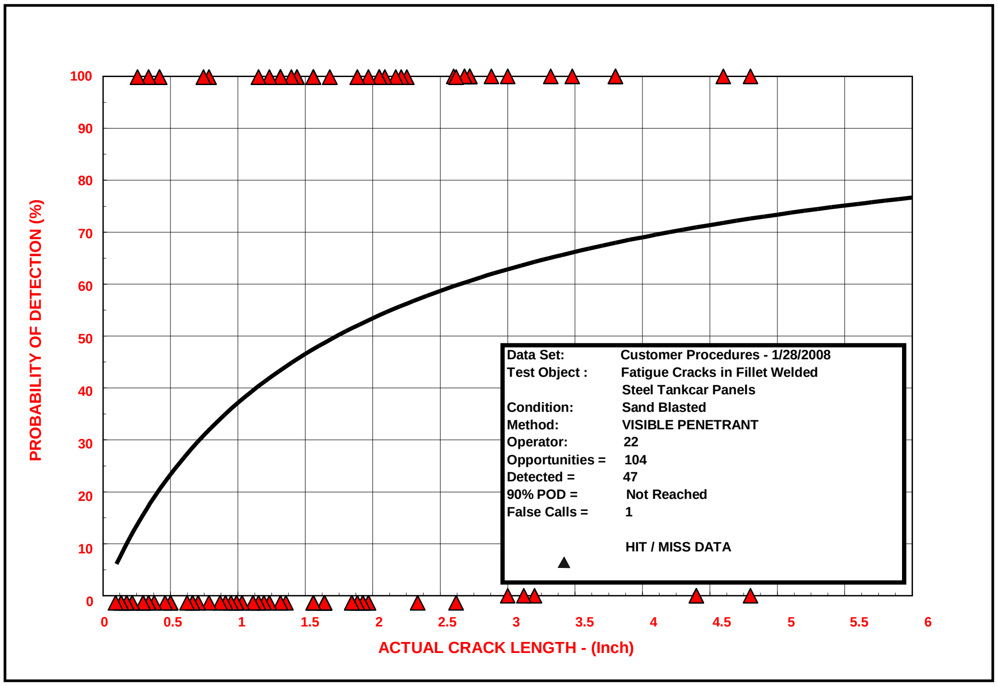


Figure 58. PT Fillet Weld POD Results for Operator 22

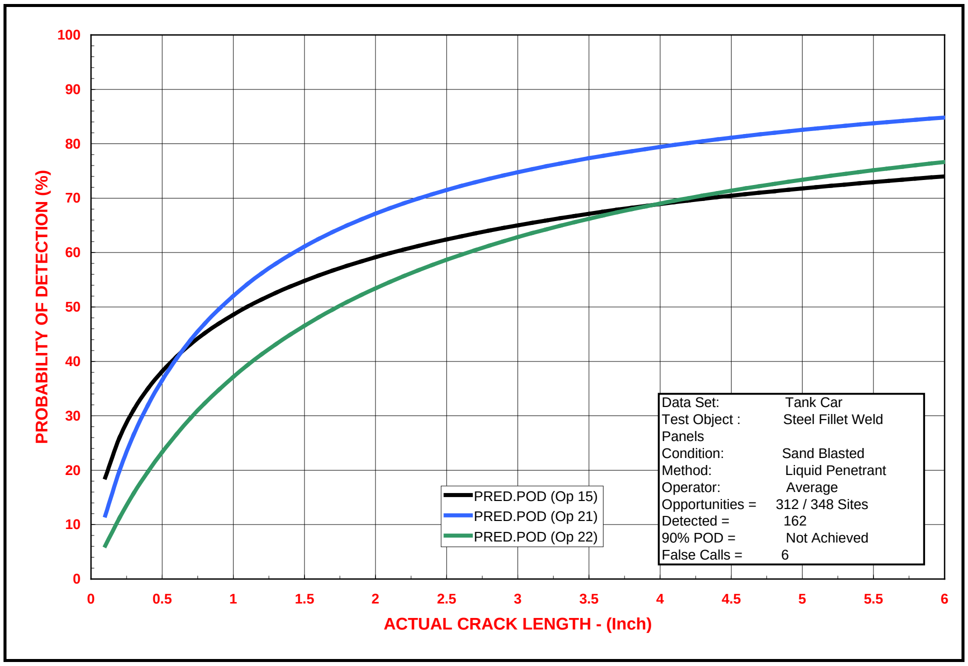


Figure 59. PT Fillet Weld POD Company B Operator Comparisons

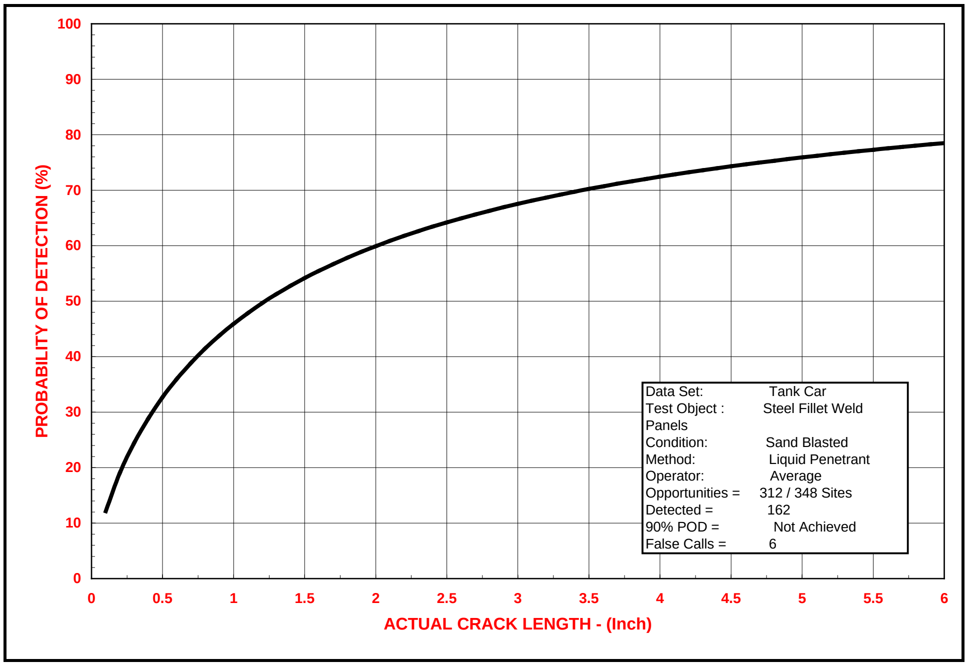


Figure 60. PT Fillet Weld POD Combined Average for Company B

2.3.4 Magnetic Particle (MT)

MT is a nondestructive test method that uses magnetic leakage fields and indicating materials to disclose surface and near surface discontinuities. It can reveal surface discontinuities that may be too small or too tight to be seen with the unaided eye. MT is based on the principle that magnetic flux is locally distorted by a discontinuity, and the magnetic flux leakage exits and reenters the magnetized object at the discontinuity; this leakage field attracts the magnetic particles applied to the test area forming an indication or outline of the discontinuity.

Advantages of the magnetic particle test method are that it is relatively economic and expedient, and MT equipment is considered portable. Also, MT can detect some discontinuities located slightly below the surface. However, MT requires a clean and relatively smooth surface, it is only applicable to ferromagnetic material, and it requires the use of electrical energy for most applications. The magnetic field alignment and field strength are critical.

MT provides an economical NDT method to evaluate discontinuities that are open to the surface or slightly subsurface, and it should continue to be a valuable method for tank car inspections. MT can be performed with minimal surface preparation and is relatively portable. Reliability of inspections can be enhanced through emphasis on operator training, equipment calibrations, and inspection procedures. It is important to employ MT equipment and materials with the desired sensitivity and ensure that the equipment is kept uniform from inspection to inspection. If the inspection process is changed, the operator should become familiar with the changes prior to performing further inspections.

As with other non-destructive evaluation (NDE) methods that use visual assessment to determine the integrity of inspection areas, MT can be enhanced by providing a greater contrast between the discontinuity and surrounding areas of the test article. For example the test area could be prepared for the dry MT approach by spraying a white developing powder over the area prior to inspection. If a discontinuity is present, this method will provide a greater contrast between the discontinuity and its surrounding areas.

MT Butt Weld POD Results

Butt weld MT POD evaluation results for industry participants A and B are provided below. The results show individual results of each company's participants and the combined average for the company. Before performing inspections of the test panels, the participants performed setup, which included calibration in accordance with the applicable company procedure. After calibration, the participants began their inspections of the butt weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

MT Butt Weld POD Results for Company A

Company A had three technicians participate in the MT PODs. Table 12 lists the POD results at different crack lengths. Figures 61 through 63 show the POD curves for each operator. Figure 64 compares the operator's results, and Figure 65 shows the combined average of all three operators.

Table 12. Results of Butt Weld MT POD Percentages (%)

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	14	45	23	27
0.50	31	53	34	39
0.75	44	58	41	48
1.00	55	61	46	54
1.25	63	63	50	59
1.50	69	65	54	63
1.75	73	67	57	66
2.00	77	68	59	68
2.25	80	69	61	70
2.50	82	70	63	72
2.75	84	71	65	73
3.00	86	72	66	75
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	5	10	0	15

