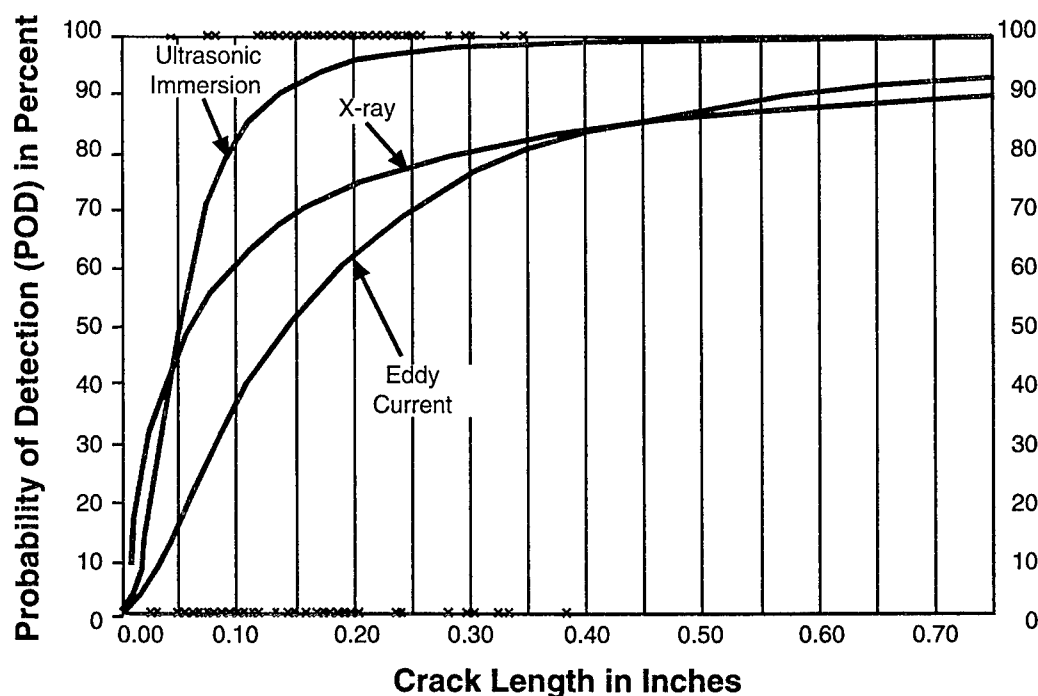


A Technology Assessment of

Probability of Detection (POD) for Nondestructive Evaluation (NDE)

NTIAC-TA-00-01



Comparison of POD for Three NDE Methods



Nondestructive Testing Information Analysis Center
A DoD Information Analysis Center Sponsored by the
Defense Technical Information Center (DTIC)

August 2001

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20020124 401

This document was prepared by the Nondestructive Testing Information Analysis Center (NTIAC), TRI/Austin, Inc., 415 Crystal Creek Drive, Austin, TX 78746-4725. NTIAC is a full service information analysis center operated under Contract SPO700-97-D-4003 for the U.S. Department of Defense, serving the information needs of the Department of Defense, other U.S. Government agencies, and the private sector in the field of nondestructive testing.

NTIAC is sponsored by the Defense Technical Information Center (DTIC) and administered by the Defense Information Systems Agency (DISA). Technical aspects of NTIAC operations are monitored by the Office of the Deputy Undersecretary of Defense (S&T).

Additional copies of this document may be obtained from:

NTIAC
415 Crystal Creek Drive
Austin, TX 78746-4725
Phone: (512) 263-2106 or (800) NTIAC 39
Fax: (512) 263-3530
E-mail: info@ntiac.com

ISBN 1-890596-20-5

This document was prepared under the sponsorship of the U.S. Department of Defense. Neither the United States Government nor any person acting on behalf of the United States Government assumes any liability resulting from the use or publication of the information contained in this document or warrants that such use or publication of the information contained in this document will be free from privately owned rights.

Use of trade names of manufacturers in this publication does not constitute an official endorsement of such products or manufacturers, either expressed or implied by the Department of Defense or NTIAC.

Approved for public release; distribution unlimited

All rights reserved. This document, or parts thereof, may not be reproduced in any form without written permission of the Nondestructive Testing Information Analysis Center.

Copyright© 2001, Nondestructive Testing Information Analysis Center.

Cover Acknowledgement: Rummel, W. D. and G. A. Matzkanin. "Nondestructive Evaluation (NDE) Capabilities Data Book, Third Edition." NTIAC-DB-97-02, Nondestructive Testing Information Analysis Center, 1997.

NTIAC-TA-00-01

A Technology Assessment of

**Probability of Detection (POD) for
Nondestructive Evaluation (NDE)**

By

George A. Matzkanin

And

H. Thomas Yolken



Nondestructive Testing Information Analysis Center
A DTIC-Sponsored DoD Information Analysis Center
Operated by TRI/Austin, Inc., *A Texas Research International Company*

August 2001

Approved for Public Release; Distribution Unlimited

Preface

This Technology Assessment was prepared under NTIAC Contract No. SPO700-97-D-4003 which is funded by the Defense Technical Information Center (DTIC). Preparation of Section 4 of this report was partially supported by the Air Force Research Laboratory Materials and Manufacturing Directorate under an NTIAC Technical Area Task (D.O. 0015). This Section was included as Appendix B in the Final Report for that project.

CONTENTS

	Page
Preface.....	iii
List of Figures.....	vii
List of Tables.....	xi
1.0 INTRODUCTION.....	1
1.1 Background	1
1.2 Scope of Technology	2
2.0 NDE CAPABILITIES QUANTIFICATION	3
2.1 Probability of Detection (POD)/Probability of False Alarms (POFA).....	3
2.1.1 NDE Process Output	3
2.1.2 Accept/ Reject Decisions From NDE Processes.....	5
2.2 Summary of Experimental Data in the NDE Capabilities Data Book, Third Edition.....	8
3.0 CURRENT STATUS OF EXPERIMENTAL DATA FOR POD/ POFA.....	11
3.1 Ceramic Inclusions in Ti Castings	11
3.2 Rail Flaws	12
3.3 Use of Actual In-Service Data	13
3.4 Transfer of POD Data from Simple Shapes to Complex Shapes	15
3.5 Recent Analytical Efforts	16
3.6 Modular Approach	17
3.7 Neural Networks	18
4.0 MODELS AND COMPUTER SIMULATION FOR POD/ POFA FOR NDE.....	21
4.1 Historical Perspective From the Literature	21
4.2 Recent Modeling Efforts	36
5.0 CONCLUSIONS AND PROGNOSIS	43
6 REFERENCES.....	45
7 BIBLIOGRAPHY.....	51

List of Figures

	Page
Figure 2-1	Signal responses for a single anomaly measurement3
Figure 2-2	Signal (Plus Noise) and noise response distributions4
Figure 2-3	Flaw detection at a threshold signal level4
Figure 2-4	Failure to detect smaller flaws at the same threshold signal level 115
Figure 2-5	Threshold decision level results in false calls6
Figure 2-6	Decisions from signal responses at low discrimination levels.....6
Figure 2-7	Conditional probability in flaw detection7
Figure 3-1	Probability of Detection (POD) as a function of shell diameter as obtained for the 0.904 inch titanium thickness (Four operators)12
Figure 3-2	Probability of detection from six rail flaw inspections.....13
Figure 3-3	Hit/miss inspection data from six damage locations15
Figure 3-4	Relative POD curve for the flat plate case and the adjusted curve for the complex shape case using equivalent threshold discrimination level criteria16
Figure 4-1	Probability of detection for circular cracks versus crack radius for three different thresholds, using a 10 MHz center frequency transducer.....22
Figure 4-2	Diagram of probability of detection model and its application to NDE system qualification and optimization and to computer- aided design for inspectability. POD, probability of detection23
Figure 4-3	Calculated POD curves for circular cracks at these depths below a cylindrical surface. (a) axial and circumferential seam increments of 2.5 mm (0.1 in.); (b) axial and circumferential seam increments of 5 and 1.3 mm (0.2 and 0.05 in.), respectively.....24

		Page
Figure 4-4	Example of a CAD-generated display of inspectability (POD) as a function of position within the cross section of a simulated disk	25
Figure 4-5	The predicted relative operating characteristics of eddy current NDE for crack openings of 1 mm, 0.75 mm, and 0.25 mm for curves 1, 2, 3, and 4, respectively 31.....	26
Figure 4-6	Eddy current inspection of the T37 front wing-spar lower cap lug. The probe is scanned in a circular arc over the beveled edge of the part. Cracks tend to occur in the region indicated.....	27
Figure 4-7	Calculated relative operating characteristics for an automated eddy current system.....	27
Figure 4-8	A sample POD prediction for an automated eddy current system	28
Figure 4-9	Schematic diagram a finite element based POD modeling approach	29
Figure 4-10	Probability density functions of eddy current signals with and without a flaw	30
Figure 4-11	A general Unified Life Cycle Engineering Environment.....	30
Figure 4-12	Conceptual model for NDT reliability. The information required to estimate reliability (POD, PFI) and the added value from an inspection is shown	32
Figure 4-13	Absolute amplitude and phase comparisons between model and experiment for a typical #4 flat bottom hole at normal-incidence using a transducer focused on a hole.....	33
Figure 4-14	Proposed ETC POD estimation methodology	34
Figure 4-15	UT measurements and theoretical model prediction. Experimental ultrasonic measurements and theoretical model predictions for cracks in heat exchanger tubing.....	37
Figure 4-16	Calculated POD curves for different ultrasonic detection threshold levels	37

		Page
Figure 4-17	General strategy of POD methodology. Shaded are as correspond to POD and PFA, as deduced from noise and signal distributions, for a particular threshold.....	38
Figure 4-18	General steps in developing modeling methodology for POD predictions	39
Figure 4-19	POD curves for three levels of NDE reliability.....	40

List of Tables

	Page
Table 4-1	PC-Based models for inspection reliability available in AEA Technology, showing physical model used to calculate signal in calculating POD and PFI35
Table 4-2	POF calculations for three different NDE reliability levels and three crack populations39

1.0 INTRODUCTION

1.1 Background

Assuring the integrity and reliability of engineering structures continues to challenge the technical community. Structures and materials deteriorate with time and accumulate a variety of defects such as cracks, corrosion, disbonds and delaminations. Exacerbating the situation, reduced federal and commercial procurement budgets are requiring the operational life of various aging assets to be extended. This is placing great importance on the ability to find, characterize and address the deleterious effects of operation in a wide variety of environments. Successful life management of these aging systems substantially depends on the ability of NDE to identify and quantitatively characterize defects and changes in the materials and structures throughout their lifetime.

