DOT/FAA/AR-07/63

Air Traffic Organization
Operations Planning
Office of Aviation Research
and Development
Washington, DC 20591

Update of Default Probability of Detection Curves for the Ultrasonic Detection of Hard Alpha Inclusions in Titanium Alloy Billets

February 2008

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AR-07/63		

4. Title and Subtitle	5. Report Date
UPDATE OF DEFAULT PROBABILITY OF DETECTION CURVES FOR THE ULTRASONIC DETECTION OF HARD ALPHA INCLUSIONS IN TITANIUM ALLOY BILLETS	February 2008
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
R. Bruce Thompson[1], Bill Meeker[1], Mike Keller[2], Jeff Umbach[3], Thomas Chiou[1], Yurong Wang[1], Dick Burkel[2], Waled Hassan[4], Kevin Smith[3], Thadd Patton[2], and Lisa Brasche[1]	

9. Performing Organization Name and Address		10. Work Unit No. (TRAIS)
[1]Iowa State University 1915 Scholl Road Ames, IA 50011	[3]Pratt & Whitney 400 Main Street MD Q-8 East Hartford, CT 06108	11. Contract or Grant No. DTFA0398FIA029
[2]General Electric 1 Neumann Way Cincinnati, OH 45215	[4]Honeywell 111 S. 34th Street Phoenix, AZ 85034	

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
U.S. Department of Transportation Federal Aviation Administration Air Traffic Organization Operations Planning Office of Aviation Research and Development Washington, DC 20591	Final Report
	14. Sponsoring Agency Code ANE-100

15. Supplementary Notes

The Federal Aviation Administration Aiport and Aircraft Safety R&D Division technical monitors were Rick Micklos, Paul Swindell, and Cu Nguyen.

16. Abstract

This report describes efforts to update the default probability of detection (PoD) curves for the ultrasonic detection of hard alpha inclusions in titanium alloy billets, with the results intended to replace the curves currently found in Advisory Circular 33.14-1, "Damage Tolerance for High Energy Turbine Engine Rotors." This report describes the data on which the update is based, the statistical approaches used to analyze those data, the results of the PoD analysis, technical issues encountered during the analysis, and conclusions. The original Default PoD curves were based on what is known as the Default PoD-3Dimensional database, which documents the results of inspections conducted before 1995 using conventional inspection techniques. Two additional sets of data that became available after 1995 are described in this report and greatly extend the information available. The updated PoD curves are found to be considerably different from the previous curves, obtained in 1995, with the PoD of the Multizone technique shown to be clearly superior to that of the conventional inspection technique.

17. Key Words	18. Distribution Statement
Probability of detection, Ultrasonics, Engines, Inspection	This document is available to the U.S. public through the National Technical Information Service (NTIS) Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	152	

ACKNOWLEDGEMENTS

The work described in this report would not have been possible without the important input of the following individuals.

Jon Tschopp, Cliff Shamblen, Dave Copley (retired), Jonathan Blank and Andy Woodfield (General Electric)
James Hall and Bill Knowles (Honeywell)
Dave Nissley and David Snow (Pratt & Whitney)
John Moyers (RMI Inc.)
Tim Gray (Iowa State University)
Floyd Spencer (Sandia/AANC)
Tim Mouzakis and Jon Bartos (consultant) (Federal Aviation Administration)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AANC	Airworthiness Assurance NDI Validation Center
AC	Advisory Circular
AIA	Aerospace Industries Association
CBS	Contaminated Billet Study
CBS_{10}	Subset of CBS database with flaw, area known from detailed metallography
DAC	Distance Amplitude Correction
DPD-3D	Default Probability of Detection-3Dimensional
EFBH	Equivalent Flat-Bottom Hole Area
ETC	Engine Titanium Consortium
FAA	Federal Aviation Administration
FBH	Flat-bottom hole
FSH	Full-screen height
ISU	Iowa State University
JETQC	Jet Engine Titanium Quality Committee
LEV	Largest extreme value
MOOT	Method of Optimal Truncation
ML	Maximum “loglikelihood”
OEM	Original equipment manufacturer
PoD	Probability of Detection
Re	Effective reflectivity
RISC	Rotor Integrity Sub-Committee
ROI	Region of interest
SDH	Side-drilled hole
SHA	Synthetic hard alpha inclusion
SNR	Signal-to-noise ratio
SNR_c	Signal-to-noise ratio for conventional inspection
SNR_{mz}	Signal-to-noise ratio for Multizone inspection
TRMD	Turbine Rotor Materials Design

EXECUTIVE SUMMARY

This report describes efforts to update the Default Probability of Detection (PoD) curves for the ultrasonic detection of hard alpha inclusions in titanium alloy billets, with the results intended to replace the curves currently found in Advisory Circular 33.14-1, "Damage Tolerance for High Energy Turbine Engine Rotors." The report is structured to provide a high-level overview of the major technical steps and results in the main body, with the extensive experimental and analytic work conducted to support these results documented in appendices. This report summarizes the strategy, the data on which the update is based, the implications of the structure of the data on the type of PoD analysis techniques needed, the statistical approaches used to analyze the data, the results of the PoD analysis, technical issues that were encountered in the study, and conclusions. Supporting details are provided in eight appendices.

The original Default PoD curves were based on what is known as the Default PoD 3Dimensional database, which documents the results of inspections conducted before 1995 using conventional inspection techniques. Two additional sets of data are described in this report that became available after that time and greatly extend the information available. Noteworthy features of these new data sets are the facts that (1) they greatly increase the amount of available data that can be used to assess PoD, (2) they include Multizone as well as conventional inspection data, and (3) for one data set, developed as a part of the Contaminated Billet Study (CBS), both Multizone and conventional inspections were performed on the same material, which allows the Multizone measurements to provide a referee technique to be used in the study of conventional inspection misses.

To provide background for the PoD analysis, some of the assumptions of the previous Default PoD analysis were re-examined in light of the information contained in the greatly expanded database. Physics-based models and experimental data on model systems were used to assist in the interpretation of the new data. A number of assumptions made in the previous analysis were found to be inappropriate based on this new information.

During the statistical analysis, a number of nontraditional techniques were developed. In the analysis of neither the conventional inspection nor the Multizone data was a simple linear regression (the approach traditionally used in PoD studies) sufficient. The nontraditional approaches are of two types. Some represent standard approaches that are well known in the statistics community but are sometimes overlooked in the analysis of PoD of nondestructive evaluation techniques. Other nontraditional approaches were special techniques that were tailored to the particular needs of the analysis of the Default PoD data sets.

There are two techniques commonly used to inspect titanium billet. The conventional technique uses a single transducer, often cylindrically focused near the billet surface. The Multizone technique uses a set of transducers, each designed to produce a beam focused at a different depth in the billet and is used to inspect a neighboring zone.

The updated PoD curves are found to be considerably different from the previous curves obtained in 1995. In the analysis of the conventional inspection data, it was found that the structure of the data, including known misses, and the nontraditional analysis techniques that had

to be developed to analyze that data, had a profound influence on the estimate of PoD, which is significantly less optimistic than the previous estimate. For small flaw sizes, the new curve is comparable to the original curves. However, the new estimate of PoD as a function of flaw size has a considerably lower slope, and hence, the PoD is predicted to be much less for large flaw sizes. The core areas at which the new curve reaches 10%, 50%, and 90% PoD are greater than those of the old curve by factors of approximately 1, 5, and 50, respectively.

For the case of the Multizone, the previous PoD estimate was based on a number of assumptions since no Multizone data were actually available at that time. This report provides the first true estimate of the Multizone PoD, taking into account a dual reject criteria based on amplitude and signal-to-noise ratio. This estimate is much closer to the previous one than was the case for the conventional inspection. Overall, the new estimated PoD curve has a lower slope. The core areas that the new curve reaches (10%, 50%, and 90% PoD) are less than those of the old curve by factors of 2.1, 2.1, and 1.4, respectively. However, for flaws of large size (having estimated PoD values greater than about 96%), the new curve flattens out and falls below the old curve. A full knowledge of misses was not available in the Multizone analysis. Only 25% of the CBS material was subjected to assessment for misses, and the referee technique used did not have a dramatically higher sensitivity than the Multizone. The PoD results represent best estimates based on the data at hand.

The PoD of the Multizone technique is shown to be clearly superior to that of the conventional inspection technique. The core areas at which the new Multizone curves reached values of 10%, 50%, and 90% PoD are less than those of the new conventional inspection curve by factors of approximately 4.6, 23, and 173, respectively.

1. INTRODUCTION.

1.1 PURPOSE.

The probability of detection (PoD) quantifies the effectiveness of nondestructive examinations. More specifically, the PoD is the probability that a particular nondestructive examination technique will detect a flaw of a given size. It is influenced by many factors such as the physical setup of the inspection (which can be understood in terms of the physics of the measurement), the performance of the operator setting up and interpreting the inspection (often called human factors or operator performance), and the reflectivity of the flaw of interest (generally influenced by its morphology) [1].

PoD plays a major role in the management of the lives of rotating components of aircraft engines. Figure 1 schematically shows how the knowledge of PoD is combined with estimates of the density of defects initially present (a priori defect distributions) and flaw growth laws to develop probabilistic estimates of component lifetime. In this context, PoD plays an important role in assessing the effectiveness of improved inspection techniques. PoD curves are also used as a step in estimating the a priori defect distributions, as quantified by exceedence curves, from knowledge of defects found during manufacturing [2]. The exceedence is defined as the probability (per million pounds of titanium) of having a hard alpha inclusion of a given size or larger [3].

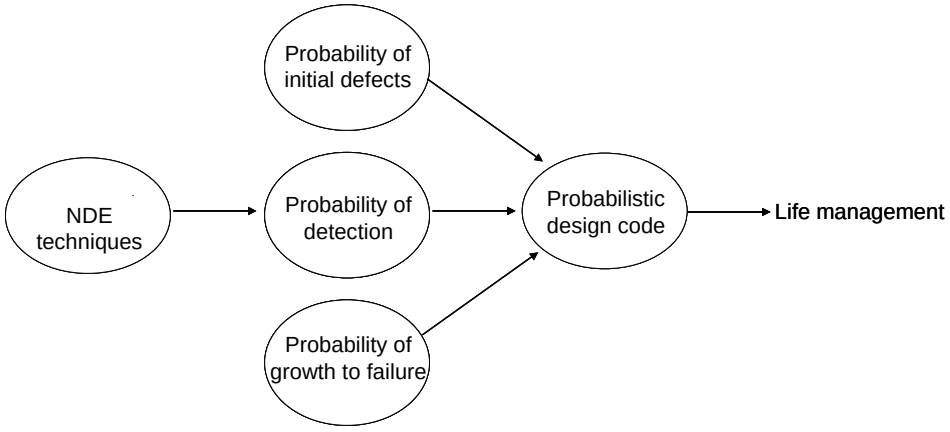


Figure 1. Schematic View of Role of Nondestructive Evaluation, as Quantified by PoD, in Probabilistic Life Management

A set of Default PoD curves appears in the Federal Aviation Administration (FAA) Advisory Circular (AC) 33.14-1, “Damage Tolerance for High Energy Turbine Engine Rotors” [3]. This report describes the results of an effort to update those curves based on new data and refined analysis. The way in which the original Default PoD curves were produced will be reviewed in the next section.

1.2 BACKGROUND.

1.2.1 The DPD-3D Database.

PoD can be estimated from data describing the response versus size for a representative set of defects. The original estimates of the Default PoD were based on such information, captured in what is known as the Default PoD-3Dimensional (DPD-3D) database. These data were obtained by recording the response levels of a set of naturally occurring hard alpha inclusions that were detected by suppliers during normal production inspections using conventional inspection techniques (to be defined more precisely in a subsequent section). All defects were carefully evaluated metallographically to obtain three-dimensional estimates of their sizes so a database, including response and size, could be constructed. More specifically, the database had 24 usable entries representing defects found in billet and bar. The material (alloy, billet diameter, melt type), the inspection results (calibration reflector, indication amplitude, noise levels), and the flaw size (axial, X_1 and X_2 dimensions for both the diffusion zone and core) were documented. Explanation of these terms follows.

The flaw response was measured in terms of % full-screen height (FSH) after calibration with a flat-bottom hole (FBH) of a specified size. Since data were present that were calibrated to both a #3 FBH (3/64 inch diameter) and a #2 FBH (2/64 inch diameter), a transformation was required to convert the FSH values to values that could be directly compared as discussed below. A deficiency of the data set was that, according to standard practice in the manufacturing environment, an indication that was off scale was often recorded as 100% or >100%, with no effort made to quantify the exact indication amplitude when the response was off scale. The exact value of a response >100% is of no value in the disposition of the defect, since such amplitudes are above the reject criteria. However, information of the exact value of the response is of value in the PoD analysis, and there are statistical tools available that can treat data sets in which this information is missing. There are five entries in the DPD-3D data set that show responses of >100% and another seven that show responses of 100%. It seems quite likely that most, if not all of the latter, had responses >100%, but there is no information as to how much larger.

The defects tend to be highly elongated with a major axis parallel to the axis of the billet: the axial dimension (A_x) that was recorded. Unfortunately, during the metallographic analysis, the rotation angle of the defects about this axis, with respect to the radial and circumferential directions in the billet, was lost. The dimensions X_1 and X_2 denote the dimensions in two orthogonal directions perpendicular to each other and to the axis but of unknown rotation about the axis of the billet.

Given this raw data, two transformations were used to determine the response and area that were used in the original Default PoD analysis. The % FSH was converted to a response measure known as the equivalent FBH area (EFBH). This was taken to be the size of a FBH that would produce the same signal amplitude as that of the naturally occurring flaw. The EFBH is computed according to the relation

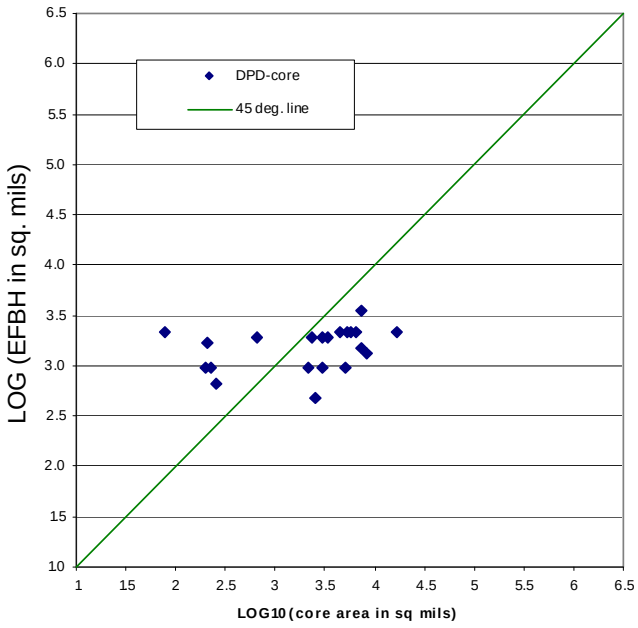
$$\text{EFBH} = (S/S_c)(\pi/4)(C/64)^2 \qquad (1)$$

where S is the flaw signal strength (in units of % FSH), S_c is the calibration signal strength (in units of % FSH), and C is the size of the calibration hole (in units of 1/64 inch diameter).

The area of the flaw was determined from the measured dimensions A_x, X_1, X_2. The goal was to use the area of the flaw in the axial-circumferential plane, the cross-section that would be interrogated by the ultrasonic inspection. However, since knowledge of the rotation of the X_1 and X_2 axes about the axial direction was lost, an estimate had to be made. The defect area was estimated to be

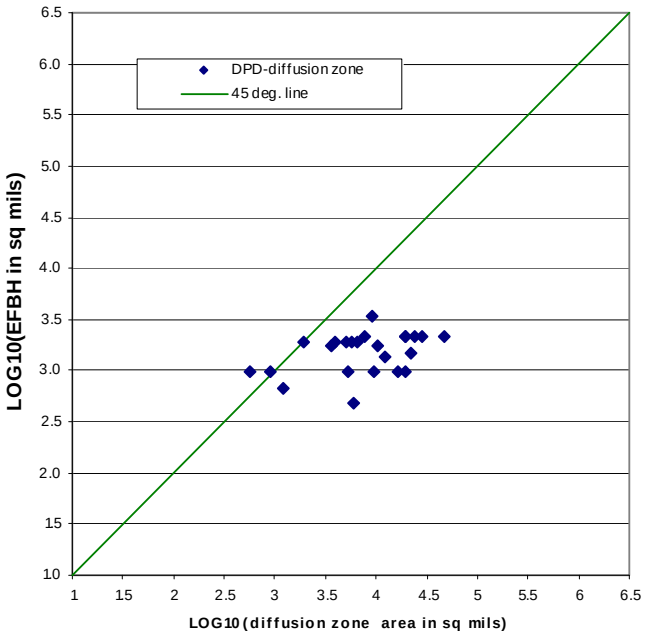
$$\text{Defect Area} = (\pi / 4) A_x \sqrt{X_1 X_2} \tag{2}$$

Figure 2 shows the response versus area (on a logarithmic scale) when the area is taken to be that of the (a) diffusion zone and (b) core. The terms diffusion zone and core will be defined more precisely in the discussion of the Contaminated Billet Study (CBS) database in appendix C.



(a)

Figure 2. Log EFBH Response Versus Log Metallographic Area for DPD-3D, Database, Conventional Inspection (a) Versus Log Diffusion Zone Area and (b) Versus Log Core Area



(b)

Figure 2. Log EFBH Response Versus Log Metallographic Area for DPD-3D, Database, Conventional Inspection (a) Versus Log Diffusion Zone Area and (b) Versus Log Core Area (Continued)

1.2.2 Effective Reflectivity Method.

The original Default PoD curves were based on the effective reflectivity (*Re*) method. As originally described in 1986 [4], the basic idea was to define *Re* as the ratio of the EFBH to the metallurgical area.

$$Re = \text{EFBH/Defect Area} \tag{3}$$

In principle, defects that are strongly reflecting would have a value of *Re* close to unity, while those that are weakly reflecting would have smaller values. In this technique, *Re* is treated as a statistical distribution from which the PoD can be predicted for various threshold levels.

In 1995, a modification of the *Re* method was developed by Burkel, et al. [5]. It was recognized that the *Re* method could also be thought of as a regression analysis and had many features in common with the $\hat{a}$ versus a technique that had been developed by Berens, et al. [6], as discussed in the next section.

1.2.3 $\hat{a}$ Versus a Method.

In the $\hat{a}$ versus a approach, $\hat{a}$ represents the flaw response (in this case, EFBH as defined in equation 1) and a represents the flaw size (in this case, the defect area as defined in equation 2). A linear regression is performed in logarithmic space, i.e., one seeks to describe the data by the random function

$$\log(\hat{a}) = B_1 \log(a) + B_0 + \delta \tag{4}$$

where δ is normally distributed with constant standard deviation. The regression analysis determines B_1, the slope of the regression, B_0, the intercept, and δ, the standard deviation. Fitting the data to a regression line in log-log space is equivalent to the statement that the flaw response varies with area in accordance with the following relation

$$\hat{a} = Ka^{B_1} \tag{5}$$

where $K = 10^{B_0}$. Thus, it is assumed that the median flaw response is proportional to the flaw area raised to an unknown exponent to be deduced from the data.

1.2.4 Relationship of *Re* and $\hat{a}$ Versus a Methods.

Burkel, et al. [5] recognized that the assumptions of the original *Re* approach were equivalent to those of the $\hat{a}$ versus a approach if the slope of the regression, B_1, was constrained to be equal to unity. Equation 5 shows this to be equivalent to the assumption that the flaw response is directly proportional to the flaw area. This assumption was an integral part of the original development of the *Re* approach and was based on the well-known fact that the signal produced by a FBH in the far field of an ultrasonic transducer is directly proportional to the area of the FBH [7]. Burkel, et al. also used a different distribution to describe deviations of the data from the regression line, so that the full description of their approach was that the data was described by the random function

$$\log(\text{EFBH}) = \log(a) + \mu + \varepsilon \tag{6}$$

where ε is the largest extreme value (LEV) distribution with constant standard deviation. The regression analysis determines the intercept μ and the standard deviation of the LEV distribution.

A feature of this approach is that one needs to estimate only two, rather than three, parameters from the regression, allowing the analysis of smaller data sets, with the enabling assumption being that the flaw response is proportional to area. This was justified theoretically based on the known behavior of FBHs. In addition, a large database [5], as reproduced in figure 3, was consistent with this assumption. This database contained data from many materials and defect types, including nickel-based powder metal forgings, titanium billet, titanium forgings, steel forgings, and the different types of defects that occur in each one. The overall trend in this data seems very consistent with the assumption of a slope of unity, i.e., response is directly proportional to area.

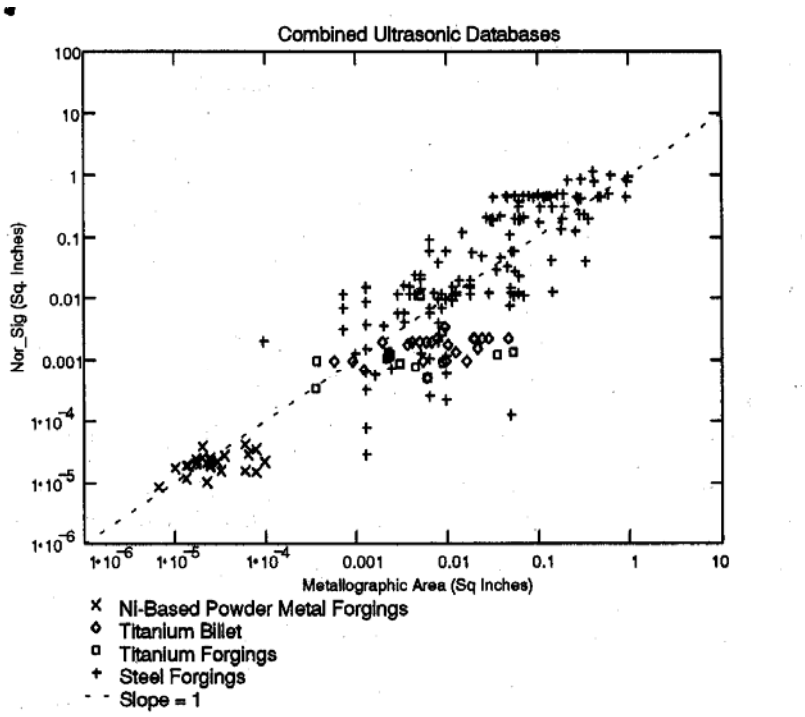


Figure 3. The EFBH Versus Metallographic Area for a Large Ultrasonic Database for Conventional Inspections

A subset of this large database was the DPD-3D database shown in figure 2(a). The DPD-3D database does not seem to have the same slope as the data in the combined database. However, as noted in section 1.2.1 and illustrated in figure 2, a large number of the DPD-3D responses were saturated. Hence, it was assumed at the time, the slope of the regression in the preparation of the original Default PoD curves should be forced to be equal to unity to be consistent with the larger database. The saturated data points were treated as right censored (meaning that one only knows that a response was present greater than or equal to the 100% value). In addition, the analysis assumed that the data was truncated (i.e., the response of all flaws in the inspected material may not have been present in the data set since there may have been misses, corresponding, for example, to flaws producing signals below a recording threshold).

1.2.5 Original Default PoD Curves.

The above procedures were used to determine the original Default PoD curves, which appear in AC 33.14-1 [3]. Thus, those curves were based on pre-1995 data obtained during conventional inspections that were conducted during manufacturing by a supplier. As noted in the above discussion, the sizes of the defects were determined from three-dimensional metallography, with the area being determined from the dimensions of the diffusion zone. There was an ambiguity regarding the rotation about the axis of the defects. A number of data points were treated as right-censored data due to saturation.

Figure 4 shows the PoD curves that would be obtained, based only on the DPD-3D database for #3 and #2 FBH calibrations, when the reject criteria was set at 75% of the calibration response

and the diffusion zone was taken as the measure of area. Related curves for reject criteria of 50% and 100% of the calibration response are found in AC 33.14-1.

At the time they were created, the curves, based on a #2 FBH, were taken to be representative of a Multizone inspection (which is calibrated at that level), although the curves were actually based on data obtained during conventional inspections.

The curves in figure 4 fall between those in figure A5-3 (#2 FBH calibration) and figure A5-4 (#3 FBH calibration) of AC 33.14-1 [3]. This is as expected, since the curves in figures A5-3 and A5-4 of AC 33.14-1 assume rejection at 50% and 100% of the signal from the calibration reflector and, hence, span the results for rejection at 75%, as shown in figure 4.

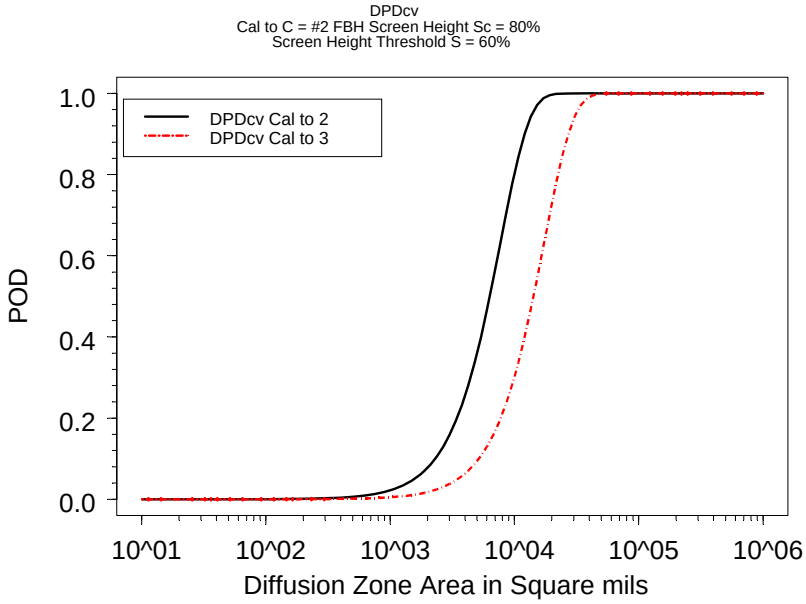


Figure 4. Original Default PoD Curve for #3 FBH and #2 FBH Calibration With Area of Diffusion Zone Taken as the Explanatory Variable for DPD-3D Database

Figure 5 shows the corresponding PoD curve that would have been obtained had the area of the core been used as the explanatory variable. When the initial Default PoD curves were produced, the lifing community felt that the area of the diffusion zone was the most appropriate measure to define initial flaw size in fracture analysis. More recent work within the lifing community, discussed further in sections 3.2 and 4.2, and in section A-4 of appendix A, suggests that the core area may be a more appropriate measure of initial size. In anticipation of the change in flaw size measure, figure 5 is provided at this point.

A casual examination suggests that the curves in figure 5 are just shifted versions of those in figure 4. However, this is not exactly true. It would be an exact shift if all the same data points were used and the areas of the core and diffusion zone were related by some fixed factor. Neither of these conditions were satisfied since some data points were lacking core dimensions and others were lacking diffusion zone dimensions (24 points were available with diffusion zone areas and 21 were available with core areas) and the areas were not related by a fixed factor.

Nevertheless, for practical purposes, a simple shift to a smaller size is a good approximation of the effect of changing the explanatory variable from a diffusion zone area to a core area.

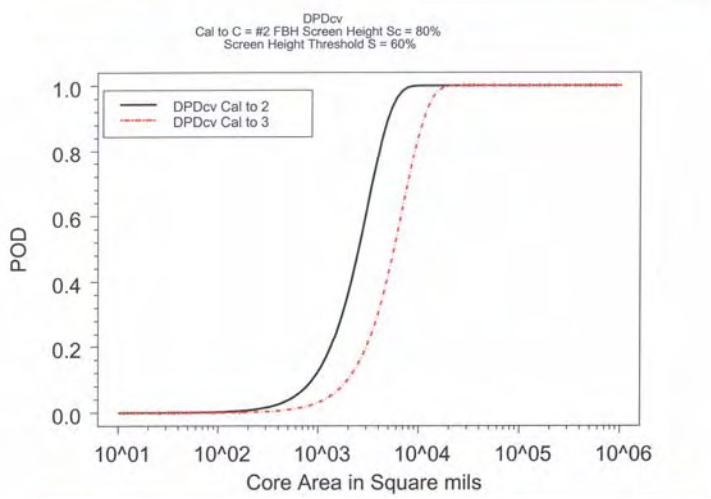


Figure 5. Original Default PoD Curve That Would Have Been Obtained for a #3 FBH and #2 FBH Calibration if the Area of Core had Been Taken as the Explanatory Variable for DPD-3D Database

1.2.6 Need for Update.

The assembly of the DPD-3D database and its analysis to produce the original PoD curves was a major advance, but a number of factors that occurred in the ensuing decade indicated the need to revisit the problem. Since 1995, a new, zoned inspection (the Multizone system) was implemented. One motivation for revisiting the Default PoD curves was to quantify the improvement offered by the higher focusing capability of multiple transducers in the Multizone system. An additional motivation was the knowledge, as discussed in section 1.2.1, of the limitations of the DPD-3D database. In addition, it was not clear that the assumptions of the statistical likelihood analysis (mean flaw response directly proportional to area with the distribution of responses having a constant standard deviation about this mean) were appropriate to this database. Two new data sets, known as Jet Engine Titanium Quality Committee (JETQC) and CBS, had also become available, which included Multizone as well as conventional inspection information. The JETQC was formed under the auspices of the Federal Aviation Administration (FAA), with memberships including all U.S. and European aircraft engine producers. The primary purpose of JETQC is the rapid dissemination of titanium alloy melt-related defect issues and data collection. Among the JETQC activities was the creation of a shared, sanitized database of melt-related defects from which data was extracted for use by the Engine Titanium Consortium (ETC) in the PoD analysis described here. The latter follows guidance provided in AC 33.15-1, “Manufacturing Process of Premium Quality Titanium Alloy Rotating Engine Components” [8]. The CBS database was made available through the efforts of the FAA-funded research program performed by ETC [9]. These new data sets provide a basis for the re-examination of the Default PoD curves.

1.2.7 Conventional Inspections.

Figure 6 shows a schematic comparison of the conventional inspection and Multizone systems. In the conventional inspection technique, shown at the top, a single, often cylindrically focused, transducer is used to inspect the billet. When a cylindrically focused transducer is used, the beam is focused near the surface in the circumferential-radial plane. There is no focusing in the axial direction, and the size of the beam in the circumferential-radial plane increases as one moves beyond the cylindrical focus towards the center of the billet.

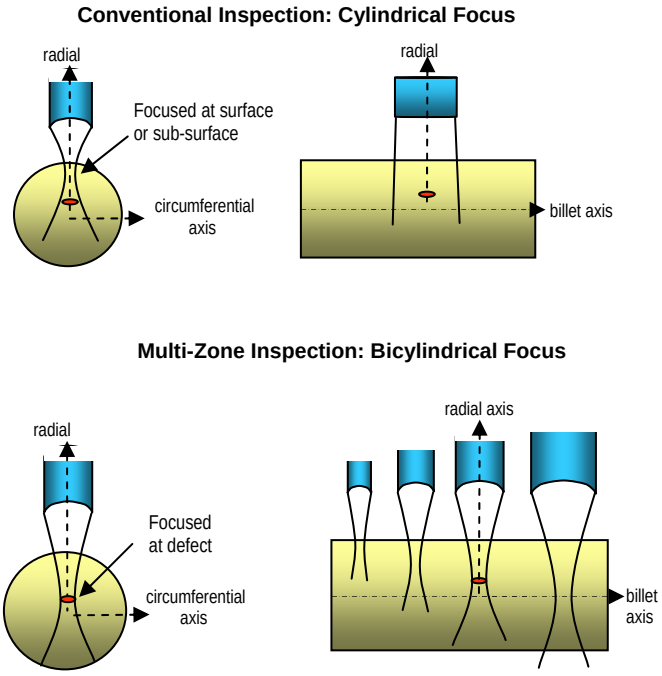
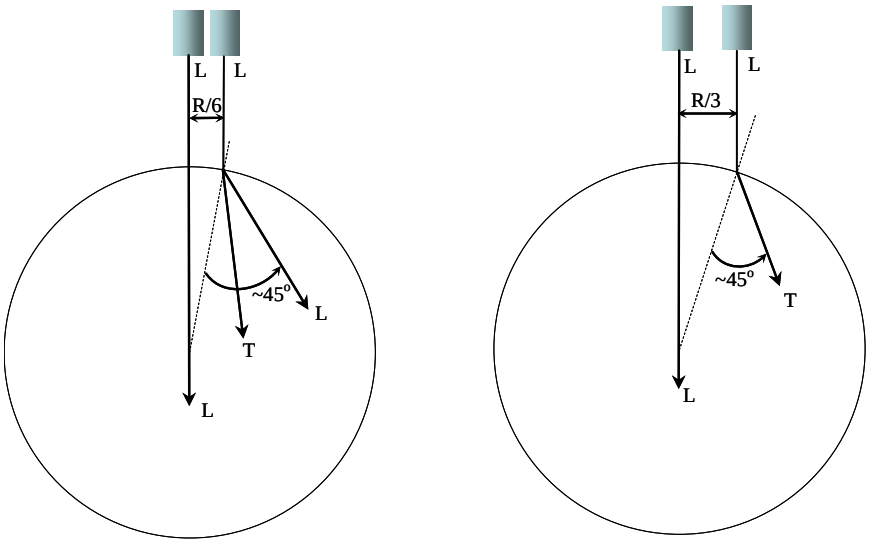


Figure 6. Schematic of Conventional Normal and Multizone Inspections

The details of the implementation of the conventional inspection technique depend on the original equipment manufacturer (OEM) requiring its use. In all cases, there is one set of measurements made with the probe positioned to direct energy normal to the billet surface, thereby producing a longitudinal beam nominally propagating along a radius of the billet. In the case of one OEM of aircraft engines, a second set of data was obtained with the probe offset such that the beam impinges at an angle to the surface normal, as shown in figure 7. As shown, depending on the degree of offset, refracted longitudinal or shear beams are produced at a variety of angles. The OEM of aircraft engines that uses angle beams requires the configuration shown in figure 7(a), involving a refracted longitudinal wave. The configuration shown in figure 7(b), used by some nonaerospace customers, was part of the original inspection of the CBS material. The conventional inspection PoD estimates to be developed in this report will only be based on the normal longitudinal data.



(a) (b)

Figure 7. Schematic of Conventional Angle Inspection (a) 45° Refracted L Wave Produced by Offset of 1/6 Radius and (b) 45° Refracted T Wave Produced by Offset of 1/3 Radius

The normal longitudinal conventional inspection technique is often calibrated to a #3 FBH, where the response is set to 80% of FSH on the display of the ultrasonic instrument, with the effects of depth being taken into account through a distance amplitude correction (DAC) curve. A typical current reject criterion, based on the amplitude of the signal, is 80% of FSH, i.e., the response of a #3 FBH at the same depth. Reject criteria of 60% FSH (3/4 of the response of a #3 FBH) or 70% FSH (7/8 of the response of a #3 FBH) have also been used in some cases. An additional signal-to-noise criterion is sometimes used. In a representative implementation, the operator will make a judgment of the average noise in the billet. Then, a flaw will be rejected if it produces a signal more than 3 dB above this noise level (as long as the signal is also more than 30% FSH). The signal-to-noise ratio for conventional inspection (SNR_c) technique is defined as

$$SNR_c = S/N_{avg} \tag{7}$$

where S is the peak flaw signal and N_{avg} is the average noise.

The refracted beam conventional inspection technique may either use the same calibration as the normal conventional inspection technique or may use a separate calibration reflector, often based on the response of a side-drilled hole (SDH). As for the normal beam, rejections can be based on absolute amplitude or signal-to-noise ratio (SNR).

During conventional inspections, digital records are not gathered. The data is sometimes collected in the form of a strip chart recording and a stop-on-defect procedure. If a candidate signal is found, as indicated by a signal exceeding an evaluation threshold, the inspection will be stopped and the operator will try to maximize the signal to see whether it exceeds the reject threshold (either amplitude or signal-to-noise). This maximization includes translating the probe along the billet axis and rotating the billet, possibly changing the probe angle.

1.2.8 Multizone Inspections.

The Multizone inspection uses a set of transducers, as shown at the bottom of figure 6, each of which is bicylindrically focused at different depths to ensure that all the material is inspected with beams that are reasonably tightly focused, a procedure that is known to enhance SNR [10]. Zones are established in which the focusing is adequate, typically having a depth of about an inch from top to bottom.

In contrast to conventional inspections, the data is gathered in the form of digital C-scan records for each zone. Accept or reject decisions can be made according to either of these two criteria. Amplitude-based rejections occur when the signal exceeds 7/8 of the response of a #2 FBH. There is no DAC because, in multiple zones, the flaw is never too far from a focal plane. There is a calibration specimen that has an FBH at the top and bottom of each zone. The smallest response is set to 80% and flaws producing a signal in excess of 7/8 of that value, i.e., 70% FSH, are rejected. In addition, there is a second reject criterion based on signal to noise. The SNR for Multizone inspection (SNR_{mz}) is defined as

$$SNR_{mz} = (S\text{-}N_{avg})/(N_{pk}\text{-}N_{avg}) \tag{8}$$

where S is the peak value of the flaw signal, N_{avg} is the mean value of the noise in some defined region of interest (ROI), and N_{pk} is the peak value of the noise in the same ROI. Flaws are rejected when $SNR_{mz} > 2.5$. This is equivalent to requiring that

$$S > 2.5\, N_{pk} - 1.5\, N_{avg} \tag{9}$$

Often, the SNR criterion is more sensitive than the amplitude criterion of $S > (7/8)$ (#2 FBH).

In early implementations of the Multizone inspection, the ROI was a window defined by the operator. In more recent versions, an automatic procedure has been developed that removes operator influence and takes into account the spatial variation of noise throughout the billet, leading to an effective SNR_{mz} threshold that varies from pixel to pixel, as described in reference 11.

1.3 PROGRAM OBJECTIVES.

The objective of this program was to update the Default PoD curves for the ultrasonic detection of hard alpha inclusions in titanium billet based on the most current information, including a consideration of both conventional and Multizone inspections and all relevant data sets.

1.4 RELATED ACTIVITIES/DOCUMENTS.

A number of related documents appear in both the open literature and technical reports issued by the FAA. Technical descriptions of the *Re* technique have been prepared by Sturges, et al. [4] and Burkel, et al. [5]. These were the basis for the original Default PoD curves that appear in AC 33.14-1 [3]. Using those curves to estimate exceedance curves (defining the distribution of defect sizes expected to be produced by manufacturing) was prepared by the Aerospace Industries Association (AIA) Rotor Integrity Sub-Committee [2]. An important source of new

data that was used extensively in the preparation of the updated Default PoD curves was the CBS, as described in FAA report DOT/FAA/AR-05/16 [9].

The use of information from signals found in field inspections is a challenging process. Some perspectives on these challenges, including the problems posed by lack of knowledge of misses, as developed by a North Atlantic Treaty Organization-sponsored group, can be found in reference 12. Emerging approaches to the determination of PoD, based on the use of physics-based models, are under development, and the initial results are described in reference 13.

2. APPROACH.

2.1 OVERVIEW.

The initial approach to the update of the Default PoD curves was to develop new databases from the information obtained as a part of the JETQC and CBS programs. The combination of these databases with the DPD-3D database was to be analyzed using the *Re* technique. The *Re* technique is based on the idea that an effective reflectivity can be defined for each flaw that is equivalent to the ratio of the signal reflected by that flaw to the signal that would be reflected by a FBH of the same size. It was anticipated that an extension to the *Re* technique would be required in the analysis of the Multizone data because only a portion of the flaw would generally be illuminated by the more tightly focused beam. Hence, it could not simply be assumed that the flaw response would increase proportional to the flaw area for all flaw sizes once the flaw had grown larger than the beam. It was hypothesized that one could assume the flaw response was proportional to the area of the flaw illuminated by the beam. This required that additional data be gathered, not available in the DPD-3D database, defining the ultrasonic inspection parameters and, hence, the size of the beam illuminating the flaw. This would allow the size of the illuminating beam to be calculated at any flaw depth.

During the course of the analysis, a number of difficulties caused a deviation from the initial plan. Analysis of known misses in the CBS database implied that this data was inconsistent with the PoD curves deduced from the DPD-3D database as analyzed by the *Re* technique. The new ultrasonic databases indicated that constraining the slope of the regression line to a value of unity, as implied by the *Re* technique, was inappropriate. As discussed in appendix D, a deeper examination of the physics of scattering from defects of complex shape provided an explanation of this result. Relaxing this constraint led to an $\hat{a}$ versus a analysis that often proved unstable for the conventional data, possibly a consequence of the greater number of free parameters to be deduced from the data.

This report discusses these and other issues that were encountered and the steps that were taken to address them. Details follow in section 3 and are supported by appendices A through H.

2.2 WORK SCHEDULE.

This work was conducted in four phases.

The first phase, consisting of the analysis of new data based on the *Re* technique, was scheduled for the period of August 2003-January 2004. This was completed on schedule, but produced

unexpected, counter-intuitive results, as noted in the previous section. This suggested difficulties with the *Re* analysis technique when applied to such data.

In the second phase, during the period February-November 2004, these difficulties were addressed. A preliminary report was submitted to the FAA in Microsoft® PowerPoint® form on November 14, 2004. PoD curves were presented that represented the best estimates of the ETC at that time for both conventional and Multizone inspections.

In the third phase, some of the predictions were still felt to be counter-intuitive, and a peer review was conducted by Floyd Spencer of the Airworthiness Assurance NDI Validation Center (AANC) operated by Sandia National Laboratory for the FAA, with the results reported to the ETC on May 5-6, 2004. This review supported many decisions made in the analysis by the ETC but also contained suggestions for refinement of certain aspects. This led to further examinations of the data sets and analysis techniques in the period of November 2004-April 2005.

In the fourth phase, during the period of April 2005-April 2006, further discussion and refinement of the results were conducted by an ETC/FAA coordination team, consisting of R. B. Thompson and W. Q. Meeker, representing the ETC; F. W. Spencer of AANC, providing an independent statistical perspective; C. Nguyen and T. Mouzakis, representing the FAA; and J. Bartos, serving as a consultant to the FAA. In addition, knowledge of a Multizone miss became available, having some impact on the final estimates of PoD. This report presents the final results of all the above work, which appear in section 4.

3. DISCUSSION OF RESULTS.

3.1 STRATEGY.

The initial plan was to apply the *Re* technique to an expanded data set, including the information in the DPD-3D, CBS, and JETQC databases. It was anticipated that a correction would have to be made, in the case of the Multizone data, to take into account the fact that the flaw was not fully illuminated. The motivation for the latter was as follows.

Recall that *Re* is defined as the ratio of the EFBH (controlled by the flaw response) to the defect area. This can be thought of as a quantification of the reduced reflectivity of a defect with respect to a planar reflector of the same area because of the flaw's irregular morphology. However, it is sometimes the case, particularly in Multizone inspections, that the flaw is larger than the illuminating ultrasonic beam, as shown in figure 8. In this case, it no longer seems reasonable to assume that the response will increase with flaw area, since those portions of the flaw extending beyond the beam would not contribute to the signal. Hence, the *Re* would tend to decrease as the flaw size grows in a way that has nothing to do with the flaw morphology or reflectivity. This result seems counter to the fundamental hypothesis of the *Re* method: *Re* is a distribution controlled by flaw morphology but not flaw size.

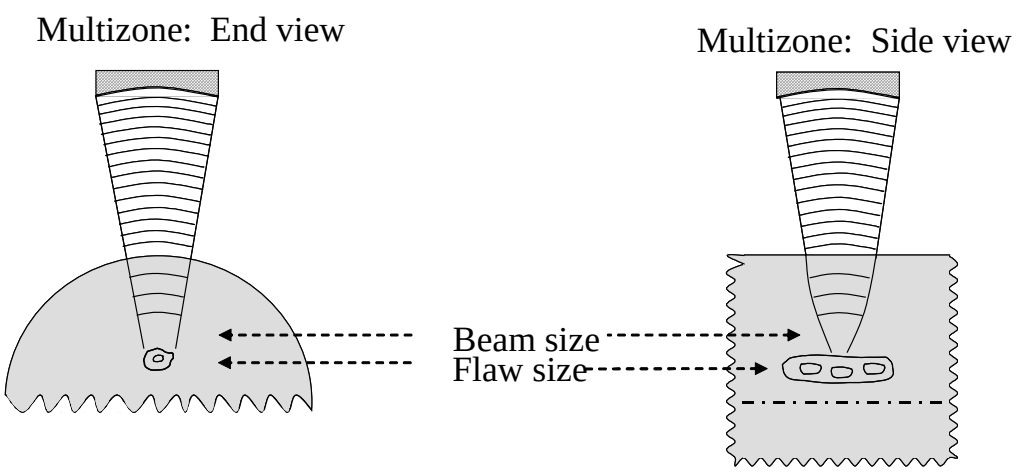


Figure 8. Effects of Finite Beam Size

One way that was suggested to overcome this limitation was to replace the defect area by the illuminated area of the flaw. When the beam is larger than the flaw, this is the full flaw area. However, when the flaw extends beyond the beam, it is just the portion of the flaw that is insonified that can contribute to the signal. This approach was considered during the first phase of the work in which attempts were made to apply the *Re* technique to an expanded data set.

During the course of that work, some conceptual problems with the illuminated area concept were encountered, which are summarized below.

If one assumes that the flaw response is governed by a distribution of effective reflectivities applied to the illuminated area, then the PoD curves would have to level off when the flaw became larger than the beam size. The area of the flaw extending beyond the beam could not contribute to the reflectivity. If flaws grew larger than this size, their PoD would not increase. Hence, PoD would never uniformly rise to 100%. The PoD curve would first rise and then reach a constant plateau related to the size of the beam.

This seemed somewhat counter intuitive, and further thought suggested a way to avoid this dilemma. Most of the billet hard alpha inclusions have an elongated shape. During a Multizone inspection, the beam will pass over an elongated flaw several times, leading to multiple opportunities for detection. It would be logical to think that the probability of detecting the flaw would increase as the number of opportunities increase, since different regions of the flaw might have different reflectivities. From the formal statistical perspective, one could say that, each time the transducer is indexed to a new position, the response is an approximately independent sampling of the *Re* distribution, where the illuminated area is taken as the size. Detection is based on the largest signal returned from the flaw. This is the maximum value obtained when the *Re* distribution is sampled N times, where N is the number of index steps in which the flaw is insonified.

This seemed to be a logical way to proceed and might be a fruitful topic for future work. However, in the short term, it was evident that this approach would require the development of some complicated statistical machinery, and so, an alternate approach was taken. Rather than try to break the analysis of the Multizone inspection down to this level of mechanistic detail to extend the *Re* approach, a simpler view was adopted. The inspection is a complex process. However, one could ignore these complexities and simply view the inspection as a “black box” with a flaw as the input and the flaw response as the output. This was the philosophy used in the

analysis described in this report. It was assumed that the inspection was more likely to obtain an optimized (large) signal level from a flaw with multiple hit opportunities. Hence, the $\hat{a}$ versus a technique was taken as a starting point.

To make this report accessible to the reader not wishing to review all of the details of the PoD analysis, a format will be adopted in which a good bit of that detail is documented in a set of appendices. The main body of this section provides a high-level overview of the major technical steps and results. Appendices, referenced in the main body, provide a detailed documentation of the extensive experimental and analytical work conducted to support those results.

3.2 SUMMARY OF NEW DATA.

Whatever statistical technique is used, the estimate of PoD can be no better than the quality of the data. Appendix A discusses in more detail the required attributes of a database, which must contain accurate values for ultrasonic response, flaw size, and position. Desirable, though not ultimately used in the current analysis, is knowledge of the size of the ultrasonic beam in the plane of the flaw.

In the Default PoD analysis conducted in 1995, the flaw size used was the extent of the diffusion zone. Improvements in the understanding of the fracture process for internal hard alpha inclusions since that time have indicated that the fracture of internal hard alpha inclusions is controlled by the size of the core rather than the size of the diffusion zone. Appendix A also discusses this. It is important to keep in mind this change in the explanatory variable when comparing the new Default PoD curves to the old Default PoD curves.