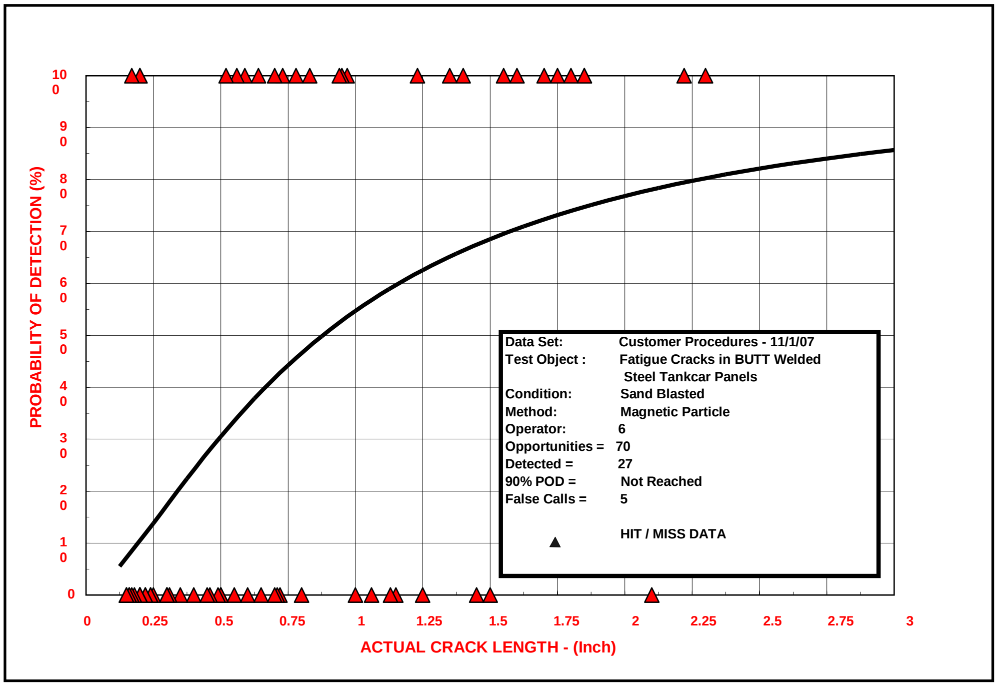


Figure 61. MT Butt Weld POD Results for Operator 6

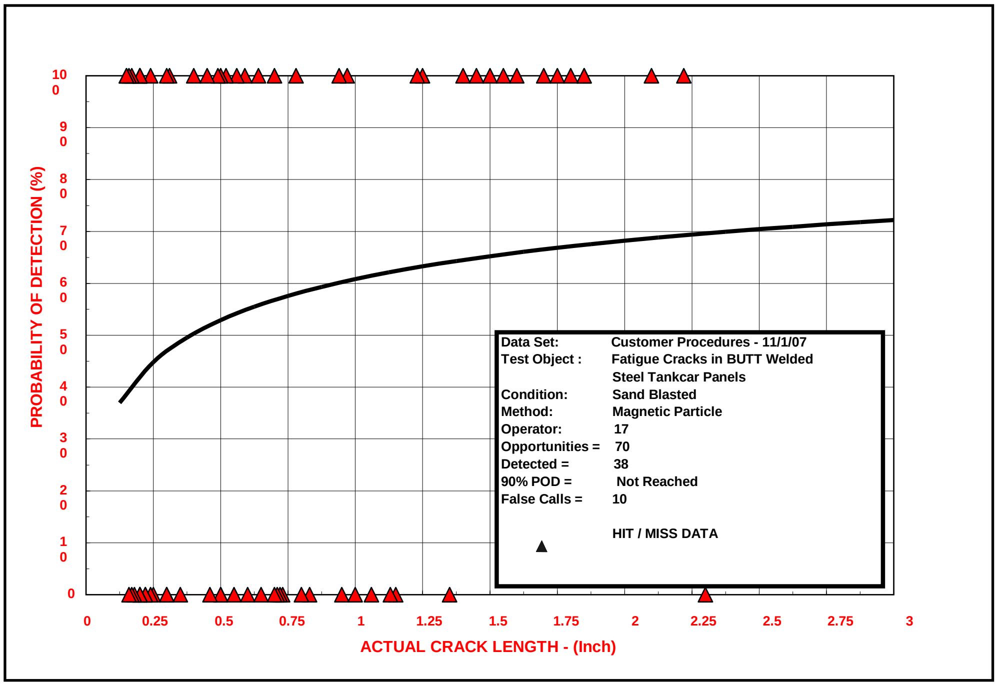


Figure 62. MT Butt Weld POD Results for Operator 17

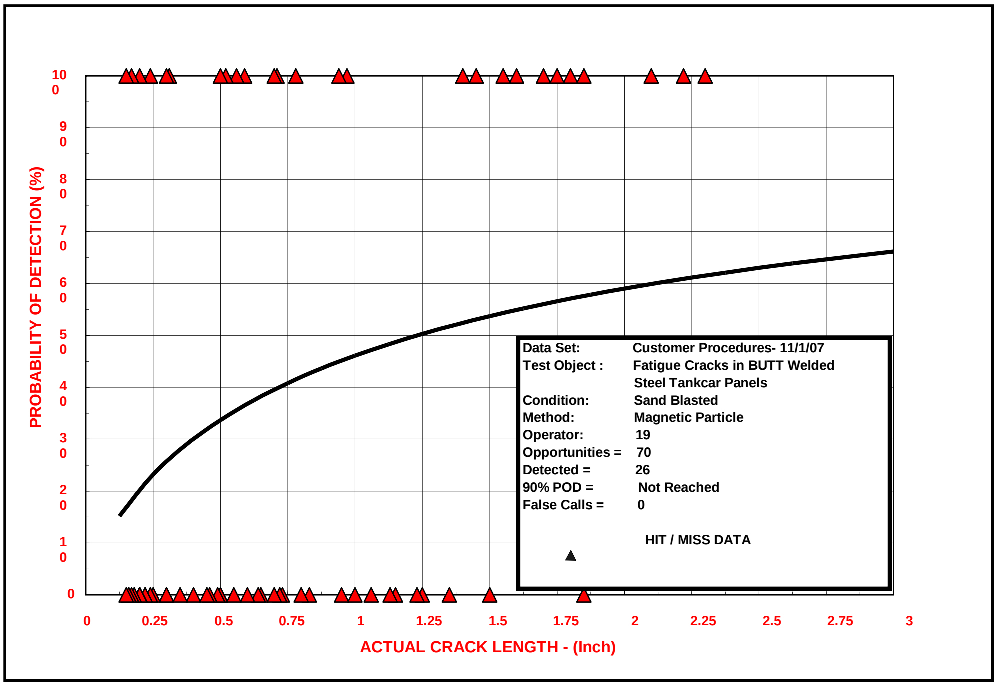


Figure 63. MT Butt Weld POD Results for Operator 19

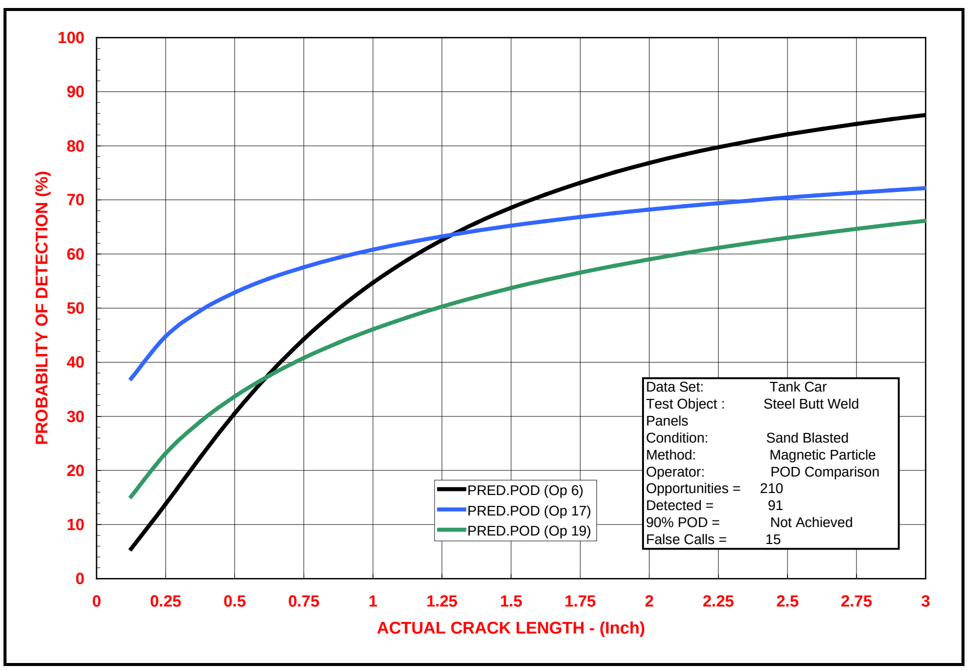


Figure 64. MT Butt Weld POD Company A Operator Comparisons

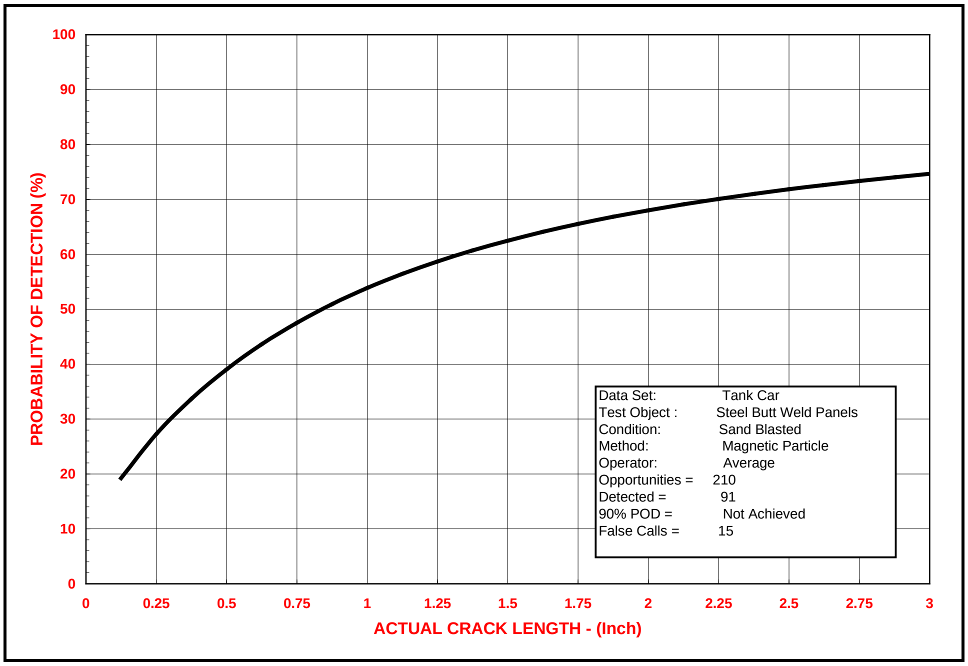


Figure 65. MT Butt Weld POD Combined Average for Company A

MT Butt Weld POD Results for Company B

Company B had three technicians participate in the MT PODs. Table 13 lists the POD results at different crack lengths. Figures 66 through 68 show the POD curves for each operator. Figure 69 compares the operator's results, and Figure 70 shows the combined average of all three operators.

Table 13. Company B Butt Weld MT POD Percentages (%)

Flaw Size (inch)	Operator 23	Operator 24	Operator 26	Combined Results
0.25	23	56	37	38
0.50	89	72	49	70
0.75	98	80	56	78
1.00	100	84	61	81
1.25	100	87	65	84
1.50	100	89	67	85
1.75	100	90	70	87
2.00	100	91	72	88
2.25	100	92	73	89
2.50	100	93	75	89
2.75	100	94	76	90
3.00	100	94	77	91
90% POD	0.50 in.*	1.63 in.	Not Achieved	2.88 in.
False Calls	32	29	7	68

Note: Grey areas depict skewed results due to a high number of false calls.
*POD for Operator 23 not valid as data did not converge.