Modern design of engineering components, structures and systems include the use of fatigue and fracture mechanics analyses to quantify damage tolerance and fitness for purpose (Rummel 2000). Application of damage tolerance assumes the presence of a flaw at all critical locations in a component/structure/ system. Structural integrity is based on the assumption that flaws of an assumed design size will not propagate to a size that could induce failure in service. The integrity of a component (fitness for service) is therefore dependent on the detection and removal of all flaws larger than the assumed design size before the component enters or re-enters service. For safety critical applications and for those designs incorporating damage tolerance as a design basis, quantification of NDE capabilities is required.

A frequently used quantitative measure of the capability of an NDE method is the Probability of Detection (POD) which gives the probability of detecting cracks of various lengths and depths (or thinning of parts due to corrosion) under various inspection conditions. The development and evolution of the POD metric has resulted in a much better understanding of NDE procedures and the sensitivity of individual procedures to changes in materials, applications and processing parameters. Knowledge of the POD provides the engineer with a useful metric for quantifying and assessing NDE capabilities.

To archive POD data obtained for common NDE procedures, in 1997, the Nondestructive Testing Information Analysis Center (NTIAC) published an engineering reference data book, the *NDE Capabilities Data Book, Third Edition* (Rummel and Matzkanin 1997). This Data Book was compiled from available quantitative NDE demonstrated capabilities data to provide a convenient single reference source for NDE engineering analysis. It offers an aid in the selection of an NDE procedure and demonstration of the capabilities of that procedure.

A very good recent overview of NDE reliability and POD can be found in Singh 2001. The author describes a generic protocol plan for NDE assessment developed under an Air Force contract. The design of experiments approach is used to quantify various influencing variables, particularly the human factors. Three human factors were addressed-job type, schedule comfort and physical comfort. A protocol demonstration study provided information on the working of the protocol as well as insight into the test process. The key objective of consistency in conducting NDE reliability studies and reporting the data was achieved. Although the protocol was developed for the U.S. Air Force, it is comprehensive and general enough for a much wider application to various other industries. The protocol is expected to help in conducting various

reliability studies with economy in execution and consistency in data acquisition and information management.

1.2 Scope of Technology Assessment

In the next section, Section 2.0, a brief discussion is presented of NDE capabilities quantification, including a summary of the data in the *NDE Capabilities Data Book, Third Edition*. In Section 3.0 information is provided on the current status of experimental POD measurements focusing on efforts reported in the literature since publication of the Third Edition of the *NDE Capabilities Data Book* in 1997.

In Section 4.0 an overview and assessment is provided of models and computer simulation of POD for NDE. It has been shown to a limited degree that NDE measurements can be modeled or simulated in a computer to determine PODs and to optimize NDE methodologies. In addition, in recent years there has been a steady increase in the ability of physical models to accurately predict the results of inspections of real parts. As a result, the opportunity now exists to incorporate such models. In Section 4.0 an extensive discussion is presented of the historical development of computational NDE and POD modeling and summaries are then provided of recent efforts in this area. An extensive bibliography of NDE mathematical modeling can be found in a recent NTIAC publication, "Overview of Mathematical Modeling in Nondestructive Evaluation" (Aldrin 2001).

The Technology Assessment concludes with a Summary and Prognosis in Section 5.0 and a List of References in Section 6.0. Additional relevant documents of interest are included in the Bibliography.

2.0 NDE CAPABILITIES QUANTIFICATION

2.1 Probability of Detection (POD)/Probability of False Alarms (POFA)

The following information has been adapted from the *NDE Capabilities Data Book, Third Edition* (Rummel and Matzkanin 1997).

2.1.1 NDE Process Output

When an NDE process is applied to a test object, the output response to an anomaly within the test object will depend on the form of detection (pattern recognition), the magnitude of the feature that is used in detection, and the relative response magnitude of the material surrounding the anomaly. For example, in an ultrasonic inspection procedure, the amplitude of the response from an anomaly within a structure may be used to discriminate the response from the grain structure (noise) surrounding the anomaly as shown in Figure 2-1. If the ultrasonic procedure (measurement) is applied repetitively to the same anomaly, a distribution of responses to both the anomaly and the surrounding material (grain structure) will be obtained as shown schematically in Figure 2-2. The measured response distribution reflects the variance in the NDE measurement process and is typical of that obtained for any measurement process. The response from the surrounding material constitutes the baseline level for use in discrimination of responses from internal anomalies. The baseline response may be termed "noise" (far different from electronic noise that is applied to the measurement instruments) and both the discrimination capability (anomaly detection) and anomaly sizing (quantification) capabilities for the NDE procedure are dependent on the relative amplitudes and the rate of change of the anomaly response with increasing anomaly size (slope). For purpose of discussion, the signal (plus noise) response will be referred to as the signal and the baseline response from the surrounding material will be referred to as noise.

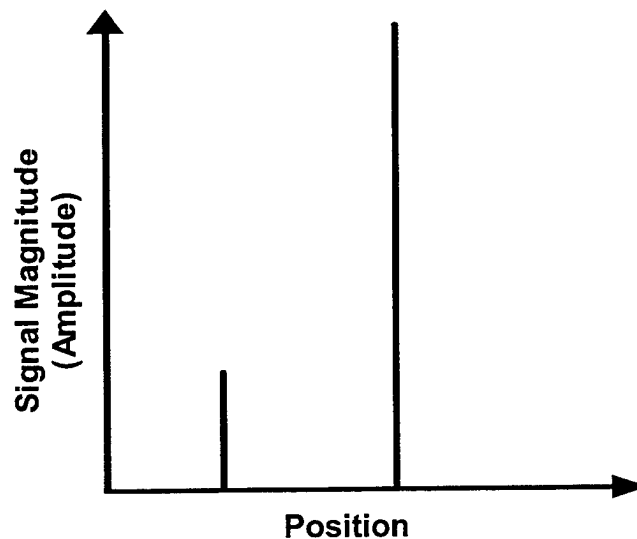


Figure 2-1 Signal responses for a single anomaly measurement
(Rummel and Matzkanin 1997)

The considerable flaw to flaw variance and variance in signal response to flaws of equal size causes increased spread in the probability density distribution of the signal (plus noise) response. If a threshold decision (amplitude) level is applied to the responses shown in Figure 2-2, clear flaw discrimination (detection) can be achieved as shown in Figure 2-3. If the same threshold decision level (acceptance criteria) is applied to a set of flaws of a smaller size (as shown in Figure 2-4), clear discrimination cannot be accomplished.

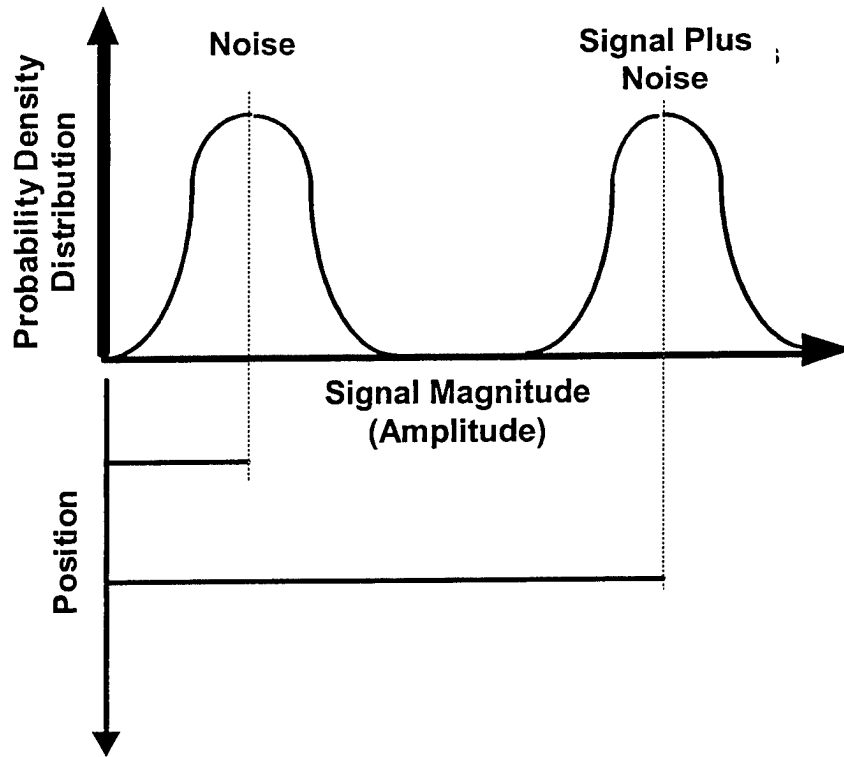


Figure 2-2 Signal (Plus Noise) and noise response distributions (Rummel and Matzkanin 1997)

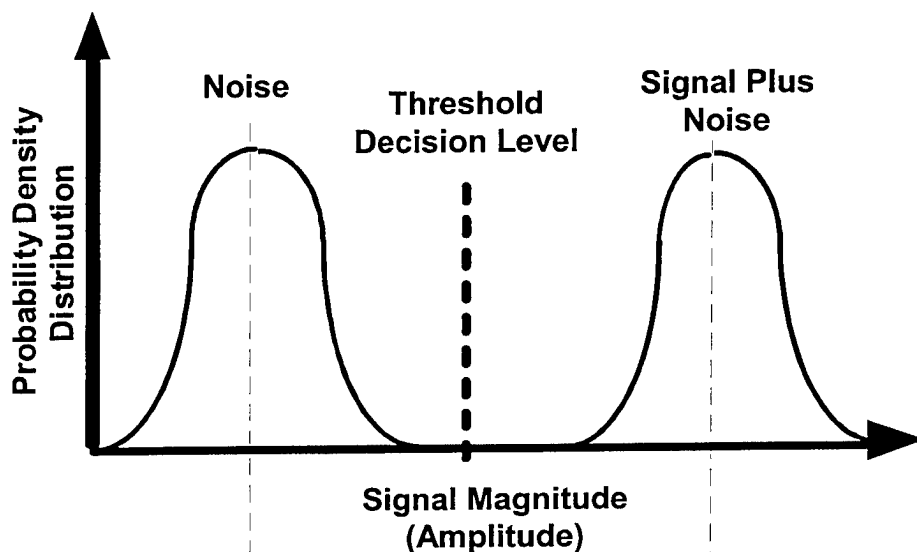


Figure 2-3 Flaw detection at a threshold signal level (Rummel and Matzkanin 1997)