The Default PoD analysis was based on three sets of data. Appendix B describes the DPD-3D and JETQC databases, each of which is based on defects found under normal production inspections. The initial motivation for creating the JETQC database was to determine root causes of hard alpha formation, with the goal of providing feedback to the manufacturing process. Thus, the information gathering process was not tailored to the needs of PoD analysis. A discussion of adjustments that were made to address situations in which this led to unreasonable results is included in appendix B.

The third data set was obtained as a part of the CBS. This study was based on a contaminated heat of titanium that occurred as a result of a power outage during production. A pressure spike related to the shutdown of the vacuum pumps before the furnace was backfilled with Argon is believed to have been the cause of the contamination. After determining that the hard alpha inclusions in the contaminated heat were representative, the heat was purchased by the ETC for a number of purposes, including gathering data for PoD analysis. The results of the CBS study, which was completed prior to the update of the Default PoD curves, are documented in reference 10. However, the exercise of determining the updated Default PoD curves required a deeper examination of a number of aspects of the data. Most important for this study was a careful determination of the sizes of the hard alpha inclusions, assessing a possible “miss” identified in the original study, and correcting some reporting errors. Appendix C summarizes the CBS data set with emphasis on this new information.

An important difference between the CBS data set and the previous DPD-3D and JETQC data sets was the fact that both conventional and Multizone inspections were performed on all CBS material. This allowed a back-to-back comparison of the two techniques and, as reported previously [10], the superior performance of the Multizone was clearly demonstrated. The existence of back-to-back inspection data has an important implication in the PoD analysis. Multizone provided explicit knowledge of a number of conventional inspection misses of flaws of relatively large size, a result that required new statistical techniques to be developed before a meaningful analysis could be conducted, and had a significant influence on the resulting Default PoD curves for conventional inspection. Unfortunately, there was no corresponding referee technique for the Multizone inspection. One candidate Multizone miss had been identified in the misses assessment activity CBS study [10], and the importance of misses on the PoD determination caused that candidate to be examined metallurgically to determine whether the signal was truly from a hard alpha inclusion. Appendix C also includes the result of that analysis, which confirmed that the signal was from a hard alpha inclusion.

3.3 SUMMARY OF IMPLICATIONS OF NEW DATA ON PoD ANALYSIS TECHNIQUES.

As discussed in section 1.2.2, the original Default PoD curves were based on the *Re* method, which assumed the flaw response is directly proportional to the flaw area. As more data became available for the PoD analysis, it was increasingly clear that this was not a good assumption. Since that assumption was based on a physical model, namely, the expected dependence of the response of an FBH (or penny-shaped flaw) on an area, there was a need to understand why this was not an appropriate assumption. Appendix D establishes the reason, based on physics-based models for the response of flaws of other shapes (spheres, cylinders, and reconstructions of the void structure of hard alpha inclusions), supported in many cases by experiment. This work supported the decision to use the $\hat{a}$ versus a technique, rather than *Re*.

Under ideal conditions, a PoD study would be based on a set of samples in which the presence of all defects were known (i.e., there were no misses of defects whose presence was not even known) and there was perfect knowledge of the signal strength produced by each defect (i.e., no saturation of large signals or misses of small signals that would have produced signals below a noise floor). In the data obtained from flaws found during manufacturing or field inspections, it is often the case that none of these conditions are satisfied. The statistical analysis must take these nonidealities in the data into account to make the best estimate of the PoD. It is often useful to conduct an analysis of simulated data to test the effectiveness of the statistical techniques employed. The idea is that one starts with a data set for which one knows the ground truth, modifies it to the form that would be found in experiment, as influenced by one or more of the above problems, and then assesses how well the analysis can correct for those deficiencies in the data. Appendix E presents the results of such a study. These strongly influenced the final selection of statistical analysis techniques used in this work.

In the $\hat{a}$ versus a technique, log flaw response, $\hat{a}$, is plotted versus log flaw area, a, and a linear regression line is fit to the result. Unfortunately, when this strategy is followed on the data sets being analyzed in this study, even taking into account the issues noted above, one often obtains very unphysical results. For example, at one point, it was predicted that there is a significant PoD for Multizone inspections for core areas on the order of square mils. This seems very unlikely based on both physical intuition and the practical fact that there are no flaws of this size

in the database. Appendix F presents an analysis based on the physics of ultrasonic scattering from defects, suggesting that one linear regression for all sizes (the assumption in $\hat{a}$ versus a analysis) is not appropriate. An upper bound is derived, below which all data should fall. This explains the data nicely, since the break in the upper bound (between a steep slope at small flaw areas and a low slope at large flaw areas) occurs very near to the point at which the data supporting the initial linear regression ends. Explicit relationships for these upper bounds are also provided in appendix F. These were used in the final PoD analysis for both conventional and Multizone inspections.

The recognition that deviations are required from a linear regression of log $\hat{a}$ versus log a data leads to the consideration of more general regression techniques. If one had enough data, one might like to fit a curved function to the log $\hat{a}$ versus log a data. For the data sets analyzed in this study, the data was not sufficient to support such a curvilinear analysis. However, the data did support a first step in that direction, a "kink" regression in which the data was fit to two line segments intersecting at a "join point." Appendix G discusses that approach in greater detail and shows how it relates to the physics-based upper bound discussed in appendix F. The kink regression was used in the analysis of the Multizone data in this report.

3.4 OVERVIEW OF STATISTICAL ANALYSIS.

3.4.1 Summary of Data.

The update of the Default PoD curves was based on the three data sets discussed in detail in appendices B and C: the DPD-3D, the JETQC, and the CBS data sets. The nature of those data sets is summarized below.

a. DPD-3D Database

- Data used in the preparation of the prior Default PoD curves
- 24 original data points based on finds by suppliers, 21 of which are suitable for analysis as a function of core area
- All obtained during conventional inspection
- Three-dimensional metallography done by an OEM
- Origins of individual points sanitized to protect supplier
- Loss of information about circumferential versus radial dimensions (orientation of metallographs with respect to billet coordinates not preserved)
- Many points saturated (typical of manufacturing practice since saturation does not affect reject decision)

- Points for which response is much greater than that of a FBH of the same size, i.e., when EFBH>>Core Area, were removed

- Final data set contained 19 points after adjustments

b. JETQC Database

- Primary purpose of JETQC database is to study root causes of hard alpha formation

- A subset of the data was selected for which the following was available:

  - Unsaturated ultrasonic response
  - Information defining the inspection conditions
  - Three-dimensional flaw size information

- 16 original conventional inspection data points based on finds by suppliers

- 14 original Multizone inspection data points based on finds by suppliers

- Metallography done by suppliers with results tabulated by JETQC in a sanitized fashion

- Loss of information about circumferential versus radial dimensions (orientation of metallographs with respect to billet coordinates not preserved)

- Points for which response is much greater than that of a FBH of the same size, i.e., when EFBH>>Core Area, were removed

- Points for which consultation with original records revealed misinterpretation in JETQC database were removed or modified

- Final data set contained 14 conventional points and 6 Multizone points after adjustments

c. CBS Database

- 64 flaws identified in the CBS

- Both conventional inspection and Multizone data
- 61 Multizone points with size information

- 51 conventional points with size information

  - Each point is either an amplitude or knowledge of a miss (left-censored data)

- Number of conventional points is less than Multizone because of regions uninspected by Conventional

- For 10 flaws, area known from detailed three-dimensional metallography (CBS_{10})

- For remaining flaws, area estimated from C-scan length using an expression validated by comparison to DPD-3D, JETQC, and CBS_{10} data

- One additional flaw, missed by Multizone and conventional inspection during CBS, but identified in subsequent analysis, added to the database

3.4.2 Synopsis of Statistical Analysis.

The technical details of the statistical analysis are presented in appendix H. Here, a qualitative overview of some of the issues and analysis techniques will be given for the reader not interested in that level of mathematical detail. Before doing so, some special statistical situations will be reviewed. Standard tools exist to handle each of these situations.

- Saturated data—The true response is greater than or equal to a certain value. Such data is treated as "right censored" data.

- Misses—A flaw was present, but not observed. Hence, the response had to be less than or equal to a noise level. This is called "left censored" data.

- Truncation—There may have been flaws present that were not known. Hence, the data that is analyzed does not correspond to the responses of all flaws. The assumption of truncation takes this into account in the analysis. It requires the specification of a truncation level below which a response would have not been recorded. There would be no knowledge of a flaw below this level, unless a referee technique provided information about misses.

The $\hat{a}$ versus a technique has been used as the basis of the analysis for the reasons discussed in previous sections. Viewing the $\hat{a}$ versus a analysis as a linear regression, it is well known how to handle censoring and truncation within this framework. However, to properly treat the rather unique data sets that are available for this study, it has been necessary to extend the $\hat{a}$ versus a technique in several ways. New statistical tools had to be developed to accomplish this.

Many challenges were encountered in the analysis of the conventional inspection data. In the conventional inspection analysis, knowledge of misses (provided by Multizone serving as a referee in the CBS data set) caused the linear regression, which is the basis of the $\hat{a}$ versus a technique to be a poor fit to the data. The assumption of a distribution of the form of equation 4 simply did not describe the observations. In particular, there were a number of misses, at large flaw sizes, which did not appear to be consistent with any reasonable set of parameters in the distribution. This implied that it was not appropriate to assume that form of the distribution. An

accommodation term had to be added to the statistical model to allow the analysis to take into account those misses, at large flaw sizes, that were not consistent with the linear regression. These were termed Type II (atypical) misses. This is in contrast to Type I misses, which are caused by data consistent with the linear regression but below the threshold. It should be noted that there is no information in the data that would identify whether a particular miss is typical or atypical.

The simulation studies (appendix E) illustrated another challenge that would have existed even without the need for an accommodation term; the importance of having a sufficient number of data points in an analysis of this type. The simulation study reported in appendix E was idealized because the synthetic data were generated in a way that was consistent with the statistical model that was used in the analysis. Hence, this problem would be expected to be more acute in the analysis of real data. Because these issues raised concerns about the accuracy of the final results, a technique known as borrowing strength was evaluated. The basic idea was to bring some of the Multizone data into the conventional inspection analysis. It is assumed that the slope of the PoD curves for Multizone and conventional inspections will be the same, with both data sets playing a role in the determination of this slope. The data is, thus, asked to primarily determine the offset of the PoD curves for the Multizone and conventional inspections. However, in the final analysis, borrowing strength was not found to have much effect and, hence, was not used.

In the analysis of both conventional and Multizone inspections, the data is considered as truncated. When there is no referee, as in Multizone data, the truncation level (the level below which a flaw signal would not be observed) is determined by the noise. However, in the analysis of the CBS conventional inspection data, there were Multizone data that provided a referee. Flaws that were found by Multizone and missed by conventional inspection were treated as left-censored data. Thus, it made sense in the analysis of the CBS data points to set the truncation level at a level comparable to the Multizone noise rather than the conventional inspection noise. Any flaw producing a signal greater than this Multizone level, but less than the conventional threshold, would be in the data set as a left-censored (miss) point.

In the Multizone inspection, rejection is based on dual criteria, an amplitude threshold and a SNR_{mz} threshold. To give credit for this, a bivariate generalization of the $\hat{a}$ versus a model was developed.

In both conventional and Multizone inspections, for the reasons cited in section 3.1, an extrapolation of the regression to small flaw sizes was not determined to be appropriate. Instead, information from the physics of scattering, as discussed in appendix F, was used to establish a small flaw limit.

More specifically, for both conventional and Multizone inspection data, the regression line relating log response to log area tended to have a small slope (on the order of 0.2) as compared to *Re* (slope of 1.0). This implied a PoD for small flaw sizes that was unexpectedly large, a result that was determined unrealistic, since flaws of those small sizes are not found in any of the databases, i.e., historical experience indicates that flaws of these sizes are not found. Physics, as described in appendix F, predicts that the slope of a plot of log response versus log area must be 1.5 for simple geometric shapes in the small flaw limit (Rayleigh scattering regime). A physics-

based model was developed to describe the response of an upper-bound void in the shape of a 7:1 prolate spheroid in this regime, as well as in the regime in which the flaw is large with respect to the wavelength. It was required that the regression model that empirically takes into account the response of flaws of larger size not be extended above this line. Where the regression line lay above this small flaw, upper-bound response, it was replaced by the upper-bound small flaw response. The variance in the small flaw limit, for lack of other information, was taken to be the same as that inferred from the regression analysis. As an approximation, the correction is the same for conventional and Multizone inspections when the response is given in terms of EFBH.

The presence of the Multizone miss and the general structure of the Multizone data set suggested that a straight line (as implied by using a linear regression) would not provide a good fit to the data, as shown in figure C-32 of appendix C. Since the physics also suggest that the true relationship between response and area must be curved (see figure G-1 of appendix G), it was decided to perform a kink-regression, as described in appendix G. Instead of fitting the data to a single line, the data were fit to two line segments, joined at a kink. This was a compromise between simplicity (just two line segments) and the likely reality of a curved relationship.

When this kink regression was used, the philosophy of the small flaw correction was the same as discussed above for the linear regression case. When the kink regression line passed above the small flaw limit, it was replaced by that limit. For the specific case of the Multizone data analyzed in this work, the kink regression lay below the response of the upper-bound, small flaw limit for nearly the entire region for which PoD has a significant value, reducing the importance of the small flaw limit calculations in the final analysis.

Technical details of the analysis of the conventional and Multizone inspection data may be found in appendix H. Included is treatment of truncation and confidence intervals.

3.5 RESULTS OF PoD ANALYSIS OF CONVENTIONAL DATA.

3.5.1 The Full Database.

In the analysis of the conventional inspection PoD, only the results of normal, longitudinal inspection were considered, and rejection was assumed to only be based on an amplitude criterion. Thus, no credit was taken for refracted angle inspection that was employed by one OEM of aircraft engines or for informal, signal-to-noise rejection criteria that are sometimes employed by industry. Hence, the analysis could be characterized as conservative. Specifically, the analysis was based on the JETQC, DPD-3D, and the 1994 CBS data with area estimated from length by an empirical formula given in appendix C for those flaws which were not destructively sectioned. Some corrections had to be made to the original data sets based on a re-examination of original records and relevant physics. Conventional misses in the dead-zone (depth <0.375) and the end burst region were removed from the CBS data set. Defects for which the EFBH was much larger than area were removed from the DPD-3D and JETQC conventional data set. No cases needing such attention were found in the CBS data sets for either conventional or Multizone inspections.

In appendices C and D, plots of $\hat{a}$ versus a are presented individually for the DPD-3D, JETQC, and CBS data sets. Figure 9 shows the combination of all three (the data set that was analyzed to obtain the PoD). A special notation was adopted to identify the different nature of the individual data points, since this nature was taken into account in the analysis. The notation is as follows:

- cv = conventional inspection data
- N = normal longitudinal data
- Right = right censored points (saturated)
- Left = left-censored points (known misses)
- Exact = data for which unsaturated responses have been observed

Note that there information on misses only for the CBS data, and hence, left-censored points. This information was developed systematically for the conventional inspection (because of the Multizone inspection acting as a referee).

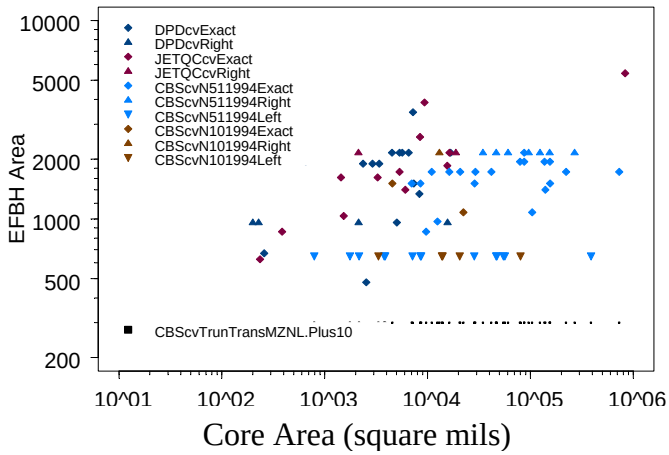


Figure 9. Complete Data Set That was the Basis for the Conventional Inspection Analysis

3.5.2 Analysis Results.

As discussed previously, the analysis included a model accommodation term to describe Type II (atypical) misses that were inconsistent with a model based on a linear regression. A physics-based correction was also introduced for small flaws. In the early stages of the analysis, strength was borrowed from Multizone amplitude versus area relationship (common PoD slope parameter). This assumption was dropped in the final analysis. Truncation was set at a low level because of the Multizone information on misses.

Figure 10 repeats the data shown in figure 9, with the linear regression line added, as determined by the maximum likelihood analysis. Note that from an intuitive perspective, this regression fits all but the left-censored (misses) data. The accommodation term required to account for Type II misses is not shown. Also shown is the line corresponding to the small flaw correction. The fact that the small flaw correction intersects the regression line very close to the point below which there is no further data (with no adjustable parameters) lends a considerable amount of confidence to the correctness of this analysis.

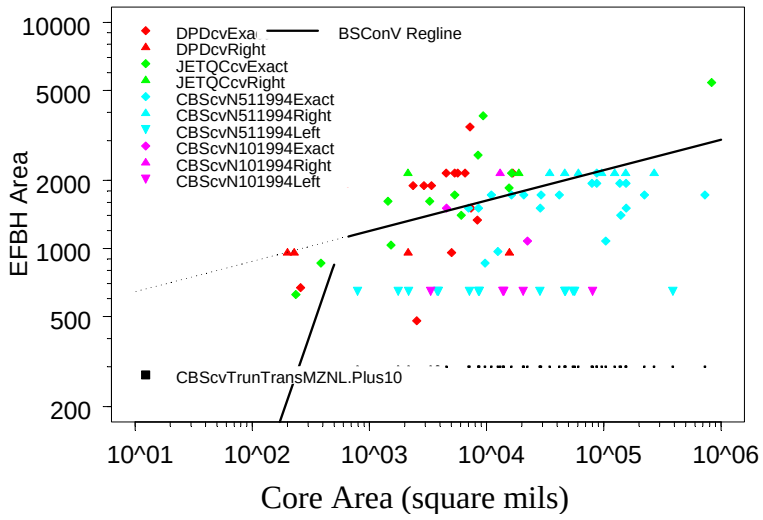


Figure 10. Experimental Data Showing Fits of Linear Regression and Small Flaw Limit

In figure 11, the conventional inspection PoD curves are shown as they would exist if there were no Type II (atypical) misses, i.e., as governed by the linear regression line and small flaw limit, as shown on figure 10, but ignoring the accommodation term needed to account for the Type II misses. The results of analysis with borrowing strength or without borrowing strength are both shown. It should be noted that the borrowing strength analysis had little influence on the results and, hence, this approach was not employed in the next stage of the analysis, which took into account the Type II misses.

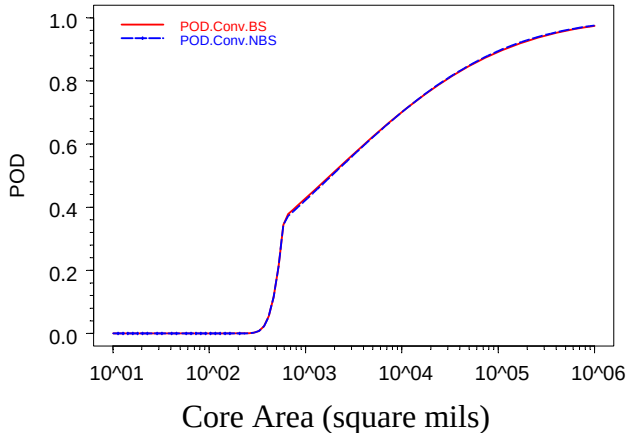


Figure 11. Results of Conventional Inspection Analysis if Type II Misses are Ignored

A separate conventional inspection analysis was conducted, ignoring all of the misses and performing a standard $\hat{a}$ versus a analysis. The results were nearly indistinguishable from those shown in figure 11.

Figure 12 shows the estimate of PoD when the Type II misses are taken into account and compares it to the probability of no Type II misses (1-probability of Type II misses). The areas

of the defects that were hits or misses are indicated by the symbol "x" plotted at an ordinate value of 1.0 or 0.0, respectively. It intuitively shows that the relatively slow increase in PoD as flaw area increases is a consequence of the misses at relatively large flaw sizes. These misses were not well described by a simple linear regression, hence, the need to introduce an accommodation term to describe unexplained Type II misses. It is clear that the physical effects described by the probability of Type II misses are controlling the PoD.

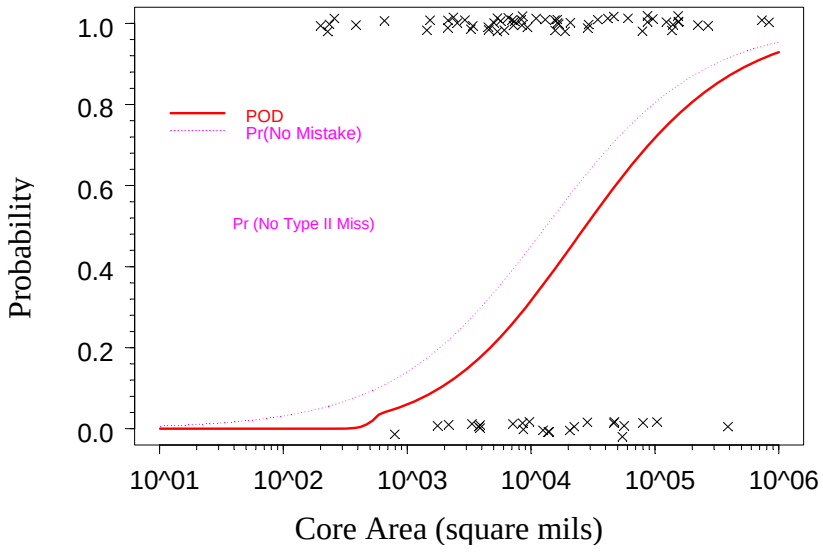


Figure 12. Conventional Inspection PoD Estimate Taking Type II Misses Into Account

3.6 RESULTS OF PoD ANALYSIS OF MULTIZONE DATA.

3.6.1 Review of Data.

From one perspective, the Multizone data are simpler than the data for the conventional inspection case in that there are only two sets (JETQC and 1995 CBS data). There was no application of a significantly higher sensitivity referee technique to provide systematic information on misses. About 25% of the CBS material was inspected with a technique that was marginally superior to Multizone as a part of the CBS misses assessment, and this led to information on one miss, as mentioned previously and discussed in detail in appendix C. From another perspective, the data is more complex than for the conventional inspection because each flaw has a SNR_{mz} as well as an amplitude, and rejections occur when either one exceeds a threshold value. Hence, the analysis must take into account the fact that there are dual reject criteria based on amplitude and SNR_{mz}.

3.6.2 Analysis Results.

As was mentioned in section 3.4.2 and discussed in detail in appendix H, a bivariate regression was used to model amplitude and SNR_{mz} simultaneously. Figure 13 shows the amplitude data and the kink regression line as determined by the maximum likelihood analysis. The line corresponding to the small flaw correction (as described in appendix F) is not shown because it lies above the kinked regression line for the domain covered by this graph. Truncation was at 10% FSH above the noise level.

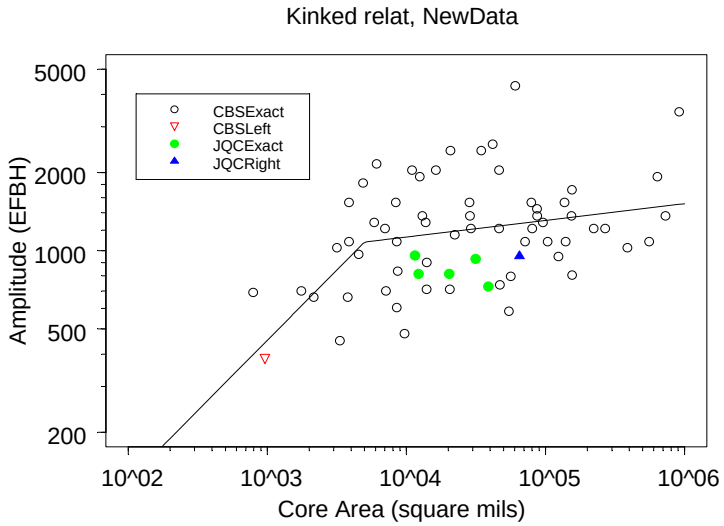


Figure 13. Experimental Amplitude Data Showing Fit to Kink Regression

To conduct a bivariate analysis, it is necessary to establish a small flaw correction for the SNR_{mz} data as well as for the amplitude data. The small flaw limit for SNR_{mz} was developed by conducting a regression analysis to relate SNR_{mz} to amplitude and then using this to translate the physics-based small flaw limit for amplitude into a small flaw limit for SNR_{mz}. Figure 14 shows the results of the regression relating amplitude to SNR_{mz}.

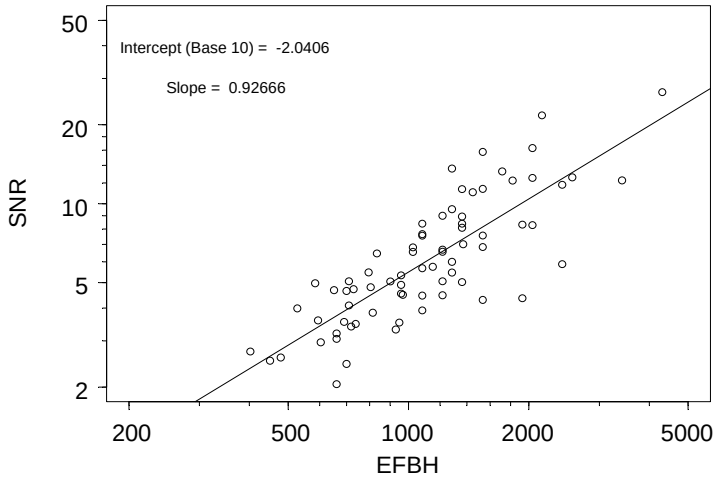


Figure 14. Regression Relating Amplitude and SNR_{mz}

Figure 15 shows the SNR_{mz} data and the kink-regression line as determined by the maximum likelihood analysis. Again, the line corresponding to the small flaw correction for SNR_{mz} lies above the kink-regression line and, hence, is not shown.

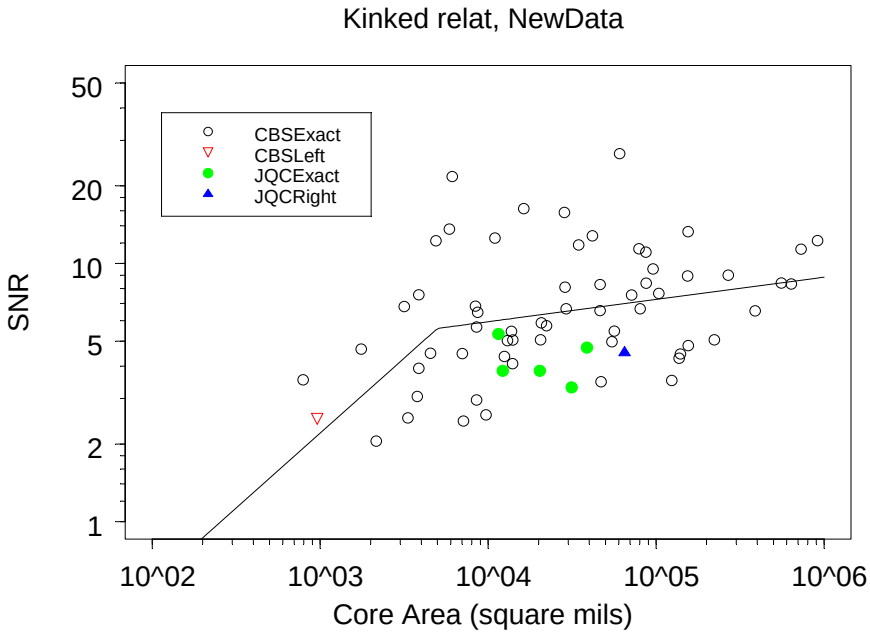


Figure 15. Experimental SNR_{mz} Data Showing Fit to Kink Regression

A bivariate model was developed based on both the amplitude and SNR_{mz} data. To fully capture the bivariate model would require a three-dimensional display. As a simplification, figure 16 shows contours of the bivariate response distribution for two different sizes of flaws for just the CBS data. Those portions of the distribution falling outside the lower, left-hand box correspond to rejects. Hence, PoD is the volume of the distribution (normalized to unity total volume) outside of this region. As expected from practical experience, the SNR_{mz} criterion rejects more flaws than the amplitude criterion.

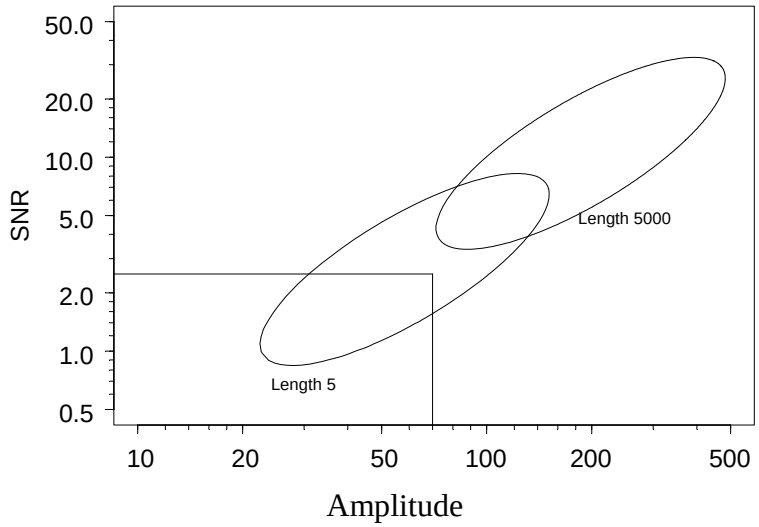


Figure 16. Contours of Bivariate Model for Two Different Flaw Sizes as Compared to Reject Criteria

The Multizone PoD curve was based on a bivariate analysis of the CBS plus the JETQC databases. The result is shown in figure 17. The estimated PoD curve has two segments joined at an area of about 5 x 10^3 square mils. These segments are controlled by the regression to the left and right of the kink point, respectively.

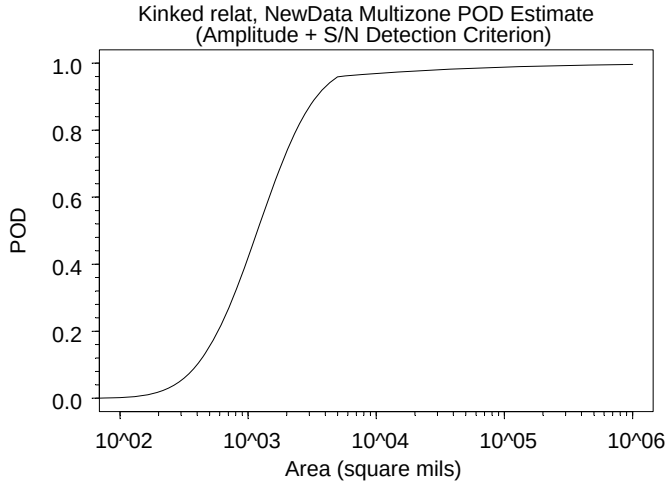
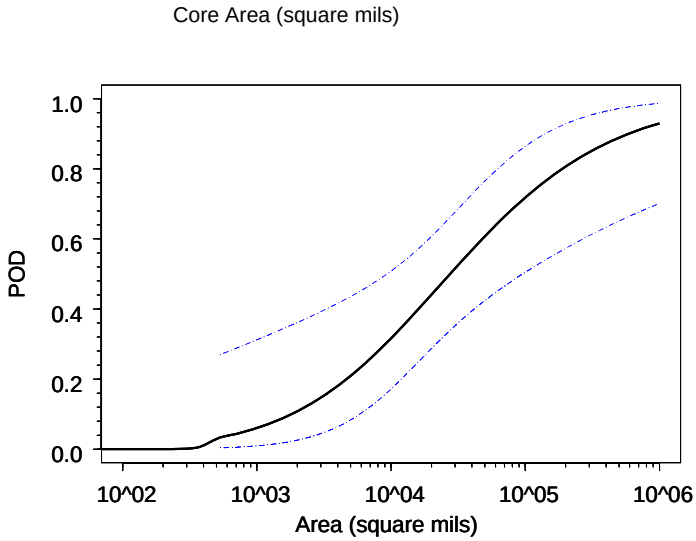


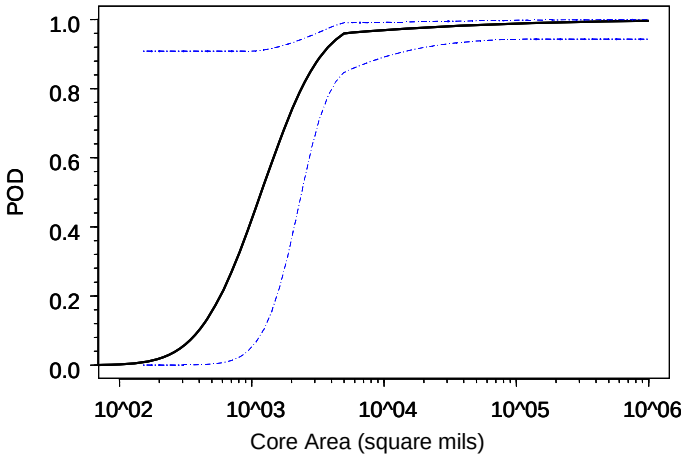
Figure 17. The PoD Curve for Multizone Inspection

3.7 RESULTS OF CONFIDENCE INTERVAL ANALYSIS.

The simultaneous 95%, two-sided confidence intervals were evaluated for both the conventional and Multizone data following the analysis techniques described in appendix H. Figure 18 shows the results.



(a)



(b)

Figure 18. 95%, Simultaneous, Two-Sided, Confidence Intervals (a) Conventional Inspection PoD Estimate (Amplitude-Only Detection Criterion) and (b) Multizone PoD Estimate (Amplitude Plus SNR_{mz} Detection Criterion)

3.8 DISCUSSION OF TECHNICAL ISSUES ENCOUNTERED IN THIS STUDY.

The determination of Default PoD curves for detection of hard alpha inclusions in titanium billet, based on data found during manufacturing, is a challenging task given the nature of the inspection and resultant test data. Among the challenges are (1) intrinsic difficulties in estimating PoD from finds information [12]; (2) the weak influence of the area of hard alpha inclusions on their ultrasonic response; (3) the existence of sources of variability such as inspection setup and flaw morphology, not fully understood, whose effects were not captured by

simple linear regression models; and (4) the need (in the Multizone case) to account for an accept or reject decision process based on more that one measured quantity. This report represents the effort of a team of experts and is believed to have produced a set of curves that are significantly superior to those that were previously available. In the process, as identified throughout this report (including the appendices), a number of technical issues were encountered. In this section, some of these issues are discussed and observations are made for consideration by groups engaged in future updates of the Default PoD curves. Obviously, the decision to take actions based on these observations will be tempered by other technical and cost or benefit factors present at the time of the update.

3.8.1 Importance of Accurate Destructive Flaw Sizing.

The accuracy of PoD estimates can be no better than the quality of the input data. Generally, the sizing data are obtained by destructive sectioning. However, obtaining accurate size data in the field requires significant effort since serial sectioning is a tedious and costly process. In the early databases, the size information obtained when hard alpha inclusions were found during manufacturing appeared, in some cases, to have significant errors. This is understandable since the cost and time pressures to make a rapid disposition of material in which a candidate hard alpha was found and the competing needs of identifying the cause of the defect, as opposed to quantifying its size, are natural forces against expending the needed time and effort. In some of the early data sets obtained in the manufacturing environment, the sectioning procedures used were designed to identify the root cause of the hard alpha inclusion rather than provide quantitative dimensional data for PoD analysis. Recent modifications of those procedures were recommended in an advisory circular [14] to make them more appropriate to the flaw sizing task. These represent significant improvements. Development and implementation of further proposed improvements to sizing procedures, which are underway, would provide better data for future Default PoD assessments.

3.8.2 Effects of Complex Flaw Morphology on PoD.

The complex morphology of hard alpha inclusions can have a significant effect on their ultrasonic response, which can be quite different from that of synthetic reflectors. For example, experimental plots of log $\hat{a}$ versus log a for naturally occurring hard alpha inclusions were typically found to have a slope on the order of 0.2 in this study, implying that the flaw response increased as the area of the flaw raised to this exponent. In other words, this implies that $\hat{a} = K\, a^n$, where K is a constant and n is on the order of 0.2. This is a significant change from the assumption that flaw response is directly proportional to flaw area, which was a core assumption in the development of the original Default PoD curves based on the *Re* analysis. The implications of the differences in the response of synthetic and naturally occurring flaws are discussed below.

The slope of the regression line (defining the exponent of the dependence of $\hat{a}$ on a), along with the variability of the response data about the regression line (as quantified by the standard deviation obtained in the regression analysis), controls the shape of the PoD curve. A slope of the regression line on the order of 0.2 is quite different from that of synthetic defects typically used in PoD studies, being relatively small in comparison. The most common synthetic defect used in PoD studies is the FBH and the experimental and theoretical studies reported in

appendix D (as well as an extensive set of references in the technical literature) show that the FBH has a regression slope (exponent of the dependence of $\hat{a}$ on a) of 1.0. This occurs because the surface of the FBH is flat so that the reflections from all points on the surface add in phase (interfere constructively so that their effects are additive). The FBH response is, therefore, proportional to the area. For the determination of the PoD of hard alpha inclusions, synthetic hard alpha inclusions (SHA) are often used as the synthetic defect rather than FBHs. The SHAs are most often right circular cylinders, although spheres have also been used. As an approximation, the response of the end of a flat-bottom, cylindrical, synthetic hard alpha inclusion has a regression slope (exponent of the dependence of $\hat{a}$ on a) of 1.0; the same as an FBH. This was also experimentally demonstrated in a number of studies including the Phase I PoD work of the ETC [13]. This is because the end is also a flat surface so all reflected signals add in phase and the response is directly proportional to area as long as the flaw is fully illuminated. Reflectors of spherical shape, including SHAs, have a slope (exponent of the dependence of $\hat{a}$ on a) on the order of 0.5, which is also well known, and has been further validated by the data in appendix D. The theory reported in appendix D, supported by theoretical and experimental references cited therein and by Phase I of the ETC [13], indicate that cylindrical synthetic hard alpha inclusions, viewed from the side, will have a regression exponent of 0.75 for SHAs of fixed aspect ratio. It would then be expected that PoD curves inferred from data obtained on such synthetics would be different (steeper) from those of naturally occurring defects.

The shape of a PoD curve is a consequence of the selectivity of the flaw response parameter that is compared to an accept or reject criteria. If the response parameter varies slowly with flaw area, the PoD curve will have a low slope, deviating significantly from the step function that is often considered to be ideal. This is the case for the current inspection of hard alpha inclusions. The measure of response (i.e., the response parameter) is the peak signal amplitude (and/or SNR_{mz} ratio) and the data in this report show that this response is proportional to flaw area raised to a power of the order of 0.2. There are considerable electronic data provided by some current inspections (e.g., Multizone) that are currently not used in the accept or reject decision. When billet inspections result in C-scans, the data can potentially provide the number, shape, intensity gradients, and other features of illuminated pixels. This is much more information about the area of the flaw than just peak amplitude and SNR_{mz}. The development of a rejection criterion, based on this information, is the subject of promising research, but this has not reached the stage of production-ready procedures. However, if these techniques were in a production-ready form and were used in the accept or reject decision, they might have a greater ability to differentiate between small and large flaws. Consequences of the use of such a procedure could be PoD curves that are significantly different in shape and location than the current curves.

3.8.3 Importance of Knowledge of Misses.

Knowledge of misses can have a significant impact on PoD predictions. The ideal way to deal with misses is through disciplined experiments with known opportunities for misses. However, that is difficult when trying to assess PoD from data on naturally occurring defects found in the manufacturing environment. If the inspection being assessed behaves in accordance with the statistical distribution assumed in the analysis (i.e., no atypical misses), this is not a major issue. In that case, treating the data as truncated in accordance with well-established statistical principles can largely eliminate the bias that would result from ignoring unknown misses, with

some sacrifice in precision. However, should the data not follow the assumed statistical model, treating the data as truncated without the knowledge of misses will not necessarily capture the effects of those misses and can lead to errors in the resulting prediction of PoD. Experimental knowledge of misses is required to obtain needed information, such as whether the miss data are consistent with the statistical model that was initially assumed. The conventional inspection data included several large Type II (atypical) misses, the presence of which was revealed by the Multizone inspection. An accommodation term was developed to take them into account in the statistical modeling of the data. No referee technique with significantly greater sensitivity than Multizone was available, and hence, the possibility of Type II (atypical) Multizone misses inconsistent with the assumed statistical model could not be assessed. Based on the data available, it is unknown whether such misses occurred. As techniques with improved sensitivity become available in the laboratory, it may become possible to assess the importance of this effect by using these advanced techniques to scan future production billets or, if possible, by rescanning the remaining CBS material. An alternative approach with the same goal would be to design and conduct a set of independent experiments aimed at verifying the results of the analysis that led to the current Default PoD curves.

3.8.4 Unique Challenges of Estimating PoD for Ultrasonic Detection of Internal Defects.

Generally accepted PoD analysis techniques, such as $\hat{a}$ versus a, were primarily designed for quantifying the PoD of eddy-current detection of surface-breaking defects such as fatigue cracks. However, as discussed in this report, there are a number of challenges that make the application of those techniques to the determination of the PoD of ultrasonic detection of internal defects problematic. Close cooperation of experts in statistics and inspections was needed to get a meaningful answer in the presence of such challenges. Erroneous results could have been produced if nondestructive evaluation personnel simply applied standard statistical packages without a full understanding of the statistical assumptions. By the same token, statisticians without a good understanding of the physics and implementation of the inspection could come up with equally misleading results. A balanced perspective, including expert knowledge of statistics, the physics of the inspection, and the practical implementation of the inspection, was required to produce meaningful results. In other complex problems of this type, efforts to assemble teams of this character could also bring value.

4. SUMMARY.

4.1 SUMMARY OF THE PoD ANALYSIS.

An update to the Default PoD curves for ultrasonic detection of hard alpha inclusions in billet was completed. Because of the effects of the complex morphology of these defects on ultrasonic response, and the sometimes limited information available from naturally occurring defects found in the manufacturing environment, a number of nontraditional steps had to be implemented in the analysis. As more information became available and the project progressed, the analysis technique evolved from the *Re* technique used in the development of the previous 1995 Default PoD curves to the $\hat{a}$ versus a technique. This technique was then modified with an accommodation term to account for Type II (atypical) misses in the conventional inspection and modified to include a kink-regression and bivariate analysis (to give credit for both the SNR_{mz} and amplitude-based reject decisions) in the Multizone inspection.

Figures 12 and 17 show the best estimates of the PoD for the conventional and Multizone inspections respectively, with confidence intervals being provided in figure 18.

4.2 COMPARISON OF NEW TO DEFAULT PoD CURVES TO ORIGINAL DEFAULT PoD CURVES.

The Default PoD curves shown in Appendix 5 of AC 33.14-1 [3] indicate the best estimates, at that time, of the PoD for ultrasonic inspection of field components. Three figures were presented in the AC for inspections calibrated to #1 FBH (1/64 inch diameter), #2 FBH (2/64 inch diameter), and #3 FBH (3/64 inch diameter).

In each of these three figures shown in AC 33.14-1 [3], two curves were shown, one for a reject criteria set equal to the response of the calibration FBH and one for a reject criteria set equal to one-half of the response of the calibration FBH. All the PoD curves were with respect to the anomaly area taken to be the area of the diffusion zone. Figure 4 summarizes those results by presenting two PoD curves, one for calibration to a #2 FBH and one for calibration to a #3 FBH, with the reject criteria set equal to 75% of the calibration level (60% FSH since calibration response is set at 80% FSH). These 75% reject curves fall between the 50% and 100% reject curves in AC 33.14-1 [3]. As in AC 33.14-1 [3], the anomaly area in figure 4 is taken equal to the area of the diffusion zone.

Since the publication of AC 33.14-1 [3], two changes have occurred. First, the anomaly area recommended for use in lifing of interior hard alpha inclusions changed from the area of the diffusion zone to the area of the core. Second, new data became available, which also required a modification of the analysis technique.

Figure 5 shows the PoD estimates that would have been made at the time of preparation of AC 33.14-1 had the core area been chosen as the anomaly area (explanatory variable). Figure 19 replots the estimated PoD curves in figures 4 and 5 together, showing the shift associated with the change in understanding of the anomaly area that should be used in fracture analysis. These PoD curves were all based on the DPD-3D, conventional inspection data only, and the curves corresponding to a #2 FBH calibration were subsequently taken to represent the Multizone capability, while those corresponding to a #3 FBH calibration were subsequently taken to represent conventional inspection capability.

In figure 19, as well as figures 20 and 21, the following convention has been adopted to facilitate an easy intercomparison of the curves:

- Inspection Technique
  - Multizone: Solid line
  - Conventional: Broken line
- Dataset
  - PoD from original DPD-3D data versus area of diffusion zone: red

- PoD from original DPD-3D data versus area of core: black
- PoD from updated data and analysis: blue

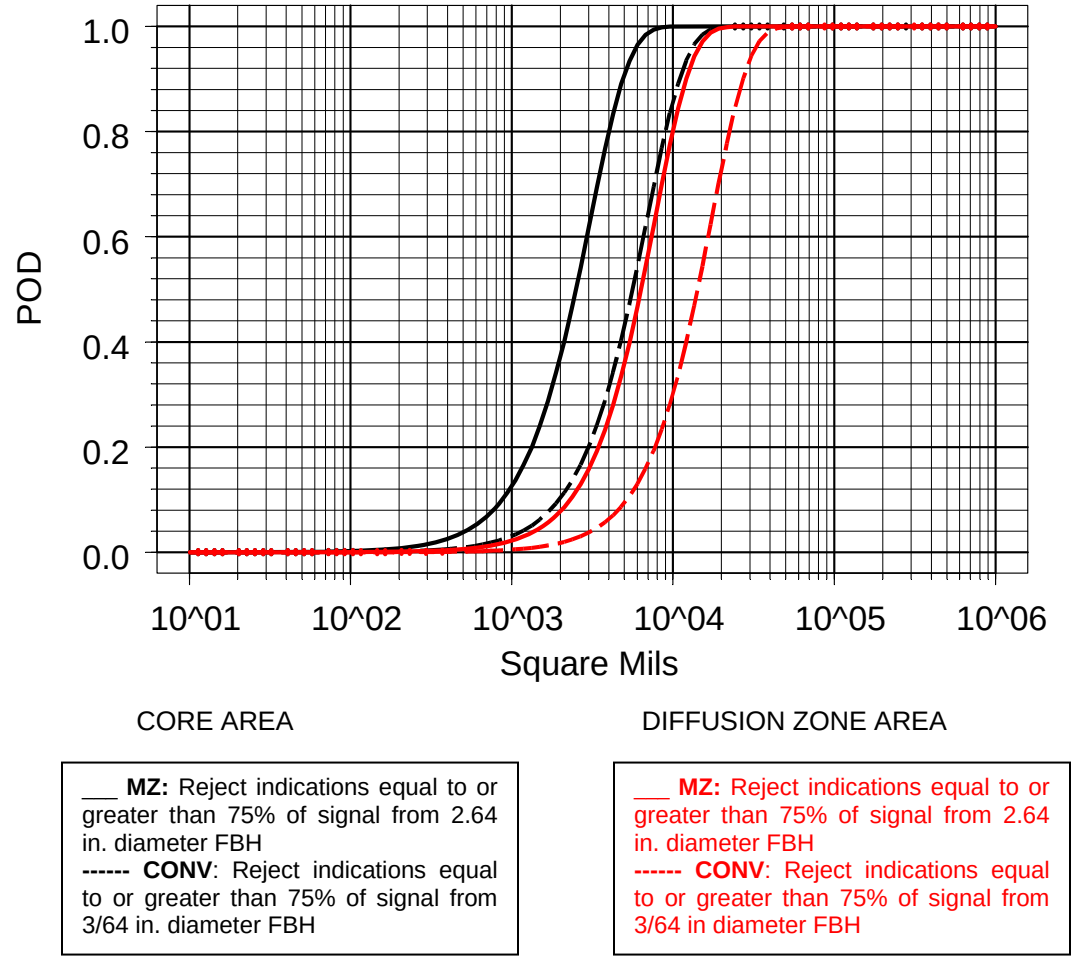


Figure 19. Changes That Would Have Occurred to the Original Default PoD Curves had the Anomaly Area Been Chosen to be the Area of the Core Rather Than the Diffusion Zone

Figure 20 compares the best estimates of the new PoD curves to the original PoD curves, both as a function of core area, i.e., after the adjustment for the change in explanatory variable, as shown in figure 19. The new Multizone curve has a different shape than the original one and is shifted to the left for most core areas. The core areas at which the new curve reaches 10%, 50%, and 90% PoD are less than those of the old curve by factors of about 2.1, 2.1, and 1.4, respectively. However, for flaws of large size (having estimated PoD values greater than 96%), the new curve flattens out and falls below the old curve. The new conventional inspection curve also has a different shape than the original one and is shifted significantly to the right for larger core areas. The core areas at which the new curve reaches 10%, 50%, and 90% PoD are greater than those of the old curve by factors of about 1, 5, and 50, respectively. These changes are due to both the additional data available and the refinement in the PoD analysis that is discussed in detail in the various appendices of this report.