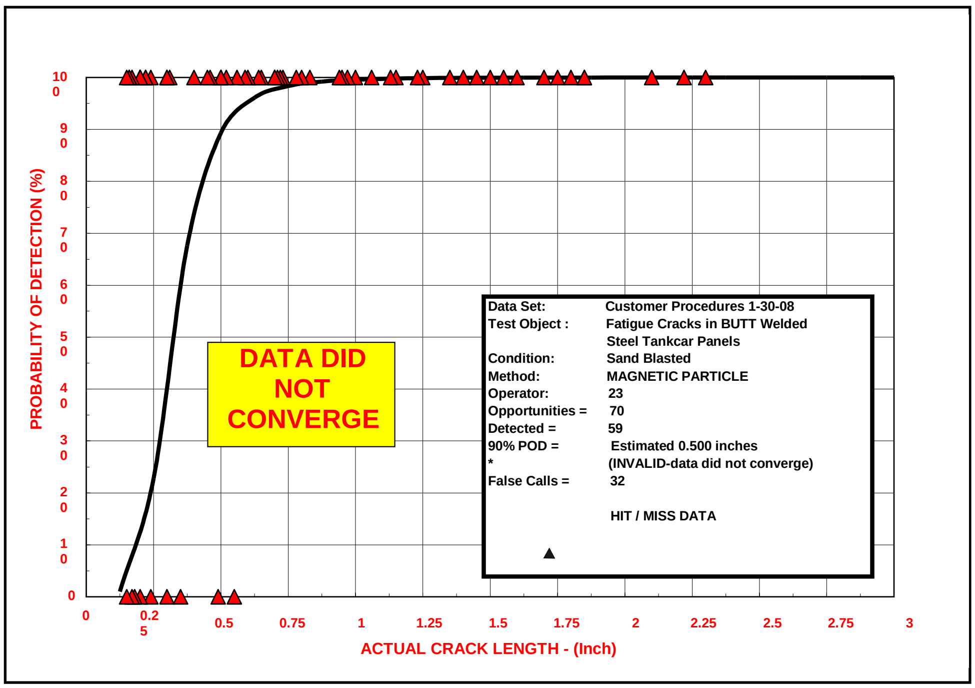


Figure 66. MT Butt Weld POD Results for Operator 23

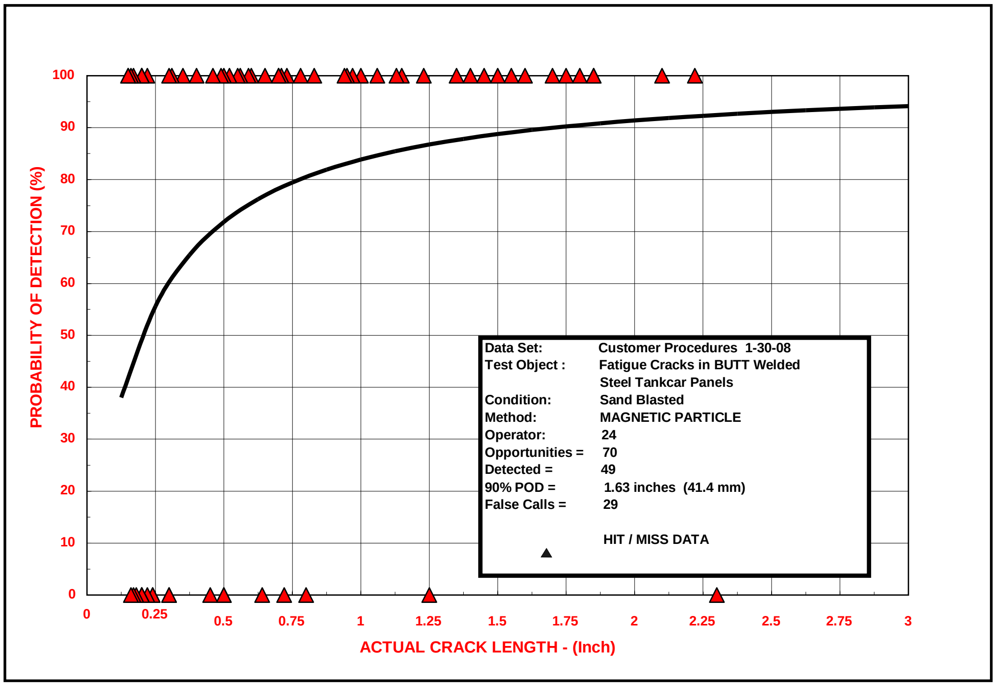


Figure 67. MT Butt Weld POD Results for Operator 24

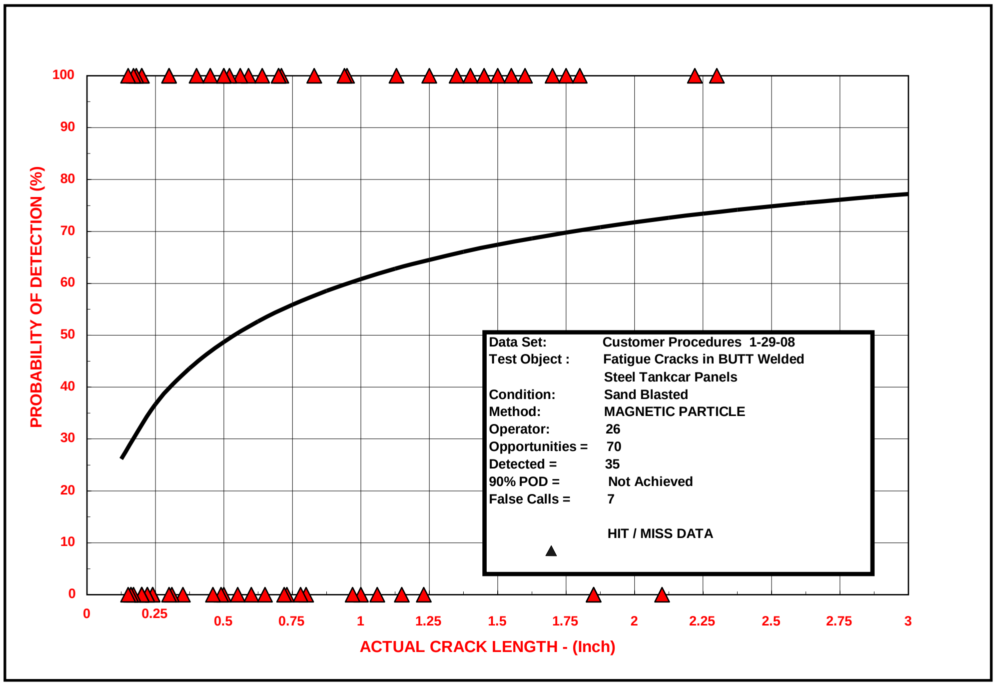


Figure 68. MT Butt Weld POD Results for Operator 26

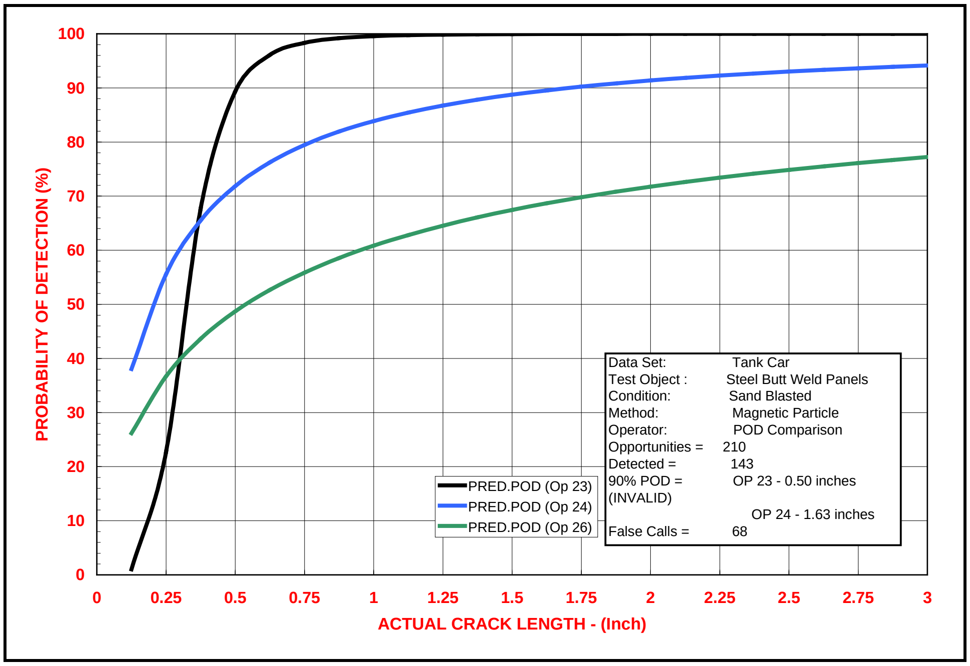


Figure 69. MT Butt Weld POD Company B Operator Comparisons

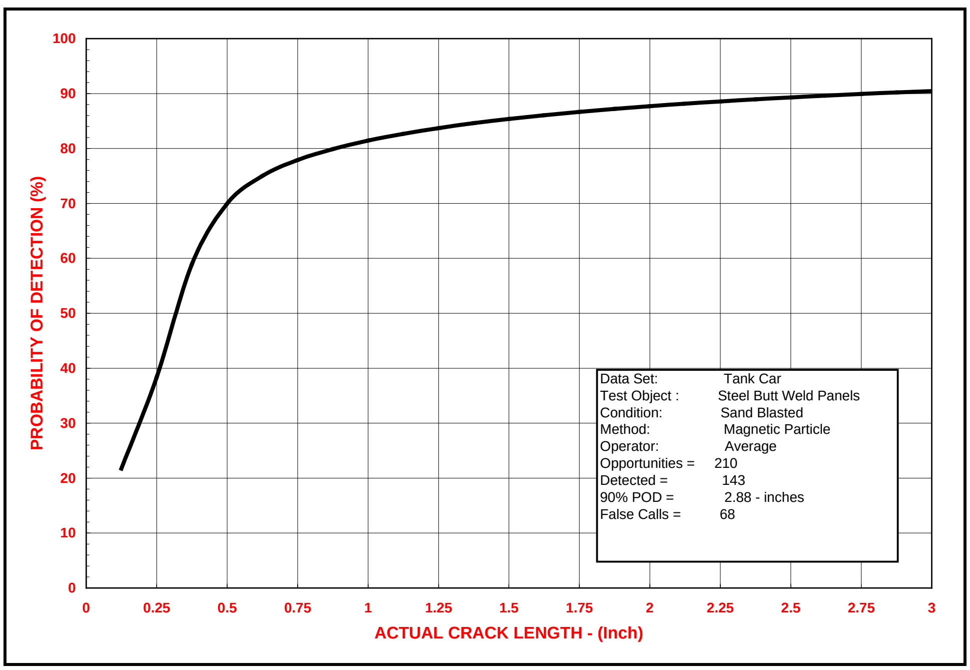


Figure 70. MT Butt Weld POD Combined Average for Company B

MT Fillet Weld POD Results

The fillet weld MT POD evaluation results for industry participants A and B are provided below. The results show individual results of each company's participants and the combined average for the company. Before performing inspections of the test panels, the participants performed setup, which included calibration in accordance with the applicable company procedure. After calibration, the participants began their inspections of the weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

MT Fillet Weld POD Results for Company A

Company A had three technicians participate in the MT PODs. Table 14 lists the POD results at different crack lengths. Figures 71 through 73 show the POD curves for each operator. Figure 74 compares the operator's results, and Figure 75 shows the combined average of all three operators.

Table 14. Company A Fillet Weld MT POD Percentages

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	67	40	38	48
0.50	71	63	44	59
0.75	72	70	47	63
1.00	73	77	49	66
1.25	74	80	50	68
1.50	75	84	52	70
1.75	75	85	53	71
2.00	76	87	54	72
2.25	76	88	55	73
2.50	76	89	55	74
2.75	77	90	56	74
3.00	77	91	57	75
3.25	77	92	57	75
3.50	78	92	58	76
3.75	78	93	58	76
4.00	78	93	59	77
4.25	78	93	59	77
4.50	78	94	59	77
4.75	78	94	60	77
5.00	79	94	60	78
5.25	79	95	60	78
5.50	79	95	61	78
5.75	79	95	61	78
6.00	79	95	61	79
90% POD	Not Achieved	2.60 in.	Not Achieved	Not Achieved
False Calls	0	0	0	0