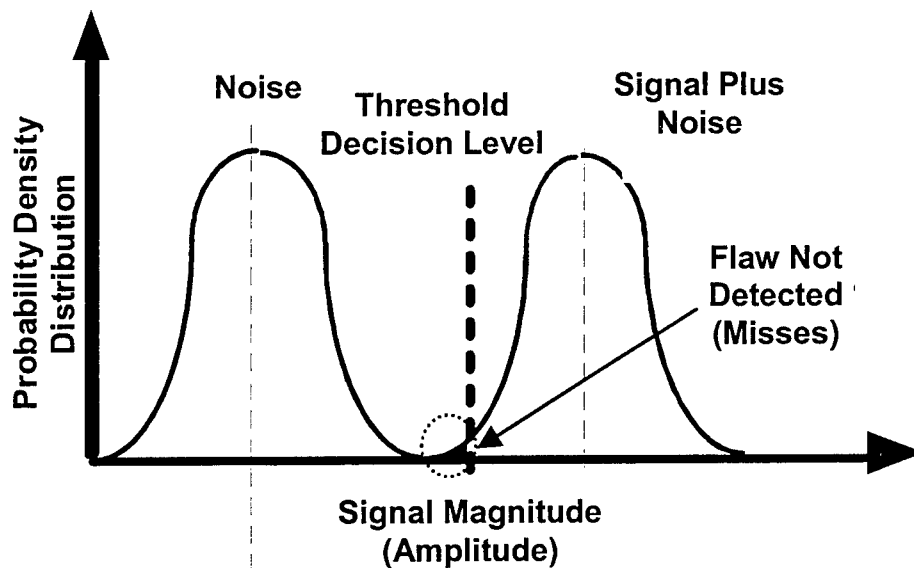


Figure 2-4 Failure to detect smaller flaws at the same threshold signal level
(Rummel and Matzkanin 1997)

In the example shown, the threshold decision level could be adjusted to a lower signal magnitude to produce detection. As the signal magnitude is adjusted downward to achieve detection, a slight increase in the noise level will result in a **false call (false alarm)**. As the flaw size decreases, the noise and signal (plus noise) responses will overlap. In such cases, a downward adjustment in the threshold decision level (to detect all flaws) will result in an increase in false calls. Figure 2-5 shows an example where the threshold decision level (acceptance criteria) has been adjusted to a level where a significant number of false calls will occur. In this example, a slight change in flaw signal distribution will also result in failure to detect a flaw. The NDE procedure is not robust and is not subject to qualification or certification for purposes of primary discrimination. The procedure may, however, be useful as a prescreening tool, if it is followed by another procedure that provides discrimination of the residuals; for example, a neural network detection process may be structured to provide discrimination at a high false call rate, but may be a useful in-line tool if other features are used for purposes of discrimination after the anomaly or variance is identified.

2.1.2 Accept / Reject Decisions From NDE Processes

It is clear that Accept /Reject decisions resulting from the application of an NDE procedure may result in both detection failures (**misses**) and false position detection (**false calls**) when the NDE procedure is operated near the limit of discrimination as shown in Figure 2-5. Theory and analysis methodologies were developed during World War II to predict the performance of radar operators in aircraft detection. The NDE discrimination / detection task is similar in nature and the same principles may be applied. From decision theory, if we assume a background signal response distribution (noise), a signal (plus noise) distribution, and a threshold decision level (acceptance criteria) as shown in Figure 2-6, it is clear that the output will be a combination of accept, reject, misses, and false calls.

The result of decisions from signal responses at low discrimination levels may be analyzed as a problem in conditional probability. Figure 2-7 is a convenient aid in visualizing a problem in conditional probability (as contrasted to joint probability - black / white).

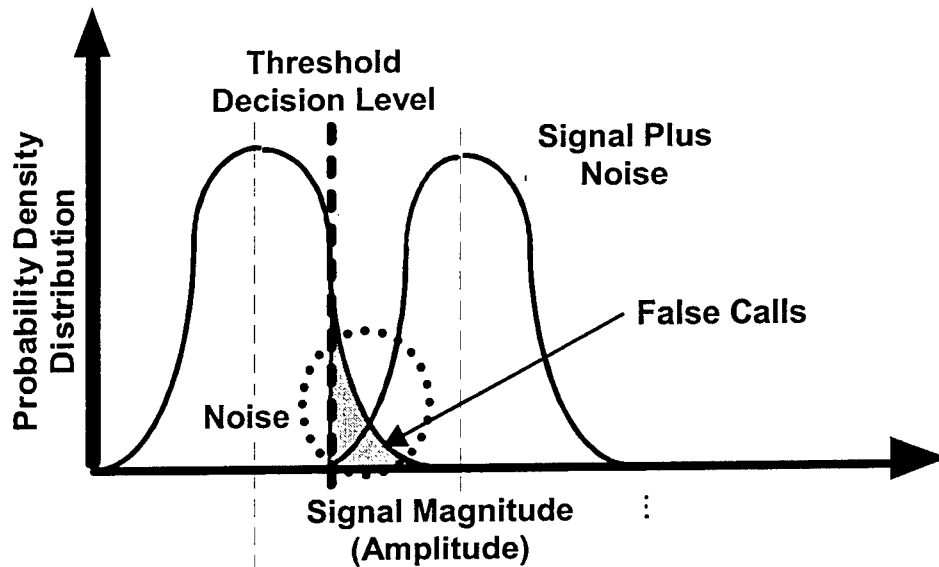


Figure 2-5 Threshold decision level results in false calls (Rummel and Matzkanin 1997)

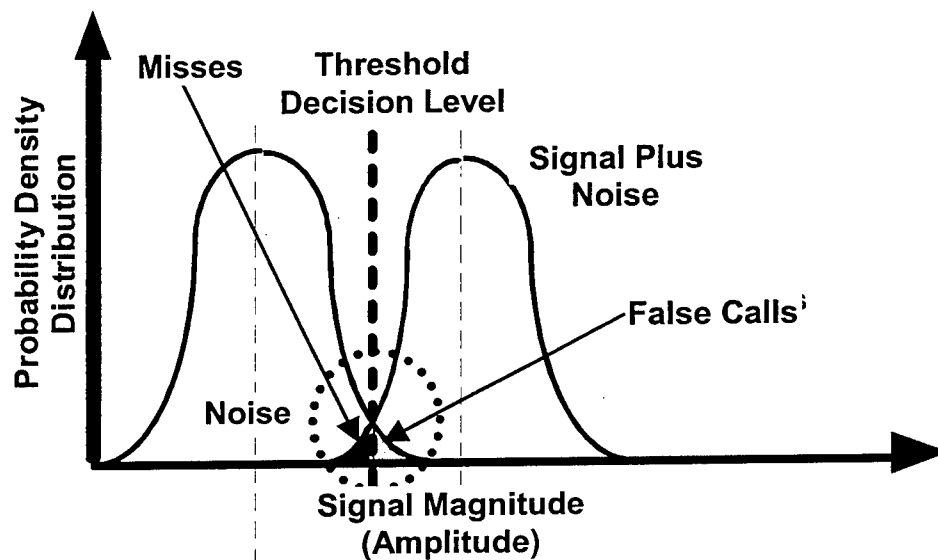


Figure 2-6 Decisions from signal responses at low discrimination levels (Rummel and Matzkanin 1997)

		Stimuli (Flaw Presence)	
		POS a	NEG n
NDE Signal (Flaw Response)	POS A	$M(Aa)$ True Positive (T.P.) (Flaw/Flaw) $P(A,a)$ (NO ERROR)	$M(An)$ False Positive (F.P.) (Flaw / No Flaw) $P(A,n)$ (TYPE II ERROR)
	NEG N	$M(Na)$ False Negative (F.N.) (No Flaw / Flaw) $P(N,a)$ (TYPE I ERROR)	$M(Nn)$ True Negative (T.N.) (No Flaw / No Flaw) $P(N,n)$ (NO ERROR)

**Figure 2-7 Conditional probability in flaw detection
(Rummel and Matzkanin 1997)**

The outcome of the NDE procedure and decision may be:

TRUE POSITIVE (T.P.),

where $M(Aa)$ is the total number of T.P. calls;
and $P(A,a)$ or $P(T.P.)$ is the probability of T.P. calls.
(Flaw Found when a Flaw is Present)
(NO ERROR CONDITION-REJECT DECISION) - **CORRECT REJECT**

FALSE POSITIVE (F.P.),

where $M(An)$ is the total number of F.P. calls;
and $P(A,n)$ or $P(F.P.)$ is the probability of F.P. calls.
(Flaw Found when no Flaw is Present)
(TYPE II ERROR CONDITION - REJECT DECISION) - **FALSE CALL**

FALSE NEGATIVE (F.N.),

where $M(Na)$ is the total number of F.N. calls;
and $P(N,a)$ or $P(F.N.)$ is the probability of F.N. calls.
(No Flaw Found when a Flaw is Present)
(TYPE I ERROR CONDITION - ACCEPT DECISION) - **MISS**

TRUE NEGATIVE (T.N.),

where $M(Nn)$ is the total number of T.N. calls;
and $P(N,n)$ or $P(T.N.)$ is the probability of T.N. calls.
(No Flaw Found when no Flaw is Present)
(NO ERROR CONDITION - ACCEPT DECISION) - **CORRECT ACCEPT**

Interdependence of the matrix quantities is denoted by:

$$\begin{aligned} \text{T.P.} + \text{F.N.} &= \text{Total opportunities for positive calls.} \\ \text{F.P.} + \text{T.N.} &= \text{Total opportunities for negative calls.} \end{aligned}$$

Therefore, only two independent probabilities need be considered in alternate inspection / decision tasks.

The specificity of an NDE procedure or the Probability of Detection (POD) of flaws may be expressed as:

$$\text{POD} = \frac{\text{T.P.}}{\text{T.P.} + \text{F.N.}} \quad \text{or} \quad \frac{\text{Total Number of Positive Calls (REJECTS)}}{\text{Total Number of Opportunities for Rejection}}$$

In like manner, the nonspecificity of an NDE procedure or the Probability of False Alarms (False Calls) may be expressed as:

$$\text{POFA} = \frac{\text{F.P.}}{\text{T.N.} + \text{F.P.}} \quad \text{or} \quad \frac{\text{Total Number of False Positive (FALSE CALLS)}}{\text{Total Number of Opportunities for Acceptance}}$$

Application of the method requires the use of flawed test objects and assumes that all flaws are of equal size and that the variance in flaw response distribution is due solely to the measurement process. Confidence limits for the respective probabilities may be calculated from the data sample size used in the test case.