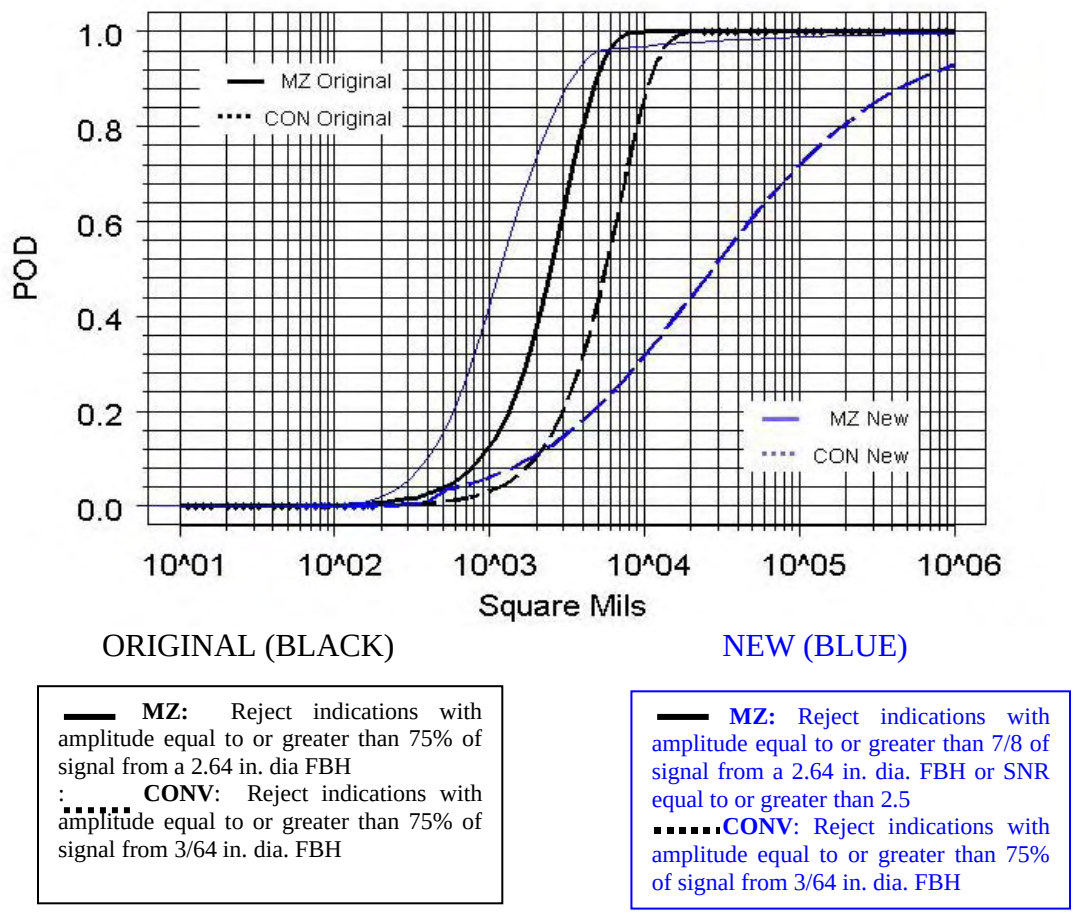


Figure 20. Comparison of Original and Current Best Estimates of Mean PoD Curve With Core Area Taken as the Explanatory Variable

The new curves are believed to be more realistic than those in AC 33.14-1 [3] for a number of reasons.

a. The core, rather than the diffusion zone, is used as the explanatory variable.

b. There is a much more extensive set of data, including three, rather than one, data set and about 5 times as many data points.

- There are Multizone as well as conventional inspection data.
- These data provide types of information that was absent in the previous analysis.
  - Actual behavior of Multizone
  - Systematic assessment of conventional inspection misses based on Multizone referee
  - Information about one Multizone miss

c. The new data provided information that guided a significant improvement over the analysis technique available in 1995 follows:

- Flaw response model bases slope of regression line on data rather than the assumption of unity slope in the *Re* model
- Accommodation term to account for Type II (atypical) misses in conventional inspection
- Kink-regression to approximate the curvature of log-log plots of Multizone response versus area
- Bivariate analysis for Multizone to accommodate the dual accept reject criteria based on both amplitude and SNR

d. The principles of physics provide an estimate of PoD capability in small flaw region where data is absent.

The conventional inspection PoD curve is based on one implementation and ignores informal SNR criteria that may sometimes be used in practice. Also ignored is the refracted angle inspection that is used by one OEM. Therefore, the curve may be a conservative estimate with respect to some implementation cases. However, the Multizone PoD curve is computed using both the amplitude and SNR detection criteria employed in implementation and is reflective of use conditions.

4.3 COMPARISON OF MULTIZONE AND CONVENTIONAL PoD CURVES.

Figure 21 directly compares the new PoD curves for Multizone and conventional inspections. The new estimates show that the PoD of the Multizone inspection is significantly higher than that of conventional inspection over a wide range of flaw sizes.

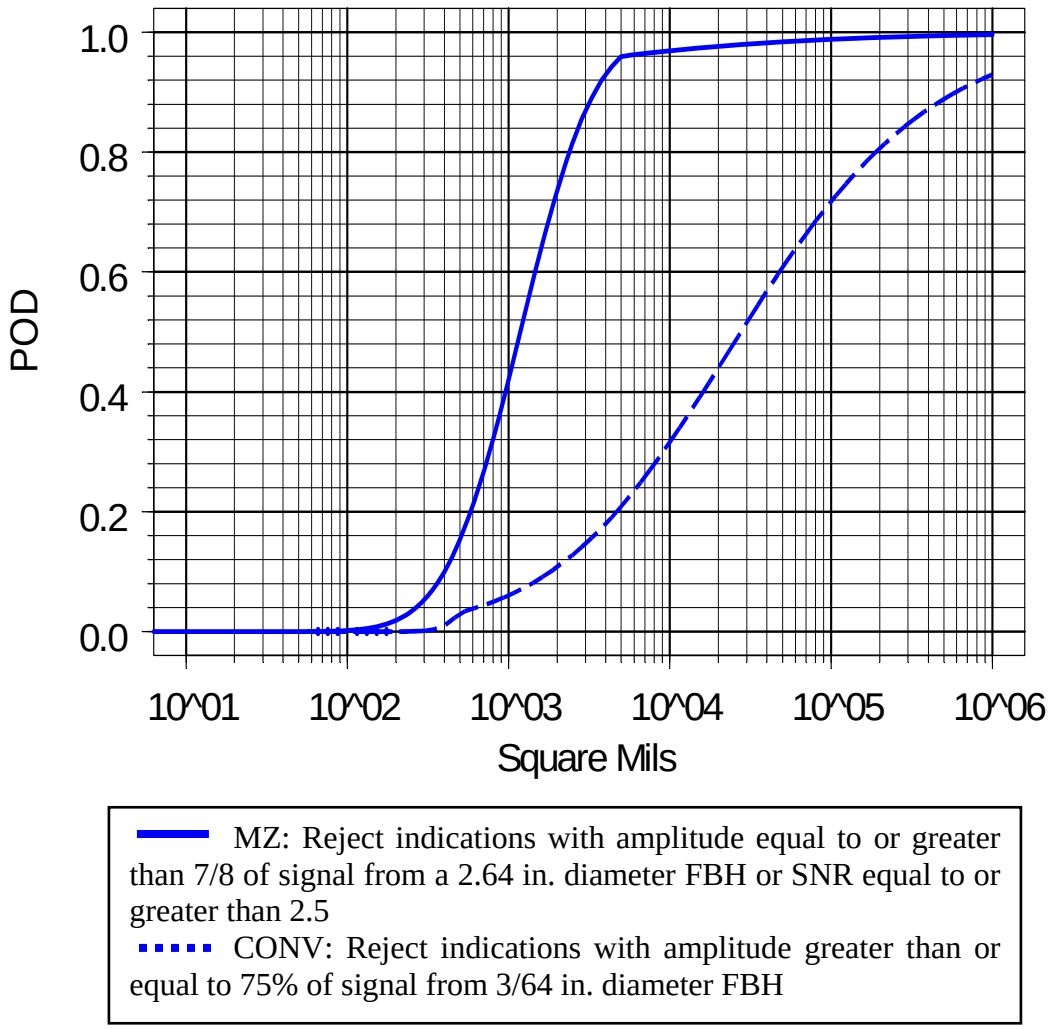


Figure 21. Comparison of Conventional and Multizone PoD Estimates

Table 1 shows how the difference is quantified by comparing the core areas that produce equal estimated PoD values for several PoD levels.

Table 1. Core Areas at Which Various Estimated PoD Values are Reached for Multizone and Conventional Inspections

PoD	Multizone Inspection (Square mils)	Conventional Inspection (Square mils)
0.1	400	1,824
0.25	669	6,643
0.5	1183	27,586
0.75	2077	126,449
0.9	3428	593,458

These results show that the estimated PoD of Multizone inspection is significantly greater than that of conventional inspection. The ratio of core areas corresponding to a given PoD for conventional and Multizone inspections increases with area as the PoD level increases. The core that can be detected with 10% PoD in a conventional inspection is estimated to be 4.6 times that for Multizone. At a PoD estimate of 50%, the ratio of flaw area of conventional inspection to Multizone has grown to 23, and at a PoD estimate of 90%, the ratio is 173.

5. REFERENCES.

1. Thompson, R.B., "Nondestructive Evaluation and Life Assessment," *ASM Handbook, Vol. 11, Failure Analysis and Prevention*, W.T. Becker and R.J. Shipley, eds., pp. 269-275, ASM International, Metals Park, OH, 2002.

2. Subteam to the Aerospace Industries Association Rotor Integrity Sub-Committee, "The Development of Anomaly Distributions for Aircraft Engine Titanium Disk Alloys," American Institute of Aeronautics and Astronautics, Inc.

3. "Damage Tolerance for High Energy Turbine Engine Rotors," FAA AC 33.14-1, ANE-110, January 8, 2001.

4. Sturges, D.J., Gilmore, R.S., and Hovey, P.W., "Estimating Probability of Detection for Subsurface Ultrasonic Inspection," *Review of Progress in QNDE*, Vol. 5A, D.O. Thompson and D.E. Chimenti, eds., Plenum Press, NY, 1986, pp. 929-937.

5. Burkel, R.H., Sturges, D.J., Tucker, W.T., and Gilmore, R.S., "Probability of Detection for Applied Ultrasonic Inspectors," *Review of Progress in QNDE*, Vol. 15B, D.O. Thompson and D.E. Chimenti, eds., Plenum Press, NY, 1996, pp. 1991-1998.

6. Berens, "NDE Reliability Data Analysis," *Metals Handbook*, Ninth Edition, Nondestructive Evaluation and Quality Control, ASM International, Metals Park, OH, 1989, pp. 689-701.

7. J. and H. Krautkrämer, "Ultrasonic Testing of Materials," Springer-Verlag, 4th Edition, 1990.

8. "Manufacturing Process of Premium Quality Titanium Alloy Rotating Engine Components," FAA AC 33.15-1, ANE-110, September 22, 1998.

9. Brasche, L., Chiou, C.-P., Thompson, B., Smith, K., Meeker, B., Margetan, F., Panetta, P., Chenail, R., Galli, F., Umbach, J., Raulerson, D., Degtyar, A., Bartos, J., Copley, D., McElligott, B., Howard, P., and Bashyam, M., "Contaminated Billet Study," FAA report DOT/FAA/AR-05/16, September 2005.

10. Nieters, E.J., Gilmore, R.S., Trzaskos, R.C., Young, J.D., Copley, D.C., Howard, P.J., Keller, M.E., and Leach, W.J., "A Multizone System for Billet Inspection," *Review of Progress in QNDE*, D.O., Thompson and D.E. Chimenti, eds., Plenum Press, NY, 1995, pp. 2137-2144.

11. Howard, P.J., Copley, D.C., and Gilmore, R.S., "The Application of a Dynamic Threshold to C-Scan Images With Variable Noise," *Review of Progress in QNDE*, Vol. 7B, D.O., Thompson and D.E. Chimenti, eds., Plenum Press, NY, 1998, pp. 2013-2019.

12. "The Use of Field Inspection Data in the Performance Measurement of Nondestructive Inspections," Final Report of the NATO Applied Vehicle Technology Working Group 051.

13. Burkel, R.H., Chiou, C.-P., Keyes, T., Meeker, W.Q., Rose, J.H., Sturges, D.J., Thompson, R.B., and Tucker, W.T., "A Tethodology for Assessment of the Capability of Inspection Systems for Detection of Subsurface Flaws in Aircraft Turbine Engine Components," FAA report DOT/FAA/AR-01/96, September, 2002.

14. Appendix 1 of Reference 8, "Minimum Recommended Procedure for 3D Inclusion Characterization."

APPENDIX A—GENERAL DATABASE CONSIDERATIONS

A.1 INTRODUCTION.

In this appendix, some general considerations regarding the databases needed for the Default Probability of Detection (PoD) analysis are discussed.

A.2 DESIRED ATTRIBUTES OF A DATABASE.

To determine the Default PoD curves for billets, a set of data from naturally occurring flaws is required. Specifically, the following information is ideally needed to fully define the inspection conditions.

- Ultrasonic response: The ultrasonic response should be determined as observed in a system that has been calibrated in a well-defined way. The flaw response should not be saturated.

- Flaw description: Flaw size should be determined in the context of a well-defined failure criterion. A three-dimensional description of the flaw size is desired. Flaw position, especially depth, is needed to allow evaluation of the degree to which PoD may depend on depth.

- Ultrasonic beam size in flaw plane: If the effective reflectivity technique is to be used, it appears to be necessary to account for the dimensions of the beam illuminating the flaw. This can be done on the basis of either calibration experiments or physics-based models. In the latter case, it is necessary to know the geometry of the part entry surface (nominally defined by the billet diameter), the parameters of the probe (e.g., diameter and focal length(s)), the setup conditions, and the flaw depth. This information is not explicitly used in some other approaches, e.g., $\hat{a}$ versus a. However, it is still valuable to provide insight that might be needed to interpret any unexpected results since the relative size of the beam and flaw clearly influence the ultrasonic response.

A.3 DEFINITION OF FRACTURE SIZE.

The appropriate explanatory variable for the PoD plots must be selected. When the initial Default PoD curves were developed, it was felt by the lifing/fracture community that the initial flaw size that should be used in the fracture analysis was the size of the diffusion zone. This choice was driven, in part, by the results of crack growth studies in air, which showed that the entire diffusion zone cracked in the first few cycles of fatigue. The result was the Default PoD curves shown in figure 4 of this report. However, subsequent work, including measurements of crack growth rates in vacuum (more appropriate to embedded flaws) as well as analysis of spin pit failure data, suggested that the size of the core was a more appropriate measure of the initial flaw size to be used in fracture. The size of the core is also believed to be more closely related to the area of the defect that reflects the ultrasonic energy since the diffusion zone is believed to be essentially transparent to the ultrasound. Hence, it was decided that the size of the core should be used as the explanatory variable in the PoD analysis. For reference, figure 5 of this report

shows the PoD curves that would have resulted from the initial analysis had the core size been chosen as the explanatory variable.

Given this decision, one must decide in which plane to determine size and what the measure of size should be. In the original Default PoD analysis, the area in the circumferential-axial plane was taken as the measure of size. Area was used since the signal was believed to be proportional to area, and the circumferential-axial plane was chosen because this is the plane that is perpendicular to the insonification direction in the billet inspection. It was recognized that, after the forging process, this plane may not transform into the plane perpendicular to the principal stress directions. Correction for this effect was to be undertaken by the lifing community as a part of their calculations.

A.4 ESTIMATES OF CIRCUMFERENTIAL DIMENSIONS.

As discussed in appendix C, axial dimensions of flaws can sometimes be estimated from C-scans, however, this is often not the case for circumferential dimensions. If defects for which only the axial length has been measured could be included in the analysis, a considerably large number of defects can be considered in the analysis, greatly increasing the stability and confidence in the statistical analysis. A transformation for converting length to area, which makes this possible, is described in appendix C.

APPENDIX B—THE DPD-3D AND JETQC DATABASES

B.1 INTRODUCTION.

In this appendix, the Default Probability of Detection-3Dimensional (DPD-3D) and Jet Engine Titanium Quality Committee (JETQC) databases are summarized. These have the common feature that they are based on information found during production finds of hard alpha inclusions, with best effort sectioning procedures.

B.2 DPD-3D DATABASE.

The DPD-3D database was discussed in the section 1.2 of this report and shown in figure 2.

B.3 JETQC DATABASE.

B.3.1 BACKGROUND.

The JETQC database is a collection of information obtained from naturally occurring hard alpha inclusions found by suppliers of engine materials during the manufacturing process. Its initial purpose was to help identify the processing causes of hard alpha inclusion formation during manufacturing. Towards this end, Federal Aviation Administration Advisory Circular 33.15-1, "Manufacturing Process of Premium Quality Titanium Alloy Rotating Engine Components," provides instructions for the gathering of data from manufacturing finds [B-1]. The original motivation for gathering this database was to track the processing root causes of hard alpha formation. However, data gathered in this fashion also has the potential of providing additional data for the purposes of probability of detection (PoD) analysis.

B.3.2 REQUEST FOR DATA FOR PoD ANALYSIS.

In December 2002, Bill Knowles, Chair of the Aerospace Industries Association Rotor Integrity Sub-Committee (RISC); Jon Tschopp, responsible for the RISC efforts to develop initial defect distributions (exceedence curves); and Bruce Thompson, coordinator of the Engine Titanium Consortium PoD efforts, made a request to Cliff Shamblen, co-chair of JETQC, for a re-examination of the JETQC database as required to support PoD analysis. In line with the initial objective of basing the analysis on the effective reflectivity technique, the letter requested information that would make it possible to determine the calibrated ultrasonic response, the size of the ultrasonic beam illuminating the flaw, and the size of the flaw.

Specifically, it was requested that the JETQC database be re-examined to obtain additional information with respect to the details of the ultrasonic inspection (aimed at helping to estimate the size of the beam in the flaw plane) and size information gained from systematic sectioning of the flaw (aimed at making as accurate an estimate of the full extent of the defect as possible). As a motivation for the latter request, it was noted that

> "It is our understanding that, although 3-dimensional 'sizes' are often given for defects, they are not always derived from a systematic sectioning of the defect.

> For example, in some cases the 'sizes' may be based on examination of micrographs in one or two transverse planes, combined with ultrasonic C-scan information (which may not provide complete information on the full extent of the diffusion zone. Although such information is quite valuable in considering the processing condition(s) that might have led to the formation of the defect, the "sizes" so obtained are not representative of the entire extent of the defect and hence cannot be used to relate the response of the defect to its fracture criticality. Hence, information that is adequate for identifying root-cause processing issues may not be adequate for PoD analysis. …"

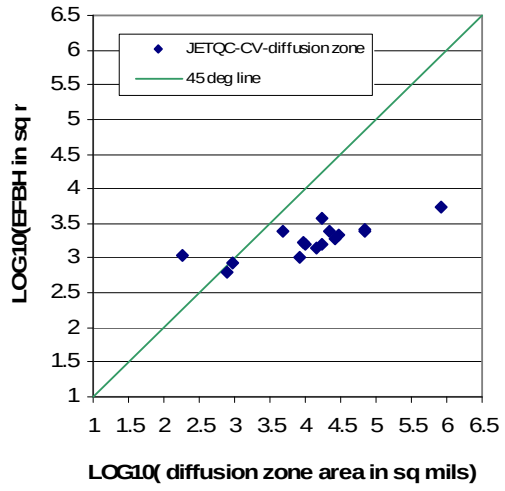
The letter concluded by noting that the original Default PoD analysis was based on information from 24 naturally occurring flaws detected prior to 1995 and requesting the JETQC files be reviewed for flaws detected after 1996 to see if a comparable number of defects could be identified for which the requested information could be provided.

B.3.3 GENERIC DESCRIPTION OF PoD DATA.

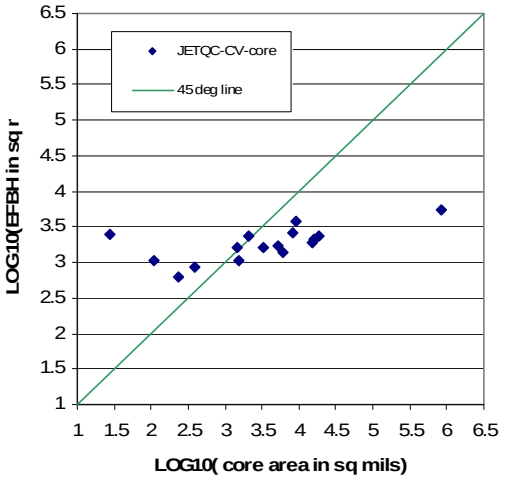
JETQC responded in August 2003 by providing a database consisting of two tables. The first table provided information on defect size. In a generic way, information was provided identifying the origin of the defect (identification number, year/quarter, alloy, billet or bar, ingot diameter, billet/bar diameter, and melting procedure). Specific information was then provided giving the signal level (% full-screen height (FSH) for a given calibration and the signal-to-noise ratio for Multizone inspection (SNR_{mz}) for some Multizone finds, estimated location from bottom and rim of the ingot, defect size (three dimensions), void size (three dimensions), and some explanatory notes. The second table provided additional sonic data. Included were detection threshold (with respect to noise and/or calibration signal), maximum noise level allowed, material noise level observed (generally one number), indication peak amplitude (% FSH), calibration information (reflector type, size, % FSH, and distance amplitude correction (yes or no)), transducer information (frequency, diameters, and focal lengths), inspection mode(s), focusing condition, and beam diameter (if known). The table contained information on 16 defects found by conventional inspections and 15 defects found by Multizone inspections, from 1996 to 2002. In subsequent clarification, it was suggested that the information provided under defect size and void size should be taken as approximate measures of the diffusion zone and core sizes, respectively. In each case, the longest dimension should be taken as A_x and the shorter two dimensions as X_1 and X_2. Rotation information appears to have been lost as in the DPD-3D database. Some of the indications were saturated.

B.3.4 GRAPHS OF RESPONSE VERSUS AREA.

In figures B-1 and B-2, graphs of response versus metallographic area are shown for the subset of the JETQC database discussed above. The data is plotted as received areas calculated using equation 2 of this report. Included are results for both conventional (figure B-1) and Multizone (figure B-2) inspections. For each inspection type, separate graphs are shown for the cases in which the diffusion zone and core dimensions are used as the basis for size determination. In total, there were 17 conventional inspection entries and 15 Multizone entries. However, not all had complete information.



(a)



(b)

Figure B-1. Log Equivalent Flat-Bottom Hole Area Versus Log Metallographic Area for JETQC Database, Conventional Inspection (a) Versus Log Diffusion Zone Area and (b) Versus Log Core Area

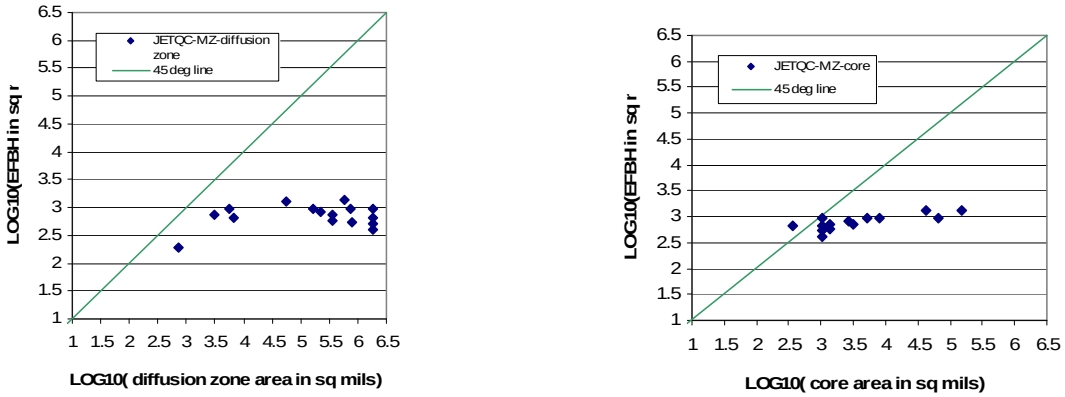


Figure B-2. Log Equivalent Flat-Bottom Hole Area Versus Log Metallographic Area for JETQC Database, Multizone Inspection (a) Versus Log Diffusion Zone Area and (b) Versus Log Core Area

B.4 ADJUSTING THE DPD-3D AND JETQC DATABASES.

Examination of the DPD-3D and JETQC conventional databases led to three concerns regarding the fidelity of the data. First, some of the entries had values of equivalent flat-bottom hole area (EFBH)>>core area. This did not seem reasonable, since it implied that a hard alpha inclusion would reflect a signal much greater than that reflected by a flat bottom hole whose area equaled that of the core. Such a situation indicated a serious sizing error, and the corresponding points were dropped from the database. Exactly how much greater the EFBH would have to be with respect to the core area to warrant dropping the data point is a judgment call. Consultation of several parties led to the decision to drop the most extreme two data from both the DPD-3D and JETQC databases. No data points with this problem were found in the JETQC Multizone database.

A second concern was the fact that several of the defects in both databases had values of one or more of the core dimensions equal to the corresponding values for one or more of the diffusion zone dimensions. In general, this was not thought to be realistic. For example, as discussed in the Contaminated Billet Study defects in appendix C, the diffusion zone would typically be two or three times the core in dimensions. Thus, there would appear to be some inconsistencies in these entries. After considerable discussion, including consultation with metallurgists at each of the original equipment manufacturer partners, it was decided to keep those flaws in the database. The reasons were (1) there are circumstances, not common but possible, in which these dimensions might be equal. For example, the diffusion zone can be separate from the core during mechanical working. (2) it was believed that the determination of core dimensions at supplier sites is more accurate than the determination of diffusion zone dimensions. Hence, if either set of dimensions is in error, it is more likely to be those of the diffusion zone, which were not used in the updated PoD analysis.

A third concern was found in analyzing the Multizone entries in the JETQC database as a function of core area. To understand this, more detailed information about the data is required.

There were 15 Multizone data points reported by JETQC. Of these, four lacked SNR_{mz} values and, hence, could not be considered in the bivariate analysis to be reported subsequently and two had sizes that were not reported in a way that could be used. The remaining nine entries were reports of what appeared to be multiple responses from two defects. Three were multiple responses for what was nominally identified as one defect and six were multiple responses for what was nominally represented as a second defect. Several of these multiple responses had the same reported void size as indicated by the presence of multiple data points of the same size, as shown in figure B-2(b). This raised the question of whether there were two independent data points, nine independent data points, or something in between. To resolve this, the original inspection records were recovered and reviewed. Without going into tedious detail, the conclusion was that there were five independent data points in these two sets of multiple responses and the true core size was sometimes considerably greater (as much as a factor of 100 in a couple of cases) than the reported void size. This was because the reported void size was defined to be the size of the largest void in the core rather than the overall size of the core that is needed for PoD analysis. The root causes of these difficulties were the facts that when the data was gathered by JETQC, PoD studies were not intended and the data was sometimes entered (in the desire to rapidly obtain processing related root-cause information) before the ultrasonic interpretation was complete. As a part of this examination of original inspection records, satisfactory data was obtained for one additional defect that was not in the database assembled by JETQC. Hence, there were six Multizone finds available for analysis based on the JETQC database. Improved procedures for sizing manufacturing defects would be beneficial to conducting these analyses in the future.

Based on those principles, figure B-3 shows plots of EFBH versus core size for the adjusted DPD-3D and JETQC databases. In these plots, circle symbols indicate unsaturated (exact) data points and upward pointing triangles indicate saturated (right censored) data. Also shown for reference is a line representing the condition EFBH = core area, which represents the expected response of FBHs.

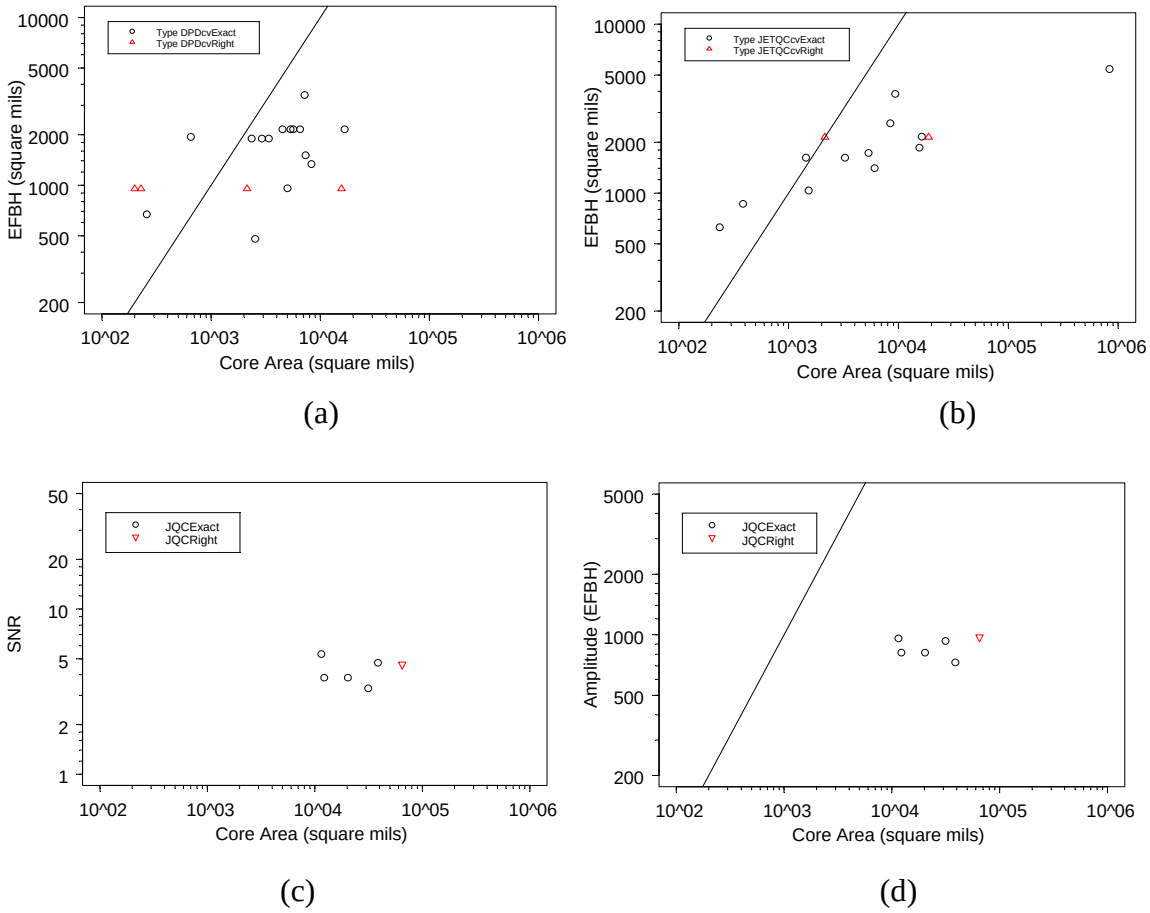


Figure B-3. Adjusted DPD-3D and JETQC Databases—EFBH Versus Metallographic Area of Core (a) DPD-3D, Conventional Inspection; (b) JETQC, Conventional Inspection; (c) JETQC, Multizone Inspection Amplitude; and (d) JETQC, Multizone Inspection SNR_{mz}

B.5 REFERENCES.

B-1 "Manufacturing Process of Premium Quality Titanium Alloy Rotating Engine Components," Federal Aviation Administration Advisory Circular 33.15-1, ANE-110, September 22, 1998.

APPENDIX C—THE CONTAMINATED BILLET STUDY DATABASE

C.1 ORIGIN OF THE CONTAMINATED BILLET.

In November 1994, while performing routine inspections for a nonaerospace customer, a titanium supplier found multiple indications with a conventional ultrasonic inspection. The large number of flaw indications (32) found during the production conventional inspection suggested that there had been a processing problem. This was subsequently attributed to a power outage and pressure spike related to the shutdown of vacuum pumps before the furnace was backfilled with argon. Upon sectioning of two of the indications and subsequent metallographic and chemical analysis, the presence of hard alpha, representative of those found in aerospace materials, was confirmed, and the existence of the heat was made known to the Engine Titanium Consortium (ETC). Because of the opportunity this heat offered, ETC proposed to the Federal Aviation Administration (FAA) William J. Hughes Technical Center to purchase the heat for inspection, processing, and life management studies as part of the ETC research and for use by the FAA Turbine Rotor Materials Design (TRMD) program performed by Southwest Research Institute. The proposal was accepted and became known as the Contaminated Billet Study (CBS) and provided a unique opportunity for study by both the inspection and lifing communities.

An FAA report describes the results of the CBS study in great detail [C-1]. It is not within the scope of the present document to reproduce the details of that document, to which the reader is referred. However, the major points are summarized in this appendix.

The exercise of determination of the updated Default Probability of Detection (PoD) curves required a deeper examination of a number of aspects of the data. Most important for this study was a careful determination of the sizes of the hard alpha inclusions, assessing a possible “miss” identified in the original study, and correcting some reporting errors. These results are also reported.

C.2 PRODUCTION INSPECTIONS.

The original production conventional inspection, called for by the nonaerospace customer, consisted of a normal longitudinal wave and angle shear wave (45°) inspection, as shown in figure 7(b) of this report. As was the industrial practice, precautions were not taken to prevent saturation of the ultrasonic signal during the conventional inspection since saturation would have no effect on reject decisions. In addition, due to the large number of flaws found, the supplier, of its own volition, performed additional Multizone scans on about half of the billets.

One of the objectives of the CBS study was to provide information for a more accurate PoD analysis. Since full amplitude information improves the accuracy of PoD determination, and since the results of Multizone inspections as well as conventional inspections were desired for all billets, the entire heat was reinspected in September 1995. The results of all of these inspections, both the 1994 and 1995 runs, are summarized in appendix C of the CBS report [C-1] and discussed further elsewhere in that report.

The conventional technique used during the reinspection was based on a normal longitudinal inspection plus a 45^{o} refracted longitudinal inspection, as shown in figure 7a of this report (in contrast to the refracted shear used in the original inspection). However, typical conventional inspection procedures were modified somewhat from those normally used to capture digital data on the flaw response in a C-scan format. This is in contrast to the strip chart read-out and stop-on-defect procedures that were generally implemented by suppliers during conventional inspections. In 1995, the Multizone and conventional inspections were run simultaneously. Since this was a small diameter billet (6 inches) requiring only four Multizone probes, there was room for two additional transducers in the Multizone transducer manifold. These were used to implement the conventional normal longitudinal and refracted longitudinal inspections. Several specific differences between the conventional inspection performed in this way (which will be called the modified conventional inspection implementation) and the normal (production) conventional inspection implementation are described below.

a. Since the inspections were done simultaneously, the conventional inspection had an index determined by the Multizone (between 0.04 and 0.05 inch per revolution), whereas a conventional inspection would normally have a much coarser index.

b. Because C-scans were available, it was possible to run automatic software to analyze the data. Thus, for both the conventional and the Multizone inspections, the following steps were performed.

- The automated image analysis software developed for Multizone was run and the amplitudes, signal-to-noise ratio for Multizone inspection (SNR_{mz}) and location of all indications were called out.
  - Note that this software uses the image-based definition of signal-to-noise ratio, given in equation 8 of this report.
  - This procedure would identify all defects with either amplitudes greater than 50% full-screen height (FSH) or SNR_{mz} greater than 2.0 (the dynamic threshold program was set to call out at 50% amplitude or 2.5 SNR_{mz}, with an option to reduce the SNR_{mz} threshold to 2.0).
- The images were visually scanned, and any additional indications not called out by the threshold program were identified.
  - Thus, any indication that had both an amplitude less than 50% and SNR_{mzx} less than 2.0 was found by visual evaluation of the image rather than by the dynamic threshold program.
  - A manual SNR_{mz} calculation was done for any indication found in the above step (in this calculation, the user defines the region of interest to be sampled for the SNR_{mz} calculation).

c. This is in contrast with the 1994 normal production implementation of the conventional inspection, in which there is no data recording, relying on a stop-on-defect procedure, as described previously, with a reporting threshold level set by the inspector based on his or her estimate of the noise level from the display. As noted previously, the signals from reported defects (causing a stop in the inspection) are generally peaked to determine if they exceeded the reject threshold. This was not done in the 1995 modified conventional inspection, as described above.

d. The reject criteria for the modified production implementation of conventional inspections for normal beams, with a flat-bottom hole (FBH) set to 80% FSH, was the following:

- Reject all signals whose amplitudes exceed 70% (7/8 of the signal from a #3 FBH).

e. The same requirements exist for modified production implementation of angle scans, however, the reference was to a 0.02″ diameter end-drilled hole rather than a #3 FBH.

It is reasonable to ask about the degree to which the modified procedure described above is representative of the results that would be obtained in a normal production implementation of conventional inspection. For example, it might be argued that the use of image data would enable one to detect indications in low-noise bands, which might not be visible using a global alarm level that is set to avoid false alarms in high-noise bands. To the extent that this is true, the indications reported on the billet maps in appendix C of the CBS Report [C-1] would not necessarily be detected by a production conventional inspection. However, this does not appear to be the case.

Table C-1 summarizes the results of the production inspections of the CBS, as reported in appendix C of the CBS report [C-1]. The depths were estimated by the operator from A-scans observed during the 1995 Multizone inspections and are believed to be accurate to no better than ±0.1 inch. The regions in the 94 MZ column that are marked with cross-hatching correspond to those billets that were not inspected by Multizone at that time. Details of the setup and the accept/reject criteria are shown in table C-2, based on extended discussion with parties involved in the inspections of the CBS material. It should be mentioned that only one engine original equipment manufacturer (OEM) requires angle scans in billet inspections. For this reason, all PoD estimates in this report were based only on the conventional, normal longitudinal data. However, the angle scan results are reported here for completeness and possible future reference.

The process of determining the Default PoD curves required a deeper analysis of the data. The results of that analysis are found throughout this appendix.

Table C-1. Original Results of Production Inspections of the Contaminated Billet

			Flaw Signal Amplitude %					
Billet	Flaw	95 MZ Depth (inches)	94 CVL	94 CVA	95 CVL	95 CVA	94 MZ	95 MZ
B1AW1	A	0.2					>100	190
B1AW1	B	0.6	45	70		42	>100	201
B1AW1	C	1.6			42	28	95	83
B1AW1	D	1.7				55	62	73
B1AW1	E	0.6	70	>100			89	127
B1AW2	A	0.3						134
B1AW2	B	0.2						225
B1AW2	C	0.8	80	80	65	>100		213
B1AW2	D	0.7			50			134
B1AW2	E	1.3	90	90	>100	>100		160
B1AW2	F	1.1	80	80	94	86		268
B1AW2	X	0.5						94
B1AW3	A	0.7	>100	90	50	60	>100	142
B1AW3	B	0.9					57	69
B1AW3	C	1.5	>100	80	62	>100	>100	253
B1AW3	D	1.1	70				91	101
B1AW3	E	1.6						47
B1AW3	Y	1.3						99
B1AW3	Z	1.3						77
B1BW1	A	2.2				68		63
B1BW1	B	1.7	70	30				113
B1BW1	C	0.6	70	>100	70	>100		142
B1BW1	D	1.0						56
B1BW1	E	0.9	30	>100				69
B1BW1	F	1.1	>100	70	179	>100		450
B1BW1	G	1.5	>100	90	100	80		213
B1BW1	H	0.4						127
B1BW2B	A	1.2			118			179
B1BW2B	B	0.3			108			160
B1BW2B	C	0.3						113
B1BW2B	D	1.6	90	90	46	116		160
B1BW2B	E	1.5	90	90	108	101		142
B1BW3	A	1.3					97	113
B1BW3	B	1.1					>100	160
B1BW3	C	0.5					>100	127
B1BW3	D	1.5					70	74
B2W1	A	0.5	80	80	60	90		142
B2W1	B	1.3	100	90	85	60		151
B2W1	C	0.9	80	80		60		127
B2W1	D	0.5						87
B2W2	A	NA (see note 1 below)						
B2W2	B	0.6						107
B2W2	C	0.2						74
B2W2	D	0.6	65	80				113
B2W2	X	0.55						61
B2W3	A	0.2					>100	107
B2W3	B	2.3				72		72
B2W3	C	2.0	40	60			49	50
B3W1B	A	2.0	70	40	35	55		84
B3W1B	B	2.2						73
B3W1B	C	2.3	>100		40	35		99
B3W1B	D	1.9	>100		50	70		134
B3W2	A	0.7		>100				357
B3W2	B	0.3	100			>100		113
B3W2	C	2.1						201
B3W2	D	1.5	80	80	60	54	>100	127
B3W2	E	1.3	80	80	74		>100	213
B3W2	F	0.2					>100	160
B3W2	G	1.3	80	80			>100	142
B3W2	H	1.6				84		253
B3W3	A	0.6	50	90		80	>100	113
B3W3	B	0.8	50	40	50	60	>100	120
B3W3	C	2.9	>100		>100	85	>100	127
B3W3	D	0.9	>100	60			>100	142

NOTE

(1) B2W2-A was reported as a "visible end burst" by GE, so can't be considered a credible flaw indication.

(2) Different amplitudes for B1AW1-B (89%), -C(56%) and -D(>100%) in 95 conv angle scan were also reported than the orignals shown in the table.

MZ = Multizone inspection
CVL = Conventional longitudinal inspection
CVA = Conventional angle inspection

Table C-2. Summary of CBS Data Sets

Quantity	95 MZ	94 CVL	95 CVL	94 CVA	95 CVA
Angle	0°	0°	0°	45°-T offset 1/6 diameter	45°-L offset 1/6 radius
Probe	5 MHz, bicylindrically focused set of 4	5 MHz, ½″ x 1″, cylindrically foc. @ 8″	5 MHz, ½″ x 1″, cylindrically foc. @ 8″	5 MHz, ¾″, spherically foc. @ 7″	5 MHz, ¾″, spherically foc. @ 7″
Liftoff	3.0″	4.5″	3.5″	3 ¾″	
Index	0.04″-0.05″	3/32″	0.04″-0.05″	3/32″	0.04″-0.05″
Record	MZ C-scan	Data sheets, no strip chart	MZ C-scan format	Data sheets, no strip chart	MZ C-scan format
Saturation present	Yes initial/Rerun no	Yes	No	Yes	No
Depth of dead zone	0.2″	0.375″	0.375″	0.4″	0.4″
Coordinates adjusted in peaking	Bridge position, billet rotation, not probe angle	Bridge position, billet rotation, probe angle	Bridge position, billet rotation, not probe angle	Bridge position, billet rotation, probe angle	Bridge position, billet rotation, not probe angle
Calibration level	#2 FBH @ 80%, weakest in each zone w/o DAC	#3 FBH @ 80%, with DAC	#3 FBH @ 80%, with DAC	0.020 SDH, weakest, @ 80% w/o DAC	0.020 SDH, weakest @ 80% w/o DAC
Calibration hole positions	2 at boundaries of each of 4 zones	0.375″, 0.75″, 1.5″, 3.5″	0.375″, 0.75″, 1.5″, 3.5″	0.4″, 0.72″, 1.0″	0.4″, 0.72″ 1.0″
Calibration billet diameter	6.0″			6.875″	6.875″
Amplitude-based evaluate level	50%	40%	50% (MZ software used)	40%	50% (MZ software)
Amplitude-based reject level	70%	60%	70%	60%	70%
Noise-based evaluate level	SNR_{mz}>2.0	A bit above peak noise (10% greater typical)	SNR_{mz}>2.0 (MZ software used)	A bit above peak noise (10% greater typical)	SNR_{mz}>2.0 (MZ software used)
Noise-based reject level	SNR_{mz}>2.5	Signals significantly above noise to be called by Level III	Not applicable, signal-to-noise not reported in conventional sense	Signals significantly above noise to be called by Level III	Not applicable, signal-to-noise not reported in conventional sense

NOTES:

1. All of the entries indicated as % are in terms of FSH.

2. The conventional specifications are in accordance with procedures, as practiced at RMI at that time. Current procedures may be different. For example, current specifications of one OEM would call for an amplitude-based reject threshold of 80% FSH and a noise-based reject threshold of 3 dB above average noise, as long as the signal is greater than 30% FSH.

3. At each site, instructions were given to evaluate and report all rejectable indications and to re-evaluate any saturated indications at reduced gain level, reporting the amplitude and gain adjustment. Reports of noise level, attenuation differences from the calibration standard, and applied attenuation corrections were requested. Summary statistics were prepared. For Multizone, the primary quantities that were statistically analyzed were amplitude and signal-to-noise (based on the definition in equation 8). For conventional, the primary quantities that were statistically analyzed were amplitude and signal minus noise for both the normal and angle inspections. Here, signal minus noise was defined as the difference in FSH height of the two quantities.

MZ = Multizone inspection
CVL = Conventional longitudinal inspection
CVA = Conventional angle inspection
SDH = Side-drilled hole
DAC = Distance amplitude correction

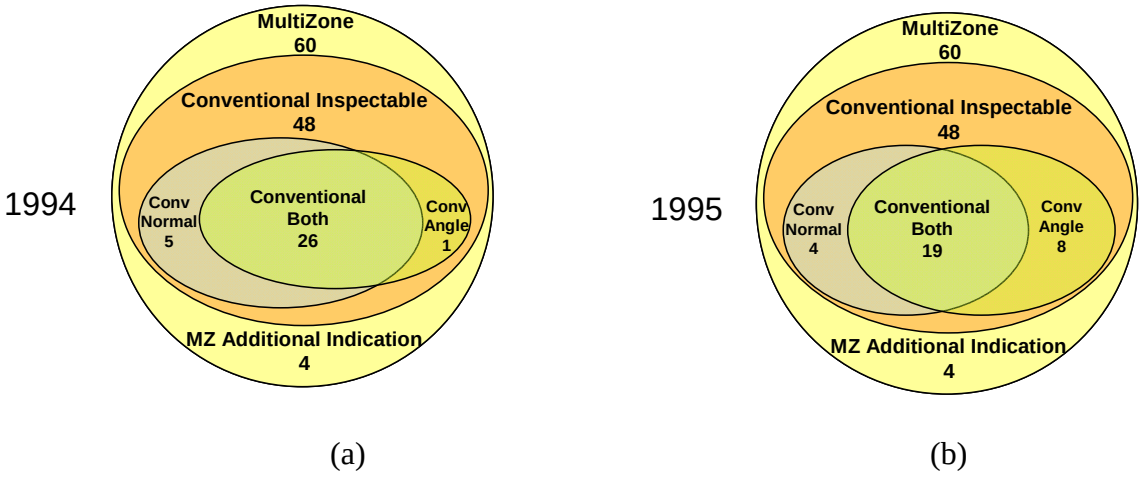
It was agreed by the ETC team, upon recommendation of the inspection manager, that the data obtained for conventional inspections in 1994 is more representative of production conventional inspections than the data obtained in 1995 using the modified procedure described above. At the latter time, there were unfamiliar procedures and many observers that may have influenced the accuracy of the setup.