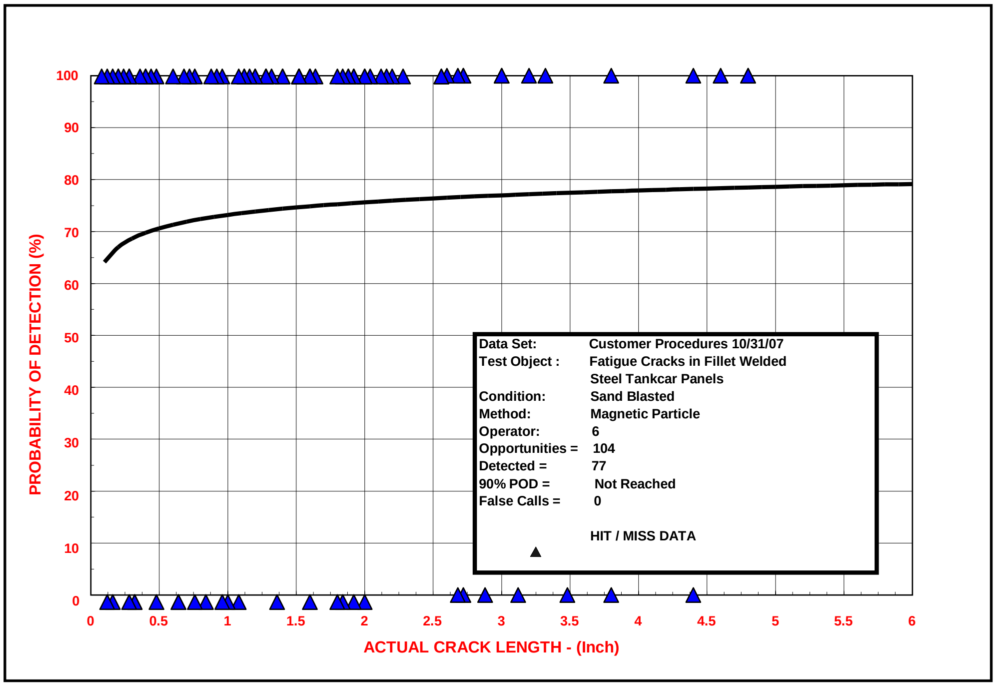


Figure 71. MT Fillet Weld POD Results for Operator 6

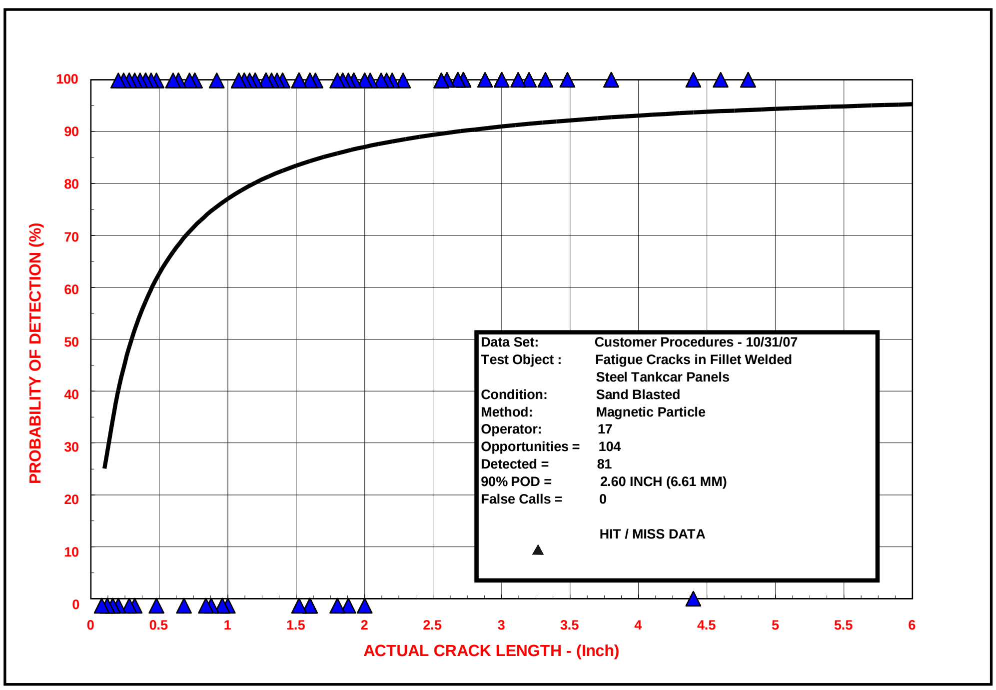


Figure 72. MT Fillet Weld POD Results for Operator 17

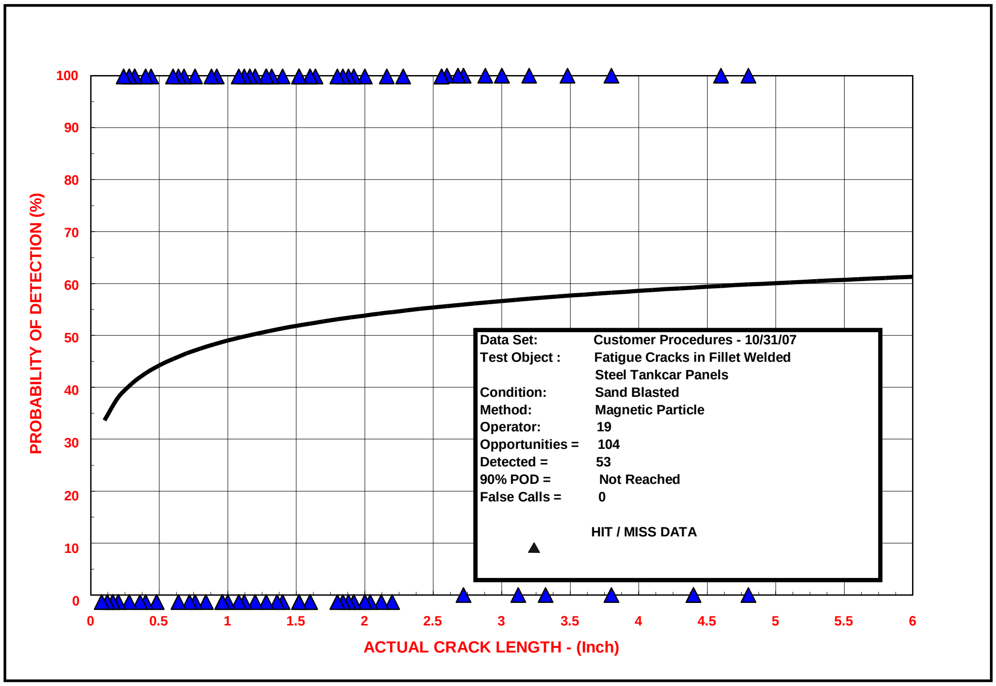


Figure 73. MT Fillet Weld POD Results for Operator 19

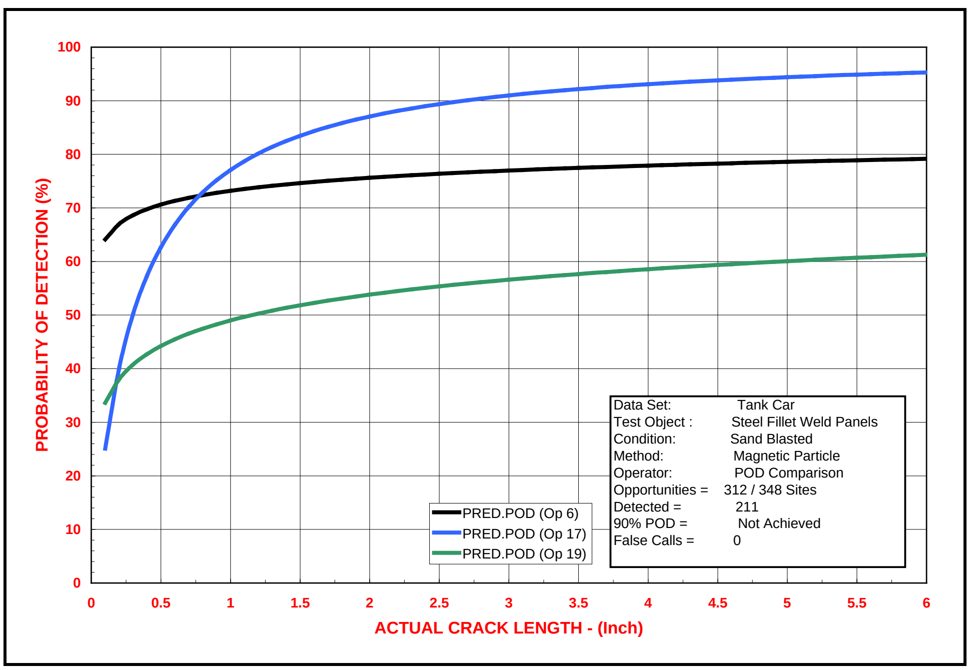


Figure 74. MT Fillet Weld POD Company A Operator Comparisons

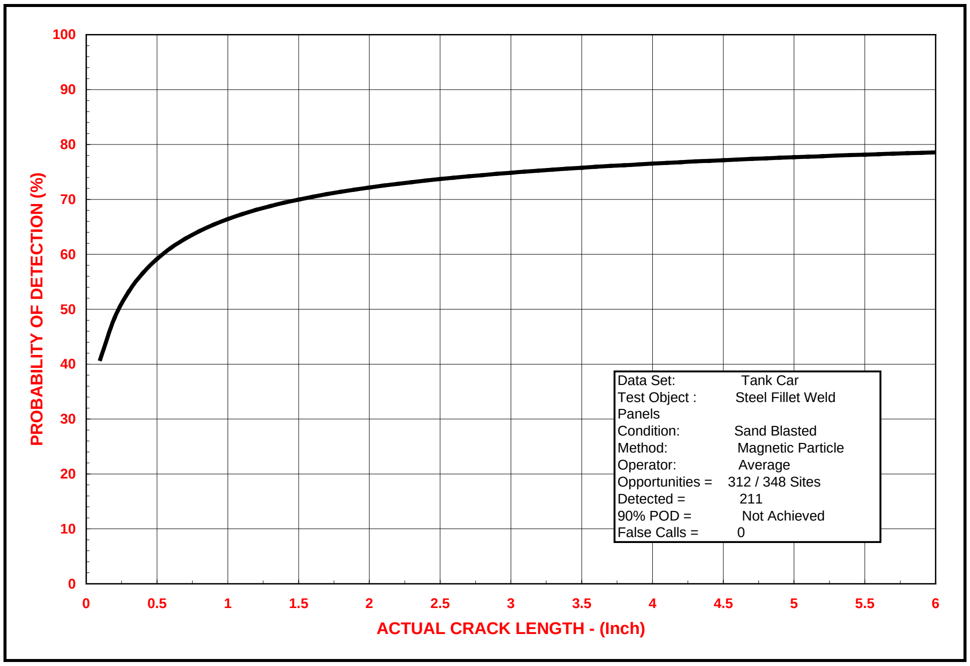


Figure 75. MT Fillet Weld POD Combined Average for Company A

MT Fillet Weld POD Results for Company B

Company B had three technicians participate in the MT PODs. Table 15 lists the POD results at different crack lengths. Figures 76 through 78 show the POD curves for each operator. Figure 84 compares the operator's results, and Figure 85 shows the combined average of all three operators.