In practical applications, the NDE process characteristic that is of primary interest is the probability of detection (POD). The acceptability of a false call rate (POFA) will be dependent on the consequence of a false call for a specific application. If false calls require significant efforts for resolution, a low level of false calls will be required. If an economical and efficient secondary method is used to resolve false calls, acceptance may become part of end to end production process requirements.

2.2 Summary of Experimental Data in the *NDE Capabilities Data Book, Third Edition*

The Third and latest Edition of the *NDE Capabilities Data Book* consolidates and organizes available reference data for demonstrated NDE performance capabilities into a single source. Guidelines are presented for selecting options for use of NDE and for assessing the potential to meet design requirements (critical flaw detection requirements). Guidelines for demonstration of specific NDE process capabilities are also presented.

Following a 65 page text (7 chapters) describing various aspects of NDE capabilities quantification, probability of detection (POD), and damage tolerance concepts, 423 POD curves are organized and presented in a series of Appendices organized by NDE method. A documentation page precedes each dataset and provides a condensed description of the test object, test artifacts, NDE procedures and results summary. The POD curves for varying test object, test artifact and data collection conditions follow the documentation page. POD data are generally presented as a function of crack length. For selected data- sets, POD data are also presented as a function of crack depth and crack depth-to-thickness ratio. POD curves are based

on hit/miss data using the log-logistic model. Original reference source information is provided for each dataset.

NDE procedure capabilities included in the Data Book are:

ET - Eddy Current Inspection	MT - Magnetic Particle Inspection
UT - Ultrasonic Inspection	VT - Visual Inspection
RT - X-Radiographic Inspection	ZT - Emerging Inspection Processes
PT - Liquid Penetrant Inspection (visible and fluorescent)	

Materials covered in the **First Edition Data Book** include:

Aluminum (2219 T-87 and 2024 T-37)
Stainless steel (AMS 355)
Titanium-6Al4V

POD curves added in the **Second Edition Data Book** for specific applications include:

4340 Steel Flat Plate Panels
Bolt Holes in J85 Seventh Stage Compressor Disks
Visual Inspection of Fatigue Cracks in Inconel 718 and Haynes 188 Flat Plates
X-Radiography of 0.060 Inch Thick and 0.250 Inch Thick 4340 Steel Flat Plates
"Edge of Light" Inspection of Bolt Holes in J85 Seventh Stage Compressor Disks

POD curves added to the **Third Edition Data Book** for specific applications include:

Aircraft Stiffened Stringer Panels
Lack of Penetration Defects in Aluminum Alloy GTA Welds
Longitudinal and Transverse Fatigue Cracks in Welds with Crowns
Longitudinal and Transverse Fatigue Cracks in Flush Welds
Water Washable Fluorescent Penetrant on Haynes 188 Flat Panels

The NDE Capabilities data Book is available in hard copy or on CD from NTIAC.

3.0 CURRENT STATUS OF EXPERIMENTAL DATA FOR POD/POFA

Since compilation and publication of the *NDE Capabilities Data Book, Third Edition* in 1997, work has continued in various laboratories to determine and measure the POD for a variety of NDE applications and to broaden the application of POD concepts. A brief list of selected recent POD determinations follows:

- Synthetic hard alpha inclusions in Ti (Meeker et al. 1998)
- Eddy current inspection of aircraft wheels (Okure and Peshkin 1995)
- Weep hole cracks in wing structures (Aldrin 2000)
- Thick section welds (Munns 2000)
- Ceramic inclusions in Ti castings (Child et al. 1999)
- Magnetic flux leakage in pipelines (Zhang et al. 1997)
- Rail flaws (Peterson et al. 2000)
- Wing splice joints (Grills 2001)

Review of recent literature on NDE reliability and POD research also reveals the following research thrust areas:

- Use of actual in-service data/extraction of POD from small data sets
- Transfer of POD from simple shapes to complex shapes
- Advanced numerical analysis
- Modular methods
- Neural networks

Brief summaries of several of these studies will be presented in the following sections. Additional information can be found in the cited references.

3.1 Ceramic Inclusions in Ti Castings

Results were recently presented of a systematic assessment to quantify the detectability of various ceramic materials in thick titanium castings (Child et al. 1999). The use of titanium-hot isostatic pressed (HIP) castings offers tremendous potential as a design option in damage tolerant aircraft structures. However, producibility of these parts necessarily includes knowledge of the capabilities and limitations of nondestructive inspection procedures used in production and acceptance. During the production process, ceramic facecoat may spall and be incorporated into the casting as a ceramic inclusion (termed "shell"). Child et al. assessed X-radiography applied to the detection of ceramic inclusions in the castings. Results were analyzed in the form of a POD as a function of shell diameter, for each anticipated shell thickness and titanium thickness. Figure 3-1 shows the results at the 0.904-inch titanium thickness level for four different facecoat (shell) formulations (A, B, C and D) from different vendors. Results of the study were used to improve facecoat formulations for increased detectability and the work was eventually expanded to seven different facecoat formulations.

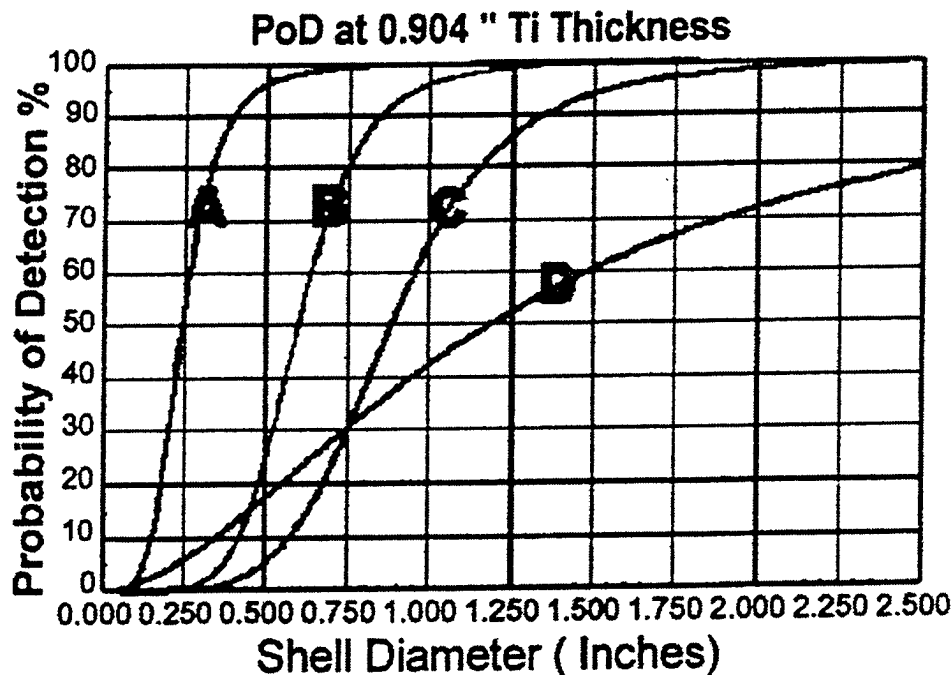
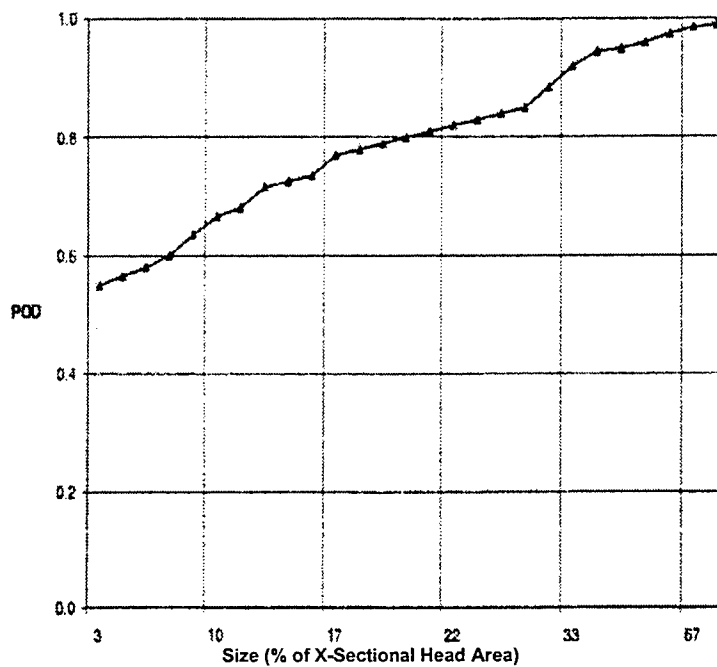


Figure 3-1 Probability of Detection (POD) as a function of shell diameter as obtained for the 0.904 inch titanium thickness (Four operators)
(Child et al. 1999)

3.2 Rail Flaws

A study to determine limitations in the size and type of detectable defects in rail flaw inspection was recently performed by Peterson, et al. (Peterson et al. 2000). The data reported is based on work performed at the Transportation Technology Center, Inc. in Pueblo, CO. Six major US railroads provided 56 flawed rail sections that had been removed from service. The rail, containing a variety of internal and surface anomalies, was placed into a test track which was then used to test and compare the performance of commercial rail flaw detection systems. The 56 flawed rails were inspected visually, ultrasonically and in some cases radiographically before they were joined in the track. In all 49 defects were cataloged on the test loop. Of these 49 defects 44 were transverse in nature and measured and sized in a similar fashion. Because common sizing was applicable, only these 44 defects were included in the statistical evaluations. Six test runs were made over the test loop with a commercial ultrasonic inspection vehicle. The results were statistically analyzed to determine significant relationships in the data.

Probability of Detection plots were obtained using a SAS statistical analysis program (Jeffrey and Peterson 1999). The overall POD using results from all six evaluations, which gives a sample size of 264 is shown in Figure 3-2. Results show that as the defect size increases the probability of detecting that defect increases. Comparison with the inspection reliability standards from the American Railway Engineering Association (AREA) showed that the POD curve is lower than the AREA inspection reliability curve for defects smaller than 30 % of cross-sectional head area. Thus, one can conclude the ultrasonic inspection technology has a lower probability of detecting smaller flaws than the expected AREA recommended industry reliability.



**Figure 3-2 Probability of detection from six rail flaw inspections
(Jeffrey and Peterson 1999)**

3.3 Use of Actual In-Service Data

As mentioned earlier, the most common approach for determining POD is to perform inspections on representative components or specimens simulating the actual parts. While this approach is practical, it is usually very expensive. A more economical approach may be to use actual field inspection data to obtain POD. This approach is particularly attractive for cases where the component or structure cannot be easily simulated, e.g., airframe inspection. However, there are difficulties with this approach. First, there is usually a very limited amount of field data which may require special statistical treatment; and second, crack growth data must exist to allow the estimation of flaw sizes at the inspection sites at inspection times before the flaws were found. Forsyth, et al. (2000) describe the problems associated with POD development from small data sets, and the sensitivity of POD models to sample size is analyzed using two different statistical functions. Inspection data from a full-scale test are analyzed with the view of developing POD, and the difficulties encountered are discussed.