Figures C-1 and C-2 show Venn diagrams summarizing various aspects of the data. Figure C-1 compares the results in different years (1994 and 1995). In each of these Venn diagrams, the numbers for the conventional inspection represent the number of entries in table C-1 (indications), not necessarily the number of rejectable defects. Figures C-1(a) and (b) are a comparison of the number of indications (entries in table C-1) found in the 1994 conventional inspection to the 1995 Multizone inspection (considered to be the referee), with the 1994 conventional-inspection data being broken down as to whether the indication was observed in normal inspection, angle inspection, or both. In comparing these numbers to the Multizone results, the following points should be noted. The 1995 Multizone found 60 indications based on the automatic signal analysis software. In addition, four additional defects were found in examination of the C-scan records, which would have been called by the automatic software had they not been very close to another defect producing larger amplitude or SNR response. Hence, there were 64 defects that were detectable by Multizone. However, as will be discussed subsequently in detail in section C.5 of this appendix, nine of these were at a depth of less than 0.4 inch below the surface. These were not accessible to the conventional inspection, whose gate starts at this depth. In addition, three more were in a region identified as having an endburst during the conventional inspection, which was not examined in detail. Finally, the four defects that were too close to a larger one to be called by Multizone were also too close to be called by conventional. Accordingly, there were 48 defects that represented inspection opportunities for the conventional inspection. Figure C-1(b) shows the analogous results from the 1999 conventional inspection adapted from the CBS report [C-1]. It is interesting to note that although the number of defects producing conventional indications was about the same in 1994 (32) and 1995 (31), the relative number of indications in the normal and angle data were considerably different. This appears to reduce the possibility that the 1995 data were possibly rendered superior by the C-scan implementation with Multizone index sizes.

Figure C-2 compares the effectiveness of the normal and angle inspections in the two years. This shows that considerably more indications were found by the normal inspection in 1994 (31) than in 1995 (23), whereas the same number of indications were found in each year by the angle inspection (27). However, any interpretation of the latter observation should be tempered by the recognition that the 1994 inspection was angle shear, whereas the 1995 inspection was angle longitudinal, i.e., a different refracted inspection mode was used.

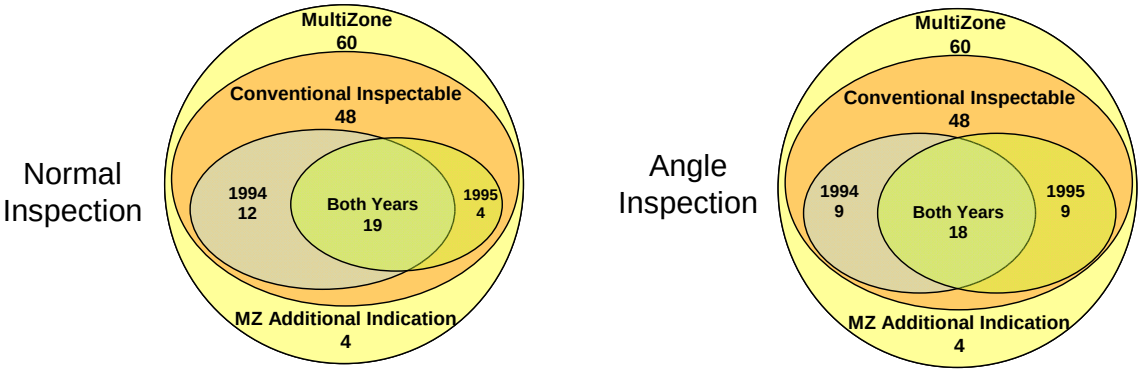
Further experiments to test the reproducibility of the Multizone and conventional inspections of billets were conducted, as described in detail in the CBS report [C-1]. An initial study was conducted by Rolls-Royce with a follow-up by the ETC. These experiments quantified the variability, due to setup, of both inspection techniques by comparing the responses of selected flaws in multiple inspections. Implicit in this was information on misses, as provided in tables 15 and 16 of the CBS report [C-1]. The information from the reproducibility study was not analyzed in the current estimation of the PoD curves since all flaws in the data set were not

subjected to multiple inspections. Inclusion of the data in the reproducibility study would have caused some defects to be weighed more heavily than others. However, the results provide insight into the question of misses, which play a key role in the current PoD analysis. In the multiple conventional inspections, there were 87 finds out of 189 opportunities. For the Multizone inspection, there were 125 rejections out of 135 opportunities. This supports the conclusion that the Multizone technique is clearly superior to the conventional technique, but it also can experience misses.



(a) (b)

Figure C-1. Venn Diagram Showing the Number of Indications in the 1994 and 1995 Conventional Inspections as Compared to the 1995 Multizone Inspection (a) 1994 (Normal Longitudinal and Angle Shear) and (b) 1995 (Normal Longitudinal and Angle Longitudinal)



(a) (b)

Figure C-2. Venn Diagram Showing the Number of Indications in the Normal and Angle Conventional Inspections in (a) Normal Inspection (1994 and 1995) and (b) Angle Inspections (1994 Angle Shear and 1995 Angle Longitudinal)

C.3 METALLOGRAPHIC RECONSTRUCTIONS AND SIZING OF THE CBS_{10}.

C.3.1 BACKGROUND.

Ten of the CBS defects were selected for detailed metallographic sectioning and reconstruction (CBS_{10}) to provide a greater insight into the factors that control their detectability. The criterion used to guide the selection of those to be sectioned was to obtain as diverse a set of responses and shapes as possible, based on the C-scan results. These were not intended to be distributed in the same way as the ultrasonic finds or misses, but rather to cover as wide a range in the parameter space of length, circumferential extent, axial extent, and ultrasonic response as possible. In addition, a second set of 10 defects with similar properties was selected for use in deformation studies of the TRMD program. The "pair-wise" selection was motivated with the hope that the reconstructed members of the pair would provide important background information for the deformation studies in the TRMD program.

The procedures that were used to reconstruct the three-dimensional shapes of the void structure from the results of serial metallographic analysis (on planes separated by 0.005 inch) are described in detail in the section "Generation of Geometrical Data for use in PoD Calculations" of the CBS report [C-1]. Figure C-3 summarizes the sectioning plan.

Each of the defects in the CBS_{10} was sequentially sectioned in 5 mil increments in a radial-axial plane, as shown in figure C-3. Composite micrographs were made in each plane, and chemical analysis via microprobe measurements were done in selected regions. From the full set of micrographs for each defect (aligned by fiducial marks), a solid model of the defect was formed.

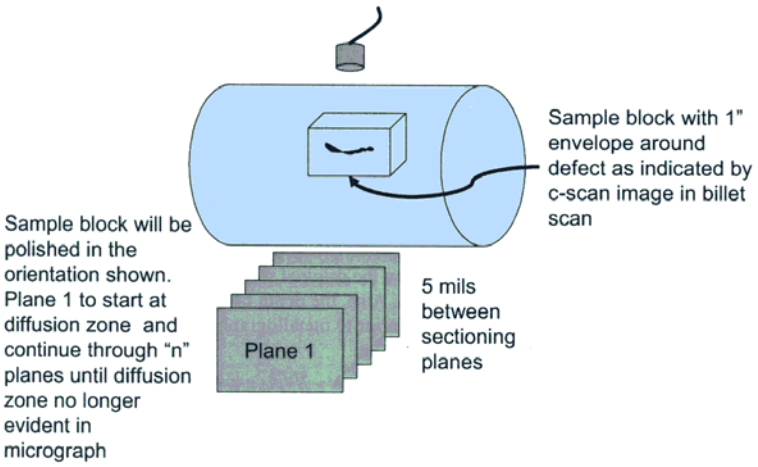


Figure C-3. Procedure for Serially Sectioning the CBS_{10}

Appendix F of the CBS report [C-1] provides detailed documentation of the results. Pictorial summaries are presented in figures C-4 and C-5, which show one perspective of the reconstruction of the void structure for each of the defects. The great variety in sizes can be seen by examining figure C-5 in which each of the images is plotted on a common scale. It is clear that the defects are highly irregular in shape. It is reasonable to assume that the dependence of the ultrasonic signals on flaw area will be quite different than for simple geometrical shapes such as FBHs.

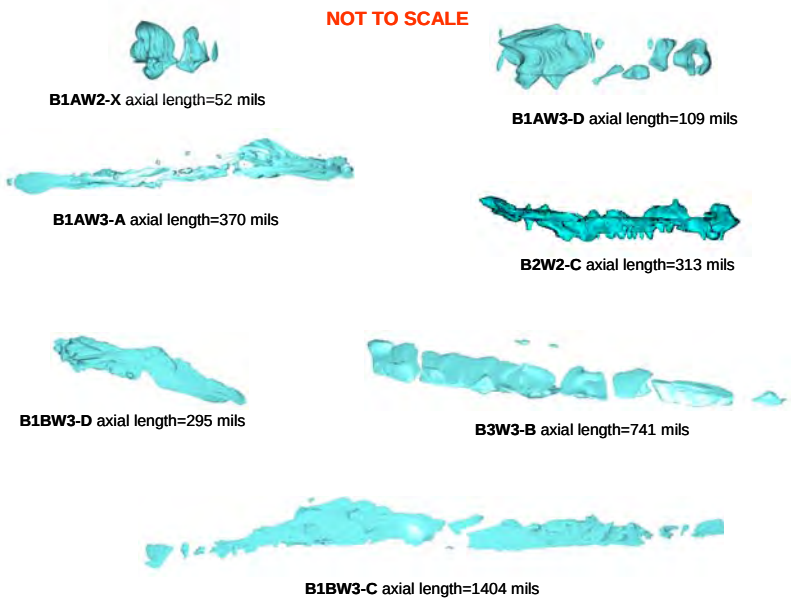


Figure C-4. One Perspective of the Reconstructed Shapes of the Void Structure in Seven of the CBS_{10}, Arbitrary Scale

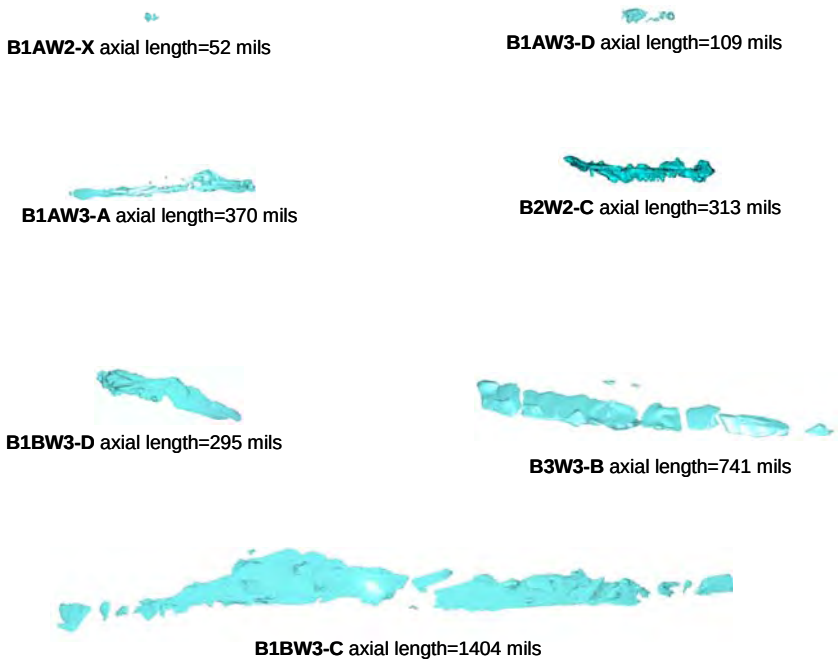


Figure C-5. One Perspective of the Reconstructed Shapes of the Void Structure in Seven of the CBS_{10}, Common Scale

C.3.2 DETERMINATION OF SIZES OF THE CBS_{10}.

Given the data contained in this set of micrographs, a round robin was conducted by the metallurgists from each of the OEMs to determine the dimensions of the diffusion zone and core for each of the ten defects.

Figure C-6 shows a typical micrograph for one plane of one defect. As shown, the defect consists of three regions.

- Black—A region that represents either a void or a hard alpha inclusion that has fallen out during the polishing
- White—A region that represents a single-phase region in the alpha (hexagonal) phase
- Gray—A region that indicates a mixed, alpha-beta phase region of the microstructure. This region changes appearance throughout the micrograph. Speaking qualitatively, the base material has a uniform appearance that changes as one enters the diffusion zone.

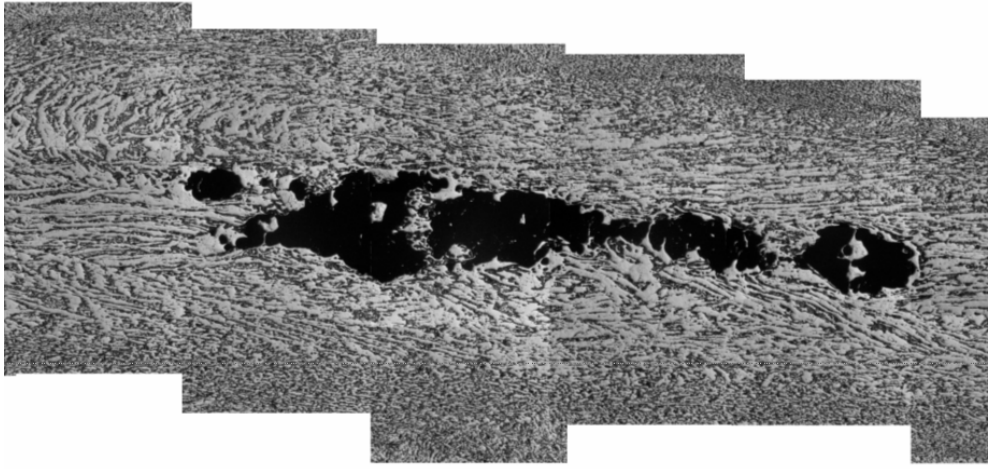


Figure C-6. The CBS Defect 7:B2W2-C-POLISH 9

The definition of these regions must be quantified to determine the size of the defects. An important element is the determination of changes in chemical composition throughout the defect region. Figure C-7 shows another micrograph from one of the CBS defects. The same three general regions are seen. Superimposed on the micrograph are three lines, along which chemical composition data was obtained by microprobe analysis.

Figure C-8 shows the variation of nitrogen along these scan lines. Note that these show regions of very low and uniform nitrogen (base material), a transition region (diffusion zone), and a region of high nitrogen (core).

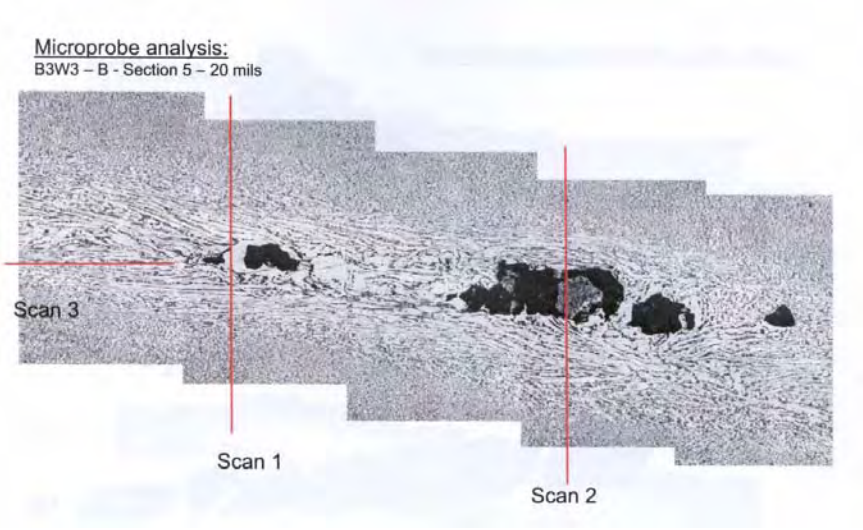


Figure C-7. The CBS Defect 1-B3W3-B-POLISH 5

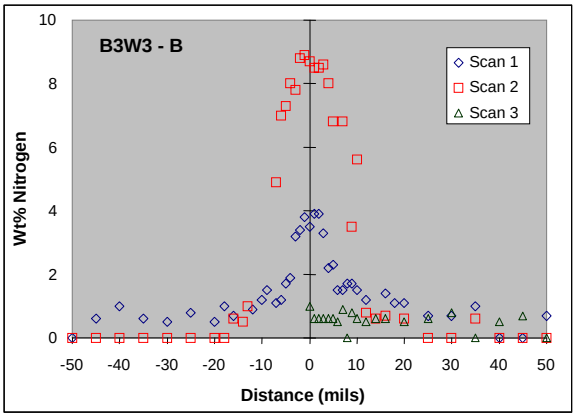


Figure C-8. Variation of Nitrogen Along the Three Scan Lines on CBS Defect 1:B3W3-B

Given information such as this, and significant additional experience based on internal studies, the metallurgists representing the three OEMs developed a definition of the diffusion zone, core, and void of a hard alpha defect.

- Diffusion zone:
  - Outer limit defined by transition from globular alpha/beta microstructure to 'stringy' alpha/beta microstructure
  - Nitrogen, if measured, ranges from minimum detectability limit to ~2-3 wt.%
  - Typically can be measured from an optical micrograph at ~100x
- Core
  - Outer limit of void plus essentially single phase alpha (or TiN)
  - Nitrogen, if measured, ranges from ~3-4 wt.% up to ~14 wt.%
  - May need to be measured from a backscattered scanning electron microscopy image at ~100-500x
- Void
  - Outer limit of void or crack
  - Typically can be measured from an optical micrograph at ~25-100x

Figure C-9 schematically shows these regions and their definitions.

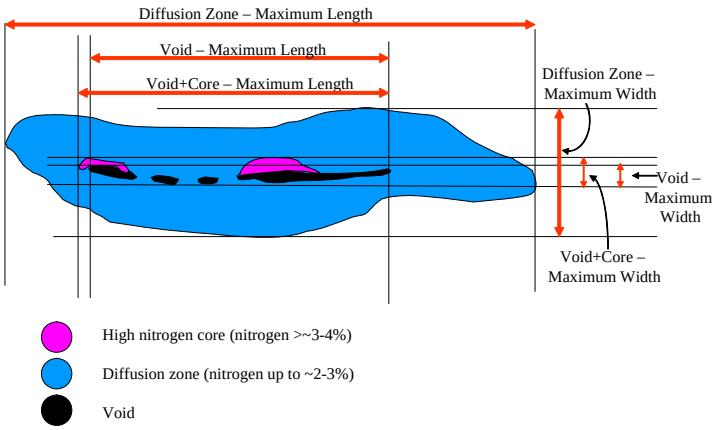


Figure C-9. Proposed Definition of Boundaries of Void, Core, and Diffusion Zone

Based on these definitions, a three-step process was undertaken to determine the three-dimensional sizes of the diffusion zone and core for the ten CBS defects characterized in detail.

C.3.2.1 STEP 1.

For step 1, the process was a determination of the planes of maximum cross-sections of the diffusion zone and core for each of the defects. It was not required that these be the same plane. The ten defects were assigned to the OEMs in groups of three, three, and four. Thus, the planes of maximum area for each defect were determined by a metallurgist from one of the OEMS.

C.3.2.2 STEP 2.

For step 2, the plane(s) defined in step 1 for the ten defects, each of the OEMs made an estimate of the sizes of the diffusion zone and core in the axial and radial directions. At the end of this step, for each of the ten defects, there were three independent OEM estimates of the axial and radial dimensions of both the core and diffusion zones in the planes of maximum cross-sectional area. The OEM's estimates were compared for each defect to determine how close, or far apart, the OEM's size estimates were to each other.

One unanticipated effect had to be addressed during step 2. Although hard alpha inclusions are generally assumed to be elongated along the axis of the billet, the defects in the CBS were rotated, one billet by about 25°, with respect to the billet axis. It can be speculated that this is a consequence of the fact that the CBS material was not intended for rotor use and had a different processing schedule. Two of the OEMs determined the size in two coordinate systems. In the original, unrotated system, the maximum extent was measured along axes parallel to the axial and radial directions of the billet. In the rotated system, the coordinates were rotated to correspond with the major and minor axes of the defect. Since the defects tended to be highly elongated along the axial direction, these rotations had little impact on the axial dimensions. However, the radial size was sometimes significantly greater in the unrotated coordinate system than in the rotated coordinate system.

In the tables and graphs that follow, all lengths are measured in mils and all areas in square mils.

Tables C-3, C-4, C-5, and C-6 show the results of the size estimates for the plane(s) of maximum size, as made by each OEM. Averages are also reported. In addition to the results obtained by the OEMs, estimates are also shown for Iowa State University (ISU). These had been made prior to the OEM estimates for the purpose of providing initial data for preliminary PoD analysis and did not benefit from knowledge of the definitions subsequently develop by the OEMs. Hence, exact agreement would not be expected and the ISU numbers were not included in the averages reported, which were based only on the OEM results. For the case of the diffusion zone, the estimates were made based on visual perception of where the disturbed region of the microstructure associated with the diffusion zone ended and the base material began. The length in the original coordinate system was determined, since rotation does not have much effect on this number. However, the radial extent was determined in the rotated system. For the case of the core, the estimates were made by software developed at ISU. This software determined the extent of the "bounding box," drawn in the original coordinates, that contains all voids. Hence, these sizes would not be expected to be identical to the OEM estimates for two reasons: (1) the OEMs included in their definition regions of single-phase alpha, and (2) the bounding box approach is based on three-dimensional information from multiple planes. If the defect is skewed, the bounding box will be larger than the dimensions in any plane.

Table C-3. Axial and Radial Sizes of the Diffusion Zone as Measured by Each Organization in the Unrotated Coordinate System

Unrotated DZ Defect ID	Axial					Radial				
	ISU	OEM 1	OEM 2	OEM 3	OEM AVG	ISU	OEM 1	OEM 2	OEM 3	OEM AVG
1-B3W3-B	999	1230	1025	NA	1127.5	NA	234	190.0	NA	212.0
2-B1BW3-C	2405	2507	2482	NA	2494.5	NA	429	407.0	NA	418.0
3-B1BW3-D	588	619	586	NA	602.5	NA	312	281.0	NA	296.5
4-B1AW3-E	2001	2621	1989	NA	2305	NA	362	288.0	NA	325.0
5-B1AW3-D	999	1065	932	NA	998.5	NA	70	61.0	NA	65.5
6-B1AW3-A	928	1273	1198	NA	1235.5	NA	149	113.0	NA	131.0
7-B2W2-C	544	728	701	NA	714.5	NA	94	75.0	NA	84.5
8-B1AW2-D	1046	1193	1076	NA	1134.5	NA	134	121.0	NA	127.5
9-B1AW2-B	622	809	808	NA	808.5	NA	73	64.0	NA	68.5
10-B1AW2-X	958	970	953	NA	961.5	NA	65	60.0	NA	62.5

DZ = Diffusion zone
NA = Not applicable

Table C-4. Axial and Radial Sizes of the Diffusion Zone as Measured by Each Organization in the Rotated Coordinate System

Rotated DZ Defect ID	Axial					Radial				
	ISU	OEM 1	OEM 2	OEM 3	OEM AVG	ISU	OEM 1	OEM 2	OEM 3	OEM AVG
1-B3W3-B	NA	1244	1035	1082	1120.3	95	119	102	105	108.7
2-B1BW3-C	NA	2503	2482	2522	2502.3	289	409	407	397	404.3
3-B1BW3-D	NA	685	642	652	659.7	191	140	148	142	143.3
4-B1AW3-E	NA	2633	1989	2174	2265.3	168	221	288	212	240.3
5-B1AW3-D	NA	1065	932	1152	1049.7	45	70	61	69	66.7
6-B1AW3-A	NA	1274	1198	1080	1184.0	112	145	113	128	128.7
7-B2W2-C	NA	729	698	708	711.7	72	86	82	83	83.7
8-B1AW2-D	NA	1200	1074	1038	1104.0	48	61	39	121	73.7
9-B1AW2-B	NA	809	808	1555	1057.3	62	75	64	62	67.0
10-B1AW2-X	NA	971	959	933	954.3	24	36	29	41	35.3

DZ = Diffusion zone
NA = Not applicable

Table C-5. Axial and Radial Sizes of the Core as Measured by Each Organization in the Unrotated Coordinate System

Unrotated Core Defect ID	Axial					Radial				
	ISU	OEM 1	OEM 2	OEM 3	OEM AVG	ISU	OEM 1	OEM 2	OEM 3	OEM AVG
1-B3W3-B	741	519	355	NA	437	119	93	67	NA	80
2-B1BW3-C	1404	1737	1732	NA	1734.5	171	221	219	NA	220
3-B1BW3-D	295	285	254	NA	269.5	106	106	101	NA	103.5
4-B1AW3-E	141	183	178	NA	180.5	15	33	29	NA	31
5-B1AW3-D	109	169	162	NA	165.5	15	24	23	NA	23.5
6-B1AW3-A	370	395	395	NA	395	60	50	37	NA	43.5
7-B2W2-C	313	424	318	NA	371	58	27	22	NA	24.5
8-B1AW2-D	275	277	275	NA	276	51	54	54	NA	54
9-B1AW2-B	135	141	137	NA	139	31	30	30	NA	30
10-B1AW2-X	52	52	50	NA	51	18	21	19	NA	20

NA = Not applicable

Table C-6. Axial and Radial Sizes of the Core as Measured by Each Organization in the Rotated Coordinate System

Rotated Core Defect ID	Axial					Radial				
	ISU	OEM 1	OEM 2	OEM 3	OEM AVG	ISU	OEM 1	OEM 2	OEM 3	OEM AVG
1-B3W3-B	NA	526	355	516	465.7	NA	55	57	55	55.7
2-B1BW3-C	NA	1738	1732	1735	1735.0	NA	204	219	222	215.0
3-B1BW3-D	NA	297	267	328	297.3	NA	58	63	67	62.7
4-B1AW3-E	NA	217	177	224	206.0	NA	18	12	26	18.7
5-B1AW3-D	NA	170	162	162	164.7	NA	17	23	22	20.7
6-B1AW3-A	NA	395	395	403	397.7	NA	41	37	52	43.3
7-B2W2-C	NA	212	152	160	174.7	NA	25	22	28	25.0
8-B1AW2-D	NA	280	278	283	280.3	NA	31	30	61	40.7
9-B1AW2-B	NA	143	138	142	141.0	NA	22	21	38	27.0
10-B1AW2-X	NA	52	50	51	51.0	NA	21	19	22	20.7

NA = Not applicable

Figures C-10, C-11, C-12, and C-13 show these data. The agreement between the estimates made by the different organizations can be qualitatively described as very good. The real metric is the degree to which any disagreements would affect PoD analysis. As shown from the data plots in section C-8, the uncertainties in sizes determined from metallographic sections are much less than the variability associated with the response of flaws of the same nominal size. Hence, size determination does not appear to be a major issue when a trained metallurgist has the opportunity to examine a set of micrographs from a serially sectioned defect with a precise definition of the different zones in hand. This conclusion does not necessarily extend to the more approximate estimates of size made under time constraints in the production environment where fewer serial sections are available.

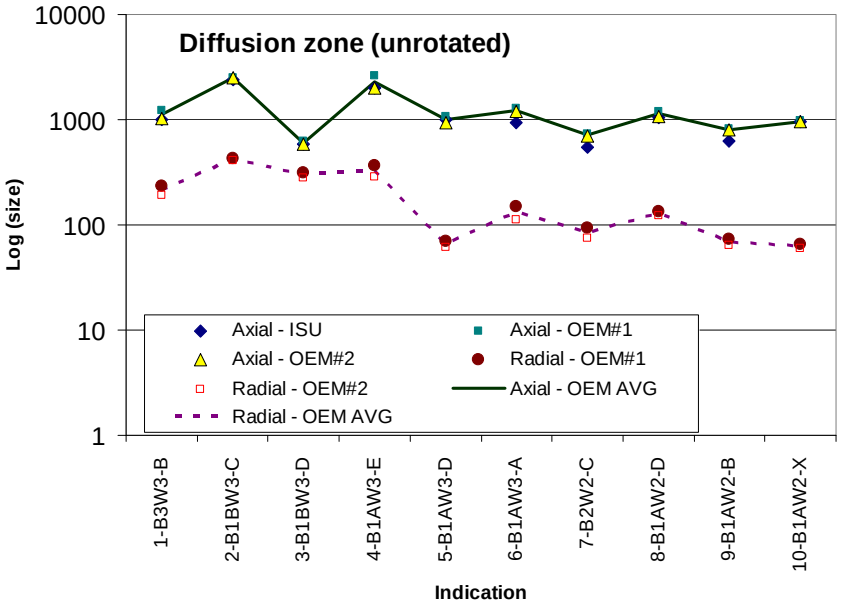


Figure C-10. Axial and Radial Sizes of Diffusion Zone as Measured by Each Organization in the Unrotated Coordinate System

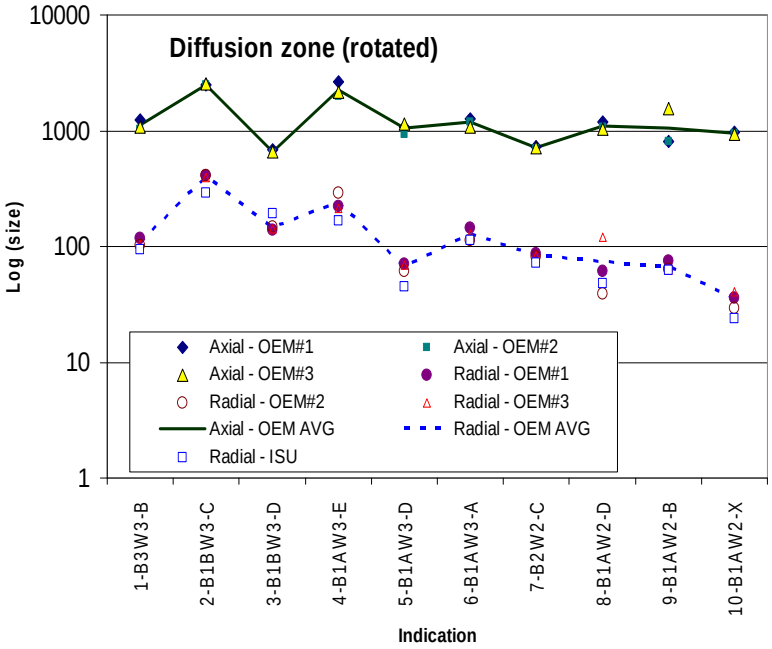


Figure C-11. Axial and Radial Sizes of Diffusion Zone as Measured by Each Organization in the Rotated Coordinate System

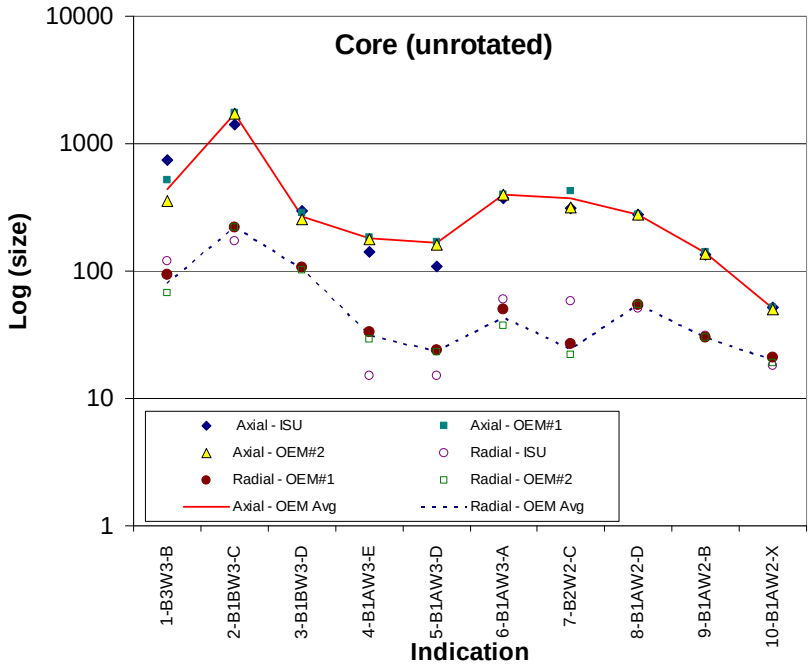


Figure C-12. Axial and Radial Sizes of Core as Measured by Each Organization in the Unrotated Coordinate System

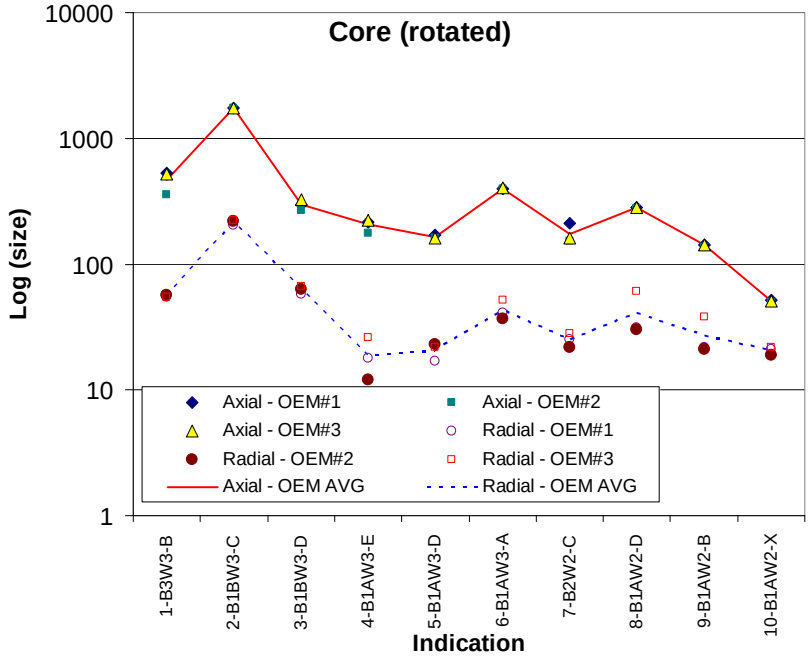


Figure C-13. Axial and Radial Sizes of Core as Measured by Each Organization in the Rotated Coordinate System

C.3.2.3 STEP 3.

For step 3, the final step was the determination of circumferential extent. Given that uniformity had been established in step 2, each OEM was asked to determine the circumferential extent of the core for the defects that they had examined in step 1. In some cases, the circumferential extent of the diffusion zone was also estimated. The estimates were made in slightly different ways. Some organizations based their estimates on the distance between the first and last polish planes in which the feature appeared. Others included information regarding the first and last plane in which the feature was not seen. Others based their estimate on plotting size versus plane position and interpolating.

Table C-7 presents the results of all OEM estimates of the circumferential size of the core and diffusion zone. Also shown is an estimate for the core made at ISU based on the bounding box approach discussed above. Averages shown are based on only the OEM estimates. Figure C-14 shows these results.

Table C-7. Circumferential Sizes of Diffusion Zone and Core

Circumferential Size Defect ID	Diffusion Zone					Core				
	ISU	OEM 1	OEM 2	OEM 3	OEM AVG	ISU	OEM 1	OEM 2	OEM 3	OEM AVG
1-B3W3-B	95	NA	95	130	112.5	80	NA	65	65	65.0
2-B1BW3-C	76	NA	71	80	75.5	59	NA	62	56	59.0
3-B1BW3-D	72	NA	NA	144	144	63	NA	NA	97	97.0
4-B1AW3-E	58	NA	NA	70	70	16	16-26	NA	26	23.5
5-B1AW3-D	113	NA	NA	145	145	32	NA	NA	35	35.0
6-B1AW3-A	84	NA	NA	92	92	32	45-57	NA	32	42.0
7-B2W2-C	64	NA	59	71	65	27	NA	43	53	48.0
8-B1AW2-D	220	NA	220	230	225	57	NA	57	70	63.5
9-B1AW2-B	97	NA	NA	130	130	64	NA	NA	56	56.0
10-B1AW2-X	138	NA	NA	160	160	34	30-42	NA	42	35.0

NA = Not Applicable

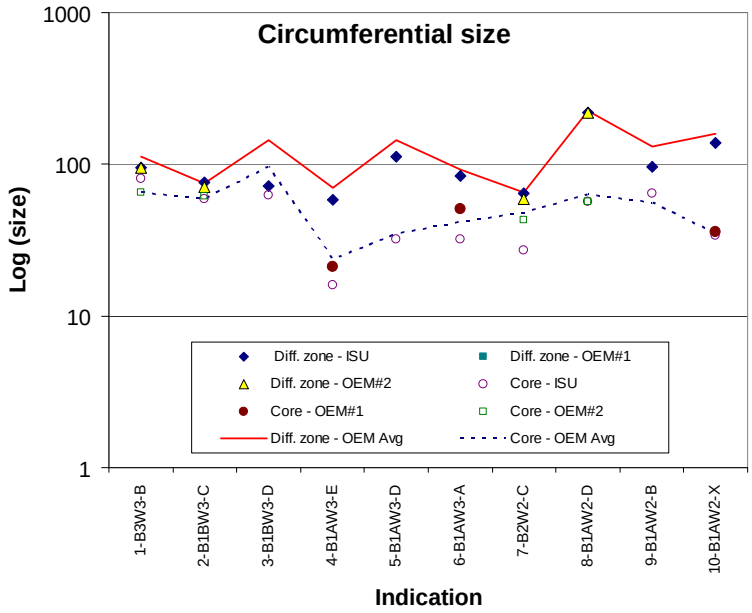


Figure C-14. Circumferential Sizes of Diffusion Zone and Core as Measured by Each Organization

Table C-8 presents a summary of the results for all defects, based on the OEM averages. In addition to the dimensions in the three directions, the area transverse to the direction of ultrasonic propagation is given, as computed by the formula

$$\mathrm{CBS}_{10}\ \mathrm{Area} = (\pi/4) A_x C \qquad \text{(C-1)}$$

where A_x is the length in the axial direction and C is the circumferential dimension. Here, the values in the rotated coordinate system are used. The area would be essentially the same in the unrotated system since the axial length is only slightly affected by the rotation. Figure C-15 shows these results.

Table C-8. Summary of Sizes of Defects

Area	Diffusion Zone				Core			
Defect ID	Axial	Radial	Circ	Area	Axial	Radial	Circ	Area
1-B3W3-B	1127.5	212	112.5	99622.8	437	80	65.0	22309.2
2-B1BW3-C	2494.5	418	75.5	147917.8	1734.5	220	59.0	80374.1
3-B1BW3-D	602.5	296.5	144	68141.1	269.5	103.5	97.0	20531.5
4-B1AW3-E	2305	325	70	126724.0	180.5	31	23.5	3331.5
5-B1AW3-D	998.5	65.5	145	113711.9	165.5	23.5	35.0	4549.4
6-B1AW3-A	1235.5	131	92	89273.1	395	43.5	42.0	13029.8
7-B2W2-C	714.5	84.5	65	36475.9	371	24.5	48.0	13986.4
8-B1AW2-D	1134.5	127.5	225	200482.7	276	54	63.5	13764.9
9-B1AW2-B	808.5	68.5	130	82549.3	139	30	56.0	6113.5
10-B1AW2-X	961.5	62.5	160	120825.7	51	20	35.0	1401.9

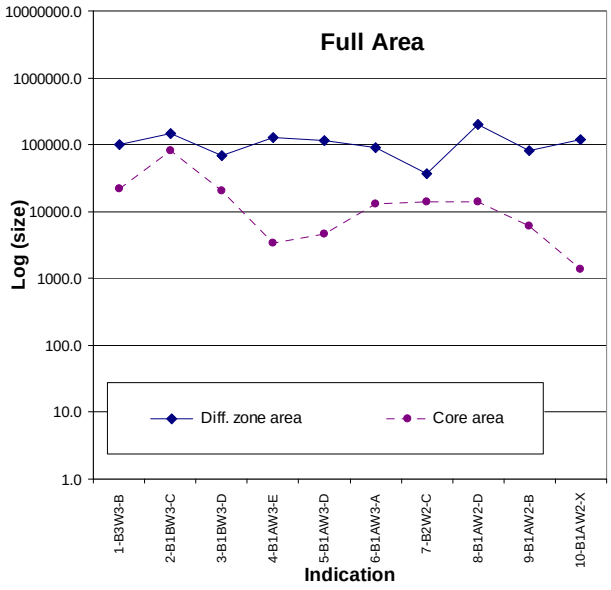


Figure C-15. Full Area of Diffusion Zone and Core Transverse to the Direction of the Ultrasonic Beam

It is interesting to note that, for some defects, the radial dimension is significantly larger than the circumferential dimension, and for others, the reverse is true. This is presumably a consequence of the fact that the billet was not processed in the standard fashion for rotor grade materials. There was a considerable amount of work in the orthogonal planes, and the material was not processed in a symmetrical fashion. For example, the way in which the billet was split apart in the final step caused material initially near mid-radius to be at the center of some of the final billet.

During the review of the data, the metallurgical team made a number of observations that are documented below.

The determination of size was not highly sensitive to the metallurgist making the determination. It was believed that adopting a uniform definition of the characteristics of the diffusion zone and core helped considerably.

There is a growing school of thought within the life management community that the core should be taken as the initial size for fracture calculations. However, there was significant sentiment among the current metallurgists that lifing algorithms should not be solely based on the assumption that a core must be present before crack growth can occur. It was believed that the diffusion zone extent can strongly affect the fatigue life of a highly stressed part and the known cases of fracture in which no voids or core were present were cited, which is called the “stealth defect.” Therefore, it was felt that the diffusion zone must also be considered in design. It was believed that it remains highly important to detect hard alpha defects with only diffusion zones. Accordingly, the metallurgists believe that it is important that research should be conducted to sort out the nondestructive signature of the diffusion zone separate from the nondestructive signature associated with the core.

The question of whether the CBS defects are typical defects that should be used in the probability of detection (PoD) analysis was discussed by the OEM metallurgists at some length. It was noted that when the decision was made to first purchase the CBS material for use in the ETC and TRMD programs, metallurgists from the OEMs made the decision that the defects were typical, despite the fact they were caused by an air leak during the processing of nonrotor grade material. These initial conclusions were, of course, based on the examination of a few micrographs along with limited microprobe analysis. Given that the currently involved OEM metallurgists have had the opportunity to examine a significantly larger set of three-dimensional data, the question of whether the defects were typical was re-examined. Some of the comments are as follows:

- There are a large number of possible causes of formation of hard alpha inclusions. The structures of the cores of the CBS defects are representative of a variety of these.
- It is not as certain that the angles of tilt of the defects or the ratio of diffusion zone size to core are representative, and these may have been influenced by the unusual processing history.
- However, since the core is expected to control the ultrasonic response, the characteristics from a nondestructive evaluation point-of-view seem representative.

These comments, made on a much larger data set, confirm the original decision that the CBS detects should be considered typical for the purposes of PoD analysis.

C.4 DETERMINATION OF SIZES FOR REMAINING CBS DEFECTS.

For the remaining CBS defects, no metallographic analysis was conducted. However, it was desirable to make estimates of their size to increase the number of flaws in the data set, which is the basis for PoD estimates. Indeed, as will be discussed in appendix E, simulations show that the stability of PoD estimates is very sensitive to the number of flaws in the data set. The uncertainty in the results would have been too great if only the CBS_{10} had been used.

The only information available on which the size of the remaining CBS defects can be estimated is the size of the images in the Multizone C-scans. To test how well that might represent the metallographically determined sizes, the CBS_{10} were examined in detail. For the CBS_{10}, the axial and circumferential sizes of the Multizone images at the 6 dB points were estimated and compared to the sizes of the core extracted from the metallurgical analysis. The results are shown in figure C-16. It can be seen that there is an excellent correlation for the axial size but a very poor one for the circumferential size. The latter is not surprising, since the circumferential size of the defects is generally less than the size of the ultrasonic beam. Based on these data, it was felt that the axial length of the remaining CBS defects could be estimated from the C-scans. However, this was not the case for the circumferential extent.

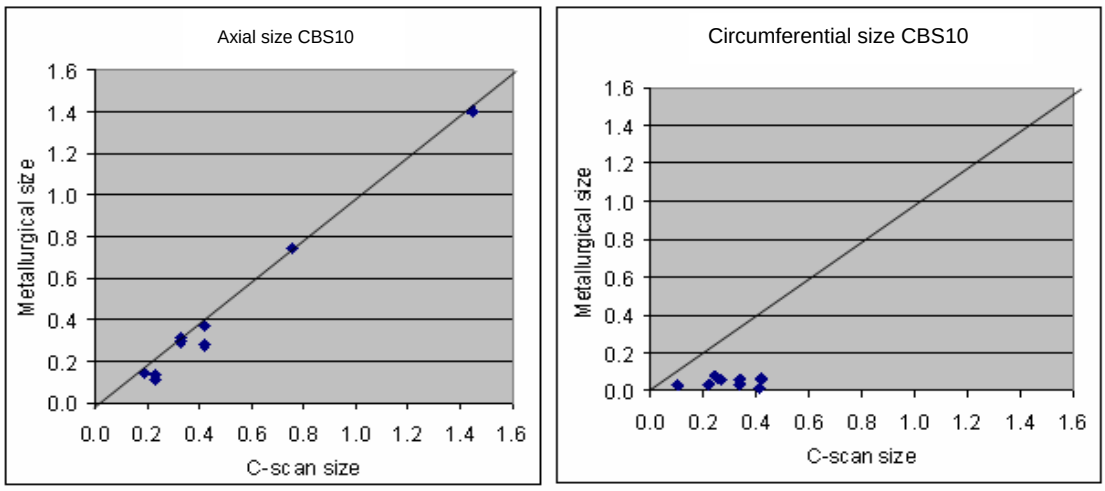


Figure C-16. Correlation of Sizes Inferred From C-Scan Images and Metallurgical Analysis for CBS_{10}

The only recourse in the determination of area was to estimate the circumferential extent from the axial extent. This would be consistent with an assumption made in the original development of exceedance curves [C-2]. Although there would clearly be scatter in such an estimate, the scatter might be acceptable given the large scatter that already exists in the data and the fact that such an assumption would greatly increase the number of data points that could be included in the analysis, greatly reducing statistical uncertainty.

A number of ad hoc relationships were examined, including linear and power law relationships. However, it was ultimately decided to rely on a relationship that had been previously developed by the Aerospace Industries Association (AIA) Rotor Integrity Sub-Committee (RISC). A subteam [C.2] recommended the following model to describe the effect of deformation on hard alpha elongation

$$L_{HA_b} = D_{HA_i} \times \left[\frac{D_i^2}{D_b^2} \right]^{0.364} \qquad \text{(C-2)}$$

where L_{HA_b} is the length of the hard alpha in the billet, D_{HA_i} is the diameter of the hard alpha in the ingot, D_i is the ingot diameter, and D_b is the billet diameter. If one additionally assumes that the volume of the hard alpha is conserved during processing,

$$\left[L_{HA_b} \times D_{HA_b}{}^2 \right] = D_{HA_i}{}^3 \qquad \text{(C-3)}$$

eliminating D_{HA_i} from equations C-2 and C-3 leads to the relation

$$\frac{D_{HA_b}}{L_{HA_b}} = \left[\frac{D_b}{D_i}\right]^{1.092} \qquad \text{(C-4)}$$

For the contaminated heat, D_b = 36 inches and D_i = 6 inches. Equation 13, thus, predicts $(D_{HA_b} / L_{HA_b}) = 0.14$ for that specific case.