Table 15. Company B Fillet Weld MT POD Percentages

Flaw Size (inch)	Operator 23	Operator 24	Operator 26	Combined Results
0.25	44	63	48	52
0.50	60	70	56	62
0.75	66	72	59	66
1.00	71	75	62	69
1.25	74	76	63	71
1.50	77	77	65	73
1.75	78	78	66	75
2.00	80	79	67	75
2.25	81	80	68	76
2.50	82	80	69	77
2.75	83	81	69	78
3.00	84	81	70	79
3.25	85	82	71	79
3.50	86	82	71	80
3.75	86	82	72	80
4.00	87	83	72	81
4.25	87	83	73	81
4.50	88	83	73	81
4.75	88	84	73	82
5.00	88	84	74	82
5.25	89	84	74	82
5.50	89	84	74	83
5.75	89	85	75	83
6.00	89	85	75	83
90% POD	Not Achieved	Not Achieved	Not Achieved	Not Achieved
False Calls	1	2	1	4

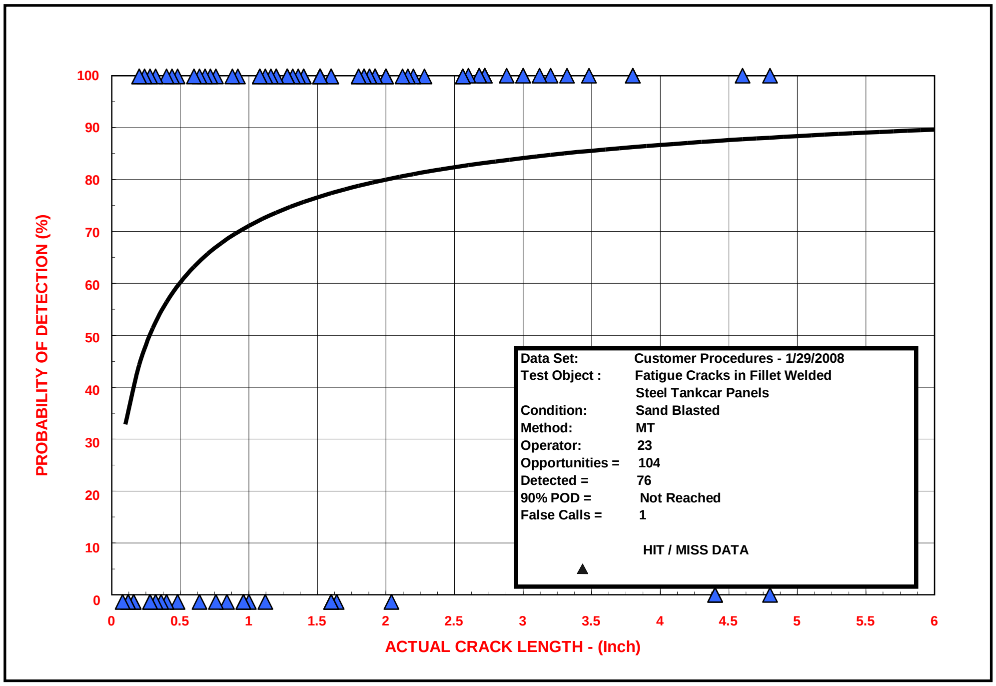


Figure 76. MT Fillet Weld POD Results for Operator 23

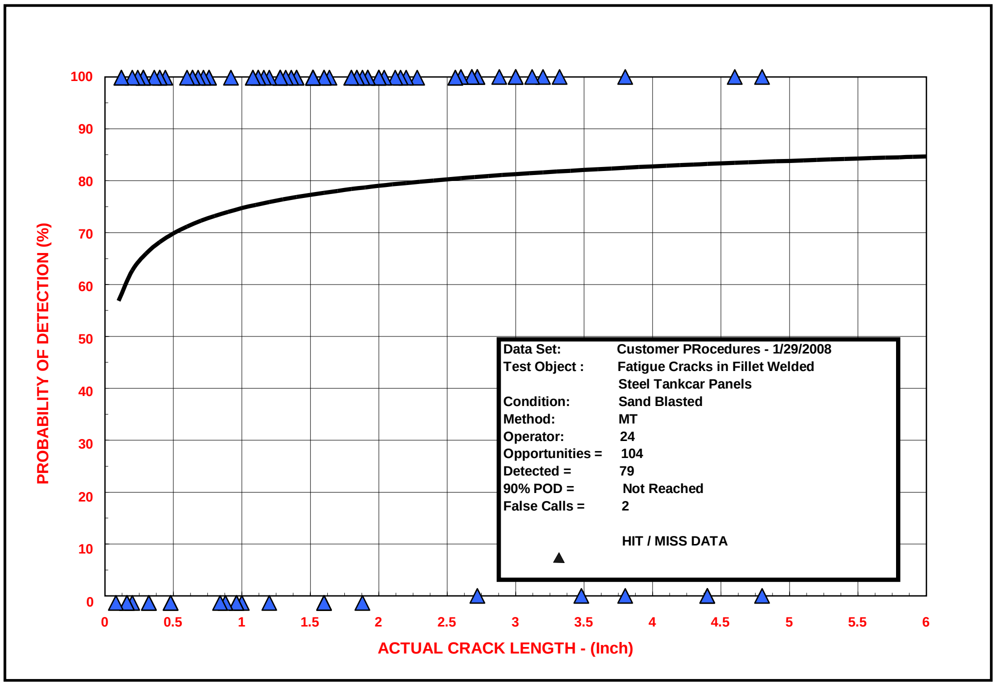


Figure 77. MT Fillet Weld POD Results for Operator 24

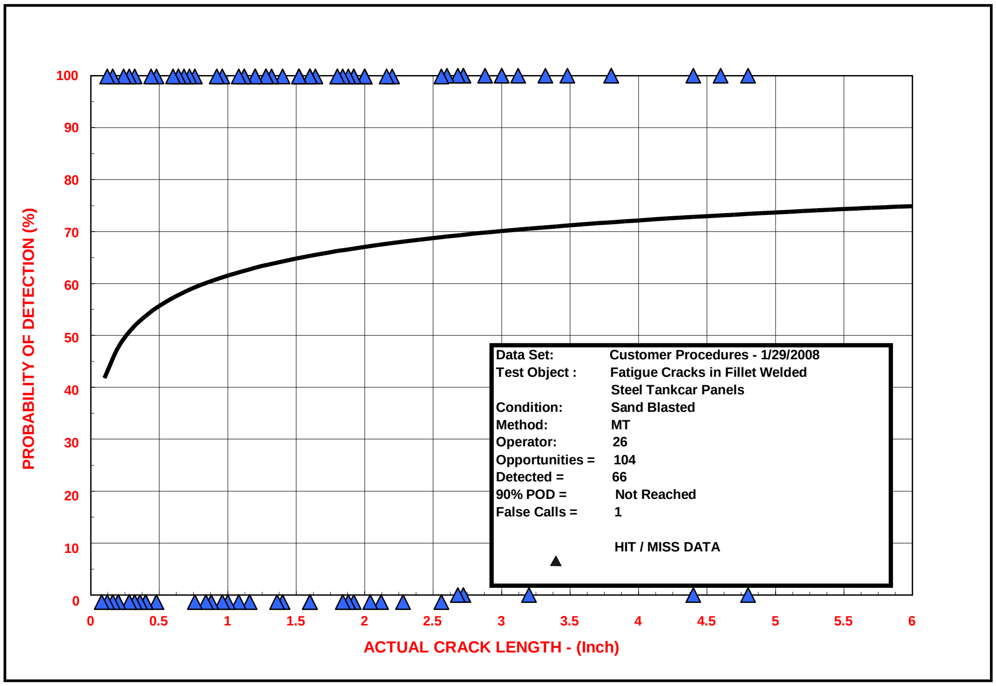


Figure 78. MT Fillet Weld POD Results for Operator 26

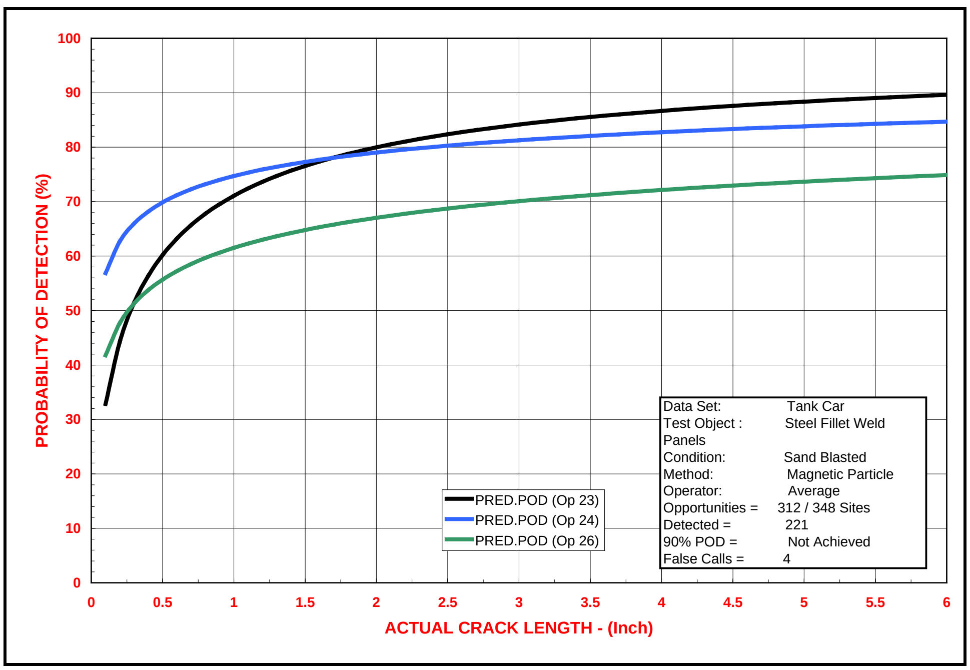


Figure 79. MT Fillet Weld POD Company B Operator Comparisons

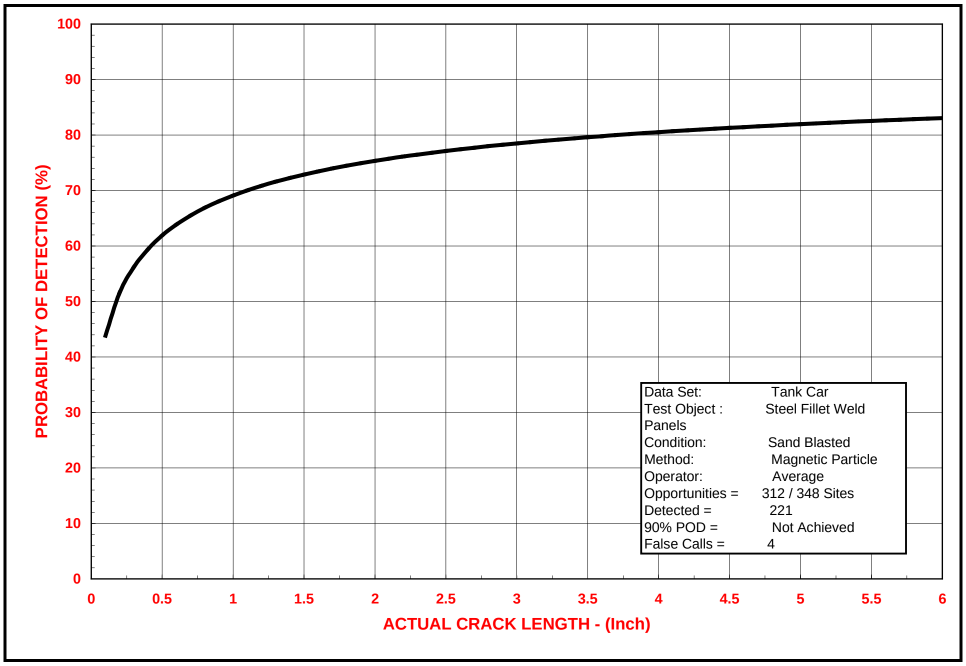


Figure 80. MT Fillet Weld POD Combined Average for Company B

2.3.5 Shear Wave Ultrasonic Testing

The UT method uses sound waves in the range of 20 to 25 MHz to generate acoustic energy for use in the interrogation of materials. It it is a versatile NDE method used to test a variety of metallic and nonmetallic materials. UT only requires access to one side of a specimen and does not present a hazard to the operator or anyone nearby during testing. UT is also capable of detecting both surface and subsurface discontinuities.

Advantages of the UT method are that it is most sensitive to planar type discontinuities and test results are known immediately (real time), it is portable, and most flaw detectors can run off batteries eliminating the need for a power supply during testing. Limitations of the method include the need for access to at least one side of the part being inspected. Conventional UT systems require a liquid coupling to transfer and receive the sound energy into and back from the part. This method requires a skilled operator to operate the equipment and to perform the inspections.