Forsyth et al. argue that field inspection information can be employed to obtain hit-miss data which can be used in turn to estimate the POD for that particular inspection. Inspections at a particular site are recorded over time or operational cycles until a flaw is found. This gives a hit, and an estimate of the flaw size is usually obtained either from the inspection result itself, or by performing a secondary inspection, or by disassembly and verification tests. The previous inspections at this test site can be used to predict miss points, by estimating the size of the flaw at the inspections performed before the flaw was found. This is a complex procedure that requires knowledge of initial flaw sizes, flaw growth rate, and the loading at the site.

Forsyth et al. identify five main problems associated with the use of field data for POD estimation. These are:

1. Uncertainty in "backcasting" flaw sizes
2. Uncertainty in crack growth
3. Flaw size estimation at time of detection
4. Uncertainty in operational conditions
5. POD model sensitivity to small sample sizes

The first four of these problems were investigated by Leemans (1998). In dealing with the fifth problem, Forsyth et al. evaluated two different POD models for their sensitivity to sample size; a traditional two-parameter curve fit model and Spencer's extension to the curve fit model (Spencer 1998). Actual data from eddy current inspections of bolt holes in a turbine disk were used to make sample data sets of varying sizes. As the sample size decreases, the spread between the mean and the 95% confidence level increases, and the effect of outliers is increased. The Spencer model showed less sensitivity to the decreasing number of data points and to the outliers than the traditional curve-fit method. However, in both cases, there was a wide variability in the mean and 95% confidence curves at small sample sizes as evidenced by large standard deviations.

The viability of using actual inspection data for POD determination was examined using inspection data from a Canadian trainer aircraft that was monitored during full-scale fatigue tests to 75,000 hours. A number of cracks were found using simple visual inspection and some were monitored periodically over time. Difficulties arose in interpreting crack growth data at many of the sites because it was found that after fastener removal, many of the holes thought to have multiple cracks actually had cracks that had joined underneath the fastener. This made it impossible to determine the total crack growth rate from measurements made with the fasteners in place. As a result, only a few flaw sites remained with valid data. Figure 3-3 shows the hit and miss points plotted from analysis at six damage locations. Due to the irregular nature of the small data set, a reasonable POD curve could not be constructed. Forsyth et al. attribute the problems in POD determination to the small data sets typical of field applications as well as uncertainty in flaw size and growth rate estimation. They argue that despite these problems, the estimation of the POD-flaw size relationship using field inspection data remains attractive, as the available alternatives are very expensive.

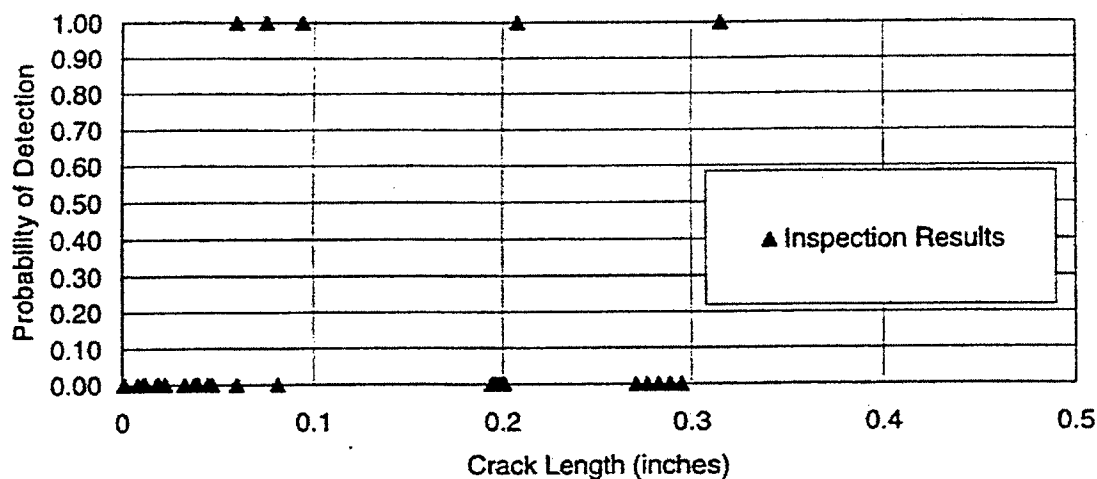


Figure 3-3 Hit/miss inspection data from six damage locations (Forsyth et al. 2000)

3.4 Transfer of POD Data from Simple Shapes to Complex Shapes

Fatigue cracks in simple test specimens are frequently used as the test artifacts for POD determination. However, simple test specimen geometries may not be representative of the NDE challenges in a complex structure or system and methodologies for transfer of the measured capability to complex shapes are required. Rummel (1999) has described an experimental procedure for transferring NDE procedure performance (POD) capabilities, which have been validated on simple specimens, to complex configurations found in field applications. The method is based on, and is an extension of, methods used in routine "calibration"/set-up of inspection procedures.

Slots are often used in establishing and applying NDE procedures used in crack detection and may be induced in large structures and complex shapes that would be difficult to crack for purposes of determining crack detection capabilities (POD). However, in most cases, the response from a slot is not equivalent to the response from a crack and the capability for crack detection must be determined experimentally. In the transfer methodology proposed by Rummel, a quantitative NDE response relationship is experimentally generated using equivalent size slots in both flat plate and complex specimen configurations, and equivalent size cracks in simple (flat plate) specimens. Transfer of the respective slot and crack responses is then a similar procedure to that used for initial "calibration"/set-up. Once a relationship between responses to slots and crack is established, and acceptance criteria are established at a point where the signal can be distinguished from the noise response, the POD can be determined. The POD for the complex shape inspection is adjusted to provide the same signal/noise discrimination as that used for the flat plate case. This may be accomplished by simple determination of equivalent signal response at the threshold POD level or a new POD curve may be generated using the signal response data and the equivalent relative acceptance threshold. A typical adjusted POD curve is shown in Figure 3-4.

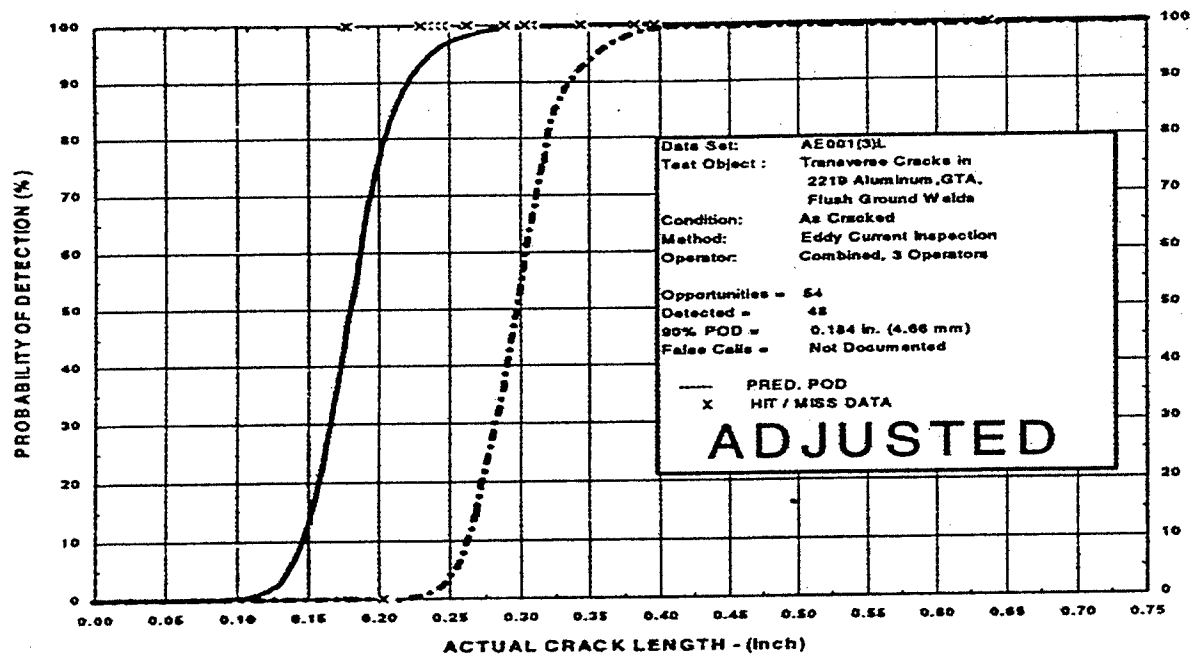


Figure 3-4 Relative POD curve for the flat plate case and the adjusted curve for the complex shape case using equivalent threshold discrimination level criteria (Rummel et al. 1999)

Rummel cautions that rigor in application of the transfer method described is required and documentation of each data acquisition and calculation step is necessary for both process control and for future re-validation. Particular care must be taken in generating equivalent slots in materials that are degraded by local heating during the slotting process. It is recommended that slot generation procedures be assessed to assure that the same size slot produces functionally equal responses in flat plate specimens, before slotting the test specimens to be used for POD adjustment.

3.5 Recent Analytical Efforts

Several recent efforts have focused on analysis of POD methodology, in particular, sources of variation and the effect of repeated inspections. Spencer (2001) reviews the fundamental concepts of using regression type data to establish POD curves and discusses the use of, and some common misconceptions about, confidence bounds. He also discusses components of variation and confidence statements, and how these concepts are related to the number of discontinuities in specimen sets and the number and nature of repeat tests. His analysis includes three main sources of variation that arise from the process of gathering experimental data: 1) discontinuities of the same size will produce different signals; 2) different inspectors will obtain different signals on the same discontinuity, even when using the same technique; 3) upon re-testing, the same inspector will obtain different signals on the same discontinuity, especially if the process of equipment setup is repeated with each test.