Finally, solving equation C-4 for D_{HA_b} and inserting that result into the expression $\text{Area} = \left(\frac{\pi}{4}\right) L_{HA_b} D_{HA_b}$, one finds

$$\text{Area} = \left(\frac{\pi}{4}\right)\left[\frac{D_b}{D_i}\right]^{1.092} L_{HA_b}^2 . \qquad \text{(C-5)}$$

For the specific case of the contaminated heat, this reduces to $\text{Area} = 0.11\, L_{HA_b}^2$. As a test of this relationship, its predictions were compared to areas measured from metallographic analysis for the Default PoD-3Dimensional, Jet Engine Titanium Quality Committee, and CBS_{10} defect for the core area. The results are shown in figure C-17. Here, the predicted area is plotted versus area determined from full metallography data. A regression analysis was performed to determine the best fit, resulting in the solid line. The dotted line shows where the points would fall if agreement between measured and predicted areas were perfect. Equation C-5 is conservative, since, for large flaw sizes, the predicted size (discrete data points) consistently falls above the actual size (given by the dotted line). Based on this result, equation C-5 will be used throughout this report to determine the area of the remaining CBS defects.

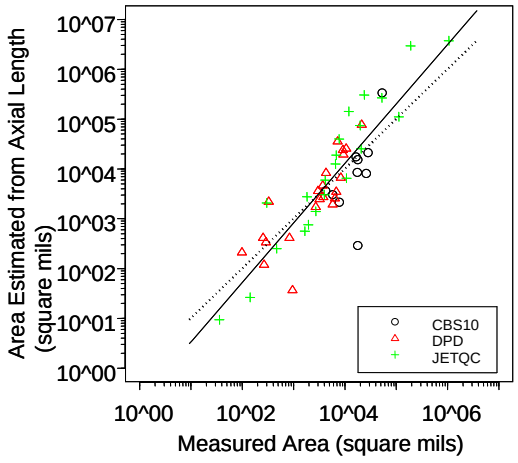


Figure C-17. Comparison of Area Predicted by Equation C-5 to Metallographically Determined Area for all Defects That Have Been Sectioned (Solid line is the result of regression and broken line is the result that would be expected for perfect agreement.)

C.5 UNINSPECTED ZONES.

In the course of the PoD analysis, a number of large flaws were identified that had been detected by Multizone but missed by conventional. These caused great difficulty with the PoD analysis since they did not fit the traditional statistical model used to describe such data, e.g., the regression line in an $\hat{a}$ versus a analysis. Two observations allowed this difficulty to be significantly reduced.

Any inspection technique has a near surface dead zone, where flaws cannot be detected. It was learned that this is 0.2 inch for Multizone, but a greater distance for conventional inspection, depending on the implementation. The physical reason has to do with the smaller, near-surface dead zone of spherically focused, Zone 1 (near surface) transducer used in Multizone compared to the cylindrical transducer used for all depths in conventional inspection. For the tests on the CBS, these dead zones had nominal values of 0.2 inch for Multizone and 0.375 inch for normal conventional inspection. An examination of table C-1 shows that nine of the defects had reported depths of less than 0.4 inch. In only two cases did these produce a conventional inspection response (B1BW2-B, 0.3 inch, 108% in 1995 normal conventional inspection; B3W2B, 0.3 inch, 100% in 1994 normal conventional inspection and >100% in 1995 angle conventional inspection). Recognizing that there is a ±0.1 inch accuracy in depth determination from Multizone A-scans, these apparent exceptions could be the result of a flaw that was actually at a depth of 0.4 inch. For the purposes of the PoD analysis, all points with nominal depths less than 0.4 inch were not considered in the conventional inspection analysis.

It must be emphasized that this implies care in the use of the resulting Default PoD curves. For them to be valid, an OEM must take into account the larger conventional dead zone in the development of forging plans to ensure that no uninspected material finds its way into the final part.

In addition, one of the billets had a condition known as an endburst. When the conventional inspection was done in 1994, “no test, visible hole” was noted on the data sheets for the first 5″ of billet B2W2. The summary indicated that this was associated with a condition known as “endburst.” The indications from these flaws were clearly visible in the C-scan records of the 1995 conventional inspection. Hence, defects B3W2-A, B, and C were not considered for the purpose of the conventional PoD analysis since, although they produced large signals, they were in a region for which no quantitative conventional records were kept. (Note that an incorrect axial position had been assigned to flaw B2W2-C at some point in the analysis; based on the 1995 C-scan, it is within the above endburst regime.)

C.6 OTHER CORRECTIONS.

In the course of this analysis, it was also noted that B3W2-H does not exist since it is the same as B3W2-G. Also, B1BW2B-A, originally indicated as a miss in the 1994 data, had actually been found. A hand correction in the original datasheets had not found its way into the final tabulation. This defect actually produced a signal of 90% for both conventional normal and angle inspections. Finally, some typos in the responses of B1AW1-D and B1AW1-E were noted.

As noted previously, these observations were taken into account in the preparation of the Venn diagrams in figures C-1 and C-2 to indicate there were 48 rather than 60 defects in regions that were actually examined during the conventional inspection.

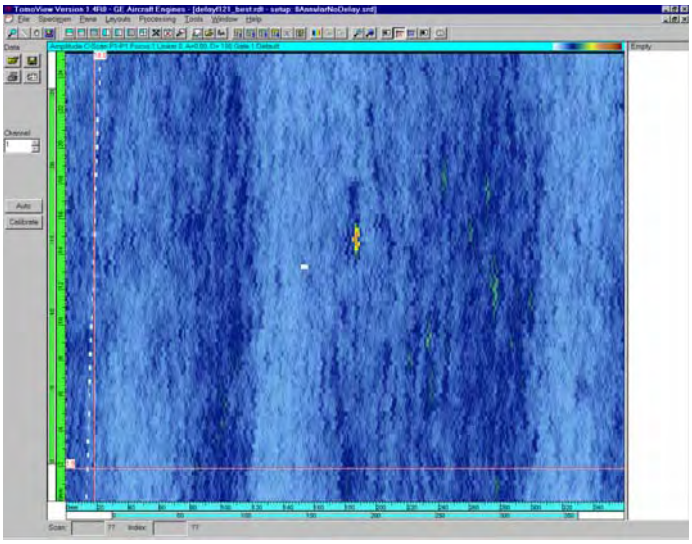
C.7 MULTIZONE MISS.

C.7.1 BACKGROUND.

In the CBS data set, there are a significant number of conventional misses whose existence is known because they were found in the Multizone inspection, which was acting as a referee. However, no Multizone misses were known since there was no referee for the Multizone inspection. Under ideal conditions, this should not be a problem since the assumption of truncation in the statistical analysis takes into account the fact that not all flaws present may have produced detectable signals. However, the accuracy of the truncation analysis is contingent upon the accuracy of the form of the statistical model assumed in the analysis, e.g., the assumption that the distribution of effective flat-bottom hole area (EFBH) responses for a given flaw size are governed by a log-normal distribution about a simple regression line. Knowledge of misses can verify this assumption or suggest that it is in error. The analysis of the conventional data provides an example in which knowledge of misses showed that such an assumption was not valid and new steps have to be taken in the statistical analysis to take this into account. The effects on the resulting PoD curves were very significant, implying a much poorer capability than would have been predicted had those misses not have been known. This caused the ETC to fully review the records of the CBS study.

C.7.2 MAJOR RESULT OF MISS ANALYSIS.

Section 3.8 of the CBS report [C-1] describes a Misses Assessment activity. As one part of this, higher sensitivity inspection was applied to 11 mults of various lengths, as detailed in table 20 of that report. These mults represented about 25% of the CBS material. The higher-sensitivity inspection was performed with an annular phased array probe with a 10-MHz center element surrounded by rings operating at 7.5-MHz (7-element Digital-Fased Array Testing System®). The results of that inspection were that all previous indications were confirmed and one additional indication was found in billet section B1BW3-CB. This new indication was then re-inspected by the standard Multizone system, producing a nonrejectable indication. However, re-inspection with a 10-MHz phased array probe produced a clear indication. Figure C-18 reproduces the results of that phased array inspection, as shown in figure 74 of the CBS report [C-1]. The top image is the C-Scan and the bottom image is the waveform. Further details may be found in the CBS report [C-1]. At that time, the ETC recommended that the material producing this flaw indication be destructively characterized to determine whether it had been produced by a hard alpha inclusion and, hence, constituted a Multizone (and conventional inspection) miss and, if so, to determine its true size. Unfortunately, contractual changes precluded obtaining this data at that time.



10 MHz phased array inspection results

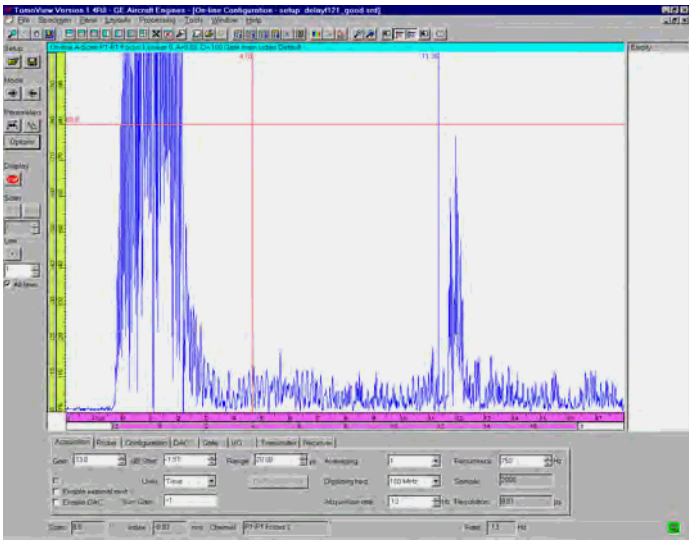
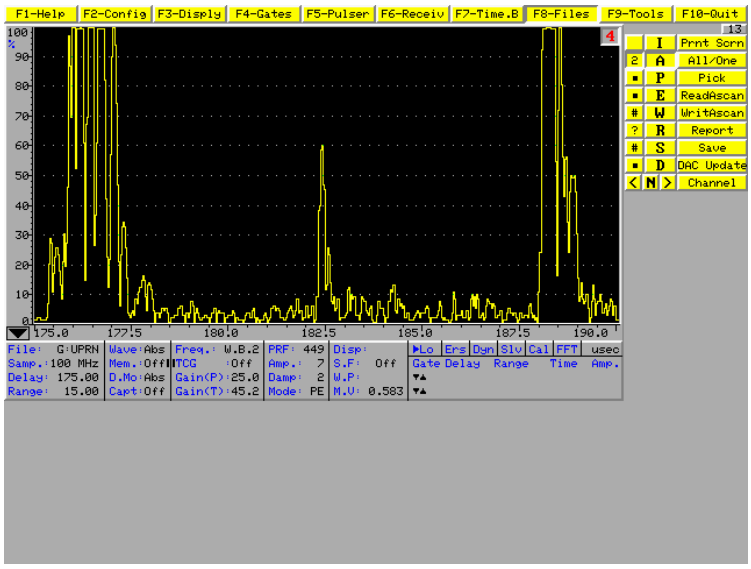


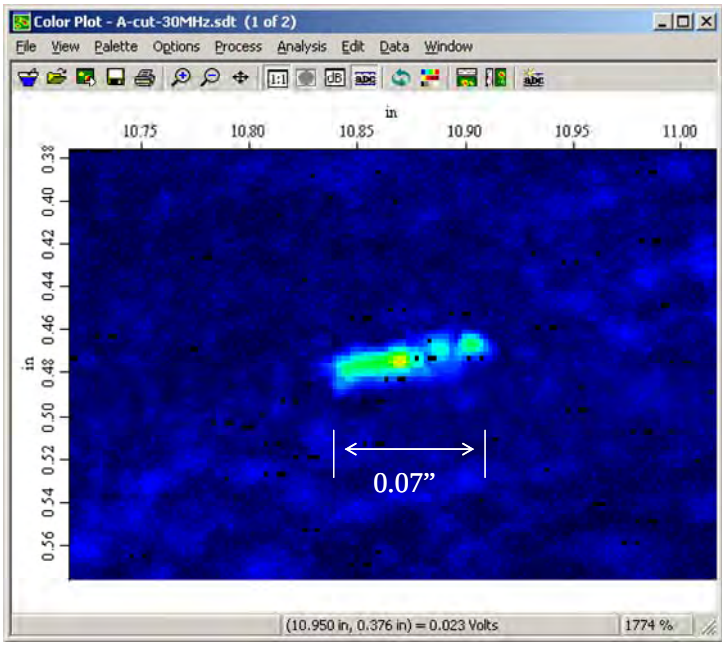
Figure C-18. The 10-MHz Phased Array Inspection Results for Multizone Miss

In light of the potential significant impact of knowledge of misses, this indication has been examined further. Included was additional ultrasonic inspection at a higher frequency to better define the extent of the defect.

Material containing the indication was first cut into a cube shape and then further sectioned to place the flaw 0.21″ below the surface. This was scanned with a focused, 30-MHz probe. Figure C-19 shows the results of ultrasonic scans of the indication in the cube form, an A-scan at 10 MHz in the initial cube and the 30-MHz C-scan in a reduced cube. An indication is clearly visible in both, and from the C-scan, the flaw appears to have a length of 0.07 inch.



(a)



(b)

Figure C-19. Scan of the Missed Indication in Cube Form (a) 10-MHz A-Scan and (b) 30-MHz Focused C-Scan

Based on the clearness of the indication, it was determined that metallurgical characterization should be undertaken using the same procedure that was used in the evaluation of the CBS_{10} and described in detail in the CBS report [C-1]. The purpose was to determine whether this indication was produced by a hard alpha inclusion and make a more precise determination of its size. The results of that determination are reported below.

Serial sectioning was conducted in accordance with the procedures used to section the CBS_{10}, as shown in figure C-3, with the exception that the planes were separated by distances of 2 mil (50 micrometers) rather than 5 mil. The size of the core was then determined following the procedures described in section C.3.2. The determination of the dimensions, based on that procedure, was made by one metallurgist from the team with metallurgists from the other organizations acting in a review mode.

The major results of the metallography and microprobe measurements are presented below. Full details may be found in section C.7.3.

Figure C-20 shows the optical metallograph of the flaw at its approximate midplane. The entire set of images will be shown in the section C.7.3. Several aspects of the image in figure C-20 are important because

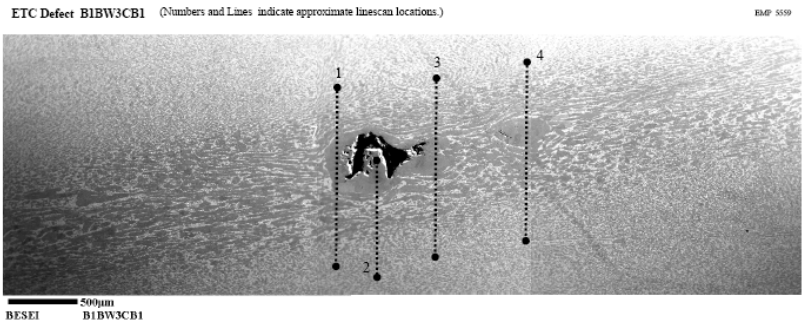
- the image structure strongly suggests that the flaw was a hard alpha inclusion. The typical void (black region), core (uniform white region corresponding to single phase alpha), and diffusion zone (mixture of alpha (white) and beta (grey) phases with graduated sizes) is observed.

- the features are qualitatively similar to those of the ultrasonic image shown in figure C-19, i.e., two strongly reflecting regions separated by a weakly reflecting region.

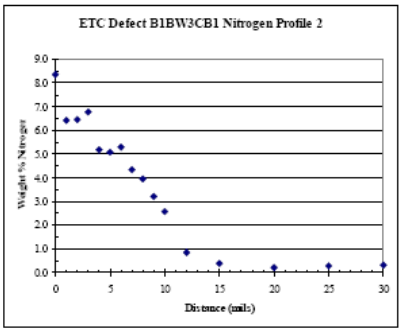
- there is not an exact one-to-one correspondence between the two regions. This should not be surprising since the ultrasonic image is made in the axial-circumferential plane, while the optical image is made in the axial-radial plane, in accordance with the sectioning procedure, as illustrated in figure C-3. In addition, one image appears to be rotated 180 degrees with respect to the other, also not surprising since no effort was made to distinguish left from right when the images were made in the cube form.

Figure C-20. Image of Center of Multizone Miss

At another plane, ion microprobe measurements were made to determine the chemical composition. Figure C-21 shows the four scan lines along which microprobe measurements were made and a graph of nitrogen content for scan line 2. In this scan line, the peak nitrogen content is 8.5%, confirming this to be a hard alpha inclusion, which was designated defect "mm." Full details may be found in section C.7.3, including graphs of the nitrogen content along the other scan lines and tables of compositions at each of the measurement points, including a set that was randomly selected at positions other than the scan lines. For scan lines other than 2, the peak nitrogen content was between 1.5% and 2.5%. The latter values are less than the nominal values of ~3-4% that are suggested to be used to define the high-nitrogen core in figure C-9. However, the fact that the optical micrograph clearly indicates that these scans pass through regions that are a uniform color, and hence, an "essentially single-phase alpha," combined with the fact that extensions of these regions are voided in other nearby image planes, leads to the identification of these regions as part of the core. It should be noted that similar regions of relatively low nitrogen content were identified as core regions in the CBS_{10} (B1AW3-E:1.5% and B1BW3-D:1.5%). It is also worth noting that the scans are taken in representative locations. Because of the cost of microprobe analysis, scans cannot be made at every slice. It is possible that higher nitrogen content regions exist in other locations of the defect.



(a)



(b)

Figure C-21. Results of Ion Microprobe Measurements of Composition (a) Scan Lines Along Which Measurements Were Made and (b) Graph on Nitrogen Content Along Scan Line 2

Given the confirmation that this was a hard alpha inclusion, the size was systematically determined following the general procedures used for the CBS_{10}. Figure C-22 shows the optical

image in the plane, which was judged to have the largest dimensions in the axial (horizontal) and radial (vertical) directions. The area identified as core has been colored. The circumferential direction was inferred from the distance between planes in which core was first and last seen. In the successive polishing in 2-mil (50 micrometer) intervals, there were 16 mils between the planes in which the core was first and last seen and 20 mils between the nearest planes in which the core was not seen. Since the true core dimension must be between these values, the average, or 18 mil, was taken to be the circumferential extent of the core.

Figure C-22. Image of Defect in Plane of Largest Dimensions With Core Region Highlighted by Color

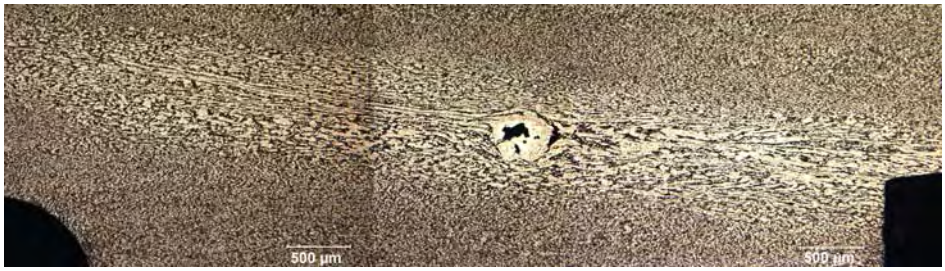
Based on the above analysis, this defect was assigned the core dimensions of 67.8 mil (axial), 24.4 mil (radial), and 18 mil (circumferential). The area of the core (of an ellipse with major and minor axes equal to the axial and circumferential dimensions, respectively) is 958.5 square mil.

These results provide evidence of one Multizone miss. It would be reasonable to ask whether there are more, and if so, what impact would the knowledge of their number and sizes have on PoD estimates (which could be negligible or large depending on the details of the answer). In this context, three points should be noted. (1) only 25% of the CBS was scanned during the misses assessment, so there is no knowledge of the presence (or lack) of misses in the remaining material, (2) this 25% was scanned with a technique that did not have a dramatically higher sensitivity than Multizone, and (3) review of the CBS report [C-1] provides indication of the possibility of additional misses since, in the reproducibility study, two of the defects found in the 1995 Multizone scan were not rejected in all subsequent scans. B1BW1A-D was rejected on 9 out of 15 opportunities (six misses) and B2W3C was found on 11 out of 15 opportunities (four misses). However, since only some of the billets were examined in the reproducibility study, there was no way to take this information into account in the development of the Default PoD curves. Hence, the curves presented in figures C-17, C-18, C-20, and C-21 are based on the knowledge of one miss only, although the possibility exists that there may have been more.

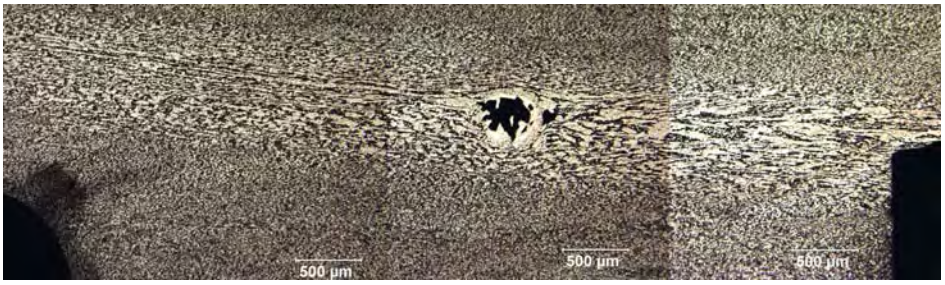
Inspection technology has evolved significantly since the end of the CBS study. It would be of value to rescan the remaining mults from the CBS with a laboratory system with the highest sensitivity available to determine if other defects that were missed in the 1995 Multizone scan can be identified, and then use this information to further assess the PoD curves for both Multizone and conventional inspections. However, since that information is not available at this time, the PoD curves to be reported will include only the one miss described above in the data set. They represent the best estimate of PoD based on the data available.

C.7.3 DETAILED METALLOGRAPHY AND CHEMICAL ANALYSIS OF MULTIZONE MISS.

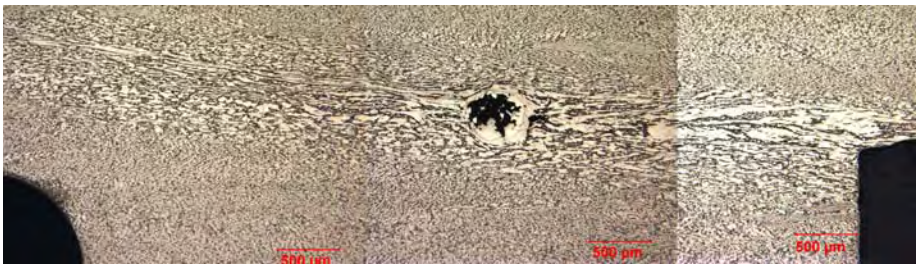
The full results of the destructive analysis of the Multizone miss are included in this section. Figure C-23 shows nine metallographic sections, 2 mil apart, with the first slice being closest to the billet outer diameter and the last slice being closest to the billet center.



Slice 1



Slice 2

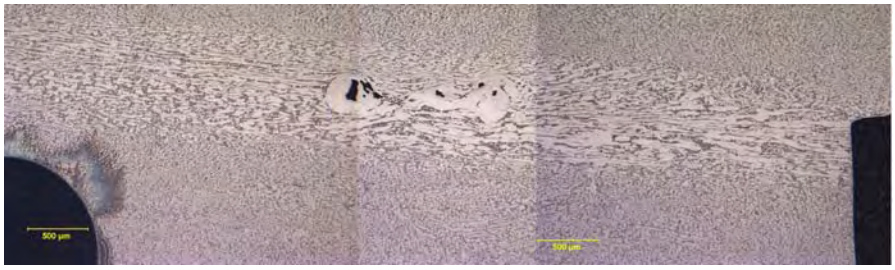


Slice 3

Slice 4

Figure C-23. Full Results of Serial Sectioning of Suspected Multizone Miss

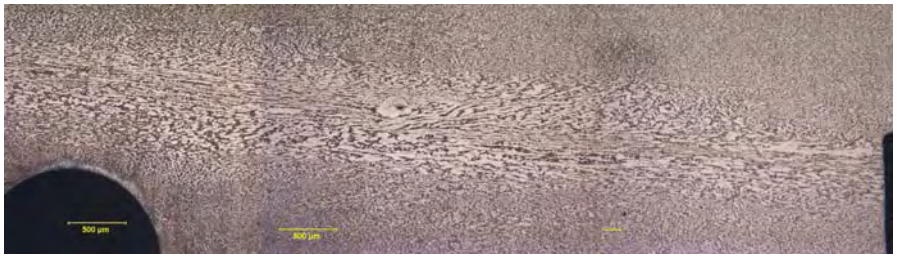
Slice 5



Slice 6

Slice 7

Slice 8



Slice 9

Figure C-23. Full Results of Serial Sectioning of Suspected Multizone Miss (Continued)

At slice 4, a microprobe analysis was performed, including both point measurements and line scans. The random points from the microprobe analysis are shown in figure C-24 and table C-9, which shows the results of chemical analysis at each point.

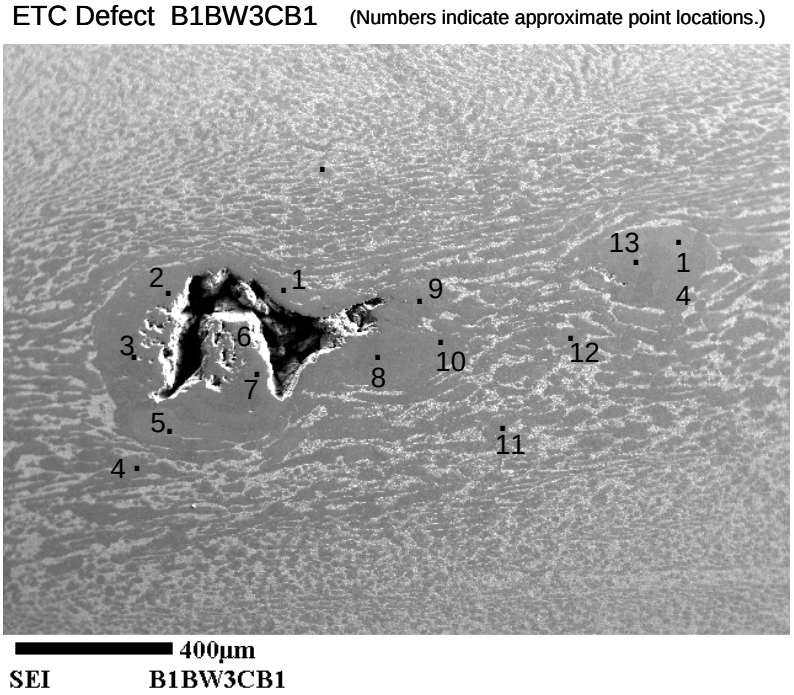


Figure C-24. Location of Point Microprobe Measurements on Slice 4 of Suspected Multizone Miss

Table C-9. Results of Chemical Analysis at Randomly Selected Points on Slice 4 of Suspected Miss Data for Areas Around the Defect

B1BW3-CB1	Weight %			
Point Number	Ti	Al	V	N
1	92	5.3	1.9	0.7
2	93	0.4	1.7	4.0
3	91	0.4	1.3	3.6
4	88	6.5	1.7	0.5
5	90	4.7	1.5	1.1
6	89	0.6	1.7	6.6
7	91	0.7	1.8	4.5
8	92	0.3	1.6	3.5
9	91	4.5	1.8	1.2
10	91	5.5	1.9	0.6

Table C-9. Results of Chemical Analysis at Randomly Selected Points on Slice 4 of Suspected Miss Data for Areas Around the Defect (Continued)

B1BW3-CB1	Weight %			
Point Number	Ti	Al	V	N
11	89	6.8	2.0	<0.5
12	88	5.9	1.9	0.5
13	94	1.1	1.5	2.0
14	92	3.8	1.6	1.1

In addition to the point chemical analysis, scans of chemical composition were also made along the lines shown in figure C-25.

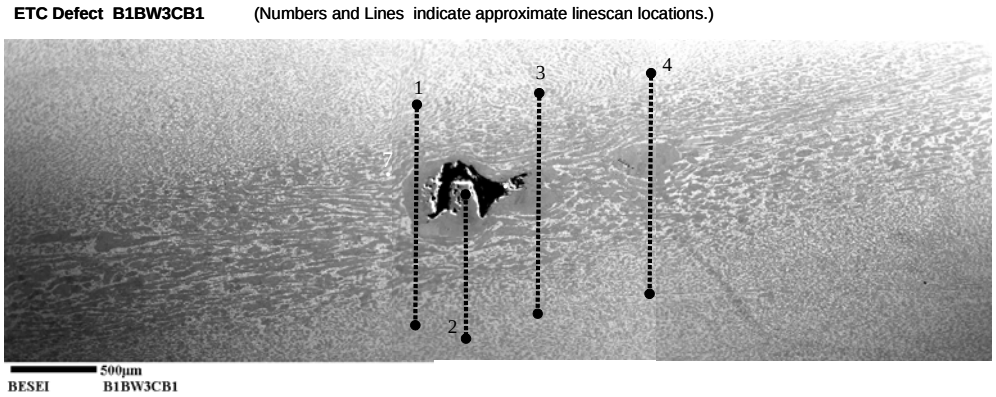


Figure C-25. Location of Line Scan Microprobe Measurements on Slice 4 of Suspected Multizone Miss

Table C-10 shows the results of the chemical analysis along line 1 and figure C-26 shows the results of the nitrogen analysis. Table C-11 and figure C-27 show the corresponding results for line 2, table C-12 and figure C-28 show the corresponding results for line 3, and table C-13 and figure C-29 show the corresponding results for line 4.

Table C-10. Results of Chemical Analysis Along Line Scan 1

Concentration Profile Data			B1BW3-CB1	
Profile 1			Weight %	
Distance (mils)	Ti	Al	V	N
0	92	7.3	1.8	<0.5
5	91	7.1	1.7	<0.5
10	83	4.5	8.8	<0.5
15	92	6.8	2.0	<0.5
17	92	6.4	1.8	<0.5
19	91	5.5	1.9	0.6

Table C-10. Results of Chemical Analysis Along Line Scan 1 (Continued)

Concentration Profile Data			B1BW3-CB1	
Profile 1			Weight %	
Distance (mils)	Ti	Al	V	N
20	91	5.3	1.7	0.8
21	91	4.2	1.4	1.3
22	91	4.5	1.3	1.1
23	91	4.0	1.2	1.6
24	92	1.4	1.1	2.2
25	93	0.5	1.0	2.5
26	91	1.8	1.1	2.3
27	93	1.6	1.1	2.2
28	93	1.0	1.1	2.5
29	93	2.7	1.4	1.7
30	92	3.8	1.4	1.4
31	91	4.6	1.4	1.1
33	91	5.2	1.6	0.9
35	91	5.1	1.7	0.6
40	89	6.6	1.7	<0.5
45	89	6.9	1.6	<0.5
50	87	7.1	1.6	<0.5

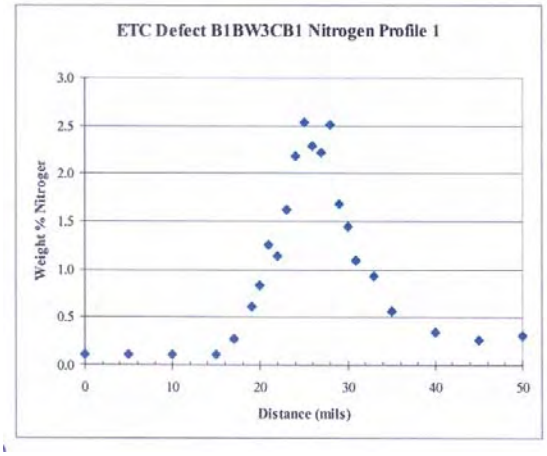


Figure C-26. Variation of Nitrogen Content Along Line Scan 1

Table C-11. Results of Chemical Analysis Along Line Scan 2

Concentration Profile Data			B1BW3-CB1	
Profile 2			Weight %	
Distance (mils)	Ti	Al	V	N
0	89	0.1	1.4	8.4
1	90	0.6	1.7	6.4
2	89	0.5	1.8	6.4
3	89	0.4	1.5	6.8
4	90	0.4	2.0	5.2
5	91	0.5	1.8	5.1
6	90	0.3	1.7	5.3
7	91	0.3	1.6	4.3
8	91	0.3	1.7	4.0
9	92	0.3	1.7	3.2
10	93	0.4	1.6	2.6
12	91	5.0	1.8	0.8
15	90	6.5	1.9	<0.5
20	89	6.9	2.0	<0.5
25	87	7.1	2.0	<0.5
30	87	7.2	1.7	<0.5

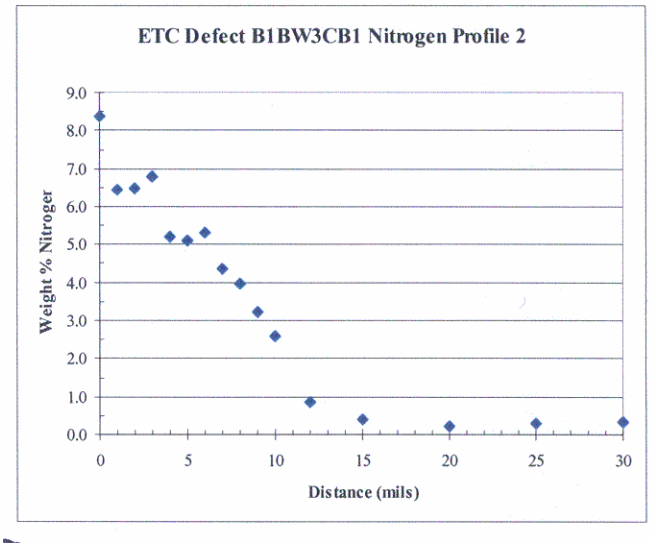


Figure C-27. Variation of Nitrogen Content Along Line Scan 2

Table C-12. Results of Chemical Analysis Along Line Scan 3

Concentration Profile Data			B1BW3-CB1	
Profile 3			Weight %	
Distance (mils)	Ti	Al	V	N
0	85	5.0	6.6	<0.5
5	88	7.2	1.9	<0.5
10	91	6.6	2.6	<0.5
15	93	6.8	2.2	<0.5
17	93	6.7	2.1	<0.5
19	92	6.2	2.0	<0.5
20	92	5.7	2.2	0.6
21	92	5.3	2.3	0.8
22	95	0.9	1.5	1.6
23	93	4.5	1.8	1.3
24	92	4.4	1.7	1.3
25	91	3.8	1.6	1.6
26	93	0.6	1.4	2.7
27	94	0.6	1.4	2.6
28	94	1.0	1.3	2.3
29	91	4.7	1.6	1.2
30	91	5.2	1.6	0.8
31	91	5.5	1.9	0.7
33	90	6.0	2.0	0.6
35	90	6.4	1.9	<0.5
40	89	6.4	2.7	<0.5
45	88	7.0	1.8	<0.5
50	87	7.2	1.8	<0.5

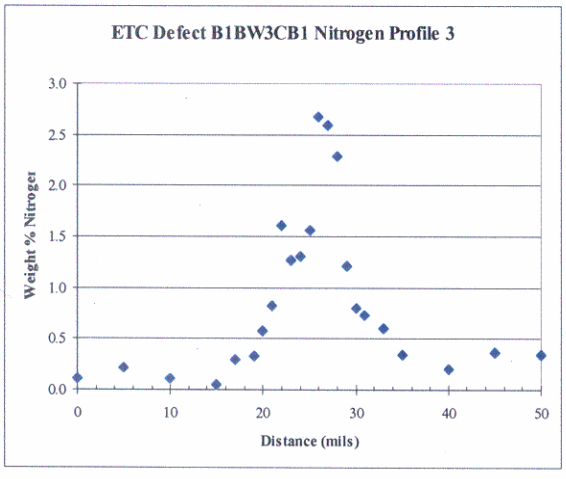


Figure C-28. Variation of Nitrogen Content Along Line Scan 3

Table C-13. Results of Chemical Analysis Along Line Scan 4

Concentration Profile Data			B1BW3CB1	
Profile 4			Weight %	
Distance (mils)	Ti	Al	V	N
0	93	7.3	1.8	<0.5
5	92	7.3	1.7	<0.5
10	93	7.2	1.6	<0.5
15	93	7.0	1.8	<0.5
17	92	6.9	1.8	<0.5
19	92	6.3	1.8	0.5
20	93	6.0	1.7	0.5
21	92	5.5	1.9	0.5
22	93	5.3	1.8	0.5
23	93	4.2	1.5	0.9
24	94	2.2	1.1	1.3
25	96	1.8	1.4	1.5
26	95	3.5	1.4	1.2
27	94	4.6	1.4	0.9
28	94	5.1	1.5	0.8
29	94	5.3	1.9	0.6
30	94	5.2	1.9	0.7
31	93	5.4	2.0	0.6
33	93	6.0	1.9	0.5
35	86	4.3	7.6	<0.5
40	94	6.7	1.7	<0.5
45	91	7.0	1.9	<0.5
50	91	7.1	1.8	<0.5

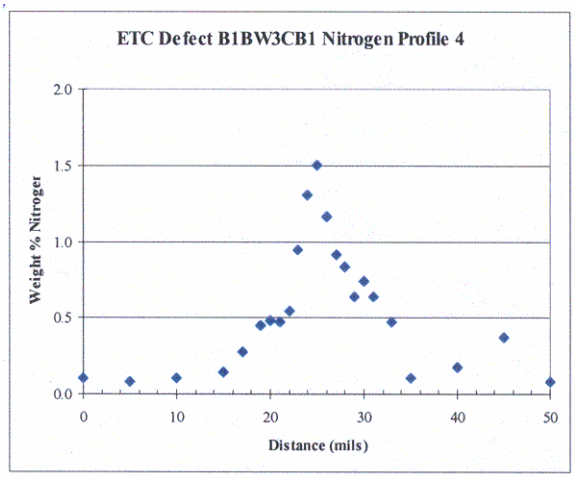


Figure C-29. Variation of Nitrogen Content Along Line Scan 4

C.8 GRAPHS OF RESPONSE VERSUS AREA.

Figures C-30, C-31, and C-32 show the response versus area for the full CBS data set. The normal conventional inspection results obtained in 1994 are shown in figure C-30. This data was chosen over the 1995 data because the 1994 runs were done in accordance with typical, stop-on-defect, industrial practice. It should be noted, however, an analysis of the 1995 data was also performed, but did not lead to substantially different results. As indicated on the figure, different symbols are used to represent different types of data. Upward pointing triangles indicate saturated data that are statistically treated as right censored. These are plotted at the level corresponding to 100% FSH and indicate that the defect in question produced a signal at this level or higher. Downward pointing triangles indicate misses (as indicated by Multizone finds), which are statistically treated as left censored. These are plotted at the noise level and indicate that the defect in question would have produced a signal at this level or lower. Circles indicate unsaturated (exact) points. For the statistical techniques employed, it is also necessary to specify a truncation level, information that allows the analysis to take into account the fact that there may have been defects present in the billet for which no ultrasonic information was obtained. The truncation level, EFBH = 278, is below the level of the left-censored points because the truncation is determined by the Multizone rather than conventional noise level. In actuality, this should be different for each flaw location. However, the data available did not allow these differences to be taken into account, so the same average value was used for all flaw locations. Figure C-31 shows the corresponding results for the conventional angle inspection. These ordinates are % FSH rather than EFBH, since the latter does not apply to inspections calibrated with side-drilled holes. Figure C-32 shows the corresponding results for the Multizone inspection, for both amplitude and SNR_{mz}.

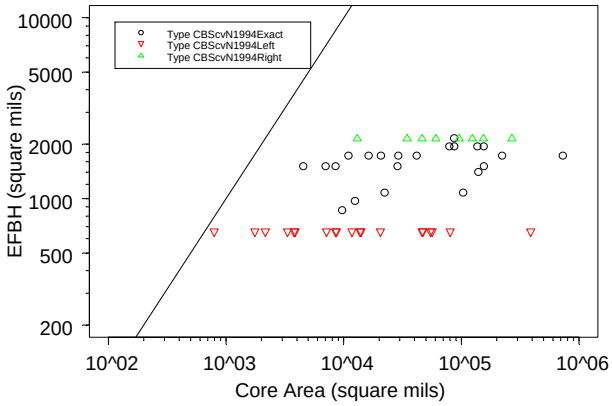


Figure C-30. Response Versus Core Area for Normal Conventional Inspection of CBS

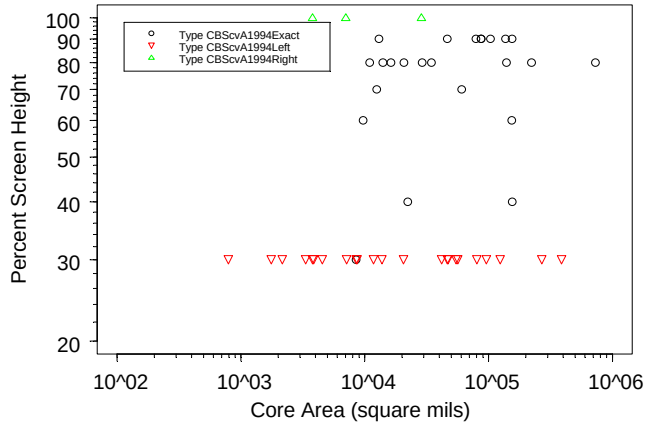
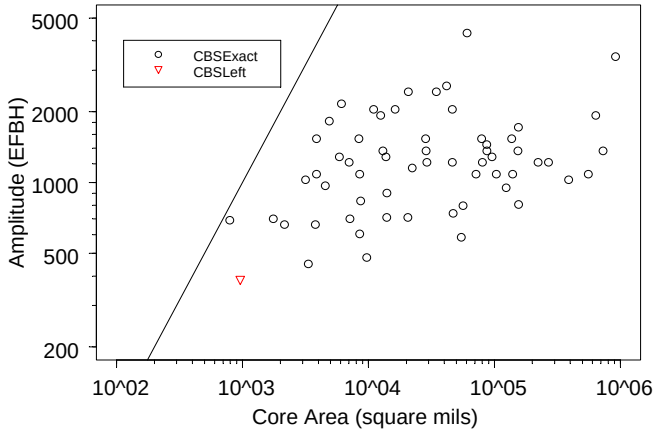
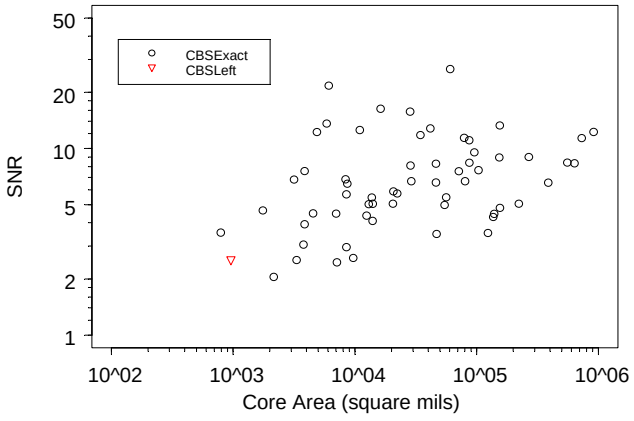


Figure C-31. Response Versus Core Area for Angle Conventional Inspection of CBS



(a)



(b)

Figure C-32. Response Versus Core Area for Multizone Inspection of CBS (a) Amplitude Response and (b) SNR_{mz} Response

C.9 FINAL FINDS DATA.

Table C-14 shows the final finds data set that was used in the analysis. This differs from table C-1 in six ways:

1. The defects that were not considered in the conventional analysis because they were in the dead zone are highlighted in yellow.

2. B2W2-A (considered as an "end burst") and B3W2-H (the defect that did not exist) have been removed from the table with the proper information entered for B3W2-G.

3. The proper information was entered for B1BW2-B-A.

4. Some typos were corrected for B1AW1-D and B1AW1-E.

5. Size information was added to the table. Included are the axial length inferred from the Multizone C-scans, the metallographically determined area for the CBS_{10}, and the estimated area for the remaining CBS defects based on equation C-5.

6. The sixth change requires some discussion. There are a number of new entries in the 1995 conventional amplitude columns, which are placed in parentheses to differentiate them from the data that was reported in the CBS report [C-1] as well as table C-1. These were obtained by two ETC team members in March 2005 while re-examining the 1995 conventional inspection records (digitally recorded C-scans) to try to quantify the signal strength of apparent misses. For each case, the C-scan records were examined at the point of which the defect was known to exist based on the Multizone record. If both agreed that there was a well-defined indication, its value was recorded, even if below the reject threshold. If there were no such indication, "NA", meaning indication level not available, was indicated. If a cell has no entry, that case was not examined.

This was a diagnostic test with two goals, and the results are reported for completeness. The first goal was to provide more information for PoD analysis to see whether a specific value of a small signal would significantly modify the results obtained when that flaw's response was treated as left censored (less than a recording level). In general, the signals recovered were all quite small, although there were a few exceptions. B2W2-B and B1BW3-C are examples of those exceptions, having angle responses of 90% and 74%, respectively, despite the fact that they were not identified as indications in the 1994 or 1995 inspections. The second goal was to determine how well the conventional misses in 1994 correlated with the digital C-scan records in 1995. It was generally observed that if a signal was missed in 1994, it produced a small signal on the 1995 C-scan records, suggesting that flaw morphology was playing an important role. The exceptions were for the angle scan data, but this may not be surprising since the 1994 and 1995 angle scans used a different mode.