Technological advances in UT provide a large number of UT processes available for use in tank car inspections. The portability of the equipment, along with some of the memory and storage capacities of the ultrasonic instruments allows for faster and more efficient calibration processes. The variety of transducers available with different sizes, angles, frequencies, and material design introduce the possibilities for inspection at most locations of the tank car. Reliability of inspections, as with the other NDT methods, can be enhanced through emphasis on operator training, equipment calibrations, and inspection procedures.

UT Butt Weld POD Results

The butt weld UT POD evaluation results for industry participants A and B are provided below. The results show individual results of each company's participants and the combined average for the company. Before performing inspections of the test panels, the participants performed setup, which included calibration in accordance with the applicable company procedure. After calibration, the participants began their inspections of the butt weld test panels. Calibration checks were made at the beginning, middle, and ending of the panel inspections.

UT Butt Weld POD Results for Company A

Company A had three technicians participate in the UT PODs. Table 16 lists the POD results at different crack lengths. Figures 81 through 83 show the POD curves for each operator. Figure 84 compares the operator's results, and Figure 85 shows the combined average of all three operators.

Table 16. Results of Butt Weld Shear Wave UT POD Percentages

Flaw Size (inch)	Operator 6	Operator 17	Operator 19	Combined Results
0.25	29	9	13	17
0.50	48	15	12	25
0.75	60	20	11	30
1.00	68	24	10	34
1.25	73	28	10	37
1.50	77	31	10	39
1.75	80	34	9	41
2.00	83	37	9	43
2.25	84	39	9	44
2.50	86	41	9	45
2.75	87	43	9	46
3.00	88	45	8	47
90% POD	Not Achieved	Not achieved	Not Achieved	Not Achieved
False Calls	11	1	4	16

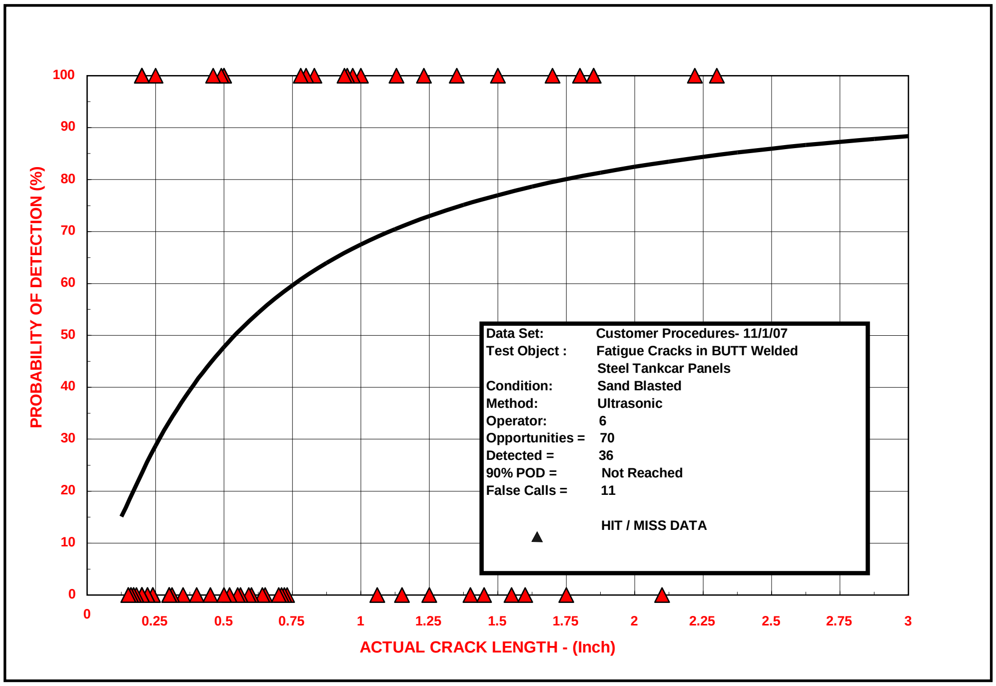


Figure 81. Shear Wave UT Butt Weld POD Results for Operator 6

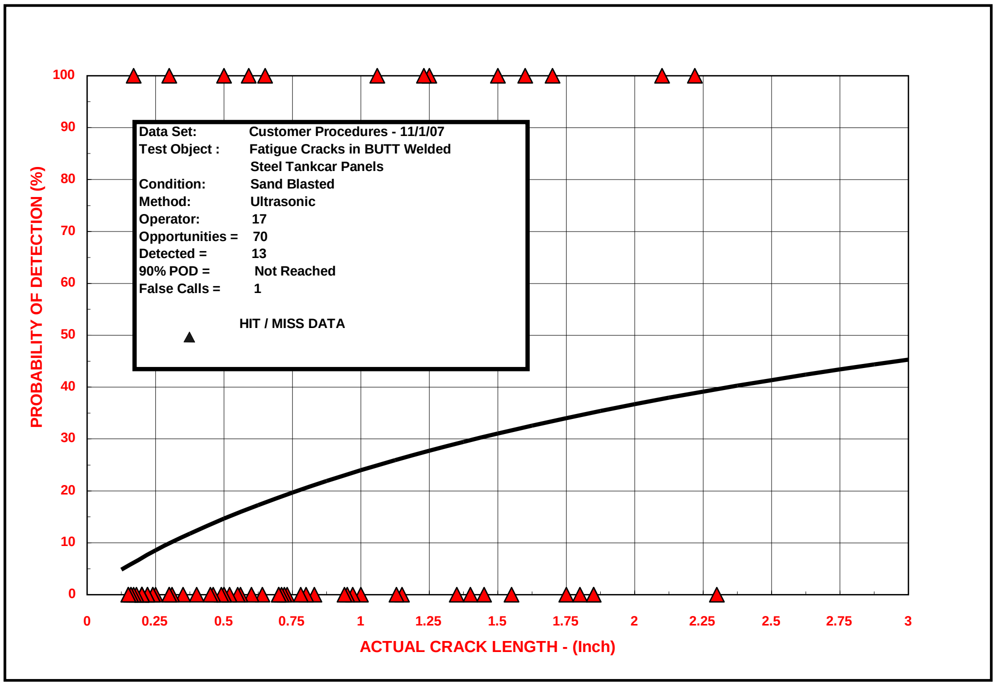


Figure 82. Shear Wave UT Butt Weld POD Results for Operator 17

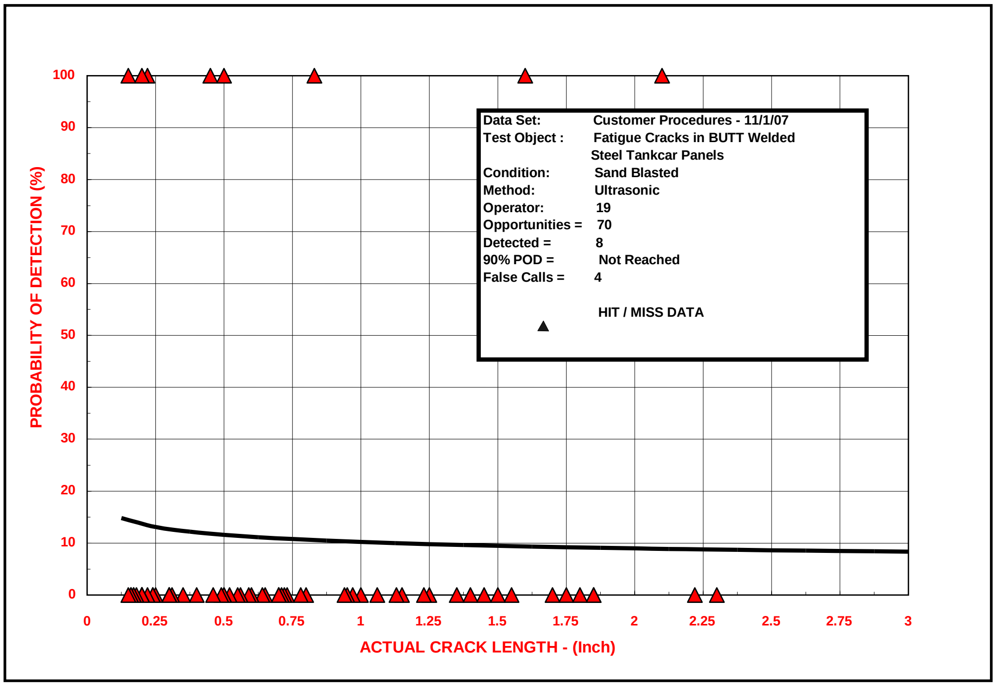


Figure 83. Shear Wave UT Butt Weld POD Results for Operator 19

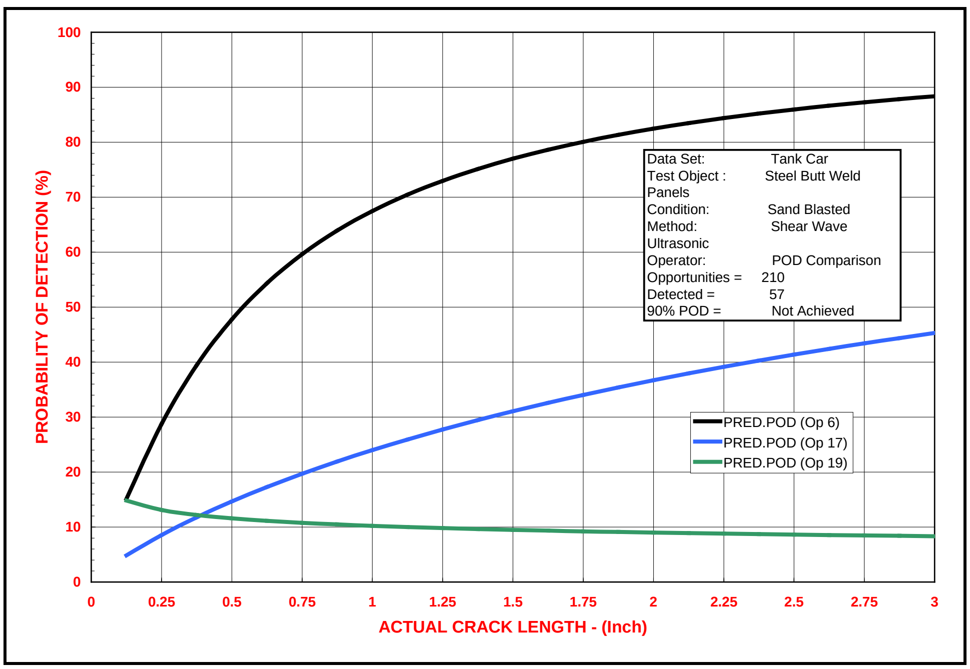


Figure 84. Shear Wave UT Butt Weld POD Company A Operator Comparisons

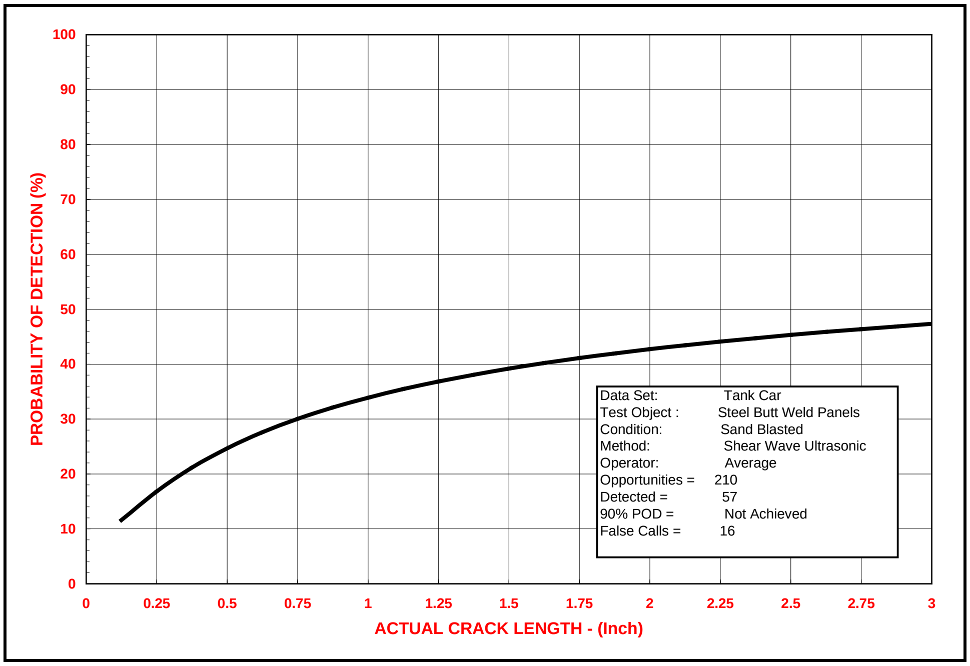


Figure 85. Shear Wave UT Butt Weld POD Combined Average for Company A

UT Butt Weld POD Results for Company B

Company B had three technicians participate in the UT PODs. Table 17 lists the POD results at different crack lengths. Figures 86 through 88 show the POD curves for each operator. Figure 89 compares the operator's results, and Figure 90 shows the combined average of all three operators.