Spencer concludes that the fundamental input for the POD estimation is the number of discontinuities. Regardless of how many tests are performed on the same set of discontinuities, it is the discontinuity to discontinuity variation that should determine a minimum difference between \hat{a}_{90} and $a_{90/95}$. This difference can be reduced by including more discontinuities in the

specimen set, thereby decreasing the uncertainty of estimating model parameters, but not by increasing the tests on the same discontinuity set. Variations other than between discontinuities are incorporated into the analysis only to the extent that it is reflected in a way that data are gathered. If between inspector variation is a significant factor, and one gathers data with a single inspector, the risk is taken that the estimated pod curve is overly conservative or overly pessimistic with respect to the average system response across all inspectors. Inclusion of multiple inspectors will ensure a better estimate of the average POD curve if these inspectors properly reflect the population of users. However, to characterize the inspector variation, the estimation of POD curves needs to be done for each inspector. Thus if the intent is to obtain information on the average POD across an inspector population, then multiple inspectors should be included. However, there should not be an imbalance of tests by each inspector. That is, don't include four tests by one inspector and only one test by another. This warning against imbalances carries over to other factors as well.

Forsyth and Fahr (2000) present experimental POD data demonstrating the marginal benefit in performing multiple inspections and errors in assuming independence of repeated inspections are explained. Two repeated inspection scenarios were examined. In the first case, several inspections were carried out to determine the POD of a liquid penetrant technique applied to the detection of cracks in the bolt holes of turbine disks. There were 381 inspection opportunities in the data set with 320 flaws. These data were used as the basis for simulating multiple penetrant inspections. In the second case, a test of the independence of multiple inspections was performed by combining the results of three different eddy current procedures. The procedures were manual inspection and human interpretation, automated inspection and human interpretation, and automated inspection and interpretation using pattern recognition. Both the simulated and actual inspection data demonstrated for very different inspection techniques that although the combined inspection data resulted in a higher POD value than the best individual inspection, independence and thus the benefit of repeated inspections are much less than that predicted by the assumption of full independence.

3.6 Modular Approach

A modular approach to NDE reliability has been proposed by Mueller et al. (2001). In this approach, a system is divided into appropriate modules and the discrete reliability of each module is determined based on the nature of each module's information. The knowledge gained within each of the modules allows an optimization of the total system. The reliability of the total system is then determined by combining the individual reliabilities of the modules, including their possible correlation. The modular approach facilitates direct integration of the reliability formula derived at the American-European Workshop.

$$R = f(IC) - g(AP) - h(HF)$$

where R is the total reliability, IC is an intrinsic capability describing the physics and basic capabilities of the devices, AP consists of factors of industrial application such as restricted access in the field, and finally HF describes the human factors.

Mueller et al. discuss two examples in applying the modular approach to determine system reliability. In the first example, the POD of radiographic detection of thermally induced cracks in welds of ferritic tubes in nuclear power plants is examined. As a metric for each module, POD is used as a function of crack depth. The POD for the IC function was modeled based on

the physics of the source and the interaction of the x-rays with the material and the discontinuities. The POD associated with the AP + HF function was determined from experimental results of film interpretation by seven inspectors. In the second example, the POD was determined for the detection of porosity in electron beam welds of titanium alloy aircraft engine parts using automated ultrasonic testing with focussed probes. Both examples demonstrate the feasibility of the modular approach for practical applications and illustrate how this approach provides an opportunity to study the principal factors influencing optimization of the NDE system.

3.7 Neural Networks

Recognizing that flaw characterization and discontinuity classification are important in POD determination, several workers have recently examined use of neural networks for this application. Schmerr and Song (2000) point out that trying to determine the type of discontinuity that produces a particular measured NDE response is a difficult NDE classification problem and one that is ideally suited to the application of modern techniques such as neural networks. The various approaches used in the past for solving NDE discontinuity classification problems have resulted in little difference in performance. Schmerr and Song argue that probabilistic neural networks (PNNs) have a number of characteristics that make them an ideal choice for solving NDE discontinuity classification problems. The authors describe the general architecture of a PNN and the use of PNNs in solving a variety of ultrasonic and eddy current discontinuity classification problems. They discuss the important role that feature determination and selection play in obtaining good classification results and the role that modeling in conjunction with PNN can have in improving those features and in obtaining high performance discontinuity classification results.

A systematic approach to flaw characterization was used by Song et al. (1999) for inversion of eddy current flaw signals using a novel combination of neural networks and finite element models. Specifically, finite element models were developed to predict eddy current testing signals from axisymmetric flaws in tubes and the accuracy of the models was experimentally verified. The tubes were Inconel, 19 mm OD with a wall thickness of 1.3 mm. Using the finite element models, Song et al. constructed a database composed of 216 eddy current flaw signals simulated from 108 axisymmetric circumferential flaws of four types. From each flaw 30 features were extracted. Of these, a set of 11 features was selected for flaw classification and another set of 10 features selected for flaw sizing. As a robust inversion tool for eddy current flaw signals, an intelligent flaw classification system was used that can determine the flaw type and flaw size parameters from the selected features. This is a hybrid system composed of two different paradigms of neural networks: probabilistic neural networks for flaw classification, and back propagation neural networks for flaw sizing. Song et al. suggest that the excellent results obtained demonstrate the good possibility for this system as a robust tool for practical flaw characterization in tubes.

In a recent report, Aldrin et al. (2000) describe a neural network assisted, automated ultrasonic inspection technique for application to detection of cracks in fuel tank weep holes in C-141 aircraft wing structures. A numerical technique was developed and validated to calculate simulated waveforms for reflection and backscattering of ultrasound from top or bottom cracks emanating from the weep holes. From simulation and experimental studies, algorithms were developed to determine the number and location of weep holes in a scan, and to detect the presence of either or both bottom and top cracks for each weep hole. The bottom crack

algorithm incorporates three successive neural networks for refined crack detection. Both top and bottom crack neural networks incorporate data from multiple transducer locations for robustness to geometric noise and weep hole size. The automated ultrasonic procedure using neural networks was successfully implemented and validated on a broad data set. Weep holes in a section of a C-141 wing containing 25 holes with bottom and top EDM notches were correctly classified as well as samples with and without saw cuts. Preliminary studies with fatigue crack samples showed that the automated procedure is also valid for real cracks. The feasibility of the automated system for top and bottom crack detection was shown to exceed the desired capability of detecting cracks down to 0.070 inch. Work is in progress to determine the POD for the completely automated procedure.

4.0 MODELS AND COMPUTER SIMULATION FOR POD/POFA FOR NDE

It is an arduous and expensive task to experimentally determine POD curves for the wide variety of available NDE methods. Each method must be individually investigated for each specific material or application, and for a variety of flaw types and sizes. The expense is also rapidly increasing due to the introduction of new or improved NDE methods and new and improved materials.

It has been shown to a limited degree that NDE measurements can be modeled or simulated in a computer to determine PODs and to optimize NDE methodologies. In addition, in recent years there has been a steady increase in the ability of physical models to accurately predict the results of inspections of real parts. As a result, the opportunity now exists to incorporate such models into new computer simulation procedures for POD determinations that reduce time-cost constraints. Enhanced, computational POD and PFA capabilities have a number of additional payoffs. These models can be used during part and process design to help optimize NDE capability for finished parts and systems. The models can also be used to optimize NDE measurement methodology and procedures, to help validate NDE systems, and to develop simulations which can be used to help train inspectors. Physics-based POD models have reached a level of maturity in which they are the basis for accurate simulations of the results of inspections of parts of complex geometry. This allows them to be considered as a component of POD determinations and a number of scenarios are under consideration. The savings in cost and time provide strong motivation for developing and adopting these approaches. At this stage modeling techniques are approaching the point where they can be used in a major way with conventional experimental procedures being used for validation. In the following section an extensive discussion is presented of the historical development of computational NDE and POD modeling and summaries are then provided of recent efforts in this area.

4.1 Historical Perspective From the Literature

In a 1983 paper, Fertig and Richardson (1983) at the Rockwell International Science Center developed and described an integrated model for assessing the performance of an ultrasonic inspection system for detecting internal flaws. Their model is able to calculate the performance of the ultrasonic system in terms of probability of detection and other related quantities. Their work was based on efforts by other investigators that were able to properly account for the effects of real part geometry on sound propagation, and to measure noise spectra due to various noise mechanisms. The authors were able to incorporate the results of these earlier efforts by others into their model, which computes a signal-to-noise ratio for any given transducer configuration and flaw state. One of the benefits of this effort is the ability to choose the optimum transducer configuration guided by the model's calculations.

The authors also dealt with the important consideration of detecting flaws having attributes that range over an extensive class. In order to evaluate the performance of a transducer configuration, one must average over a distribution of flaws using a specified a priori distribution. In order to be able to do this for general flaw distributions and noise processes, the authors developed an ultrasonic simulation computer code. The simulation code has four basic sub-models:

- 1) Energy Transfer (interface losses, attenuation losses, refraction and diffraction effects, flaw scattering of ultrasound)
- 2) Flaw State (size, shape, orientation, material, closure state, roughness, etc.)...The authors note that they only worked with planar elliptical cracks with no correction for closure or roughness.
- 3) Noise Process (electrical, grain scattering, spurious scatterers, transducer ring down, near by surface scatter, etc.)
- 4) Decision Algorithm (A binary decision, flaw present or not present).

Fertig and Richardson described, in detail, these sub-models and calculated their basis. They gave, as an example, a wide variety of numerical results from their model for an engine bore manufactured out of IN-100. In Figure 4-1, they showed the results from an 800-micron crack for three thresholds, 0.1, 0.5, and 0.9 fraction of the time found or accepted. The visionary work by these authors was part of the beginning of a slow revolution in the way POD is determined. The revolution has been gaining momentum over the past 18 years, and modeling is now being viewed as a very promising way to determine POD with judicious use of experimental data for confirmation.

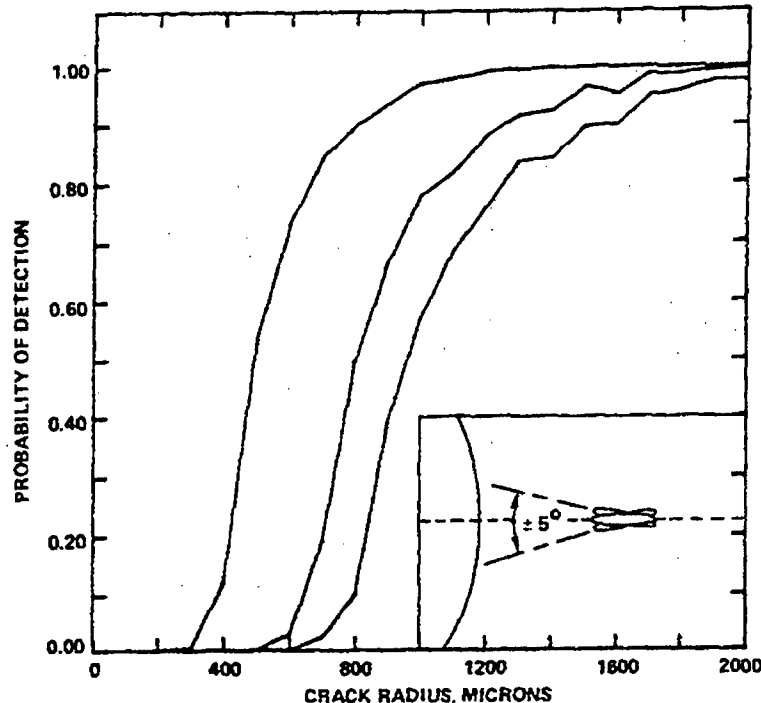


Figure 4-1 Probability of detection for circular cracks versus crack radius for three different thresholds, using a 10 MHz center frequency transducer (Fertig and Richardson 1983)

In the 1970s and 1980s other early researchers described their efforts to develop physical models that described an NDE measurement system that could be or were utilized to model and obtain POD information. In 1977, Segal, et al. (1977) reported on a physical model for x-ray radiography. Other early work on ultrasonic NDE physical modeling was also reported (Coffey, et al., 1982; Thompson and Achenbach, 1985).