Table C-14. Final Results of the Production Inspection of Contaminated Billet

				Flaw Signal Amplitude % FSH						Signal-to-Noise Ratio		Flaw Area in sq in.	
Billet	Flaw	95 MZ Depth (inches)	95 MZ C-scan axial size (inches)	94 CVL	94 CVA	95 CVL	95 CVA	94 MZ	95 MZ	94 MZ	95 MZ	full core area from OEM metallographic analyses	full core area based on length determined from C-scan images
B1AW1	A	0.2	0.2098					>100	190	6.5	12.24		0.005
B1AW1	B	0.6	0.3356	45	70	(30)	42	>100	201	4.8	4.36		0.012
B1AW1	C	1.6	0.7132			42	28	95	83	4.06	5.47		0.056
B1AW1	D	1.7	0.1259			(NA)	(NA)	62	73	3.4	4.65		0.002
B1AW1	E	0.6	0.2517	70	>100	(NA)	(>100)	89	127	4.2	4.47		0.007
B1AW2	A	0.3	0.2300						134		13.59		0.006
B1AW2	B	0.2	0.1390						225		21.7	0.006	0.002
B1AW2	C	0.8	0.3834	80	80	65	>100		213		16.29		0.016
B1AW2	D	0.7	0.2760			50	(38)		134		5.46	0.014	0.008
B1AW2	E	1.3	1.1119	90	90	>100	>100		160		4.29		0.136
B1AW2	F	1.1	0.6134	80	80	94	86		268		12.8		0.041
B1AW2	X	0.5	0.0510			(NA)	(NA)		94		5.05	0.001	0.000
B1AW3	A	0.7	0.3950	>100	90	50	60	>100	142	2.7	5.02	0.013	0.017
B1AW3	B	0.9	0.1395			(NA)	(NA)	57	69	2.6	2.05		0.002
B1AW3	C	1.5	0.5581	>100	80	62	>100	>100	253	5.8	11.79		0.034
B1AW3	D	1.1	0.1655	70		(32)	NA	91	101	5.5	4.49	0.005	0.003
B1AW3	E	1.6	0.1805			(NA)	(NA)		47	NA	2.52	0.003	0.004
B1AW3	Y	1.3	0.3256			~~34~~	~~32~~		99	NA	5.05		0.012
B1AW3	Z	1.3	0.6511			(NA)	(NA)		77	NA	3.48		0.047
B1BW1	A	2.2	0.2772			(30)	68		63		2.96		0.008
B1BW1	B	1.7	0.2772	70	30	(46)	(NA)		113		5.67		0.008
B1BW1	C	0.6	0.5082	70	>100	70	>100		142		8.09		0.028
B1BW1	D	1.0	0.1848			(NA)	(NA)		56		2.79		0.004
B1BW1	E	0.9	0.1848	30	>100	(NA)	(NA)		69		3.05		0.004
B1BW1	F	1.1	0.7392	>100	70	179	>100		450		26.6		0.060
B1BW1	G	1.5	0.6468	>100	90	100	80		213		8.28		0.046
B1BW1	H	0.4	0.6468			(NA)	(NA)		127		6.56		0.046
B1BW2B	A	1.2	1.1819	(90)	(90)	118	(NA)		179		13.27		0.154
B1BW2B	B	0.3	0.5065			108	(NA)		160		15.74		0.028
B1BW2B	C	0.3	0.8020						113		7.55		0.071
B1BW2B	D	1.6	0.8442	90	90	46	116		160		11.38		0.078
B1BW2B	E	1.5	0.8864	90	90	108	101		142		8.38		0.086
B1BW3	A	1.3	0.1868			(30)	(NA)	97	113	3.99	3.92		0.004
B1BW3	B	1.1	0.1868			(36)	(58)	>100	160	5.78	7.56		0.004
B1BW3	C	0.5	1.7345			(40)	(74)	>100	127	3.52	6.67	0.080	0.331
B1BW3	D	1.5	0.2695			(NA)	(NA)	70	74	3.85	5.06	0.021	0.008
B1BW3CB	MM											0.001	0.001
B2W1	A	0.5	2.5641	80	80	60	90		142		11.35		0.723
B2W1	B	1.3	0.8858	100	90	85	60		151		11.05		0.086
B2W1	C	0.9	0.5128	80	80	(NA)	60		127		6.67		0.029
B2W1	D	0.5	0.2797			(NA)	(36)		87		6.46		0.009
B2W2	B	0.6	1.8700			(40)	(90)		107		6.55		0.385
B2W2	C	0.2	0.3710						74		4.09	0.014	0.015
B2W2	D	0.6	1.1220	65	80	(42)	(52)		113		4.46		0.138
B2W2	X	0.55	0.7013			(NA)	(38)		61		4.97		0.054
B2W3	A	0.2	0.1690					>100	107	6.5	6.81		0.003
B2W3	B	2.3	0.0845			(NA)	72	49	72	NA	3.54		0.001
B2W3	C	2.0	0.2957	40	60	(NA)	(NA)	36	50	2.9	2.59		0.010
B3W1B	A	2.0	1.1838	70	40	35	55		84		4.8		0.154
B3W1B	B	2.2	0.2537			(30)	(NA)		73		2.45		0.007
B3W1B	C	2.3	1.0570	>100		40	35		99		3.52		0.123
B3W1B	D	1.9	0.9302	>100		50	70		134		9.52		0.095
B3W2	A	0.7	2.8704		~~>100~~				357	NA	12.25		0.906
B3W2	B	0.3	2.2412	~~100~~			~~>100~~		113	NA	8.39		0.553
B3W2	C	2.1	2.3985						201	NA	8.32		0.633
B3W2	D	1.5	1.4155	80	80	60	54	>100	127	4.23	5.06		0.220
B3W2	E	1.3	0.3146	80	80	74	(NA)	>100	213	6.7	12.53		0.011
B3W2	F	0.2	0.2752					>100	160	3.1	6.83		0.008
B3W2	G	1.3	0.4325	80	80	(NA)	(84)	>100	253	NA	5.88		0.021
B3W3	A	0.6	0.9676	50	90	(30)	80	>100	113	4.27	7.65		0.103
B3W3	B	0.8	0.4370	50	40	50	60	>100	120	4.83	5.74	0.022	0.021
B3W3	C	2.9	1.5566	>100		>100	85	>100	127	6.35	8.99		0.267
B3W3	D	0.9	1.1780	>100	60	(34)	(30)	>100	142	4.53	8.93		0.153

NOTE

(1) B2W2-A (considered as an "end burst") and B3W2-H (considered not exist) have been removed

(2) Different amplitudes for B1AW1-B (89%), -C(56%), and -D(>100%) in 1995 conventional angle scan were also reported than the orignals shown in the table.

(3) 94 Multizone scans for billets B1AW2, B1BW1, B1BW2-B, B2W1, B2W2, B3W1-B were not performed as indicated by the shaded areas.

(4) The defects that were not considered in the conventional analysis because they were in uninspected regions are highlighted in yellow.

(5) March 22-23, 2005, meeting at GE-QTC: conventional entries in green parentheses were updated by R. B. Thompson and D. Copley.
(NA) denotes no indication visible on the C-scan images and B1BW2B-A was "re-identified" in 94 conventional by taking into account a 14.5" billet cut-off.

(6) Billet and flaw names bordered in green color denote the 10 cutoffs.

(7) Billet ID and Flaw ID in bold light blue color denote the four additional indicaitons added by D. Copley,
after he further examined the C-scan maps and reported on January 12, 1996.

(8) Entries shaded horizonally are data not reliable and/or in uninspectable regions.

(9) B1BW3CB-MM is the latest indication found in 2006 by re-inspecting that billet.

(10) Entries in parentheses were not a part of the original record but were added in 2005 as a part of re-examination of the digital C-scan records. "NA" means a value was not available, i.e., no indication could be identified. A blank cell means that the digital record for the flaw in question was not examined.

MZ = Multizone inspection
CVL = Conventional longitudinal inspection
CVA = Conventional angle inspection

C.10 REFERENCES.

C-1. Brasche, L., Chiou, C.-P., Thompson, B., Smith, K., Meeker, B., Margetan, F., Panetta, P., Chenail, R., Galli, F., Umbach, J., Raulerson, D., Degtyar, A., Bartos, J., Copley, D., McElligott, B., Howard, P., and Bashyam, M., "Contaminated Billet Study," FAA report DOT/FAA/AR-05/16, September 2005.

C-2. Subteam to the Aerospace Industries Association Rotor Integrity Sub-Committee, "The Development of Anomaly Distributions for Aircraft Engine Titanium Disk Alloys," American Institute of Aeronautics and Astronautics, Inc.

APPENDIX D—EXAMINATION OF THE *Re* ASSUMPTION THAT RESPONSE PROPORTIONAL TO AREA

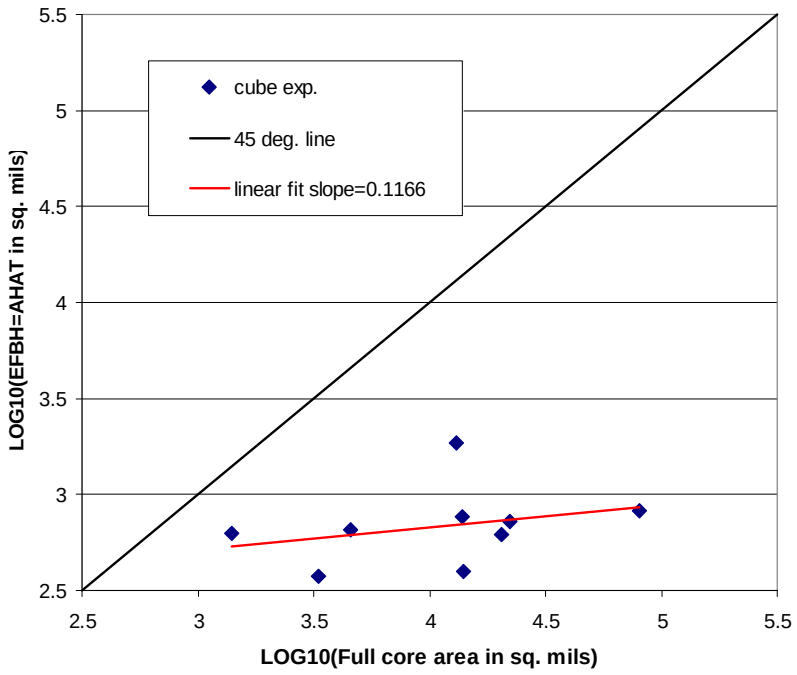
D.1 INTRODUCTION.

In conducting probability of detection (PoD) analyses, one is essentially fitting a regression line through the data relating response to area, properly taking into account such issues as censoring (one may have saturated data or known misses which simply bound the possible response) or truncation (one may not have information from all defects in the population). The question is whether to constrain the slope of the regression line to be unity (effective reflectivity (*Re*) technique), let it be determined by the data ($\hat{a}$ versus a technique), or constrain it to some other value based on independent knowledge of the physics.

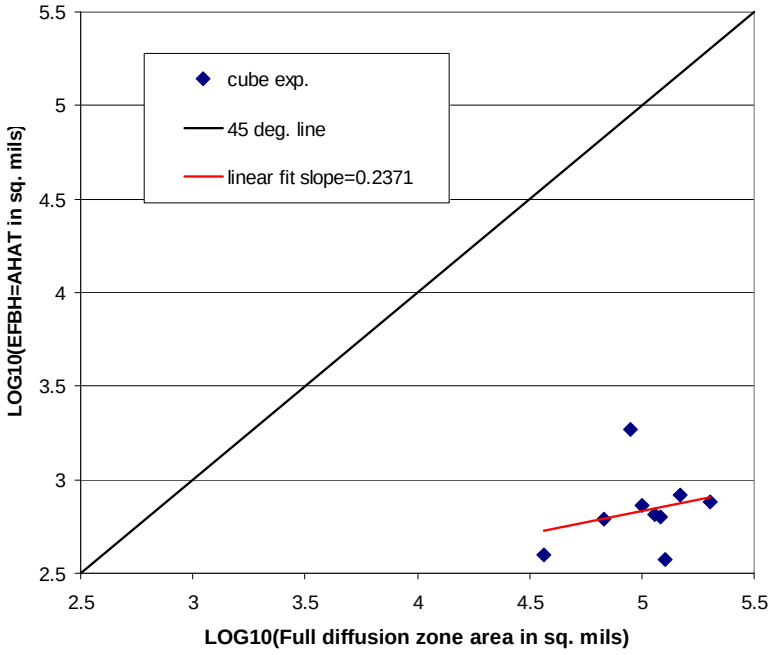
D.2 EXPERIMENTAL CONFIRMATION THAT HARD ALPHA RESPONSE IS NOT PROPORTIONAL TO AREA.

The *Re* technique (slope constrained to be unity) does not seem to be appropriate to these particular data sets. Examination of the data in figures 2, B-1, B-2, B-3, C-30, C-31, and C-32 all suggest that the slope of plots of log $\hat{a}$ versus log a is much less than unity.

Figure D-1 shows further experimental confirmation that the slope is less than unity for the subset of the Contaminated Billet Study (CBS) defects for which the area is known from detailed metallography (CBS_{10}) as examined with a spherically focused beam in small laboratory samples that contained the defects (a step in the sectioning process). The regression indicates a slope of 0.12 when plotted versus core area and 0.24 when plotted versus diffusion zone area.



(a)



(b)

Figure D-1. Log Plots of Equivalent Flat-Bottom Hole Area Versus Metallurgy Area as Measured With Focused Beam in Laboratory Cube Geometry (a) Versus Diffusion Zone Area and (b) Versus Core Area

D.3 INTERPRETATION BASED ON EXPERIMENTAL AND THEORETICAL RESPONSE OF REFLECTORS OF SIMPLE SHAPE.

To better understand the slope, the theoretical literature for the ultrasonic response of reflectors of simple shape was examined [D-1 and D-2]. Table D-1 summarizes the results [D-3]. In this table, "r" is taken to be a radius of a crack, pore, or truncated cylinder and "L" the length of the truncated cylinder. It can be seen that if one assumes a flat crack, the signal would be expected to increase directly proportional to area. However, this is the only shape for which a unity slope is obtained. For example, for a spherical shape, the signal would increase as the radius, or the square root of area. This is equivalent to a slope of 0.5 in the log-log plots of $\hat{a}$ versus a.

Table D-1. Theoretical Expectations of Dependence of Signal on Size for Simple Targets*

Reflector	Far Field, Short Wavelength Response	N
Flat crack	r^2	1
Spherical pore	r	0.5
Truncated cylinder	$Lr^{1/2}$	0.75
Random distribution of M pores of constant size	$M^{1/2}$	0.5
Tilted flat crack	Complex function	Undefined

*It is assumed that the signal is proportional to $(\text{area})^N$, with N tabulated in the final column

For the case of a truncated cylinder, the situation is more complicated. The signal is proportional to $Lr^{1/2}$, but cross-sectional area, as seen by the ultrasonic beam, is Lr. Simple arithmetic shows that the signal is proportional to $(L/r)^{1/4}$ $(Lr)^{3/4}$, with the former factor being the aspect ratio raised to the 1/4 power and the latter factor being the area raised to the 3/4 power. Hence, for a fixed aspect ratio truncated cylinder, the signal is proportional to $(\text{area})^{0.75}$, i.e., N = 0.75. However, if the aspect ratio changes, the regression line will translate.

More complex cases can also be considered. Consider the pore structure of a hard alpha inclusion to consist of pores of a fixed size, randomly positioned about a line segment and randomly displaced from that line so that the signals from the individual pores add incoherently. In this model, an increase in hard alpha size corresponds to an increase in the number of pores, and it can again be shown that N = 0.5. If the pores also change size as the size of the overall hard alpha grows, N would have yet a different value. Even greater complications arise if one thinks of the response of the pore structure of a hard alpha inclusion in terms of that of a tilted flat crack.

Based on these arguments, it is clear that there is no theoretical justification for constraining the slope of plots of log $\hat{a}$ versus log a to be unity. Given the importance of this conclusion to the current and future analyses of PoD, experimental confirmation of these theoretical arguments was sought.

Figure D-2 shows the results of this confirmation based on data available to the Engine Titanium Consortium (ETC). Each of the original equipment manufacturer partners in the ETC contributed internal data of response versus area. Included were data for flat-bottom holes (FBH) (Pratt & Whitney), spherical pore and tungsten carbide inclusions (Honeywell and General Electric), and spherical voids (General Electric). For the FBH data provided by Pratt and Whitney, the slope was 1.061. For the spherical pore and tungsten carbide inclusion data provided by GE, the slopes were 0.601 and 0.475, respectively. For the tungsten carbide inclusion data provided by Honeywell, the slope was 0.462. These data are in excellent agreement with the theoretical expectations of 1.0 and 0.5, respectively, validating the arguments summarized in the first two rows of table D-1. Experiments reported in the literature have also validated the model that leads to the third row in table D-1. For the case of cavities in the form of side-drilled holes, published results of benchmark experiments have provided validation [D-4-D-8]. Work by the ETC has produced a limited validation for the case of a truncated cylindrical synthetic hard alpha inclusion [D-9].

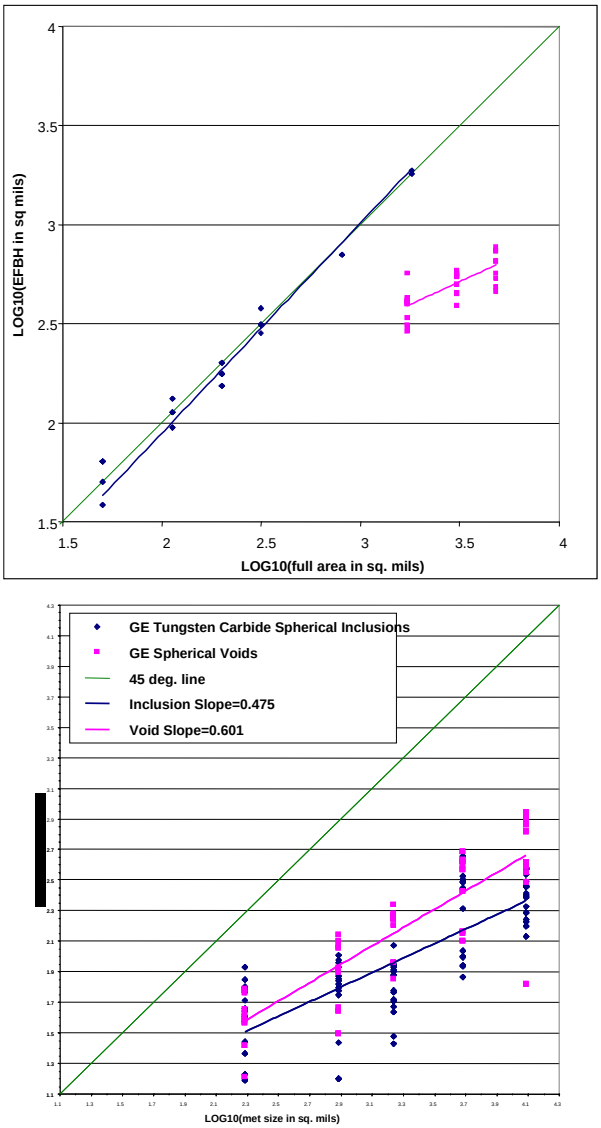


Figure D-2. Experimental Slopes of Responses of FBHs and Spherical Reflectors (Log $\hat{a}$ Versus Log a Plots)

D.4 INTERPRETATION BASED ON THEORETICAL PREDICTIONS OF HARD ALPHA RESPONSE.

For the CBS_{10}, three-dimensional reconstructions of the defect surfaces were made and used to drive theoretical descriptions of the response. This capability was used as a tool to estimate what the slope of response various area would be (in log-log space) for defects with the complex geometry of hard alpha inclusions. Figure D-3 shows the results. In these simulations, the defect was assumed to lie in a homogeneous material and was examined with a spherically focused transducer. Simulations were conducted for four different, spherically focused beam sizes ranging from a small beam equivalent to the beam produced in a billet by a Multizone probe (probe 1) to a much larger beam such as one produced by a conventional inspection (probe 4). Results for a fifth, planar probe, are also included for comparison. For these five

cases, the slopes of plots of log $\hat{a}$ versus log a vary from 0.2783 to 0.4509. It should be emphasized, however, that these calculations are missing two sources of variability that exist in practice, the influence of microstructure on the signals and the effects of run-to-run inspection variability due to factors such as setup, calibration, transducer, surface, and noise. It is not known whether these effects would change the slope for similar plots of field data. It seems safe to say that the slope is no more than 0.5, and probably less. These simulation results also clearly indicate the trend that the slope increases with larger beam diameter (i.e., broader and more uniform beam illumination over the flaw surface).

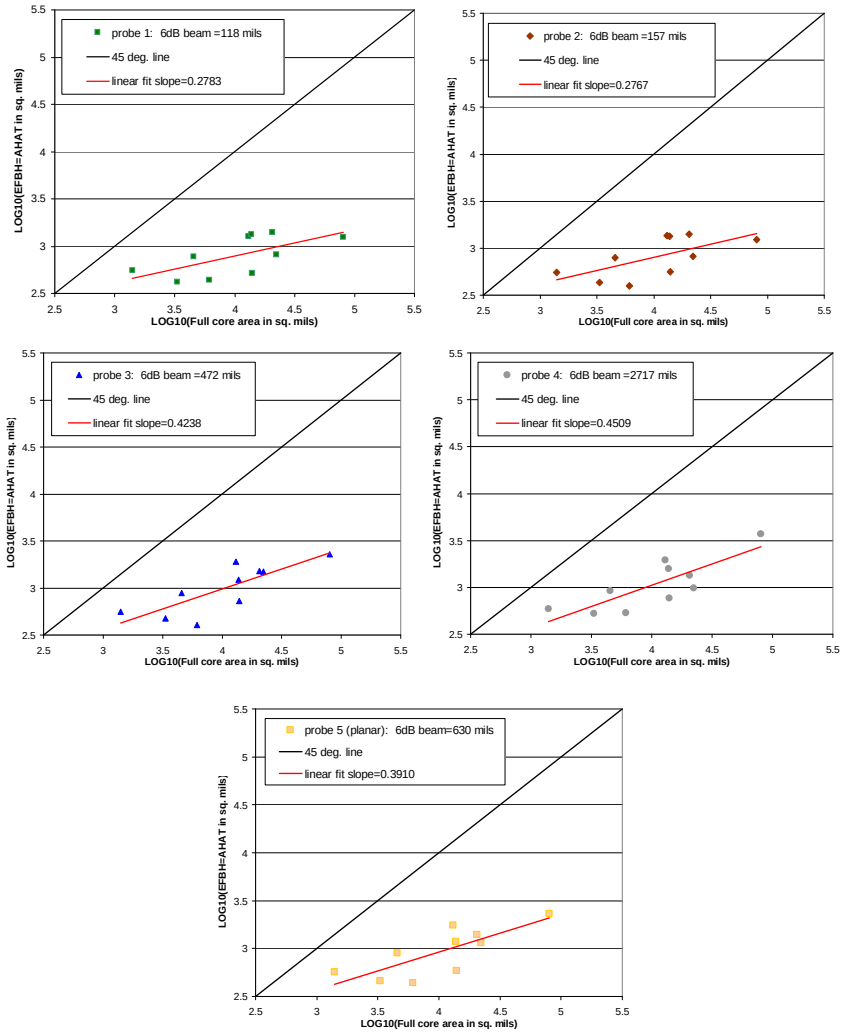


Figure D-3. Simulation of $\hat{a}$ Versus a for the CBS_{10}, as Illuminated by Spherically Focused Beams of Various Diameters

D-5 REFERENCES.

D-1. Schmerr, L.W., *Fundamentals of Ultrasonic Nondestructive Evaluation—A Modeling Approach*, Plenum Press, NY, 1968.

D-2. Schmerr, L.W. and Song, S.-J., *Ultrasonic Nondestructive Evaluation Systems— Models and Measurements*, Springer, Berlin (in press).

D-3. Thompson, R.B., "Ultrasonic Benchmarking: Past Progress and Future Goals," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 23B, D.O. Thompson and D.E. Chimenti, eds., AIP, NY, 2004, pp. 1529-1536.

D-4. Diligent, O., Chatillon, S., Darmon, M., Lhemery, A., and Mahaut, S., "Results of the 2004 UT Modeling Benchmark Obtained with the CIVA Software Developed at the CEA," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 24, D.O. Thompson and D.E. Chimenti, eds., AIP, NY, 2005, pp. 1843-1850.

D-5. Spies, M., "Prediction of Ultrasonic Flaw Signals and Model-to-Experiment Comparison," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 24, D.O. Thompson and D.E. Chimenti, eds., AIP, NY, 2005, pp. 1851-1858.

D-6. Bostrom, A. "Ultrasonic Benchmarking With UTDefect," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 24, D.O. Thompson and D.E. Chimenti, eds., AIP, NY, 2005, pp. 1859-1863.

D-7. Krishnamurthy, C.V., Shankar, M., Balasubramaniam, K., "Patch Element Model for the Evaluation of Displacement Fields Within an Elastic Solid From a Non-Contact Immersion Transducer: Application to the 2004 Ultrasonic Benchmark Problem," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 24, D.O. Thompson and D.E. Chimenti, eds., AIP, NY, 2005, pp. 1864-1871.

D-8. Song, S.-J., Park, J.-S., Choi, Y.-H., Kang, S.-C., and Kim, K.-J., "Model Predictions to the 2004 Ultrasonic Benchmark Problems," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 24, D.O. Thompson and D.E. Chimenti, eds., AIP, NY, 2005, pp. 1872-1879.

D-9. Burkel, R.H., Chiou, C.-P., Keyes, T., Meeker, W.Q., Rose, J.H., Sturges, D.J., Thompson, R.B., and Tucker, W.T., "A Methodology for Assessment of the Capability of Inspection Systems for Detection of Subsurface Flaws in Aircraft Turbine Engine Components," FAA report DOT/FAA/AR-01/96, September 2002.

APPENDIX E—STUDIES OF EFFECTIVENESS OF PoD ESTIMATION PROCEDURES ON SIMULATED DATA SETS

E.1 STRATEGY.

Early attempts to analyze the various data sets discussed in section 1.2.1, 3.2, appendix B, and appendix C to produce probability of detection (PoD) results yielded many confusing results, including regression lines with negative slopes. Dramatically different results were obtained as the data sets and/or analysis procedures changed. To provide guidance with respect to how best to proceed with the statistical analysis, simulation studies were conducted.

The idea was to start with a synthetic data set consistent with a known PoD and then see how accurately that known PoD could be reproduced by analyzing the synthetic data in various ways. The strategy was to mimic the current data sets to as great an extent as was possible, including the presence of misses, which could be treated as known or unknown. For example, in some of the Contaminated Billet Study (CBS) data sets, both conventional and Multizone inspections were performed on the same defects. If there was a miss in a conventional inspection and a corresponding find in a Multizone inspection, the miss became known, and this information could be used in the analysis of the conventional inspection data. However, had a Multizone inspection not been done, there would have been no knowledge of conventional misses, as is the case for the Default PoD-3Dimensional and Jet Engine Titanium Quality Committee databases. The way that these two situations are handled statistically is significantly different, as is illustrated by the simulations reported in this appendix.

More specifically, a simulation tool was developed to generate synthetic data consisting of flaw sizes and the corresponding responses, leading to synthetic Multizone and conventional inspection data (corresponding to assumed $\hat{a}$ versus a relationships for each inspection type). The purpose was to experiment with different statistical models and to evaluate different estimation procedures as a guide to reconciling some of the ambiguities that were encountered in the analysis of the real inspection data being used to generate updated PoD curves. It is helpful to know the "ground truth" when comparing different methods of analysis.

One important caution should be noted. In simulations such as these, the simulated data is generated based on the same statistical model used in the regression. Hence, the data is consistent with the analysis tools. However, it is not necessarily the case that real data follow the statistical model used in the analysis. An example in this work was the presence of Type II (atypical) misses in the analysis of the conventional inspection data. Tests are available to assess the consistency of data with a given statistical model, but the results are not always fully conclusive. Hence, the results in this appendix only provide guidance regarding what would occur under optimal conditions. As noted above, the real data is not optimal.

E.2 CREATION OF SIMULATED INITIAL FLAW DISTRIBUTION CONSISTENT WITH CURRENT EXCEEDANCE CURVES.

The analysis started by constructing a hypothetical distribution of flaw sizes. This was generated to be roughly consistent with the exceedence curve that was developed by the Aerospace

Industries Association Rotor Integrity Sub-Committee team [E-1]. Figure E-1 shows the simulated distribution. To avoid an infinite number of very small flaws, this distribution was truncated at a lower limit of 200 square mils, a size that would be too small to contribute to PoD curves.

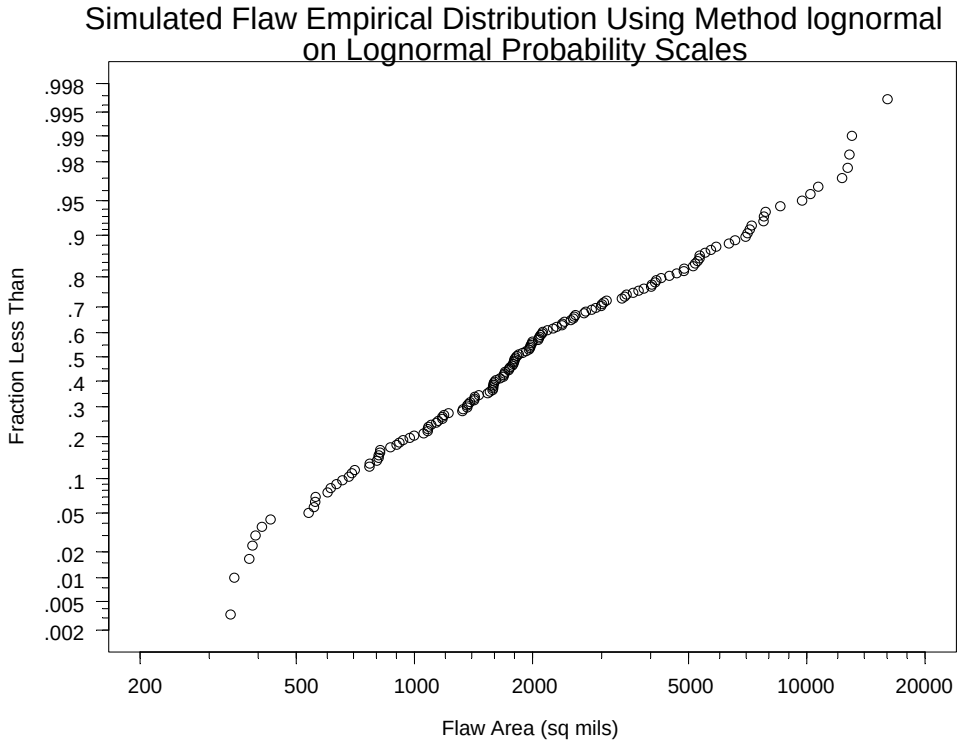


Figure E-1. Distribution of Flaw Sizes Assumed in the Simulation

Figures E-2 and E-3 simulate the results of Multizone and conventional inspections based on this postulated distribution of flaws sizes. It is assumed that the Multizone flaw response data (figure E-2) are governed by the following relation, with the flaw and inspection variability taken to be equal.

$$\log(\hat{a})=3 + .5\log(a) + FlawError + InspectionError \qquad \text{(E-1)}$$

Thus, one assumes that the "true" slope of the regression curve, from which PoD is derived, is 0.5. The Multizone values of percent screen height are based on an assumed calibration to a #2 flat-bottom hole (FBH). The conventional inspection data (figure E-3) were generated in an analogous way, assuming calibration to a #3 FBH. Hence, the points tend to be lower on a percent screen height scale. In agreement with expectations, there is a significantly greater number of defects in the Multizone data set that will be detected than there are in the conventional inspection data set. As will be seen, this has an impact on the uncertainty in the PoD analysis as well as the estimated PoD itself.

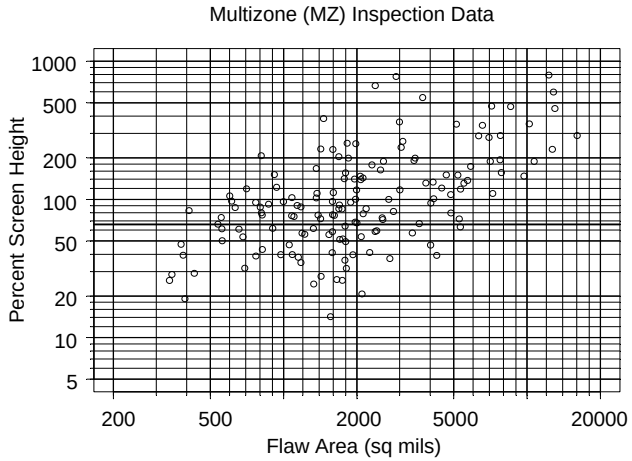


Figure E-2. Synthetic Multizone Data

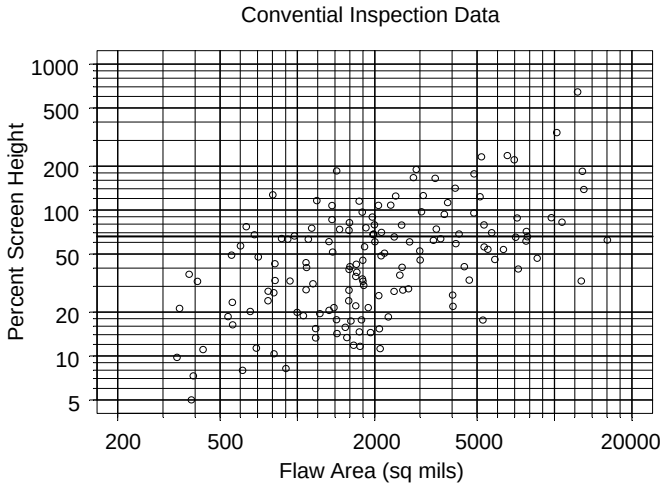


Figure E-3. Synthetic Conventional Data

E.3 SIMPLE REGRESSION ANALYSIS ASSUMING THAT ALL DATA IS AVAILABLE.

Figures E-4, E-5, E-6, and E-7 show the results of the PoD analysis of the above Multizone data when it is assumed that all data is recovered in the inspection, i.e., the signals from all flaws were captured, irrespective of the strengths of those signals. Figure E-4 is a plot of log equivalent flat-bottom hole area versus log area. The data tend to fall below the 45 degree line (unity slope) since the slope built into equation E-1 is 0.5. The result of a regression analysis of this data is shown in figure E-5. Since all the data are available, no truncation is assumed in the analysis. It shows that the slope estimated by the regression is 0.537, not far from the true value of 0.5. Figures E-6 and E-7 show the PoD curves corresponding to a #2 and #3 FBH, respectively.

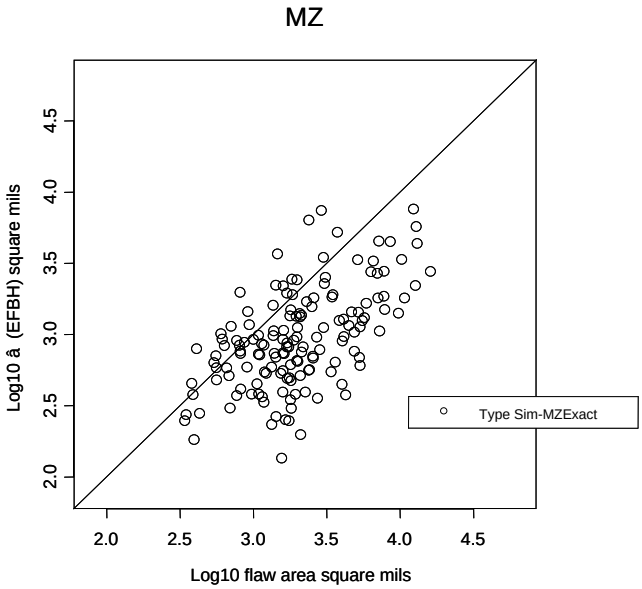


Figure E-4. Synthetic Multizone Data (Assuming perfect inspection with no misses)

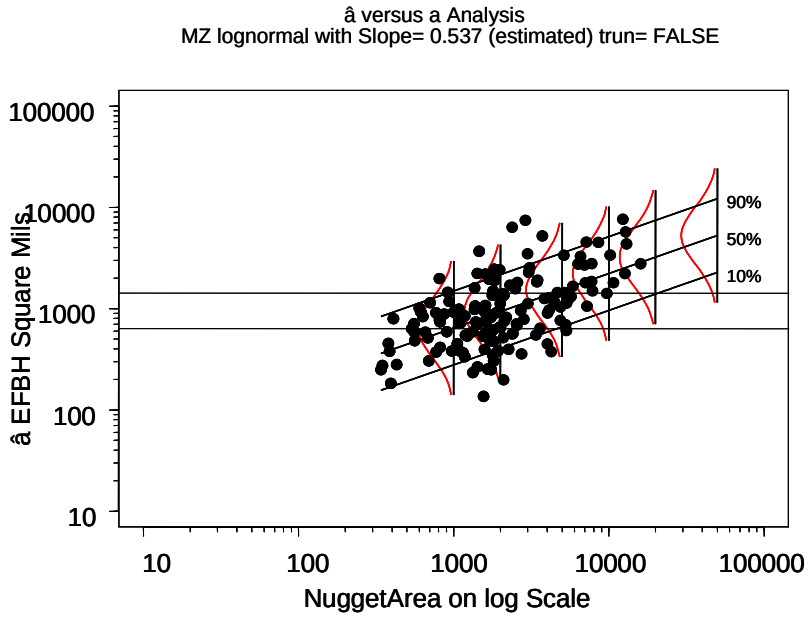


Figure E-5. Result of Regression Analysis of Full Multizone Data. (The horizontal lines are detection thresholds for #2 (lower) and #3 (upper) FBH. The PoD is the area in the densities above the thresholds. Plots of PoD follow for this example)

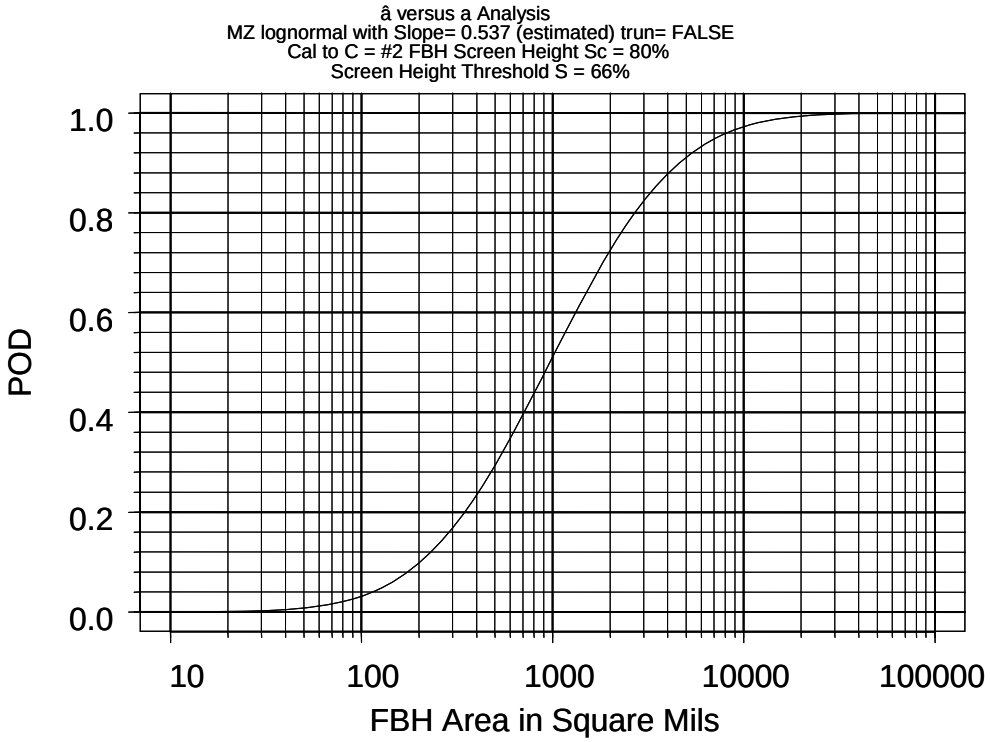


Figure E-6. Resulting PoD Curve Using $\hat{a}$ Versus *a* Method With Calibration to #2 FBH

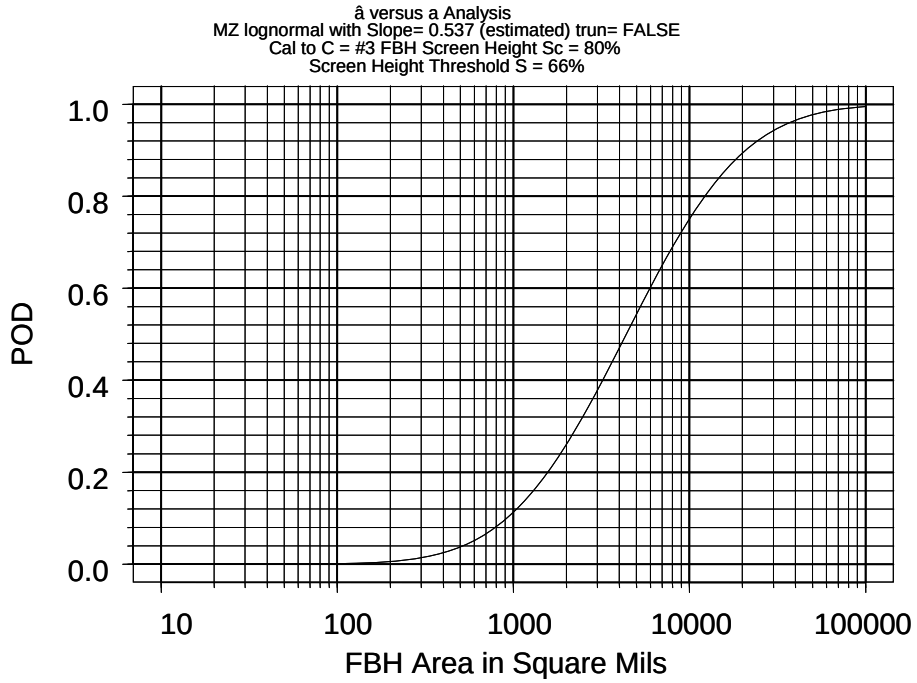


Figure E-7. Resulting PoD Curve Using $\hat{a}$ Versus *a* Method With Calibration to #3 FBH

E.4 REGRESSION ANALYSIS WITH KNOWN MISSES BELOW THRESHOLD.

The analysis was repeated assuming that there was a known number of misses. This situation occurs when there is an independent reference technique. In the real world, there is no such referee technique for Multizone, but Multizone data provides an approximate referee for conventional inspections. Figure E-8 shows the synthetic data set with the misses plotted at an assumed noise level. The flaw responses that were below threshold are treated as being left censored at the noise level. In other words, the statistical analysis is based on the knowledge that a flaw is present, but its signal strength is equal to or less than the indicated noise level. To see the effects of truncation, compare figure E-4 with figure E-8.

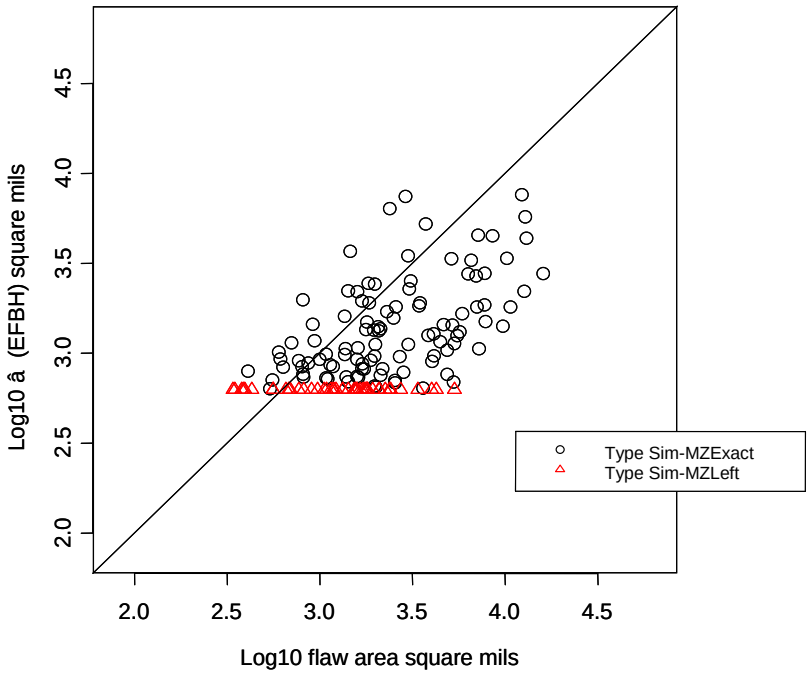


Figure E-8. Left Censored Multizone Data, Which Would Arise if One Knew About the Flaws Below Threshold (#2 FBH calibration used here for the Multizone inspection)

Figure E-9 shows the result of a regression analysis of this data. The estimated slope is 0.562 (in comparison with 0.537 for the complete data and the truth of 0.5). With left censoring, there is some, but not much, loss of information. The PoD curves would be similar to the complete data example shown in figures E-6 and E-7 (because the fitted model is similar).

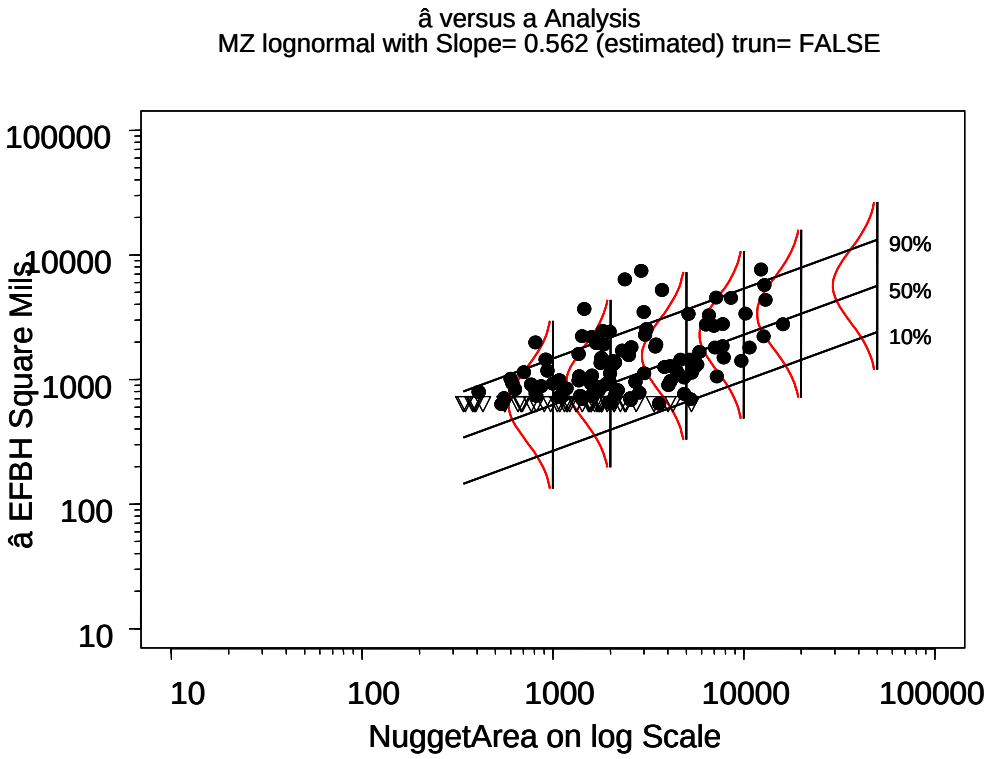


Figure E-9. Regression Analysis $\hat{a}$ Versus a Data Using the Left-Censored Multizone Data (Assuming the existence is known, but not the signal response, for all flaws below threshold)

E.5 REGRESSION ANALYSIS WITH TRUNCATION.

It was next assumed that the misses were not known. Figure E-10 shows the truncated Multizone data. In this case, mimicking an actual field inspection without any reference technique, the number of misses is not known, although it is acknowledged that they may exist. Note that in comparison with figure E-4, the truncated data gives an impression of little or no slope (as shown in real data). The fitted model in figure E-11 (based on a truncation model that adjusts for the unknown misses) obtains an estimated slope of 0.742, a significant deviation from the true value of 0.5. This corresponds to the analysis of field Multizone data with no reference technique available. This raised the possibility that a similar deviation may occur in the analysis of the production data; the best estimate will be made of the PoD, but the extent to which this deviated from the real production performance cannot be quantified.

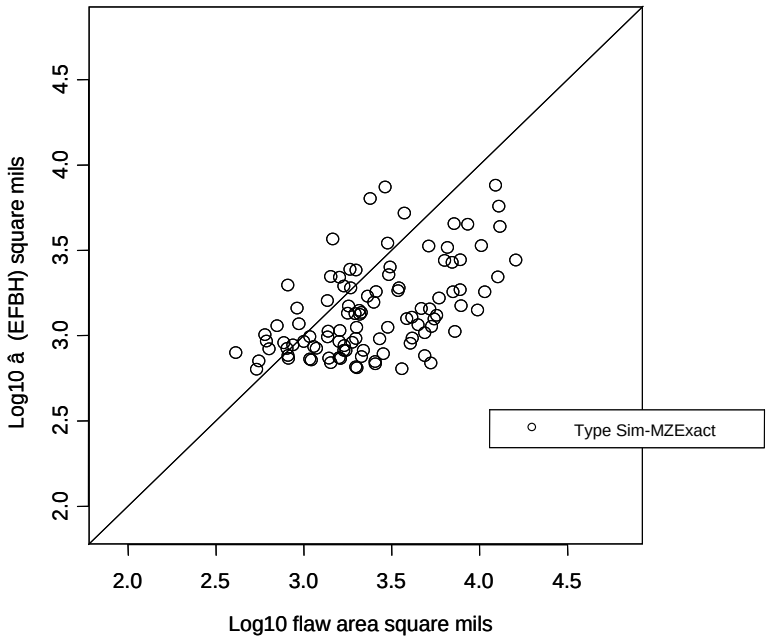


Figure E-10. Truncated Multizone Data (Compare with figures E-4 and E-8)

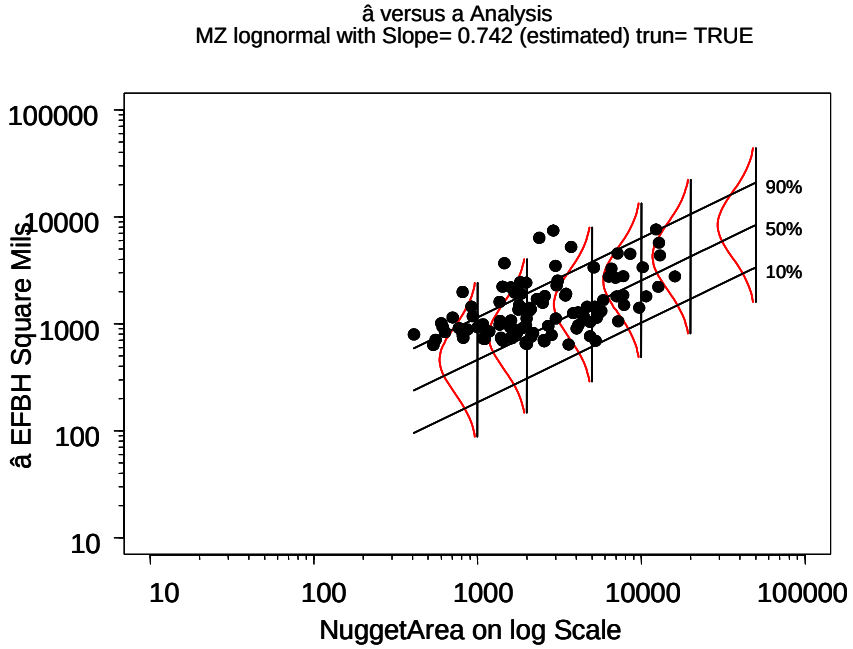


Figure E-11. Truncation Model Fitted to the Truncated $\hat{a}$ Versus a Data

The above analyses were conducted based on the same initial data set, as shown in figure E-4, before any censoring or truncation. To test the sensitivity of the results to sampling, the simulation of inspection errors and analysis was repeated 50 times for each of the test cases.