Table 17. Company B Butt Weld Shear Wave UT POD Percentages

Flaw Size (inch)	Operator 15	Operator 25	Operator 26	Combined Results
0.25	55	33	16	35
0.50	66	65	38	56
0.75	72	80	55	69
1.00	76	88	67	77
1.25	79	92	75	82
1.50	81	94	80	85
1.75	82	95	84	87
2.00	84	96	87	89
2.25	85	97	89	90
2.50	86	98	91	91
2.75	86	98	92	92
3.00	87	98	93	93
POD Achieved	Not Achieved	1.13 inches	2.38 inches	2.25 inches
False Calls	54	6	0	60

Note: Grey area depicts a high number of false calls that skew the results.

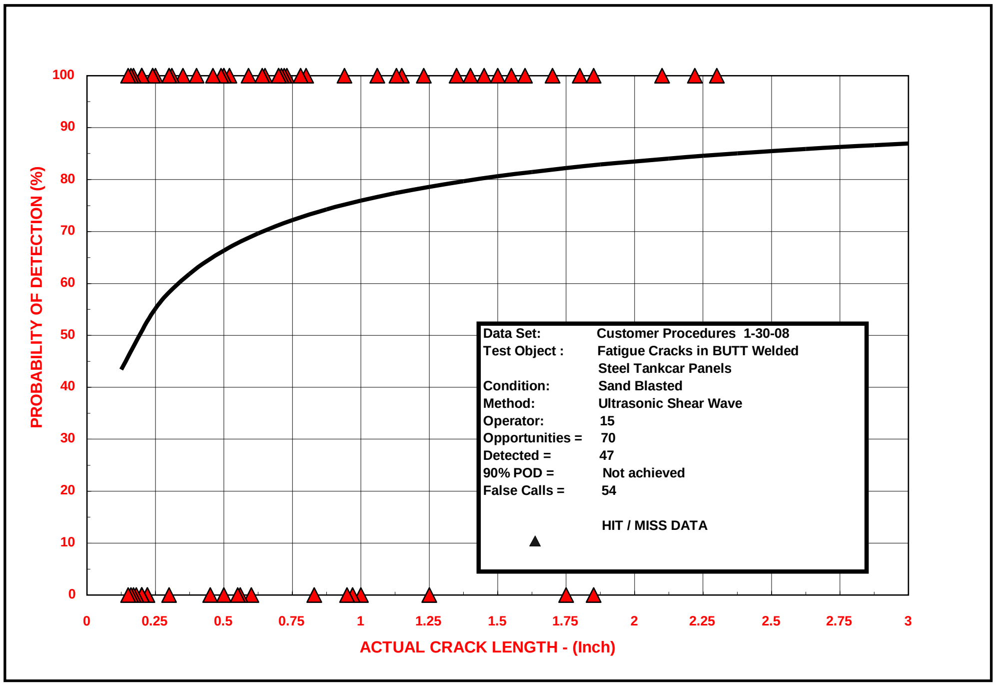


Figure 86. Shear Wave UT Butt Weld POD Results for Operator 15

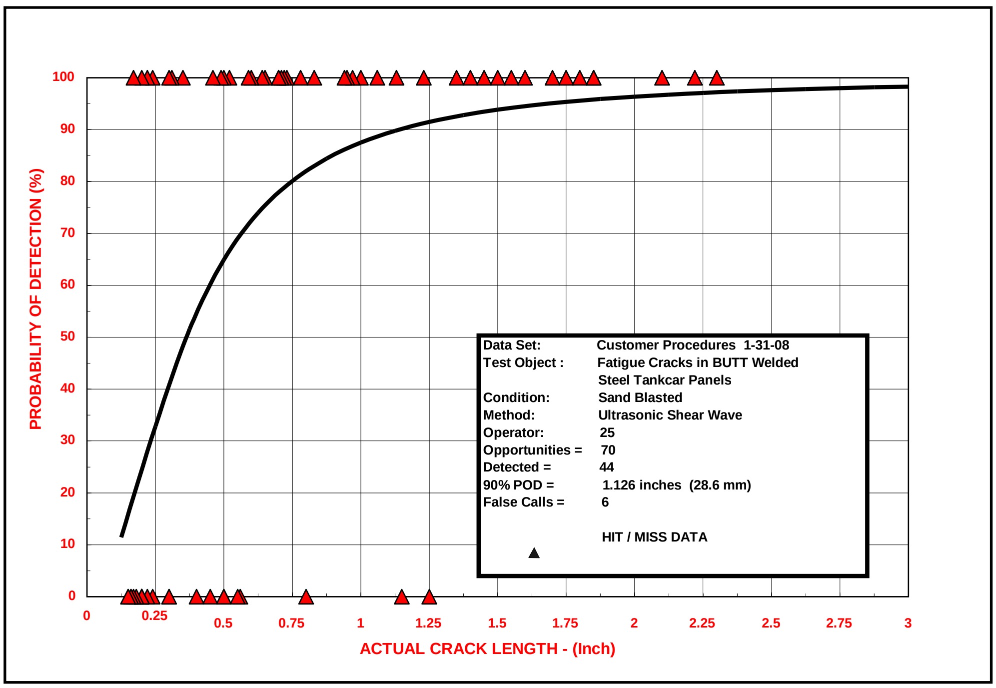


Figure 87. Shear Wave UT Butt Weld POD Results for Operator 25

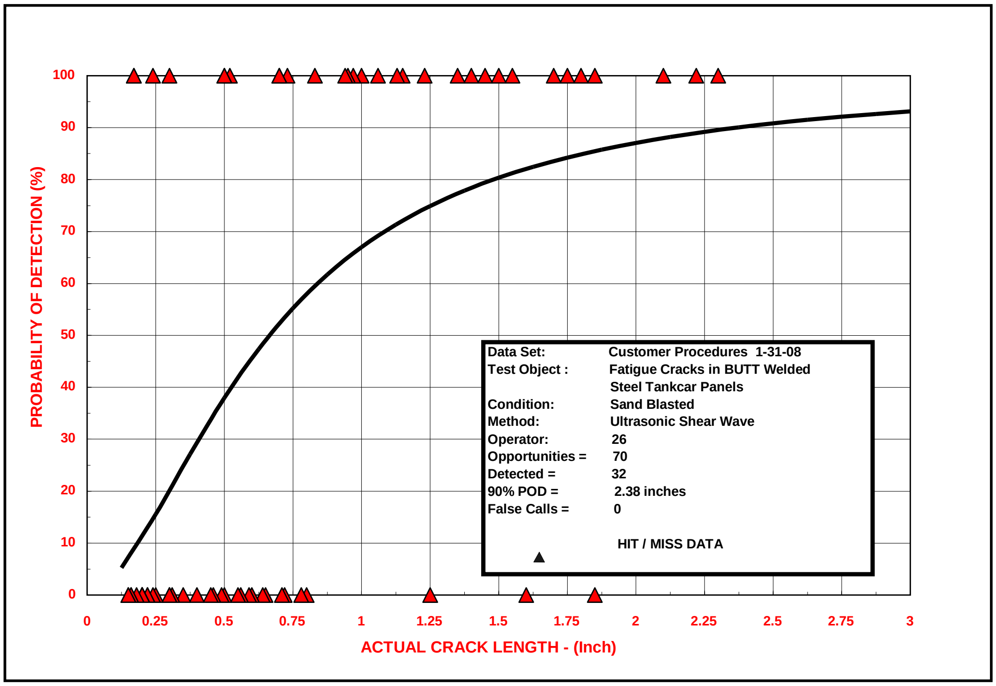


Figure 88. Shear Wave UT Butt Weld POD Results for Operator 26

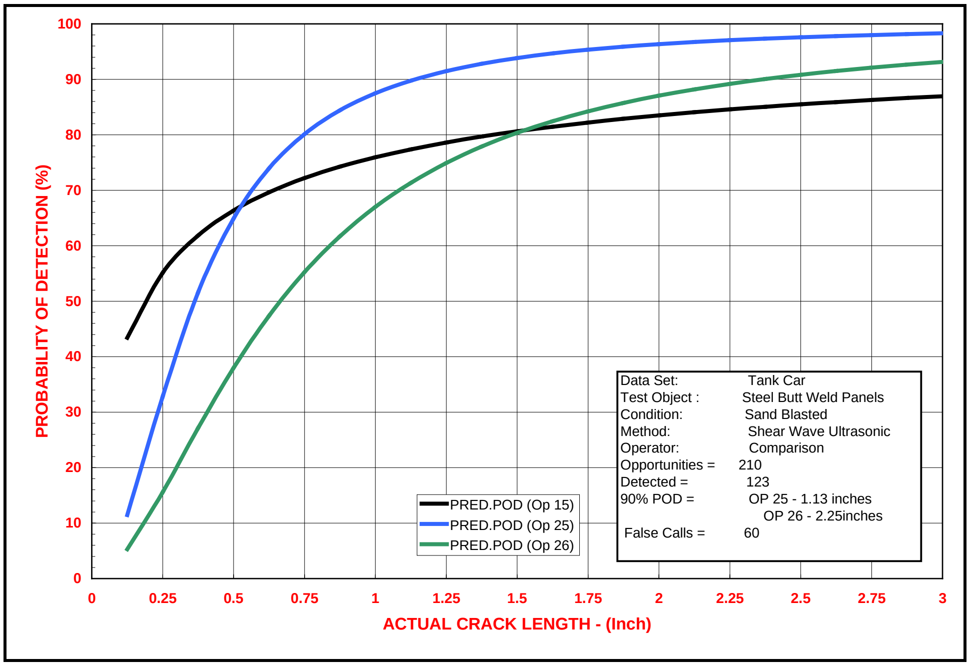


Figure 89. Shear Wave UT Butt Weld POD Company B Operator Comparisons

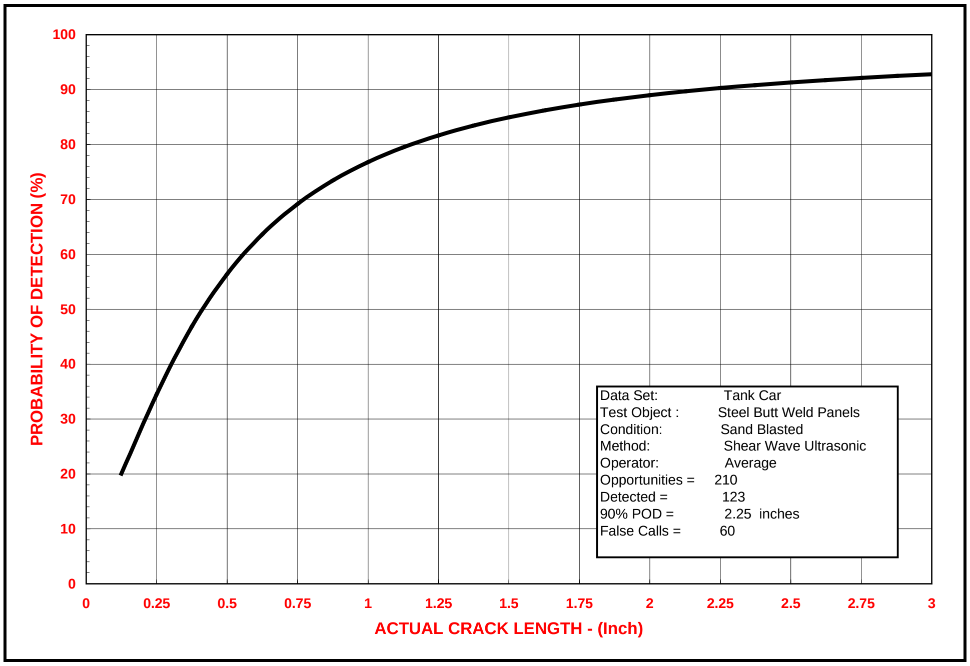


Figure 90. Shear Wave UT Butt Weld POD Combined Average for Company B

3. Conclusion

The POD evaluations validated that detection is a function of multiple variables associated with the detection and measurement processes. The capability (sensitivity) of the test method has a direct influence on flaw detection. In the case of the tank car PODs performed, the dry powder magnetic particle method provided a higher percentage of detection than the direct visual method for both butt weld and fillet weld inspections. The POD method is a useful tool for assessing variations in detection capabilities of both different NDT procedures and different NDT operators.

Figures 91 and 92 show typical views for the direct visual and magnetic particle inspection methods performed on tank car panels. When a white background and darker colored magnetic powder is used, the increased level of contrast provided by the magnetic particle method gives the operator a greater opportunity to discriminate between a cracked and noncracked area at the weld.

Figure 91. Fillet Weld Termination — DVT

Figure 92. Fillet Weld Termination — MT

The POD curve shown in Figure 93 represents the results of one operator and three different inspection methods. For this operator, the liquid penetrant and magnetic particle methods revealed a higher level of detection capability than the direct visual method. This same result is also shown in Figures 94 and 95, which present the current industry average for both butt and fillet weld inspections using the methods that were applied to each set of test specimens. The industry averages come from all the tests done in previous dates, including the ones done in this report.