In 1989, Gray and colleagues at Iowa State University (Gray, et al., 1989) published a review of models for predicting NDE reliability. Their review focused on the need for NDE measurements as part of a fracture control philosophy that depends on a damage-tolerate design. The essential feature of such an approach is the incorporation of redundant load paths so that, even if local fractures do occur, the structure will be safe for a period of time and can either be removed from service or repaired. Implementation of this damage-tolerate design approach rests on three methodologies: stress analysis, nondestructive analysis, and failure modeling.

The authors pointed out that extensive capabilities were in place for modeling stresses and failures, and these are widely used in the design process. However, the modeling of nondestructive evaluation was not nearly as widely available or accepted. Instead, frequent use was made of empirical rules based on extensive actual NDE measurement demonstration programs. For both economic and time reasons, the authors stated that there was a significant need to develop a model based system for estimating NDE reliability which is often measured in terms of the probability of flaw detection at a given confidence level. Figure 4-2 shows the systems approach that is envisioned by the authors for POD models.

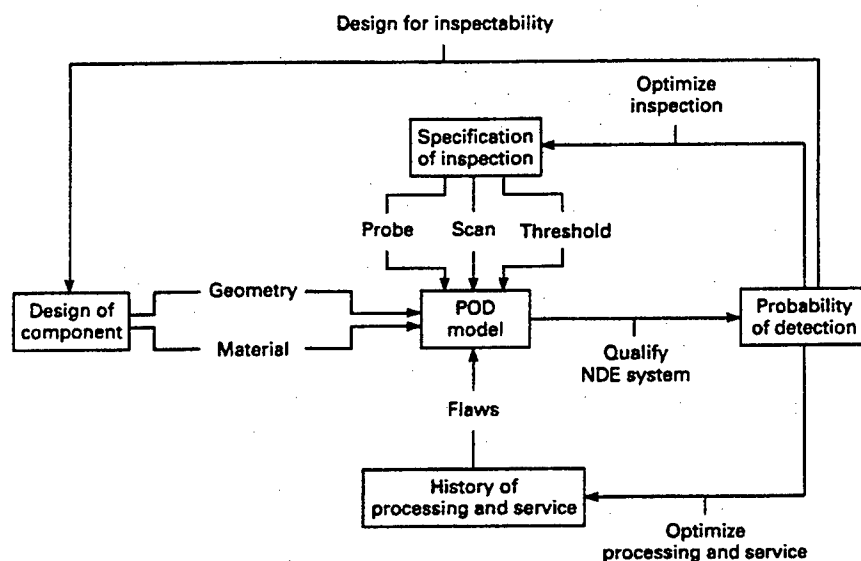


Figure 4-2 Diagram of probability of detection model and its application to NDE system qualification and optimization and to computer-aided design for inspectability. POD, probability of detection (Gray, et al., 1989)

The purpose of their review article was to present the current status and future directions of efforts to develop such a modeling capability. The article was not a comprehensive review of the subject, but rather a summary of the authors' experience in modeling of NDE inspection of

aerospace components with emphasis on engine components. Their review focused on ultrasonic, eddy current and x-ray radiography, and they discussed the physically based measurement models that have been developed for these NDE techniques.

The ultrasonic NDE simulation models described were based on the formalism of the electromagnetic reciprocity relationship, a measurement model, an ultrasonic beam model, and scattering approximations. This then leads to POD models for ultrasonic NDE that were used by the authors as shown in Figure 4-3 for circular cracks at three different depths below a circular metal surface. In a similar manner, the authors surveyed their past efforts on eddy current modeling and radiographic inspection modeling. From these results the authors conclude that these or similar models may:

- Replace costly and time-consuming experimental programs for the prediction of NDE reliability
- Improve the validation and optimization of inspection procedures
- Improve component design and the definition of the life cycle.

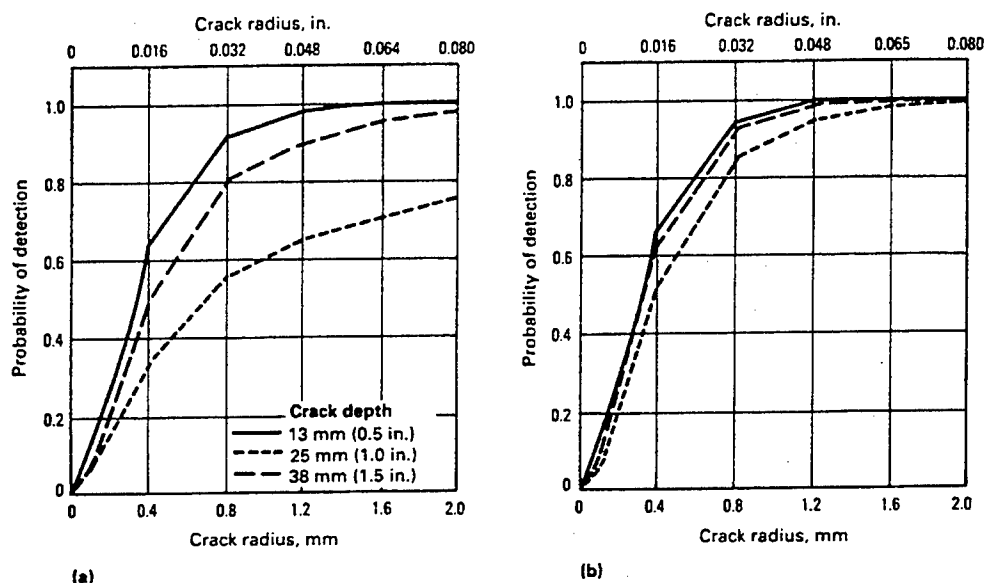


Figure 4-3 Calculated POD curves for circular cracks at these depths below a cylindrical surface. (a) axial and circumferential seam increments of 2.5 mm (0.1 in.); (b) axial and circumferential seam increments of 5 and 1.3 mm (0.2 and 0.05 in.), respectively (Gray, et al., 1989)

Finally, Gray et al. presented some examples of how NDE reliability models can be integrated with CAD systems. Figure 4-4 shows the results of a computer simulation of ultrasonic POD for an oblate spheroidal inclusion in a jet engine turbine disk. POD is shown as a function of position in the disk. The POD scale ranges from black for the lowest to white for the highest POD values. The low POD in the fillet region is due to the combination of a course scan mesh and a tight fillet region.

During the 1990s, the efforts on modeling of POD for NDE continued to advance slowly. In 1990, Nakagawa, at Iowa State, and Beissner, at Southwest Research Institute (Nakagawa and Beissner, 1990), reported on a generalized model for eddy current detection of a wide variety of flaws. This model was based on a more limited model developed earlier by one of the authors (Beissner, et al., 1988).

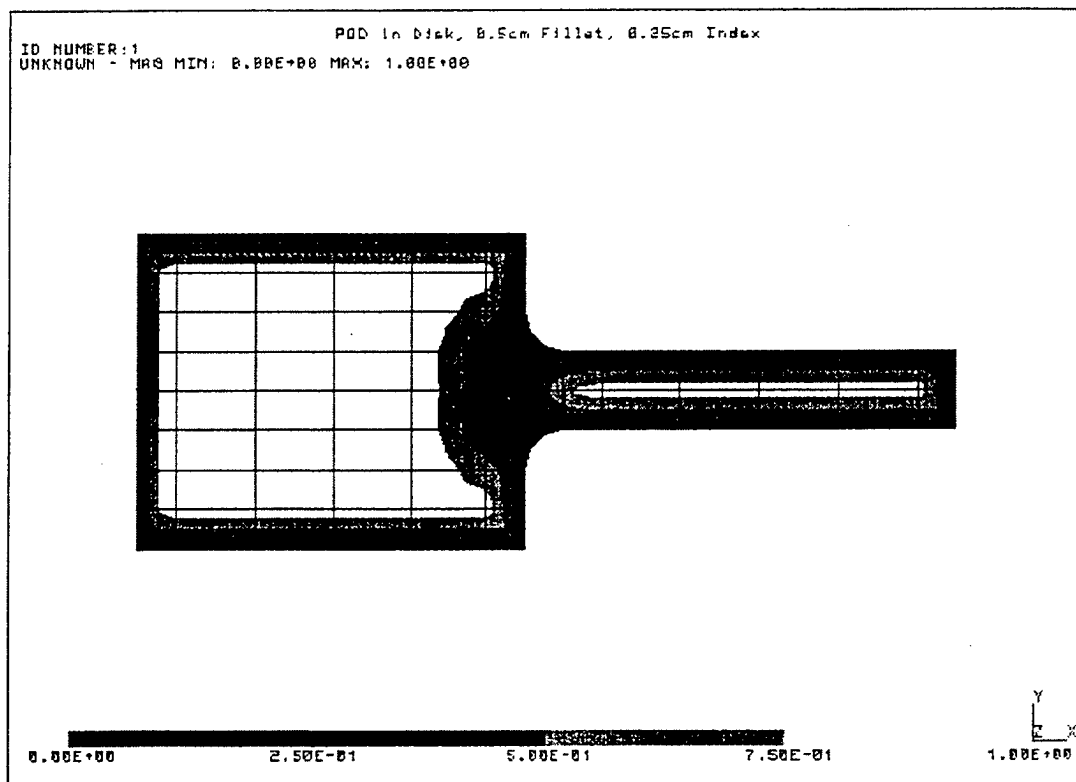


Figure 4-4 Example of a CAD-generated display of inspectability (POD) as a function of position within the cross section of a simulated disk (Gray, et al., 1989)

The authors used a theoretical physical modeling approach to replace experimentally expected flaw signals and their variabilities. The physical model approach basically is a half-space problem. Namely, a metal specimen is occupying the half-space below a flat surface, on which is a surface breaking flaw. In addition, they used some experimental data to adjust noise parameters and an overall calibration constant. The authors concluded from their work that a computer simulation could be used to replace a large number of experimental eddy current measurements to determine POD for tight fatigue cracks. Figure 4-5 shows the results of their model used to predict the operating characteristics of eddy current NDE for various crack openings. It should be noted however, that the application of this modeling effort is limited to flat plate samples and simple circular eddy current coils.