Figure E-12 shows the regression results of repeating the simple analysis in which all of the data is recovered. This is a summary of 50 data-generation/estimation simulations using various exact Multizone data (assuming no misses) with the simple (fit a line to the data) estimation method, showing the 50% line for each trial. Note that the estimate is approximately unbiased (gives the correct answer on the average) and that the sampling variability is small.

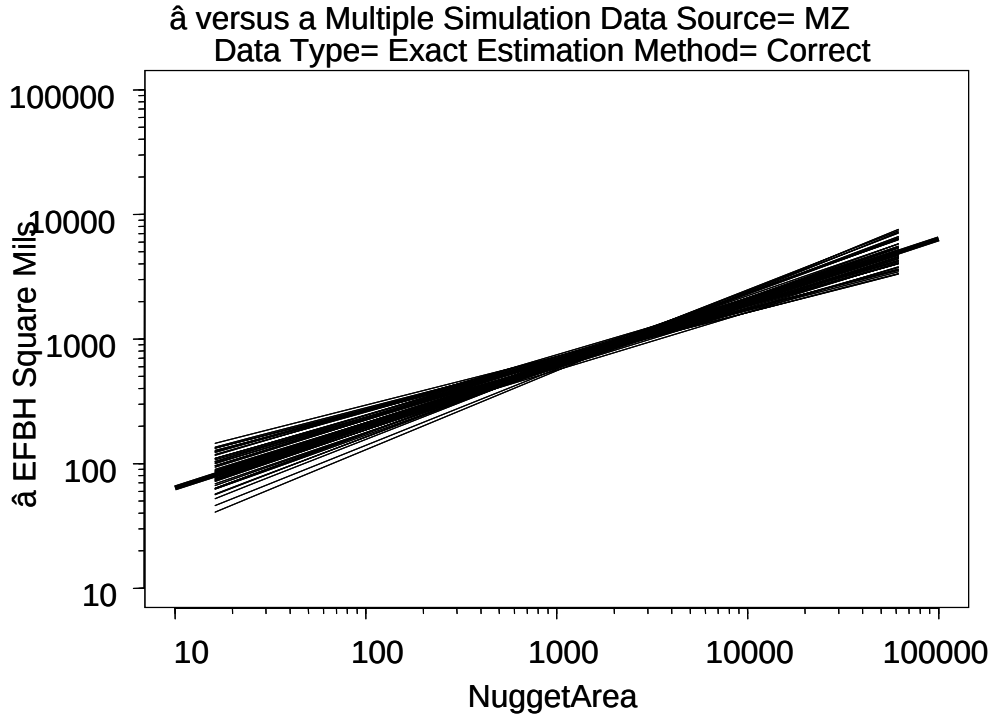


Figure E-12. Sensitivity of Simple Regression Analysis to Sampling (Multizone)

Figure E-13 shows the corresponding regression results for the censored case (known misses). This is a summary of 50 data-generation/estimation simulations using the censored Multizone data with the appropriate censored data estimation method, showing the 50% line for each trial. The longer darker line is the 50% line for the truth. Note that the estimate is again approximately unbiased (gives the correct answer on the average).

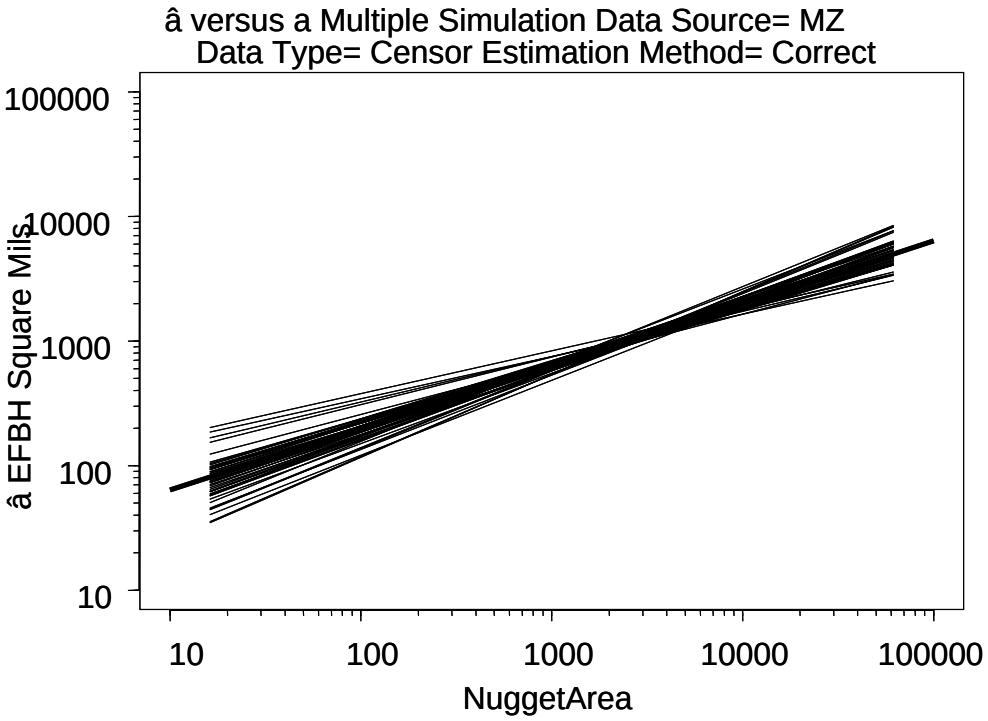


Figure E-13. Sensitivity of Censored Regression Analysis to Sampling (Multizone)

Figure E-14 shows the corresponding regression results for the truncated data (unknown misses). This is a summary of 50 data-generation/estimation simulations using the truncated Multizone data with the appropriate truncation estimation method, showing the 50% line for each trial. The longer darker line is the 50% line for the truth. Note that the estimate is still approximately unbiased. However, there is a large amount of sampling variability arising from the lack of knowledge about the possible number of misses compared with the previous graphs that show the sampling variability if there were no misses (exact data) and if one knew how many misses and their sizes (censored data). This corresponds to the analysis of field Multizone data with no referee technique available.

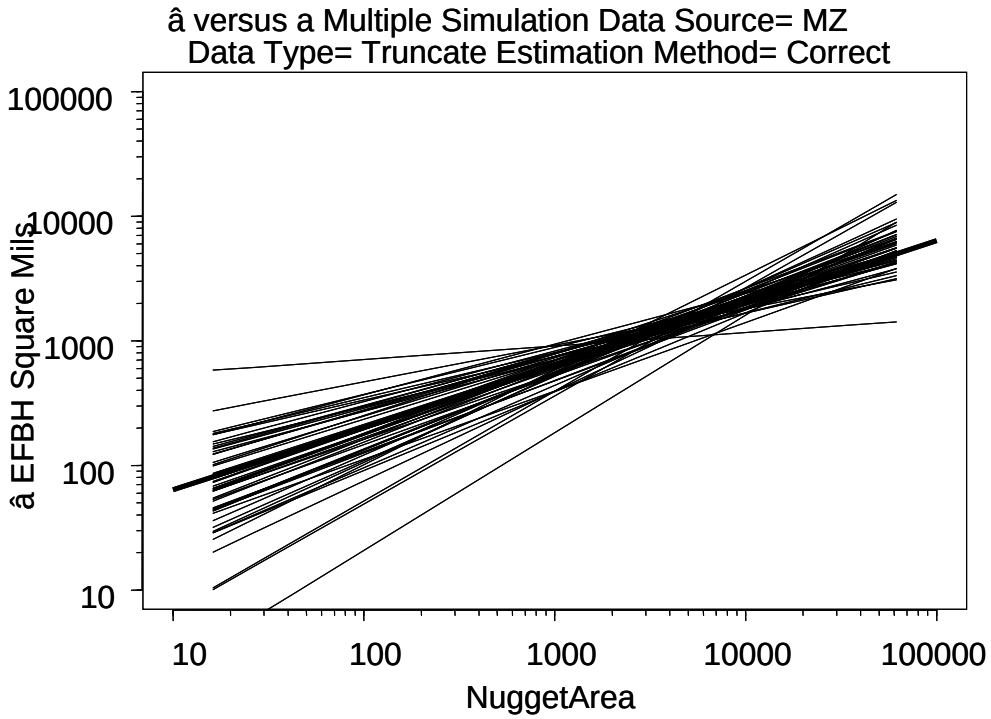


Figure E-14. Sensitivity of Truncated Regression Analysis to Sampling (Multizone)

The importance of properly accounting for truncation in the analysis is shown in figure E-15. This is a summary of 50 data-generation/estimation simulations using the truncated Multizone data, as in the previous case, showing the 50% line for each trial. However, now the simple estimation method is used, in which truncation is not taken into account in the regression. Note that the procedure is seriously biased (relative to the truth). This demonstrates why it is not possible to just fit a line to the data.

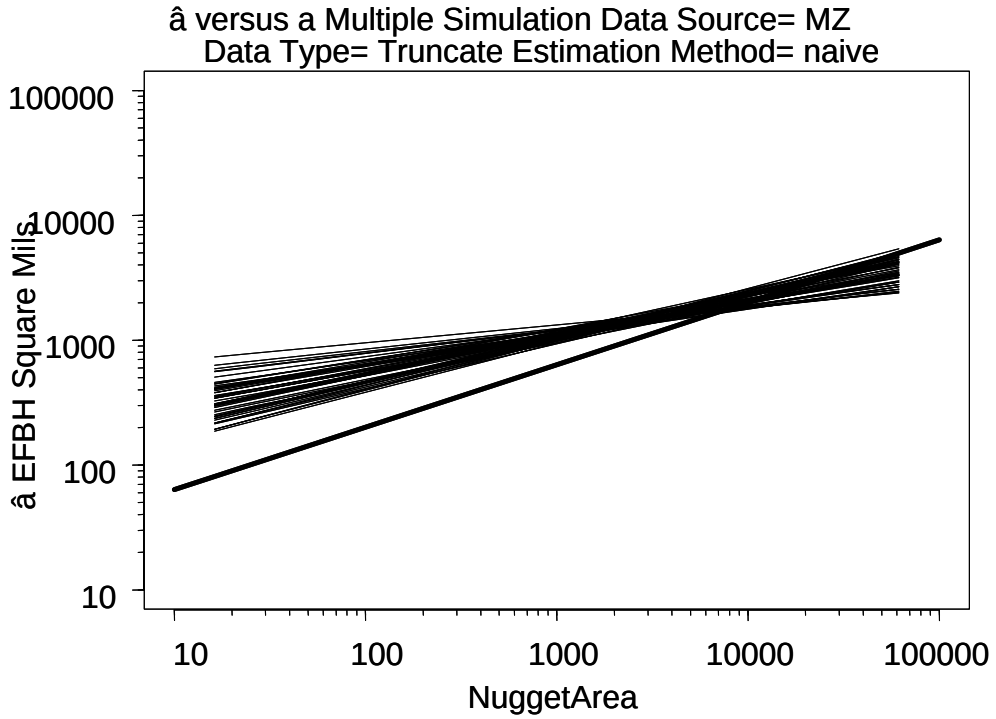


Figure E-15. Errors in Fitting Simple Regression Line to Truncated Data (Multizone)

E.6 EXAMINATIONS OF UNCERTAINTIES.

The above calculations were repeated with emphasis on the uncertainties in PoD that are produced by the statistical uncertainty in the regression line, as affected by the various analysis techniques. Figure E-16 shows the results of a different realization of the same simulation that led to figure E-14, in which a regression line is fitted to simulated Multizone data under the assumption of truncation. Again, the assumed truth is the longer, thicker line. The other lines are estimates that would have arisen in the analysis of a Multizone inspection like that performed in the CBS, based on different samples of the distribution.

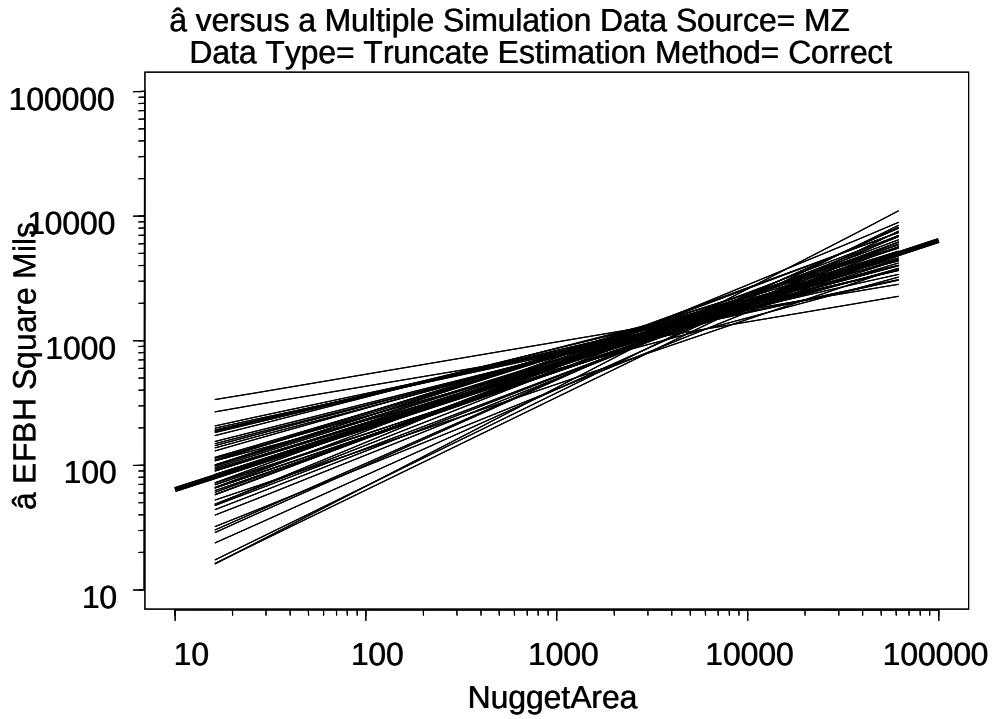


Figure E-16. Sensitivity of Truncated Regression Analysis to Sampling (Multizone)

Figure E-17 shows the corresponding PoD curves. This demonstrates that there is a significant amount of statistical uncertainty in the determination of the PoD for this Multizone case.

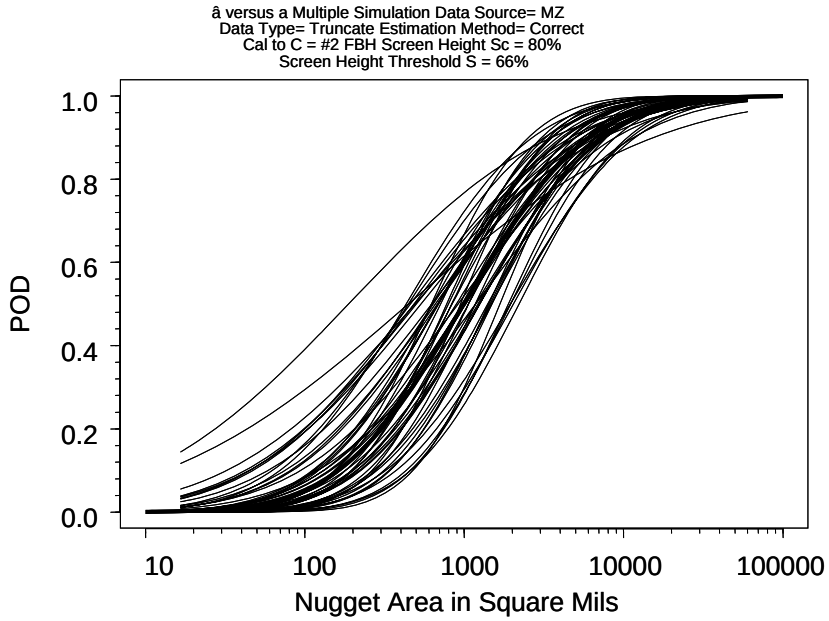


Figure E-17. Sensitivity of PoD to Sampling for Truncated Regression Analysis of Multizone Data

Figure E-18 shows the result for a simulated conventional inspection, illustrating the large amount of sampling error that would arise. A regression line is fitted to simulated conventional inspection data under the assumption of truncation. The statistical variability is larger than for the Multizone case because of the larger number of misses (i.e., there is less data available from a conventional inspection because of the greater sensitivity of the Multizone inspection). In these particular simulations, out of a population of 150 flaws, there were, on average, 42 conventional finds compared to 98 finds, on average, in the Multizone simulations.

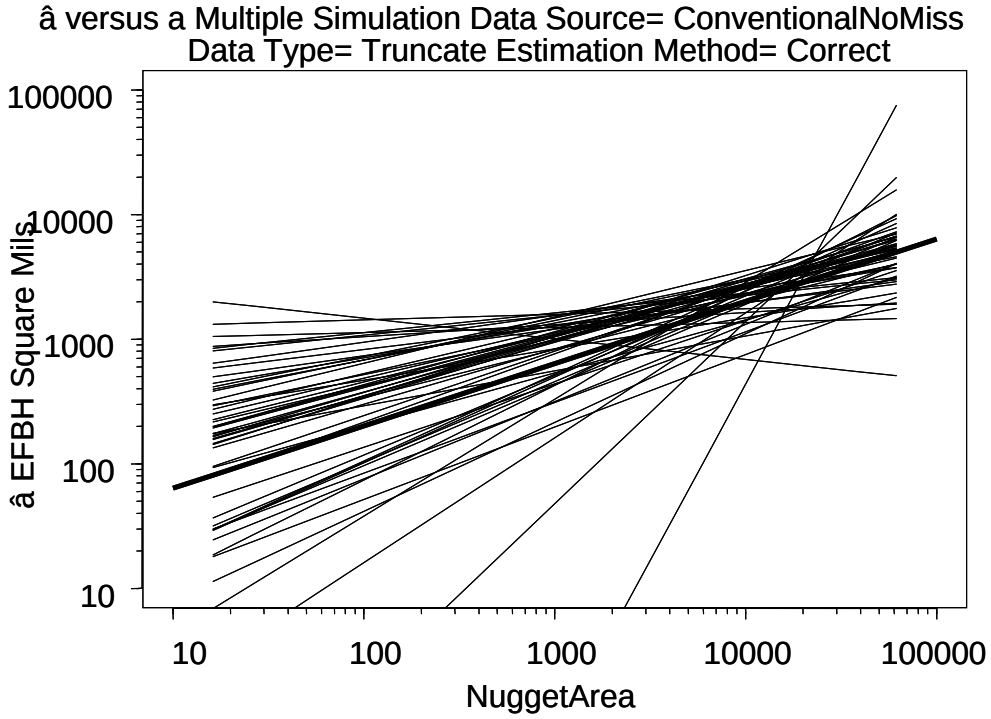


Figure E-18. Sensitivity of Truncated Regression Analysis to Sampling (Conventional Inspection)

Figure E-19 shows the corresponding PoD curves. This demonstrates that there is an enormous amount of statistical uncertainty in the estimated PoD of the conventional inspection, due to the smaller number of amplitudes that were recorded.

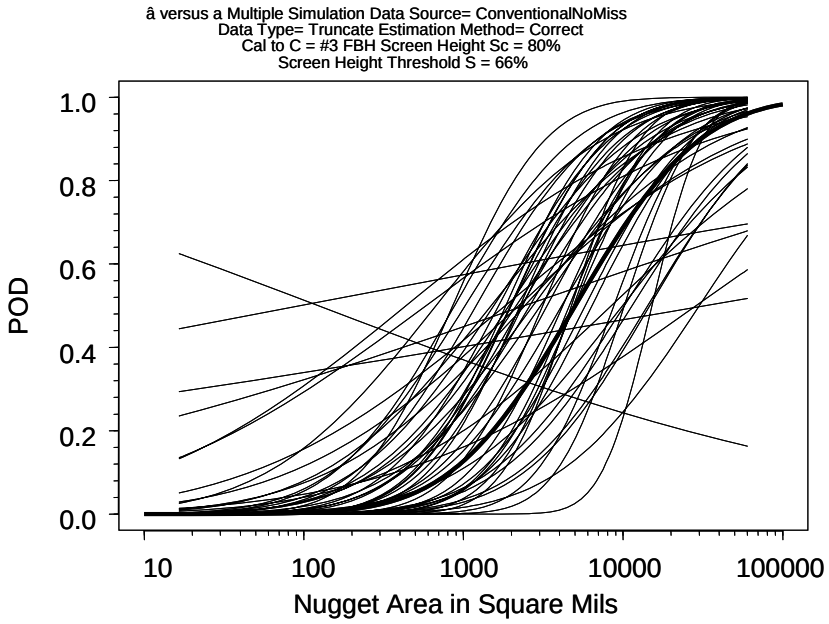


Figure E-19. Sensitivity of PoD to Sampling for Truncated Regression Analysis of Conventional Data

Figure E-20 shows how the knowledge of misses will stabilize the conventional analysis. These are estimates based on the conventional data, augmented with the information about the number of misses (treated as left censored observations) that were discovered through the Multizone inspection. The additional information improved the estimation considerably.

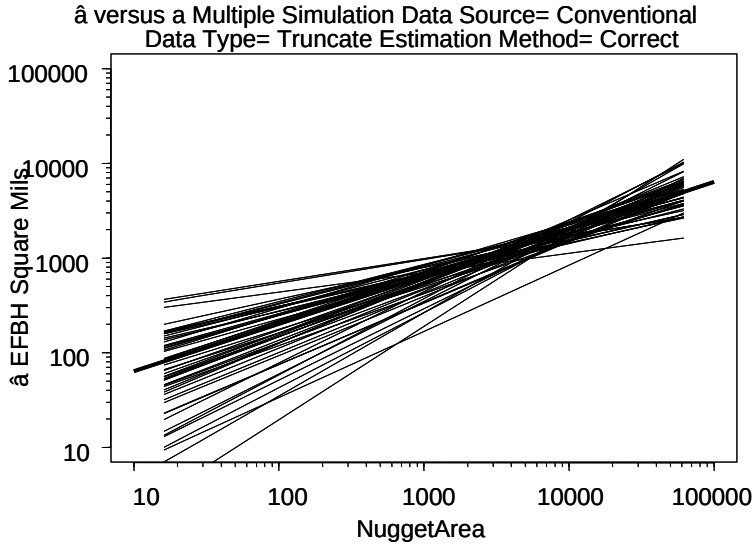


Figure E-20. Sensitivity of Truncated Regression Analysis With Known Misses to Sampling (Conventional)

Figure E-21 shows the corresponding PoD curves. The uncertainty is greatly reduced by the knowledge of misses.

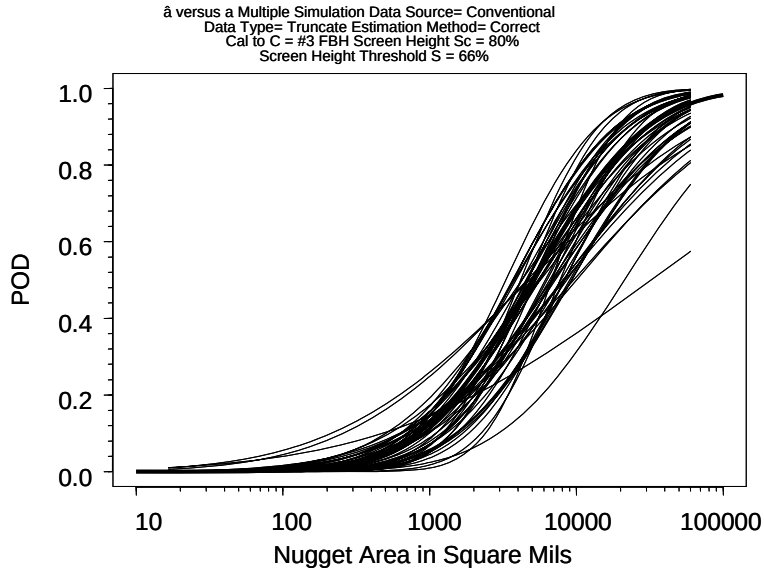


Figure E-21. Sensitivity of PoD to Sampling for Truncated Regression Analysis of Conventional Data With Knowledge of Misses

Figure E-22 shows the PoD that would result by treating the same data set (conventional with known misses) using a hit or miss analysis method (logistic model). For larger flaws, there is more variability in the estimates because information about actual amplitude of the signal is not being used in the analysis. For smaller flaws, there is evidence of bias in the estimate, arising because of truncation (unknown misses also missed by Multizone) being ignored.

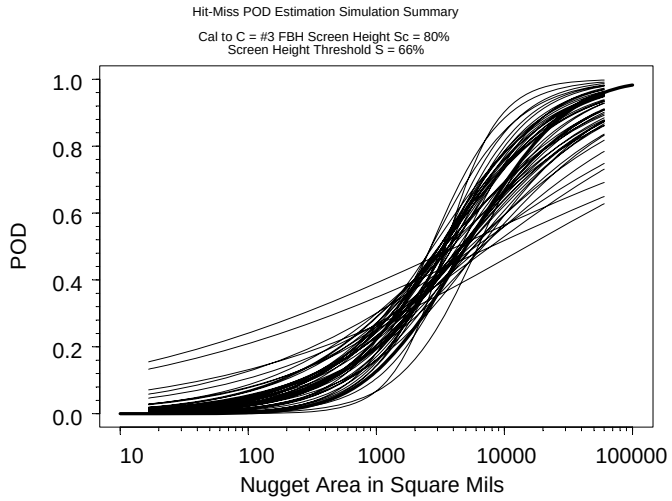


Figure E-22. Sensitivity of PoD to Sampling for Hit or Miss Analysis of Conventional Data With Misses

E.7 CONCLUSIONS OF SIMULATION STUDIES.

Major conclusions from the studies of the simulated data sets include the following major points:

- The needed truncation analysis can be unstable (large sampling variability leading to a large uncertainty), particularly for small data sets.
- If the slope in the estimation is known (as was assumed in the *Re* analysis), the uncertainty could be reduced.
- However, the answer obtained will be sensitive to the assumed value of the slope.
- Simply fitting a line to the data to determine the slope (without accounting for truncation) can lead to significantly biased results.

It is emphasized again that, in this simulation, the model used in the analysis of the synthetic set was the same as the one used to generate the data. Further complications arise when the actual data does not perfectly fit the assumed model used to guide the regression, a case that was encountered in the analysis of the conventional inspection data from naturally occurring hard alpha inclusions.

E-8 REFERENCES.

E-1. Subteam to the Aerospace Industries Association Rotor Integrity Sub-Committee, "The Development of Anomaly Distributions for Aircraft Engine Titanium Disk Alloys," American Institute of Aeronautics and Astronautics, Inc.

APPENDIX F—LIMITATION ON POSSIBLE RESPONSE LEVELS FOR HARD ALPHA INCLUSIONS BASED ON THE PHYSICS OF SCATTERING

F.1 MOTIVATION.

In the $\hat{a}$ versus a approach to determining Probability of Detection (PoD), plots of the log flaw response ($\hat{a}$) versus log flaw area (a) are developed based on experimental data. Regression techniques are then used to fit a straight line to the data and define the variance of the data about this line (equation 4 of this report). For a given threshold to which the response is compared, the PoD is then given by the fraction of the distribution above the threshold for the flaw size in question.

An explicit assumption of this approach is that the data can be fit, in log-log space, to a single straight line of slope β_1. This is equivalent to the assumption that

$$\text{Flaw Response } \alpha \text{ (Flaw Area)}^{\beta_1} \tag{F-1}$$

This is often true, at least to a first approximation, over a limited range of flaw sizes. However, the physics of the scattering of ultrasound from flaws indicates that equation F-1 cannot be a general result for all flaw sizes for a fixed value of β_1. Table F-1 shows the values of β_1 for two types of simple defects, a circular crack (or flat-bottom hole) insonified perpendicular to its face and a spherical void. Here, r is the radius of the crack or void and λ is the ultrasonic wavelength, and it is assumed that the beam is sufficiently large to capture the full scattering from the crack or void. Since the area is proportional to r^2, it is straightforward to calculate β_1, as shown. Experimental data in support of the predictions when $r/\lambda >> 1$ have already been presented in appendix D.

Table F-1. Dependence of Ultrasonic Response on Radius for Some Simply Shaped Reflectors

Site	Circular Crack		Spherical Void	
	Response proportional to	β_1	Response proportional to	β_1
$r/\lambda << 1$	r^3/λ^2	1.5	r^3/λ^2	1.5
$r/\lambda >> 1$	r^2/λ	1.0	r	0.5

The behavior of an ellipsoidal void is qualitatively similar to that of the spherical void. The response now depends on the three semimajor radii, as will be discussed in detail below. However, the value of β_1 is the same 1.5 when the flaw is small, with respect to the wavelength, and 0.5 when the flaw is large, with respect to the wavelength.

It is clear that the value of β_1 is not the same for large or small flaws with respect to the wavelength and the values are also different for different flaw shapes.

When simple linear regression (in log-log plots) is used to analyze the responses of naturally occurring flaws as a function of area for data such as that presented earlier in this report, the values observed for β_1 were typically on the order of 0.2. Although this is a good empirical fit to the data for the flaw sizes appearing in the databases under analysis, it can not apply for all flaw sizes in light of the limits specified in table F-1. For example, one would expect that in the limit of small flaws with respect to the ultrasonic wavelength, $\beta_1 = 1.5$. Hence, there must be a significant change in slope from that derived from the empirical analysis of the data. Roughly speaking, the transition occurs when kr is on the order of unity, where $k = 2\pi/\lambda$ and λ is the wavelength. At 5 MHz in titanium alloys, this occurs for flaw sizes on the order of 0.025 inch in diameter, about a #1.5 flat-bottom hole (FBH).

From another perspective, such a change in slope is necessary to avoid PoD predictions that would be counter intuitive. If the empirically determined regression line of low slope were assumed to extend to very small flaw sizes, it could imply a significant PoD for flaw sizes that are much smaller than a wavelength, a result that would not be consistent with practical experience. The situation was encountered in this work.

Gaining empirical insight into this small flaw limit is very difficult. The flaw response drops very rapidly (as the cube of the radius) as the flaw size decreases. This has the practical consequence of making flaws in this size range very difficult to detect. Hence, there is often little or no evidence of the small flaw behavior in the empirical databases (since flaws in this size range are generally not found) even though the existence of this small flaw limit is required by the physics of the scattering process.

F.2 CONCEPT OF AN UPPER-BOUND FLAW.

These ideas have been generalized to that of an upper-bound flaw. Specifically, it is assumed that the response of a naturally occurring, hard alpha inclusion will always be less than that of a circumscribing ellipsoidal void. The concept is schematically illustrated in figure F-1. The solid line shows the response of an ellipsoidal upper-bound flaw, with values of $\beta_1 = 1.5$ in the small flaw limit and 0.5 in the large flaw limit. Also shown is a beam size plateau, which is required to take into account the fact that the flaw response cannot continue to rise after the flaw becomes sufficiently larger than the beam. In particular, the equivalent flat-bottom hole area (EFBH) can never be larger than the beam cross-sectional area. It seems reasonable to assume that the response of naturally occurring hard alpha inclusions, as a function of area, will be described by a curved line, somewhat below the response of the upper-bound flaw, as indicated by the broken curve in figure F-1. This response will start with some initial slope, rise, and then curve to reach a maximum response when the flaw extends well beyond the beam. Over limited ranges of area, the analysis of empirical data is often simplified by approximating this curve by a straight line. However, it is clear that no single straight line can describe the expected response of naturally occurring hard alpha inclusions for all areas.

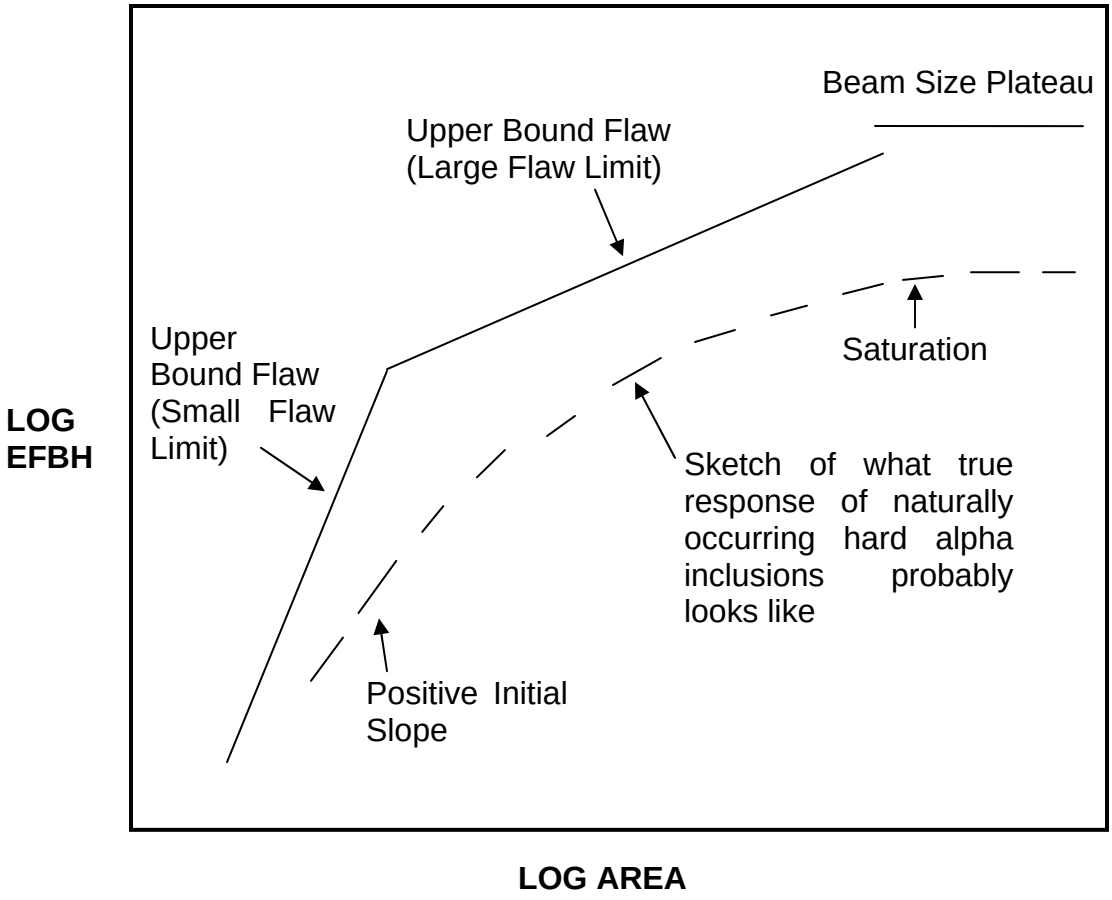


Figure F-1. Schematic of Relative Responses of Upper-Bound Flaw and Naturally Occurring Hard Alpha Inclusion

In the development of the Default PoD curves, the upper-bound flaw was taken into account by the use of scattering theory. The starting point is the assumption, noted above, that a "cigar-shaped" (prolate spheroid) void circumscribing the core of the hard alpha inclusions would produce a response greater than the true inclusion. Put in other words, the response of the circumscribing spheroid would provide an upper bound to the true response. The assumption is that the empirically determined regression line should never be above the response of this upper-bound flaw. In this work, whenever the physics-based response of this upper-bound flaw falls below the regression line, the response is taken to be that of the upper-bound flaw rather than the regression line. This would not be expected to happen along the portion of the regression line supported by experimental data. Primarily, the physics-based limit is intended to provide a guide to avoid incorrect extrapolation of the empirical curve to small flaw sizes.

In implementing this approach, it was assumed that the circumscribing void should be a prolate spheroid with an aspect ratio of 7:1, i.e., a cigar-shaped defect oriented parallel to the axis of the billet. The ratio of 7:1 arises from equation C-4, which describes the expected elongation of hard alpha inclusions, applied to the case of a billet that was reduced in diameter by a factor of six from the ingot. The reduction in diameter of 6 was chosen since this is the case for the

contaminated billet, the source of the majority of data analyzed in this study. In recognition of the fact that the direction of wave propagation during either the normal or angle inspections of the billet is perpendicular to the billet axis, the direction of insonification will be taken to be perpendicular to the axis of the prolate spheroid. Figure F-2 shows the upper-bound prolate spheroid and its illumination.

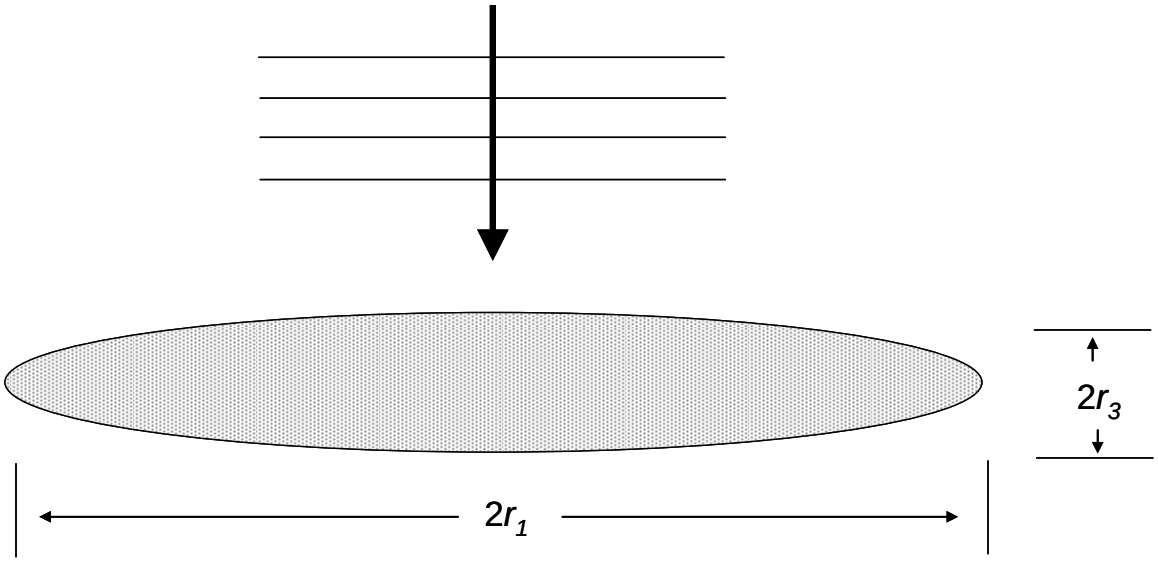


Figure F-2. Schematic of Upper-Bound Flaw (In this work, r_1 / r_3 was chosen to be 7.)

Given this physical definition of the upper-bound flaw, it is necessary to compute the strength of the scattered signal and, from this, the EFBH. The definition of the EFBH is given in equation 1 of this report. To determine the EFBH, one needs to know the signals from a calibration FBH and from the upper-bound prolate spheroid. The use of physics-based models for the scattering process to estimate these signals will be discussed below. By way of background, it should be noted that obtaining a general, exact solution for all wavelengths is a very complex problem for all but the simplest flaw geometries and was not attempted. However, much more straightforward is the solution at very long and very short wavelengths, as described in the following sections. Since there was interest in an upper bound, these high- and low-frequency limits were extended until they met, and this curve was taken as the bound.

F.3 MEASUREMENT MODEL.

In this study, a measurement model [F-1] was used that showed, when the flaw is fully illuminated by a beam that can be locally approximated by a plane wave, the flaw response is proportional to a quantity known as a scattering amplitude, A. Then, to evaluate equation F-1, there is a need to determine

$$(S/S_c) = (A/A_c) \qquad \text{(F-2)}$$

The scattering amplitude, a quantity often determined in scattering calculations, is defined as follows. Assume that the flaw is illuminated by a plane wave of displacement amplitude u_o.

Then in the far field, where the scattered fields are spreading spherically, the magnitude of the displacement of the scattered wave, *u*, is given by

$$|u/uo| = |A/d| \tag{F-3}$$

where *A* is the scattering amplitude and *d* is the distance from the scatterer. In general, *A* will depend on the relative angles of illumination and scattering. In this study, the only concern will be with the value of *A* in the backscattered direction corresponding to a pulse-echo measurement.

F.4 FLAT-BOTTOM HOLE RESPONSE.

For pulse-echo measurements at normal incidence, the response of FBHs is often taken to be equal to flat circular cracks of the same radius. The FBHs used in the calibration of Multizone and conventional inspections, #2 and #3 holes, respectively, have sizes such that *kr* is equal to 1.5 and 2.25, respectively, at 5 MHz. In this regime, the Kirchhoff Approximation (a high frequency or large flaw approximation) provides an accurate approximation to the scattering amplitude [F-2], which is simply related to the flaw area by the equation

$$|A_c| = kr^2/2 = \text{Area}/\lambda \tag{F-4}$$

where Area$=\pi r^2$. The expression is consistent with the widely accepted prediction that the response of cracks are proportional to their area [F-3]. In addition, it is properly normalized so that it can be directly compared to the response of the upper-bound, prolate spheroid, as presented in section F.5.

F.5 UPPER BOUND, PROLATE SPHEROID RESPONSE.

For the upper-bound, prolate spheroid, analytic formulae can be provided for the scattering in two regimes, small and large wavelength with respect to the flaw. In the large flaw (short wavelength) regime, known as the optical limit, Rose [F-4] has shown that the magnitude of the scattering amplitude of an ellipsoid is given by

$$|A| = \frac{r_1 r_2 r_3}{2 r_{e^2}} \tag{F-5}$$

Where r_1, r_2, and r_3 are the semimajor axes (radii) of the ellipsoid and, for illumination along a principal axis, r_e is the semimajor axis along the direction of illumination. Further details may be found in reference F-2. For the case of the upper-bound prolate spheroid, as sketched in figure F-2, $r_e=r_2=r_3$, and

$$|A| = r_1/2: \text{ upper-bound prolate spheroid, large flaw} \tag{F-6}$$

For comparison, the same result would be obtained for a sphere (a special case of a prolate spheroid), while the result would be $|A| = \frac{r_1^2}{2r_3}$ for an oblate spheroid (pancake shape), illuminated broadside, when r_3 is the semiminor axis.

It should be noted that as the flaw grows larger, it may exceed the ultrasonic beam in size, particularly for Multizone inspections. However, equation F-6 remains a good approximation, even when the flaw size is significantly larger than the beam size. The primary contribution to the scattered field occurs from a small patch on the spheroid surrounding the point where an incident wavefront would be tangent to the surface of the spheroid. The dimensions of this spot could be computed in a fairly straightforward fashion in terms of the wavelength and axes of the prolate spheroid. However, the primary interest is in the upper bound, and equation F-6 provides this, therefore, for conciseness, beam size effects will not be explored further.

When the flaw dimensions are small with respect to the wavelength, it is known as the Rayleigh scattering regime. In this case, as indicated in table F-1, the scattering amplitude will be proportional to (area)$^{1.5}$, which is equivalent to the cube of the radius for a sphere. Simple dimensional arguments show that this implies that the scattering amplitude is proportional to the square of the frequency, a well-known aspect of scattering in the Rayleigh regime.

Calculation of analytical formulae for the scattering amplitude in the small-flaw limit was a research topic of high interest in the late 1970s, and there are many papers dealing with this problem. In many ways, the work is analogous to the fracture mechanics efforts that occurred in a similar time frame to determine the stress-intensity factors of cracks of a variety of geometries. Since the small-flaw ultrasonic scattering results play an important role in the behavior of the PoD curves produced in this work, some further details are given.

Gubernatis, Krumhansl, and colleagues wrote a series of papers presenting a formal theory for scattering from inclusions and voids [F-5], applying it to the case of small flaws with respect to a wavelength [F-6] and to a number of other cases, which will not be cited here. A particular concise discussion of that work appears in Kohn and Rice [F-7] and reference will be made to specific equations in their work. All of the above work draws heavily on a classic paper by Eshelby [F-8], who determined the deformation of an ellipsoidal inclusion (including the limit of a void) subjected to a static load. This solution is of critical importance to the ultrasonic scattering problem since, in the small flaw or equivalently long wavelength limit, the deformation of the ellipsoid in the dynamic scattering problem will closely approximate that in the static problem. Further discussion may be found in Thompson, et al. [F-9].

A capsulation of the results is given below, with detailed formulae provided in section F.11.

The scattering amplitude for a longitudinal wave scattering from a void is given by

$$|\mathrm{A}_3| = \frac{k^2}{4\pi\rho}\left|\delta M - \frac{D_{3333}}{v^2}\right| \qquad \text{(F-7)}$$

where ρ is the density of the material, v is the longitudinal wave speed, δM is a mass excess (proportional to density times void volume), and $\vec{\vec{D}}$ is a tensor defined, in dyad notation, by

$$\vec{\vec{D}} = -V\vec{\vec{C}} : (I - \vec{\vec{S}})^{-1}. \tag{F-8}$$

The subscript notation has been used in equation F-8 since $\vec{\vec{D}}$ is an anisotropic tensor; a coordinate system has been chosen such that the longitudinal wave propagates in the 3 direction. Here, I is the identity matrix, V is the volume of the ellipsoidal inclusion, $\vec{\vec{C}}$ is the elastic constant tensor, and $\vec{\vec{S}}$ is a tensor relating the constrained and stress-free strains in the inclusion, as derived by Eshelby. Explicit analytical expressions for $\vec{\vec{S}}$ for some simple void geometries are provided in section F.10. This long wavelength limit is different from the short wavelength limit in that the answer depends on the elastic properties of the host through Poisson's ratio. The scattering amplitude is seen to be proportional to the square of frequency (through k^2), as would be expected for Rayleigh scattering.

For the case of the 7:1 prolate spheroid in titanium, this analysis predicts that

$$|A| = 6.664 \times 10^{-5} (\text{Area})^{1.5} f^2 : \text{ upper-bound prolate spheroid, small flaw} \tag{F-9}$$

where A is the scattering amplitude in mils, $\text{Area} = \pi(r_1 r_3)$ is the cross-sectional area in square mils and f is the frequency in MHz. In this calculation, the material properties assumed for titanium were ρ = 4.454 g/cm^3, C_L = 0.602 cm/μsec, and C_T = 0.31 cm/μsec.

F.6 OTHER SIMPLE VOID SHAPES OF PRACTICAL INTEREST.

For an oblate spheroid of major and minor axes r_1 and r_3, respectively, analytical results can also be found in section F.10.