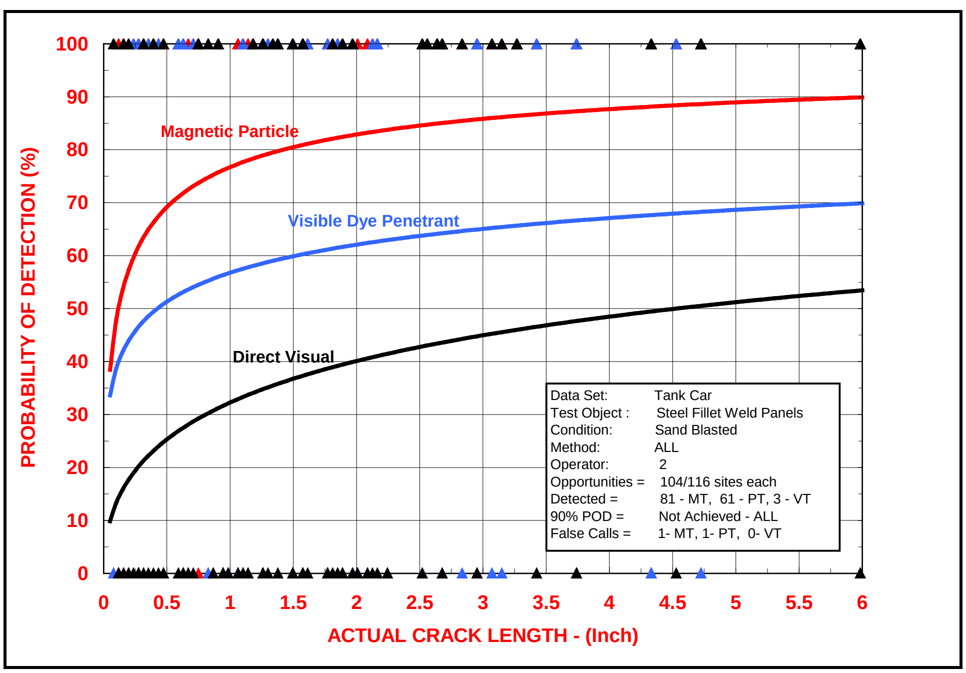


Figure 93. Results of Three NDE Methods for One Operator

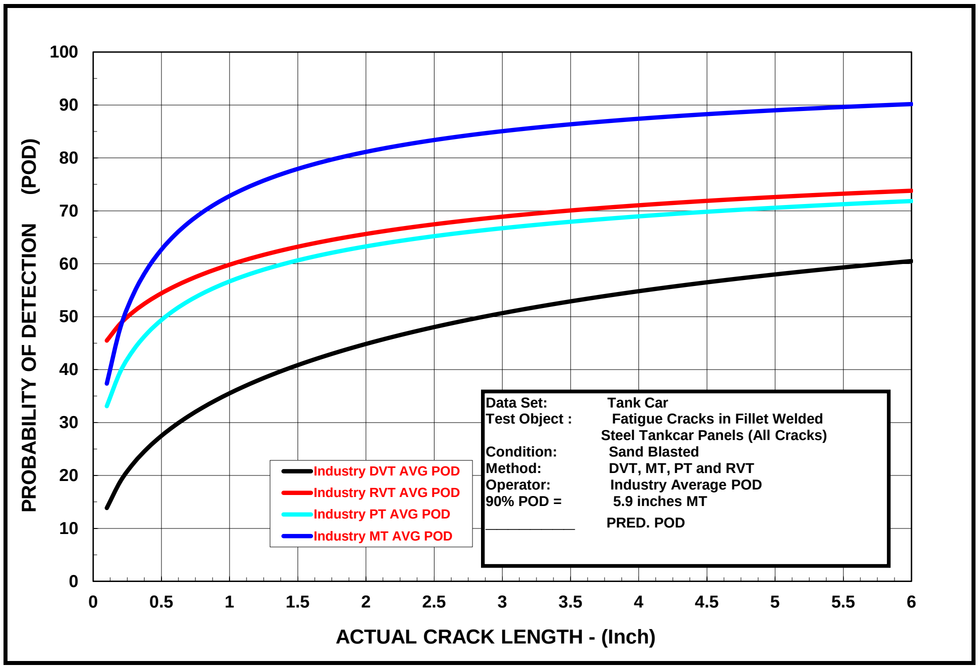


Figure 94. DVT, RVT, PT, and MT Industry Average for Fillet Welds

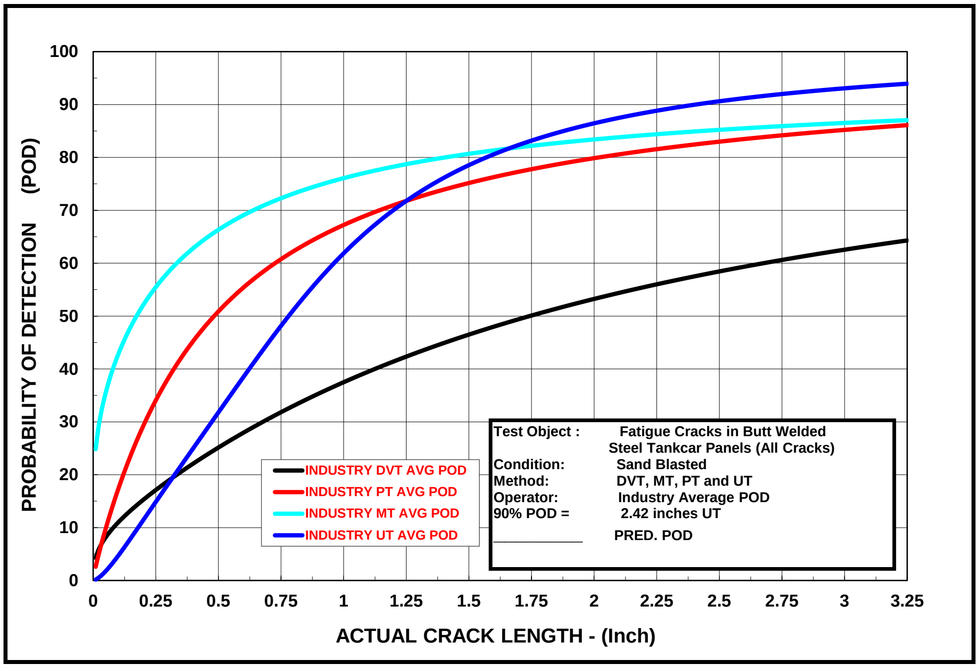


Figure 95. DVT, PT, MT, and UT Industry Average for Butt Welds

The data shows the variability from operator to operator and from one NDE method to another. The research efforts in progress focus on assisting the railroad tank car industry in developing and providing the tools and processes that will enhance both operator and method capabilities.

4. References

1. Federal Register, *Code of Federal Regulations*, Sections 179 and 180, June 1996.Washington, DC

2. National Transportation Safety Board, "Safety Recommendations," R-92-21 through R-92-24, 1992. Washington, DC

3. Rummel, Ward D., et al."The Detection of Fatigue Cracks by Nondestructive Testing Methods." NASA, (NTIS N74-17285), February 1974. Washington, DC

4. Garcia, G.A., "Railroad Tank Car Nondestructive Methods Evaluation," DOT/FRA/ORD-01/04, January 2002, Washington, DC

5. Garcia, G.A., "Quantitative Nondestructive Testing of Railroad Tank Cars Using the Probability of Detection Evaluation Approack," DOT/FRA/ORD-09/10, May 2009, Washington, DC

6. Berens, A. P and P. W. Hovey," Flaw Detection Reliability Criteria, Volume 1 – Methods and Results, AFWAL-TR-84-4022, April 1984. Dayton, OH

7. U.S. Military Standard MIL HDBK 1823, Nondestructive Evaluation System Reliability Assessment. Wright-Patterson Airforce Base, OH

8. Generazio, Edward R. (LARC-D3) edward.r.generazio@nasa.gov, NASA results on binomial point estimate testing, NASA-STD-5009. Greg Garcia greg_garcia@aar.com, September 22, 2009.

Abbreviations and Acronyms

CFR	Code of Federal Regulations
DOT	Department of Transportation
DVT	direct visual testing
FCR	false call rate
FPR	false positive rate
FRA	Federal Railroad Administration
HMR	Hazardous Materials Regulations
MT	magnetic particle
NASA	National Aeronautics and Space Association
NDI	nondestructive inspection
NDT	Nondestructive test or nondestructive testing
POD	probability of detection
PT	liquid penetrant
RT	radiography
RVT	remote visual testing
TTC	Transportation Technology Center
TTCI	Transportation Technology Center, Inc.
UT	ultrasonic
VT	visual testing