As r_3 approaches zero, the oblate spheroid approaches a penny-shaped crack. To evaluate this, the case of an eccentricity of 10^{-8} has been considered. The corresponding scattering amplitude in titanium is

$$A = 6.576 \times 10^{-5} (\text{Area})^{1.5} f^2 \tag{F-10}$$

where $\text{Area} = \pi r^2$ is the area of the crack in square mils and f is the frequency in MHz. The short-wavelength limit of the crack is the same as that of the FBH, as given in equation F-4.

By way of comparison, for the case of a 7.1 oblate spheroid, the result is

$$A = 7.314 \times 10^{-5} (\text{Area})^{1.5} f^2 . \tag{F-11}$$

Finally, for the case of a sphere, $r_1 = r_3$ and the prolate and oblate spheroid solutions approach the same limit. In titanium, this is

$$A = 1.404 \times 10^{-4} (\text{Area})^{1.5} f^2 \tag{F-12}$$

F.7 RELATIONSHIP TO EFBH.

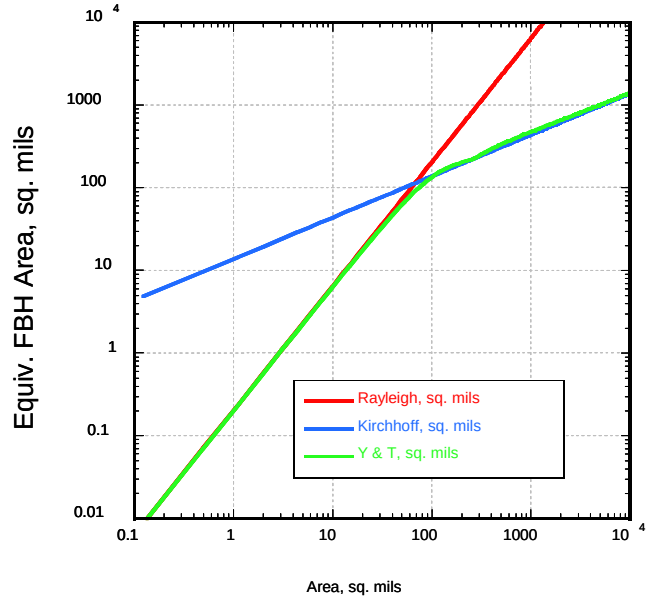
The various results for scattering amplitude, A can be used to predict the EFBH by combining the relevant equation for A with equations 1, F-2, and F-4. However, this can be considerably simplified by noting that EFBH is proportional to the area of the calibration hole. The theory for the response of that hole, equation F-4, which appears in the denominator of equation 1, is also proportional to that area. Simple arithmetic then leads to the result

$$\text{EFBH} = A\lambda \tag{F-13}$$

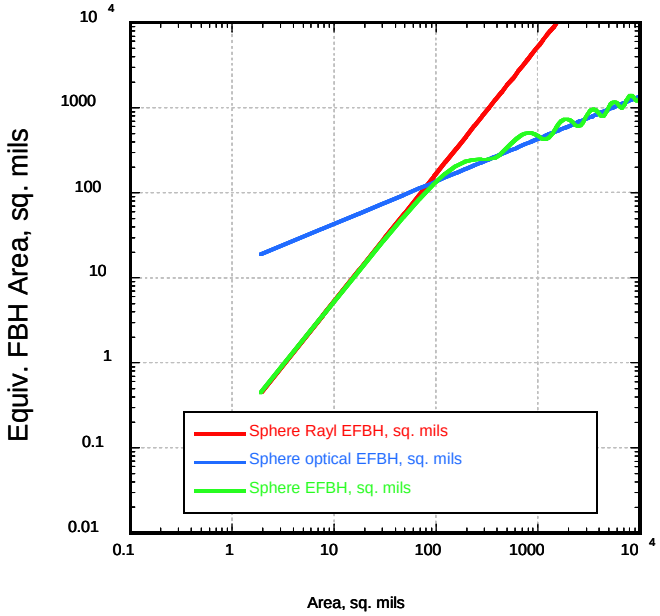
F.8 VERIFICATION THROUGH COMPARISON TO OTHER THEORIES.

To illustrate and verify these results, the predictions of the above asymptotic results to analytical or numerical results for particular geometries were compared. Figure F-3 compares the short and long wavelength asymptotic results for a spherical void (equations F-6 and F-12) with the analytical results of Ying and Truell [F-10]. In the results based on Ying and Truell, both broadband (80% bandwidth) and narrowband (single frequency) responses are presented at a frequency of 5 MHz. It shows (1) the long and short wavelength asymptotes are in good agreement with the analytical results in their respective regimes, (2) there is only a very small regime (when plotted on this log-log scale) in which the analytical theory substantially deviates from these limits for the broadband case, and (3) the single-frequency exact results show some oscillation due to resonance effects.

For prolate spheroids of particular shapes, numerical predictions of the scattering have been reported in the literature for a number of cases. Figure F-4 shows a single frequency (5 MHz) comparison of the predictions of equations F-6 and F-10 to the work of Opsal and Vischer [F-11] based on the Method of Optimal Truncation (MOOT) for a 2:1 prolate spheroid. Again, the long wavelength asymptote provides a good approximation for a significant range of flaw sizes, with some resonance effects being observed.



(a)



(b)

Figure F-3. Comparison of Analytical Theory to High- and Low-Frequency Asymptotes for a Spherical Void (a) Broadband Results and (b) Narrowband Results

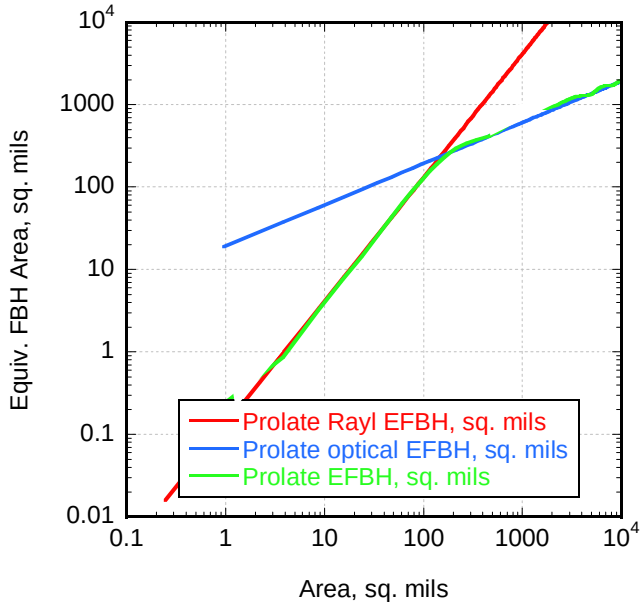


Figure F-4. Comparison of Analytical Theory to High- and Low-Frequency Asymptotes for a 2:1 Prolate Spheroidal Void

The results shown in figure F-3 provide a strong verification of the asymptotic results presented in this section via the good agreement shown with existing, well-established theories for particular geometries.

F.9 RESULTS USED IN PoD STUDIES.

Figure F-5 shows a graph of the upper-bound limit for the 7:1 upper-bound, prolate spheroidal void for titanium obtained by extrapolating equation F-10 to short wavelengths and equation F-6 to long wavelengths until they intersect. As noted above, this result is the same for conventional and Multizone inspections as long as the reflecting portion of the flaw is fully illuminated. Conventional and Multizone inspections would have different beam size plateaus, but these are likely off the scale of this plot. Based on the arguments presented above, it would be expected that no hard alpha inclusion would have a response substantially above this limit.

For comparison purposes, figure F-6 shows the analogous calculation for a FBH (or equivalently, a flat, penny-shaped crack). As before, the red and blue lines respectively show the long wavelength and optical limits. Also shown is a calculation based on MOOT [F-11], which provides the verification. It is of interest to note the different slope for large sizes (optical limit) compared to the sphere and prolate spheroid results, a consequence of the fact that $\beta_1 = 1$ for the crack, but 0.5 for the prolate spheroid in this limit.

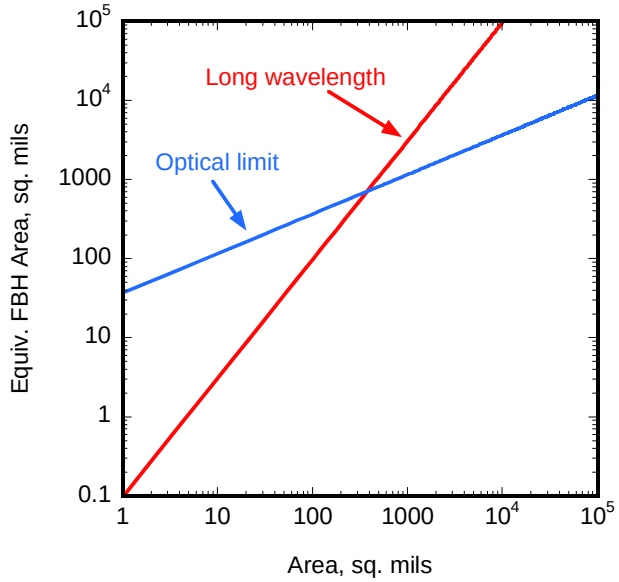


Figure F-5. Scattering From a 7:1 Prolate Spheroid Showing Different Slopes for Small and Large Flaw With Respect to the Wavelength

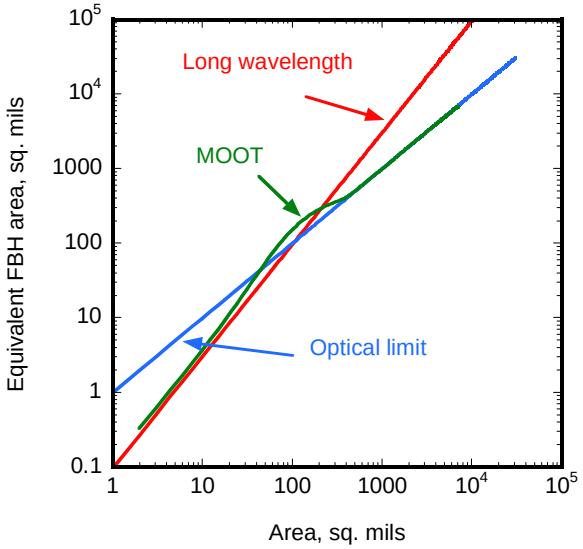


Figure F-6. Scattering From a Flat, Penny-Shaped Crack Showing Different Slopes for Small and Large Flaws With Respect to the Wavelength

These results have been used in the development of PoD curves, with predictions that have been sensible as judged by comparison to real data. It should be emphasized that there are no adjustable parameters in the predictions of EFBH for these cases.

F.10 EXPLICIT FORMS OF ESHELBY SOLUTIONS FOR SPECIAL CASES OF INTEREST IN NONDESTRUCTIVE EVALUATIONS.

The major threads of the analysis used to determine the long wavelength (small flaw) scattering amplitudes are summarized below. Here, the notation KR-n will refer to equation "n" presented by Kohn and Rice [F-7] and E-m will refer to equation "m" presented by Eshelby [F-8].

KR-5 gives an expression for the scattering amplitudes of localized defects in solids for both longitudinal and shear waves in terms of a vector, ***f,*** which relates the scattering amplitudes to perturbations in densities and elastic constants, and strain and displacement fields in an integral to be evaluated over the volume of the inclusion.

- In KR-15, the ***f*** vector is related to a mass excess, δM, and a tensor ***D***, which is determined by the deformation of the inclusion at long wavelength.

- For the case of an ellipsoidal void, KR-A25 relates ***D*** to a tensor ***S*** that appears in the classical paper of Eshelby.

- In section 3 of Eshelby, the solution for ***S*** is presented.

- E-3.7 relates ***S*** to quantities Q and R that are function only of Poisson's ratio and quantities $I_{r_1}\, I_{r_2}\, I_{r_3}$ etc. that in the general case are expressed in terms of elliptic integrals through E-3.9.

- For the case of oblate and prolate spheroids, E-3.15 and E-3.16, respectively, give expressions for particular "I" quantities in terms of simple inverse trigonometric and inverse hyperbolic trigonometric functions. It is noted that other components can be developed by cyclic interchange of various indices with E-3.10, E-3.11, and E-3.12 providing some useful relationships between the various quantities.

Equations F-7 and F-8 express the scattering amplitude for ellipsoidal flaws in terms of the $\vec{\vec{S}}$ tensor that was developed by Eshelby when he considered the static deformation of an ellipsoidal inclusion. For the convenience of the reader, explicit forms for this solution are given for spherical-, oblate spheroidal (pancake)-, and prolate spheroidal (cigar)-shaped voids. In the latter two cases, it is assumed that the illumination is from the side, presenting the largest area to the incident beam.

$S_{11}=r_1^2 QI_{r_1r_1}+RI_{r_1}$

$S_{12}=r_2^2 QI_{r_1r_2}-RI_{r_1}$

$S_{13}=r_3^2 QI_{r_1r_3}-RI_{r_1}$

$S_{21}=r_1^2 QI_{r_2r_1}-RI_{r_2}$

$S_{22}=r_2^2 QI_{r_2r_2}+RI_{r_2}$

$S_{23}=r_3^2 QI_{r_2r_3}-RI_{r_2}$

$S_{31}=r_1^2 QI_{r_3r_1}-RI_{r_3}$

$S_{32}=r_2^2 QI_{r_3r_2}-RI_{r_3}$

$S_{33}=r_3^2 QI_{r_3r_3}+RI_{r_3}$

$S_{44}=[(r_2^2+r_3^2)QI_{r_2r_3}+R(I_{r_2}+I_{r_3})]/2$

$S_{55}=[(r_1^2+r_3^2)QI_{r_1r_3}+R(I_{r_1}+I_{r_3})]/2$

$S_{66}=[(r_1^2+r_2^2)QI_{r_1r_2}+R(I_{r_1}+I_{r_2})]/2$

Oblate Spheroid:

$r_1 = r_2 > r_3$

$I_{r_1}=I_{r_2}=2\pi r_1^2 r_3/(r_1^2-r_3^2)^{3/2}[\cos^{-1}(r_3/r_1)-(r_3/r_1)(1-r_3^2/r_1^2)^{1/2}]$

$I_{r_3}=4\pi-I_{r_1}-I_{r_2}$

$I_{r_1r_3}=I_{r_1r_3}=I_{r_2r_3}=I_{r_3r_2}=(I_{r_3}-I_{r_2})/[3(r_2^2-r_3^2)]$

$I_{r_1r_2}=I_{r_2r_1}=\pi/(3r_1^2)-I_{r_1r_3}/4$

$I_{r_1r_1}=I_{r_2r_2}=3I_{r_1r_2}$

$I_{r_3r_3}=4\pi/(3r_3^2)-I_{r_3r_1}-I_{r_1r_2}$

$Q=\dfrac{3}{8\pi(1-\sigma)}$

σ=Poisson's ratio

$R=\dfrac{1-2\sigma}{8\pi(1-\sigma)}$

Sphere:

$r_1=r_2=r_3$

$I_{r_1}=I_{r_2}=I_{r_3}=4\pi/3$

$I_{r_1r_1}=I_{r_2r_2}=I_{r_3r_3}=4\pi/(5r_1^2)$

$I_{r_1r_2}=I_{r_2r_1}=I_{r_1r_3}=I_{r_3r_1}=I_{r_2r_3}=I_{r_3r_2}=4\pi/(15r_1^2)$

Prolate Spheroid:

$r_1>r_2=r_3$

$I_{r_2}=I_{r_3}=2\pi r_1 r_3^2/(r_1^2-r_3^2)^{3/2}[r_1/r_3)(r_1^2/r_3^2-1)^{1/2}-\cosh^{-1}(r_1/c_3)]$

$I_{r_1}=4\pi-I_{r_2}-I_{r_3}$

$I_{r_1r_2}=I_{r_2r_1}=I_{r_1r_3}=I_{r_3r_1}=(I_{r_2}-I_{r_1})/[3(r_1^2-r_2^2)]$

$I_{r_2r_3}=I_{r_3r_2}=\pi(3r_2^2)-I_{r_1r_2}/4$

$I_{r_2r_2}=I_{r_3r_3}=3I_{r_2r_3}$

$I_{r_1r_1}=4\pi/(3r_1^2)-I_{r_1r_2}-I_{r_1r_3}$

F.11. REFERENCES.

F-1. Thompson, R.B. and Gray, T.A., "A Model Relating Ultrasonic Scattering Measurements Through Liquid-Solid Interfaces to Unbounded Medium Scattering Amplitudes," *J. Acoust. Soc. Am.*, Vol. 74, 1983, pp. 1279-1290.

F-2. Schmerr, L.W., "Ultrasonic Modeling of Benchmark Problems," in *Review of Progress in QNDE*, Vol. 21B, D.O. Thompson and D.E. Chimenti, eds., American Institute of Physics, NY, 2002, pp. 1933-1940.

F-3. Krautkramer, J. and H., "Ultrasonic Testing of Materials," Springer-Verlag, 4th Edition, 1990.

F-4. Rose, J.H. private communication.

F-5. Gubernatis, J.E., Domany, E., and Krumhansl, J.A., "Formal Aspects of the Theory of the Scattering of Ultrasound by Flaws in Elastic Solids," *J. Appl. Phys.*, 48, 1977, pp. 2804.

F-6. Gubernatis, J.E, "Long-Wave Approximation for the Scattering of Elastic Waves From Flaws With Applications to Ellipsoidal Voids and Inclusions," *J. Appl. Phys.*, 50, 1979, pp. 4046.

F-7. Kohn, W. and Rice, J.R., "Scattering of Long-Wavelength Elastic Waves From Localized Defects in Solids," *J. Appl. Phys.*, 50, 1979, pp. 3346.

F-8. Eshelby, J.D., "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," *Proc. Roy. Soc. London A*, 241, 1957, pp. 376.

F-9. Thompson, R.B., Gray, T.A., and Meeker, W.Q., "Use of Physics-Based Models to Guide the Extrapolation of Aircraft Engine Ultrasonic PoD Data to Small Flaw Sizes," *Review of Progress in QNDE*, Vol. 25, D.O. Thompson and D.E. Chimenti, eds., American Institute of Physics, NY, 2006, pp. 1878-1885.

F-10. Ying, C.F. and Truell, R., "Scattering of a Plane Longitudinal Wave by a Spherical Obstacle in an Isotropically Elastic Solid," *J. Appl. Phys.*, 27, 1956, pp. 1986-1997.

F-11. Opsal, J.L. and Visscher, W.V., "Theory of Elastic Wave Scattering: Applications of the Method of Optimal Truncation," *J. Appl. Phys.*, 58, 1985, pp. 1102-1114.

APPENDIX G—LINEAR AND KINK REGRESSION WITH SMALL FLAW LIMIT

The discussion of appendix F clearly shows that a simple linear regression applied to plots of log $\hat{a}$ versus log a can not be valid for all flaw sizes. Hence, the response of the upper-bound flaw provides a limit above which the regression line should not be allowed to extend. As was noted in the discussion about figure F-1, the regression line should never be extrapolated to the left (i.e., above) this line. This has the further implication that the true relationship between equivalent flat-bottom hole area (EFBH) and area must be curved in log-log space, as indicated by the dashed line in figure F-1. If there were enough sufficiently high-quality data, a curved regression could be estimated. However, in practice, there generally is not enough data to support such an analysis. Then it has to be decided how to proceed.

In the Default probability of detection (PoD) analysis, which is described in appendix H, two cases were encountered, distinguished by the nature of the data. In the first case, the range and scatter of the data only supported a linear regression. There was not sufficient information to support an assessment of the curvature of the dashed curve in figure F-1. In this case, one first estimates a linear regression line from the data. The complete relationship between EFBH and area is determined by extending the regression line to the left to a point where it intersects the small flaw limit of the physics-based upper bound. This approach leads to a break in the relationship between EFBH and area, at a point determined by the intersection of the physics-based, upper-bound response and the estimate of the linear regression line, based on the data. Figure G-1(a) illustrates this case, with a hypothetical data set superimposed on the schematic relationships presented in figure F-1. This approach was used in the analysis of the conventional inspection data. This case will be called a linear regression with small flaw limit.

The second case is conceptually similar in that an estimate of the regression is joined to the physics-based small flaw limit. However, in this case, the range and scatter of the data support an assessment of curvature. As a first approximation to the curved nature of the expected flaw response curve, a “kink” regression was conducted. The curved dashed line was approximated by two straight line segments, joined at a kink. The positions of the two lines, which imply the kink point, were determined by this kink regression relationship to the data, using maximum likelihood. Figure G-1(b) illustrates this case, with a second hypothetical data set superimposed on the schematic relationships presented in figure F-1. This approach was used in the analysis of the Multizone data. Its use was motivated by the apparent curvature of the Multizone flaw response and signal-to-noise ratio data, as shown in figure C-32 for the Multizone case. This case will be called a kink regression with small flaw limit.

A “sanity” check of the second kink regression approach was to determine whether the estimated kink regression relationship lay below the response of the upper-bound flaw. In the specific example to follow, the kink regression line lay below the upper-bound response for nearly the entire region for which the PoD had a significant value.

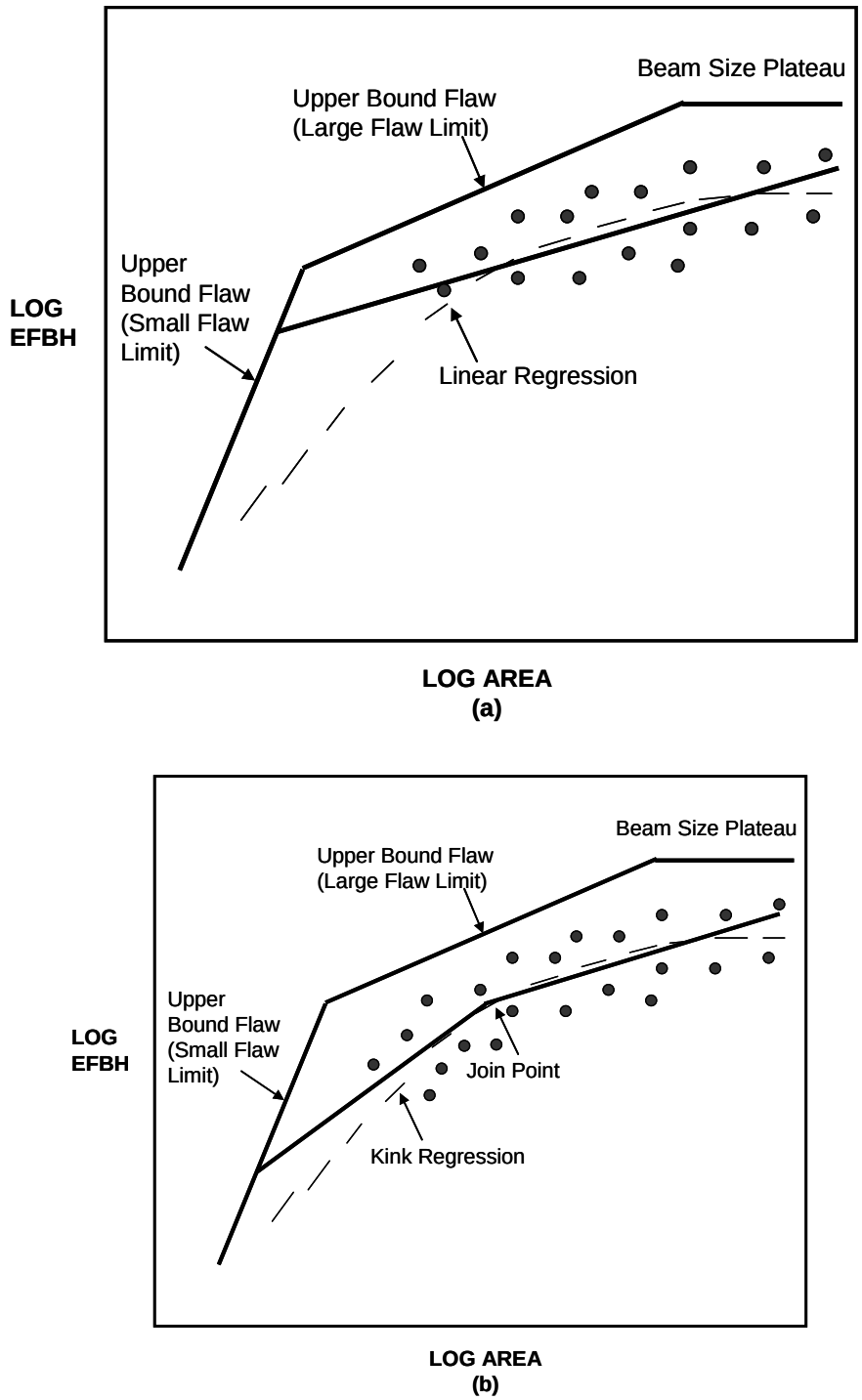


Figure G-1. Regression Bounded by Small Flaw Limit (a) Linear Regression and (b) Kink Regression

APPENDIX H—TECHNICAL DETAILS OF STATISTICAL ANALYSIS

H.1 ANALYSIS OF CONVENTIONAL INSPECTION DATA.

H.1.1 STATISTICAL MODEL.

The $\hat{a}$ versus a model for signal strength in the conventional inspection, assuming that there are no "Type II misses" in the inspection, can be written as

$$Y_{CV} = \beta_0^{CV} + \beta_1^{CV} x + \varepsilon_{CV} \tag{H-1}$$

where β_0^{CV} and β_1^{CV} are constants to be determined in the regression,

$$x = \log(a)$$

is log flaw area in square mils, and the random error term ε_{CV} is assumed to have a normal distribution with mean 0 and standard deviation σ_ε^{CV}. The response is defined as

$$Y_{cv} = \log(\hat{a}) = \log(\text{EFBH})$$

where

$$\text{EFBHSqMils} = \frac{\%FSH}{80} \times \left(\frac{\text{Cal}}{64}\right)^2 \times \frac{\pi}{4} \times 10^6$$

is the equivalent flat-bottom hole area (EFBH) (in square mils) giving the size of the flat-bottom hole (FBH) that would result in the same percent full-screen height (*%FSH)*, assuming calibration was done to a #Cal FBH. EFBH is used because some data sets have *%FSH* as a response, but a mix of calibration values were used (e.g., #2 and #3 FBHs were used to calibrate conventional inspections for bars and billets, respectively, in the Default PoD-3Dimensional data, and EFBH serves as a common measure of signal strength.

For the Contaminated Billet Study (CBS) data, there were a number of known misses and because of these misses, the $\hat{a}$ versus a model does not provide an adequate description of the data (i.e., the $\hat{a}$ versus a model does not fit the data). To have a model that fits the data, an accommodation term is used in the model. This accommodation term allows for misses other than those that would arise under the $\hat{a}$ versus a model. Such misses are called atypical misses or Type II misses. In this extended model, the probability of a Type II miss depends on flaw size. The data provide no information on which misses are Type I and which are Type II. In estimating this model, it is not necessary to identify the individual misses as Type I or Type II.

The probability of a Type II miss in the accommodation probability of detection (PoD) model is given by

$$\Pr(\text{Type 2 Miss}) = \frac{\exp(z_{\text{MissII}})}{1+\exp(z_{\text{MissII}})}$$
$$z_{\text{MissII}} = \beta_0^{\text{MissII}} + \beta_1^{\text{MissII}} x$$

H.1.2 SMALL FLAW CORRECTION.

Over the range of small flaws (where there is no data), a physics-based model was used to give an alternative regression equation. This regression relationship (from Thompson, Gray, and Meeker [H-1]), also discussed in appendix F, is:

$$\begin{aligned}\log_{10}(\text{EFBHSqMils}) &= \beta_0^{SF} + \beta_1^{SF} \times \log_{10}(\text{AreaSqMils}) \\ &= -1.0091 + 1.5 \times \log_{10}(\text{AreaSqMils})\end{aligned}$$

That is, β_0^{CV} and β_1^{CV} in (1) are replaced by β_0^{SF} and β_1^{SF}, respectively, below the point of intersection. Thus

$$\beta_0^{SF} = -1.0091$$
$$\beta_1^{SF} = 1.5$$

if base 10 logarithms are used to define the underlying lognormal distribution for amplitude. If natural logarithms are used to define the lognormal distribution for amplitude,

$$\beta_0^{SF} = -1.0091 \times \log_e(10) = -2.3235$$

The specific numerical values in these formulae come from the theory of scattering in the Rayleigh (small r/λ) regime. The numerical value of 1.5 occurs because signal is proportional to (area)$^{1.5}$ or (radius)^3. This is true for all materials. The numerical value of 1.0091 is a constant, which depends only on Poisson's ratio for an isotropic, elastic solid. It is a consequence of evaluating equation F-13, with A taken from equation F-7, $\vec{\vec{D}}$ taken from equation F-8, and $\vec{\vec{S}}$ determined following equation F-10.

H.1.3 PROBABILITY OF DETECTION.

Conditional on the event that there is not a Type II miss, the PoD is the probability that the signal is above the threshold. This probability can be expressed as

$$\Pr(Y > Y_{th}) = 1 - \Phi\left(\frac{Y_{th} - (\beta_0^{CV} + \beta_1^{CV} x)}{\sigma_\varepsilon^{\text{CV}}}\right)$$

$$= \Phi\left(\frac{x - (Y_{th} - \beta_0^{CV})/\beta_1^{CV}}{\sigma_\varepsilon^{CV}/\beta_1^{CV}}\right)$$

where

$$Y_{th} = \frac{60}{80} \times \left(\frac{\text{Cal}}{64} \right)^2 \times \frac{\pi}{4} \times 10^6$$

is the conventional detection threshold, corresponding to 60% FSH, in EFBH (square mils) units and $\Phi(z)$ is the standard normal cumulative distribution function. The quantity $\sigma_\varepsilon^{CV} / \beta_1^{CV}$ is called the PoD slope, as it describes how fast the PoD function rises.

The PoD, allowing for Type II misses, is the probability that there is no Type II miss and that the signal is above the threshold. This probability can be expressed as

$$\text{POD}(a) = \Pr(\text{No Type II Miss and } Y > Y_{th}) \tag{H-2}$$

$$= [1 - \Pr(\text{Type II Miss})] \times \Pr(Y > Y_{th}) \tag{H-3}$$

H.1.4 BORROWING STRENGTH FROM THE AMPLITUDE-ONLY MODEL FOR MULTIZONE INSPECTION.

The following similar model could be used to describe the Multizone amplitude data.

$$Y_{MZ} = \beta_0^{MZ} + \beta_1^{MZ} x + \varepsilon_{MZ}$$

where the random error term ε_{MZ} is assumed to have a normal distribution with mean 0 and standard deviation σ_ε^{MZ}.

It was suggested that switching from conventional to Multizone inspection should shift the PoD curve to the left, but not affect the shape of the PoD curve. This implies that the PoD slope should be the same for both models. This equivalence can be expressed by

$$\frac{\sigma_\varepsilon^{CV}}{\beta_1^{CV}} = \frac{\sigma_\varepsilon^{MZ}}{\beta_1^{MZ}}$$

This approach was used in some initial steps in the analysis of conventional inspection data but was abandoned in the final analysis.

H.1.5 MAXIMUM LIKELIHOOD ESTIMATES FOR THE PoD MODEL PARAMETERS.

Because of its versatility and desirable large-sample properties, the method of maximum likelihood (ML) is used to estimate the model parameters β_0^{CV}, β_1^{CV}, σ_ε^{CV}, β_0^{MZ}, σ_ε^{MZ}, β_0^{MissII}, and β_1^{MissII}. This is done by finding the values of the parameters that maximize the "loglikelihood" function. The likelihood function is proportional to the probability of the data. Taking logarithms simplifies numerical computations. The maximum of the loglikelihood

function and of the likelihood function will occur for the same parameter values. Hence, the ML estimate of parameter values, defined in terms of the maximum of the loglikelihood function, is the same as the estimate that would be obtained by maximizing the likelihood function.

The loglikelihood function can be written as the sum of the contributions for each of the n observations in the data set.

$$L(\beta_0^{CV}, \beta_1^{CV}, \sigma_\varepsilon^{CV}, \beta_0^{MZ}, \sigma^{MZ}, \beta_0^{\text{MissII}}, \beta_1^{\text{MissII}}) = \sum_{i=1}^{n} L_i \qquad \text{(H-4)}$$

and one obtains β_1^{MZ} from $\beta_1^{CV} \times (\sigma_\varepsilon^{MZ} / \sigma_\varepsilon^{CV})$. The loglikelihood contributions for the four different kinds of observations, assuming truncation at a level y_{TR}^{CV} for conventional inspection and y_{TR}^{MZ} for Multizone inspection are:

- Conventional hit with amplitude (log EFBH) y_i

$$L_i = \log\left[\frac{[1-\Pr(\text{Type II Miss})]\ \ \phi\left(\frac{y_i-(\beta_0^{CV}+\beta_1^{CV}x_i)}{\sigma^{CV}}\right)}{1-\Phi\left(\frac{y_{TR}^{CV}-(\beta_0^{CV}+\beta_1^{CV}x_i)}{\sigma^{CV}}\right)}\right]$$

- Conventional miss

$$L_i = \log\left[\Pr(\text{Type II Miss}) + \frac{[1-\Pr(\text{Type II Miss})]\ \ \Phi\left(\frac{y_i-(\beta_0^{CV}+\beta_1^{CV}x_i)}{\sigma^{CV}}\right)}{1-\Phi\left(\frac{y_{TR}^{CV}-(\beta_0^{CV}+\beta_1^{CV}x_i)}{\sigma^{CV}}\right)}\right]$$

- Conventional saturated

$$L_i = \log\left[\frac{[1-\Pr(\text{Type II Miss})]\ \ \left[1-\Phi\left(\frac{y_i-(\beta_0^{CV}+\beta_1^{CV}x_i)}{\sigma^{CV}}\right)\right]}{1-\Phi\left(\frac{y_{TR}^{CV}-(\beta_0^{CV}+\beta_1^{CV}x_i)}{\sigma^{CV}}\right)}\right]$$

- Multizone hit

$$L_i = \log\left[\frac{\phi\left(\frac{y_i-(\beta_0^{MZ}+\beta_1^{MZ}x_i)}{\sigma^{MZ}}\right)}{1-\Phi\left(\frac{y_{TR}^{MZ}-(\beta_0^{MZ}+\beta_1^{MZ}x_i)}{\sigma^{MZ}}\right)}\right]$$

The terms $1-\Phi()$ in the denominator account for the possibility of unknown misses in field inspections. That is, if a signal is below the detection threshold, one would not know of its existence. The truncation levels y_{TR}^{CV} and y_{TR}^{MZ} are explained in section H.2.

The ML estimates of the parameters are obtained by maximizing equation H-4. Then PoD is computed by substituting these estimates into equation H-2.

H.2 ANALYSIS OF MULTIZONE INSPECTION DATA.

H.2.1 STATISTICAL MODEL.

Multizone inspection uses a dual criterion to detect flaws. A detection occurs if the signal amplitude exceeds 70% FSH or if the signal-to-noise ratio (SNR) exceeds 2.5. Evaluation of PoD requires a generalization of the standard $\hat{a}$ versus a method. In particular, a bivariate regression model assuming an underlying joint distribution is used, in which the means of the marginal log signal amplitude and log SNR values depend on flaw size, but the standard deviations and correlation do not.

Let Y_{AMP}^{MZ} and Y_{SNR}^{MZ} denote the Multizone log amplitude and log SNR values respectively and let $\underline{Y} = (Y_{AMP}^{MZ}, Y_{AMP}^{MZ})$ be the observation vector. Under this model, the probability of having an inspection result contained in the rectangle $A = [y_{AMPL} \le Y_{AMP}^{MZ} < y_{AMPU} \cap y_{SNRL} \le Y_{SNR}^{MZ} < y_{SNRU}]$ is

$$\Pr(\underline{Y} \in A) = \int_{y_{AMPL}}^{y_{AMPU}} \int_{y_{SNRL}}^{y_{SNRU}} f(y_1, y_2; \mu_{AMP}, \mu_{SNR}, \sigma_{AMP}, \sigma_{SNR}, \rho) dy_1 dy_2$$

where the bivariate normal density function is

$$f(y_1, y_2; \mu_{y_1}, \mu_{y_2}, \sigma_{y_1}, \sigma_{y_2}, \rho) = \frac{1}{2\pi\sigma_{y_1}\sigma_{y_2}\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2} Q\right)$$

where

$$Q = \frac{1}{1-\rho^2}\left[\frac{(y_1-\mu_{y_1})^2}{\sigma_{y_1}^2} - 2\rho\frac{(y_1-\mu_{y_1})(y_2-\mu_{y_2})}{\sigma_{y_1}\sigma_{y_2}} + \frac{(y_2-\mu_{y_2})^2}{\sigma_{y_2}^2}\right]$$

The regression relationships

$$\mu_{y_1} = \mu_{\text{AMP}} = \beta_0^{MZAMP} + \beta_1^{MZAMP} x + \beta_2^{MZAMP} (x - x_{Join}) I_{Join}$$
$$\mu_{y_2} = \mu_{\text{SNR}} = \beta_0^{MZSNR} + \beta_1^{MZSNR} x + \beta_2^{MZSNR} (x - x_{Join}) I_{Join}$$
$$x = \log(a)$$

x_{Join} is the position of the join point where the slope changes

$I_{Join} = 1$ if $x < x_{Join}$ and 0 otherwise

express the dependency of the distribution means on flaw size. These relationships allow for a change in slope at x_{Join}. For example, the slope of the amplitude relationship is β_1^{MZAMP} when $x \geq x_{Join}$ and $\beta_1^{MZAMP} + \beta_2^{MZAMP}$ when $x < x_{Join}$. It is these relationships that introduced the kink regression, which was needed to describe the nonlinear relationship between $\log \hat{a}$ and $\log a$ for the Multizone data.

H.2.2 SMALL FLAW CORRECTION.

The small flaw corrections for Multizone amplitude/area relationship is the same as for conventional. The small flaw correction for the SNR_{mz}/area relationship is obtained from an empirical relationship between SNR_{mz} and amplitude. The empirical regression relationship between the Multizone SNR and amplitude is

$$\begin{aligned}\log_{10}(\text{SNR})_{\text{mz}} &= \gamma_0 + \gamma_1 \times \log_{10}(\text{EFBHSqMils}) \\ &= -2.0406 + 0.92666 \times \log_{10}(\text{EFBHSqMils})\end{aligned}$$

The two numerical constants in this equation are purely empirical, resulting from a regression fit to a plot of $\log_{10}(\text{SNR}_{\text{mz}})$ versus $\log_{10}(\text{EFBHSqMils})$.

The derived regression relationship between SNR and area is

$$\begin{aligned}\log_{10}(\text{SNR}_{\text{mz}}) &= \gamma_0 + \gamma_1 \times [\beta_0 + \beta_1 \times \log_{10}(\text{AreaSqMils})] \\ &= -2.0406 + 0.92666 \times [-1.0091 + 1.5 \times \log_{10}(\text{AreaSqMils})] \\ &= -2.9757 + 1.39 \times \log_{10}(\text{AreaSqMils})\end{aligned}$$

Thus, the derived regression relationship between SNR and area in terms of natural logs is

$$\begin{aligned}\log_e(\text{SNR}_{\text{mz}}) &= -2.9757 \times \log_e(10) + 1.39 \times \log_e(\text{AreaSqMils}) \\ &= -6.8518 + 1.39 \times \log_e(\text{AreaSqMils})\end{aligned}$$

H.2.3 PROBABILITY OF DETECTION.

A Multizone detection occurs if the signal amplitude exceeds 70% FSH or the SNR exceeds 2.5. The PoD can be expressed as

$$\begin{aligned}\text{PoD}(a) &= \Pr(Y_{\text{AMP}}^{\text{MZ}} \geq \log(70) \cup Y_{\text{SNR}}^{\text{MZ}} \geq \log(2.5)) \qquad \text{(H-5)} \\ &= 1 - \Pr(Y_{\text{AMP}}^{\text{MZ}} \leq \log(70) \cap Y_{\text{SNR}}^{\text{MZ}} \leq \log(2.5)) \\ &= 1 - \int_{-\infty}^{\log(70)} \int_{-\infty}^{\log(2.5)} f(y_1, y_2; \mu_{\text{AMP}}, \mu_{\text{SNR}}, \sigma_{\text{AMP}}, \sigma_{\text{SNR}}, \rho) dy_1 dy_2\end{aligned}$$

H.2.4 MAXIMUM LIKELIHOOD ESTIMATION.

The loglikelihood for the bivariate regression model is

$$L(\beta_0^{\text{MZAMP}}, \beta_1^{\text{MZAMP}}, \beta_0^{\text{MZSNR}}, \beta_1^{\text{MZSNR}}, \sigma_{\text{AMP}}, \sigma_{\text{SNR}}, \rho) = \sum_{i=1}^{n} L_i \qquad \text{(H-6)}$$

where for a Multizone inspection with log amplitude $y_{\text{AMP}i} = \log$ (EFBHSqMils) and log SNR $y_{\text{SNR}i}$

$$L_i = \log\left[\frac{f(y_{\text{AMP}i}, y_{\text{SNR}i}; \mu_{\text{AMP}}, \mu_{\text{SNR}}, \sigma_{\text{AMP}}, \sigma_{\text{SNR}}, \rho)}{1 - \int_{-\infty}^{y_{\text{TR}}^{\text{MZAMP}}} \int_{-\infty}^{y_{\text{TR}}^{\text{MZSNR}}} f(y_1, y_2; \mu_{\text{AMP}}, \mu_{\text{SNR}}, \sigma_{\text{AMP}}, \sigma_{\text{SNR}}, \rho) dy_1 dy_2} \right]$$

Again, the term in the denominator accounts for the possibility of unknown misses in field inspections.

The ML estimates of the parameters are obtained by maximizing equation H-6. Then the PoD is computed by substituting these estimates into equation H-5.

H.3 TRUNCATION LEVELS.

Following the approach used by Burkel, Sturges, Tucker, and Gilmore [H-2] in the earlier Default PoD study, the truncation level for Multizone is obtained by adding 10 dB to the observed noise level. For field finds, the noise level is recorded for a find. For the CBS data, the noise level was taken to be 20% of FSH for all observations.

The truncation level for the CBS conventional inspection data is a function of the Multizone inspection noise level because Multizone inspection information was used to identify conventional inspection misses. In particular, the Multizone noise level is converted to the EFBH scale according to Multizone calibration to a # 2 FBH. Then the corresponding % FSH under the conventional calibration to a # 3 FBH can be computed. This gives

$$\text{CalTo3}\%SH_{mzNL} = \text{CalTo2}\%SH_{mzNL} \times \frac{4.0}{9.0}$$

In particular,

$$y_{\mathrm{TR}}^{\mathrm{CV}} = \log\left(\frac{\mathrm{CalTo3}\%SH_{mzNL}+10}{80}\times\left(\frac{\mathrm{Cal}}{64}\right)^2\times\frac{\pi}{4}\times 10^6\right)$$

$$y_{\mathrm{TR}}^{\mathrm{MZAMP}} = \log\left(\frac{\mathrm{CalTo2}\%SH_{mzNL}+10}{80}\times\left(\frac{\mathrm{Cal}}{64}\right)^2\times\frac{\pi}{4}\times 10^6\right)$$

$$y_{\mathrm{TR}}^{\mathrm{MZSNR}} = \log\left(2.5\right)$$

H.4 CONFIDENCE BOUNDS.

Confidence bounds are used to quantify statistical uncertainty (i.e., uncertainty due to limited data). Confidence intervals can be viewed as a kind of statistical sensitivity analysis. Roughly speaking, one can ask what happens if one perturbs the data according to natural variability in the process that generates the data, as specified by the estimated statistical model. What kind of effect would such perturbation have on the final answers? It is important to recognize that confidence intervals reflect only statistical uncertainty and are based on the assumption that the assumed model is correct. Other forms of sensitivity analysis, wherein the model is perturbed, can be used to assess model uncertainties.

The simplest and most commonly used method to construct confidence intervals is via a normal distribution approximation where an estimate of a standard error of the function of interest (a standard error is a standard deviation of an estimator) is obtained by using a simple Taylor series propagation of error method, which statisticians call the delta method.

In general, PoD(a) for a given flaw size a can be expressed as a function of the model parameters θ and one can denote this function by $\mathrm{POD}(a;\theta)$. As explained in section H.1.5 for conventional inspection, the model parameters are $\theta = (\beta_0^{CV}, \beta_1^{CV}, \sigma_\varepsilon^{CV}, \beta_0^{\mathrm{MZ}}, \sigma^{\mathrm{MZ}}, \beta_0^{\mathrm{Miss2}}, \beta_1^{\mathrm{Miss2}})$. For the Multizone inspection the parameters are $\theta = (\beta_0^{\mathrm{MZAMP}}, \beta_1^{\mathrm{MZAMP}}, \beta_0^{\mathrm{MZSNR}}, \beta_1^{\mathrm{MZSNR}}, \sigma_{\mathrm{AMP}}, \sigma_{\mathrm{SNR}}, \rho)$. The maximum likelihood analyses provides estimates of θ, denoted by $\hat{\theta}$, and a matrix of the estimates of the variances and covariances of the parameter estimates. This estimated covariance matrix is denoted by $\hat{\Sigma}_{\hat{\theta}}$. An estimate of $\mathrm{PoD}(a;\theta)$, is obtained by evaluating at $\hat{\theta}$. Using the delta method, an estimate of the standard error of $\mathrm{PoD}(a;\hat{\theta})$ is

$$\hat{se}_{PoD(a;\hat{\theta})} = \sqrt{\left(\frac{\partial PoD(a;\theta)}{\partial \theta}\right)^T \hat{\sum}_{\hat{\theta}} \left(\frac{\partial PoD(a;\theta)}{\partial \theta}\right)}$$

where T denotes vector transpose and the partial derivatives are evaluated at $\hat{\theta}$. A transformation such as the logit transformation $\mathrm{logit}(p) = \log(p/(1-p))$ maps the 0-1 probabilities to the real line, corresponding to the range of the standard normal random variable.

By application of the delta method,
$\hat{se}_{\mathrm{logit}(\mathrm{PoD}(a;\hat{\theta}))} = \hat{se}_{\mathrm{PoD}(a;\hat{\theta})} / [PoD(a;\hat{\theta})(1-PoD(a;\theta))]$. Then a confidence interval based on the assumption that

$$\frac{\log it(PoD(a;\hat{\theta})) - \log it(PoD(a;\theta))}{\hat{se}_{\log it(PoD(a;\hat{\theta})}}$$

has approximately a normal distribution (statistical theory tells us that this approximation is good in large samples) can be computed as

$$\left[\frac{PoD(a;\hat{\theta})}{PoD(a;\hat{\theta}) + (1 - PoD(a;\hat{\theta})) \times w}, \quad \frac{PoD(a;\hat{\theta})}{PoD(a;\hat{\theta}) + (1 - PoD(a;\hat{\theta})) / w} \right]$$

where $w = \exp[z_{(1-a/2)} \hat{se}_{\text{logit}(\text{PoD}(a;\hat{\theta}))}]$.

The endpoints of the intervals constructed in this manner will always lie between 0 and 1. The confidence intervals shown in figures 18(a) and 18(b) were computed in this manner over all values of *a* that are of interest. More details on the delta method and construction of confidence intervals can be found in appendix B of Meeker and Escobar [H-3].

H.5 REFERENCES.

H.1 Thompson, R.B., Gray, T.A., and Meeker, W.Q., "Use of Physics-Based Models to Guide the Extrapolation of Aircraft Engine Ultrasonic PoD Data to Small Flaw Sizes," *Review of Progress in QNDE*, Vol. 25, D.O. Thompson and D.E. Chimenti, eds., American Institute of Physics, NY, 2006, pp. 1878-1885.

H.2 Burkel, R.H., Sturges, D.J., Tucker, W.T., and Gilmore, R.S., "Probability of Detection for Applied Ultrasonic Inspectors," *Review of Progress in QNDE*, Vol. 15B, D.O. Thompson and D.E. Chimenti, eds., Plenum Press, NY, 1996, pp. 1991-1998.

H.3 Meeker, W.Q. and Escobar, L.A., Statistical Methods for Reliability Data, John Wiley and Sons, New York, 1998.