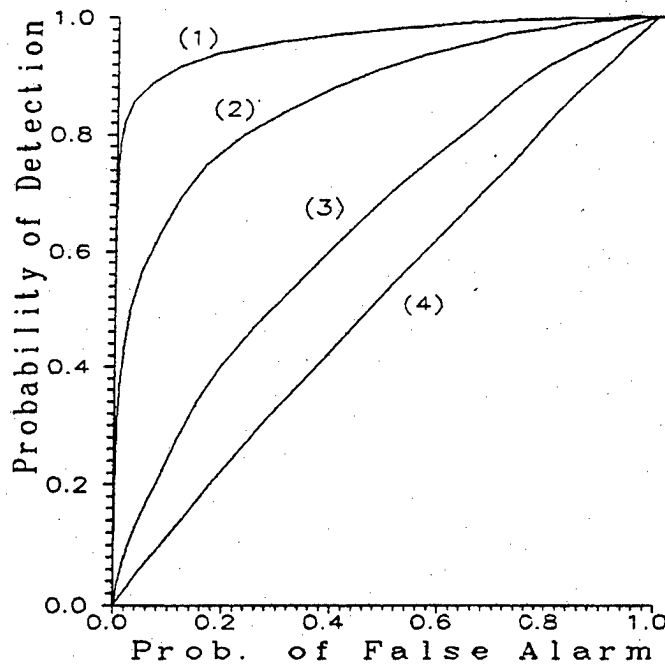


Figure 4-5 The predicted relative operating characteristics of eddy current NDE for crack openings of 1 mm, 0.75 mm, and 0.25 mm for curves 1, 2, 3, and 4, respectively (Beissner, et al., 1988)

In 1990, Beissner and Graves (1990) extended the earlier modeling work so that it would be applicable to realistic part geometries and a complex, split-D probe configuration with ferrite cores and shield. Figure 4-6 shows a model of the wing-spar lower cap lug for a T37 aircraft with the eddy current probe in the inspection position over the beveled edge at the top of the part. An inspection consists of a series of circular arc scans of the probe over the surface of the beveled edge.

The difficulty imposed by this inspection problem is that cracks occur near the edge of the beveled surface. This results in probe impedance changes due to the proximity of the edge, which tends to mask and distort the flaw signal. Because of this complex part geometry, the authors utilized a three-dimensional model. The task of modeling this inspection was accomplished in two steps. First, the authors computed the impedance loci for probe scans over flawed and unflawed parts, and secondly, they examined and interpreted these computed impedance scans for evidence of flaw signals. The authors conclude that they had successfully demonstrated eddy current POD determination by computer simulation for complex part and probe geometries. The authors cautioned that much work remained to be done before this methodology was ready for routine application. They viewed their effort as a successful feasibility study.

Work also continued at Iowa State University on eddy current modeling of POD. Nakagawa, et al. (1990) reported on their effort to develop a POD model for their automated eddy current measurement system. The heart of their automated system was a precision impedance analyzer capable of measuring impedance over a wide range of frequency ($10^2 - 10^8$ Hz). Data acquisition, processing, analysis, and display was via a personal computer, and a computer-controlled x-y positioning stage provided for either one or two dimensional scans of the specimen.

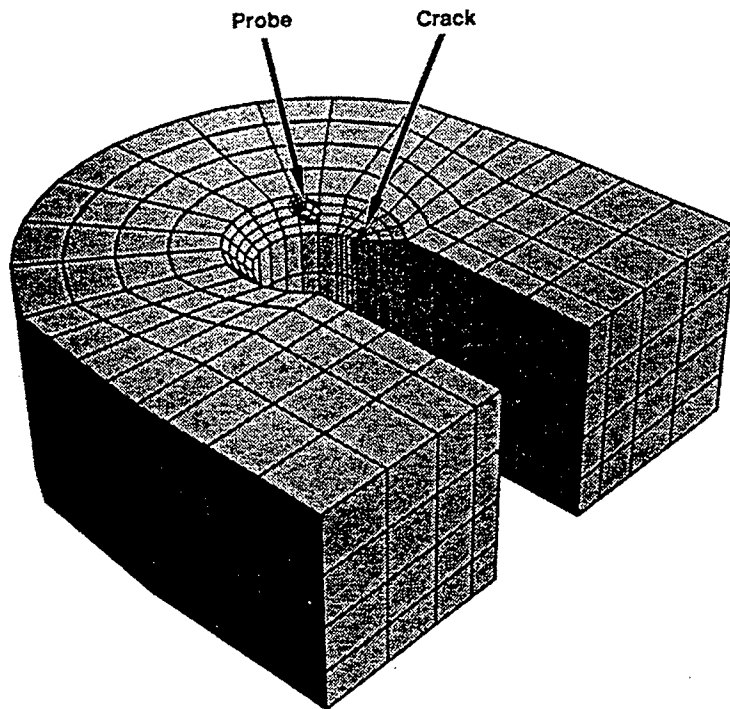


Figure 4-6 Eddy current inspection of the T37 front wing-spar lower cap lug. The probe is scanned in a circular arc over the beveled edge of the part. Cracks tend to occur in the region indicated
(Beissner and Graves 1990)

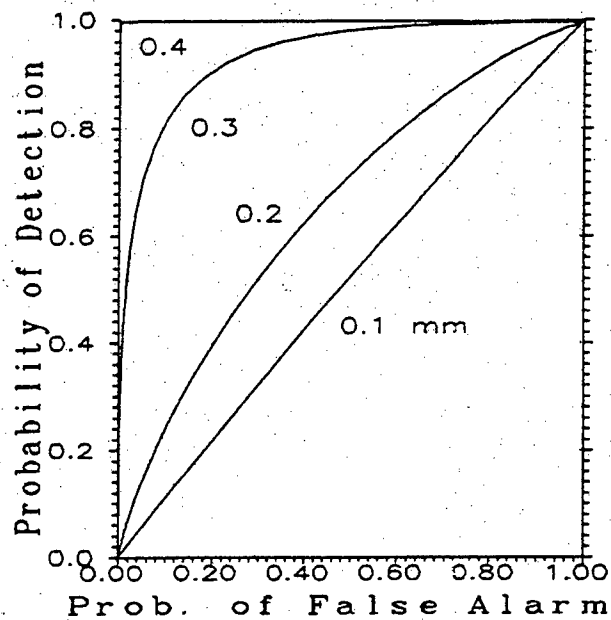


Figure 4-7 Calculated relative operating characteristics for an automated eddy current system
(Nakagawa, et al. 1990)

The authors developed a POD model for their automated eddy current system that provides a measure of inspectability in the form of relative-operating-characteristics (ROC) curves and POD curves. Their modeling results are shown in Figure 4-7 where POD and POFA are plotted for cracks of various depths in an aluminum block with a flat surface. For these calculations, the crack is assumed to be semi-elliptical, with an aspect ratio of 3. Four cracks of depths 0.1, 0.2, 0.3, and 0.4 mm were studied. The eddy current probe is taken to be a uniform-field probe operating at 100 kHz. Shown in Figure 4-8 is a sample POD prediction versus flaw depth. The assumed measurement condition is similar to the example in Figure 4-7, except that the probe frequency is 141 kHz. The dotted curve (1) was obtained by setting the threshold value to 6.5 [$m\Omega$] where the POD of a 0.3-mm crack is 50%. The solid curve (2) is for a threshold of 3.7 [$m\Omega$] where 95% POD is achieved for a 0.3-mm crack.

Nakagawa, et al. concluded that their results are of sufficient quality to allow quantitative studies of the probability of tight-crack detection by eddy current measurements.

Rajesh, et al. (1993a) and Rajesh, et al. (1993b) also modeled the POD of eddy current measurements to detect surface breaking cracks. Their work was based on a finite element approach. This paper presents a more comprehensive POD model than their earlier efforts (Rajesh, et al., 1991 and 1992). The authors pointed out that finite element models have been used extensively to model physical processes underlying NDE phenomena such as transducer response for a given specimen geometry. These models, however, are deterministic in nature and do not take into account the perturbations associated with the inspection and testing carried out in the field. As an example, sources of noise include variations in liftoff, materials properties such as conductivity and permeability, probe canting angle, scan format, surface roughness and measurement noise.

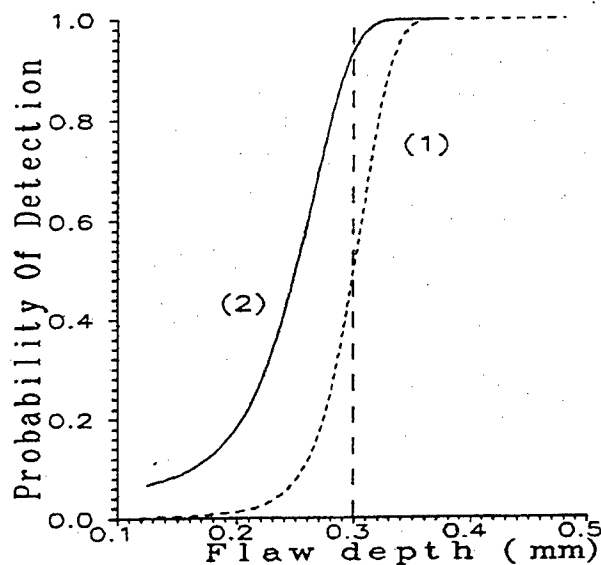


Figure 4-8 A sample POD prediction for an automated eddy current system
(Nakagawa, et al. 1990)

Figure 4-9 is a schematic diagram of the finite element based modeling approach used by the authors and Figure 4-10 shows the results of their modeling for eddy current measurement signals with and without a flaw. It can be noted that this classification can result in two types of errors, which are of significance, namely false alarm and false acceptance. The authors conclude that the feasibility of using finite element modeling in a probabilistic framework for estimating the probability of detection of a flaw had been demonstrated. Theoretical POD models allow complex defect shapes to be handled easily at low cost. Their stated future aim was to incorporate the POD model into a CAD framework.

In 1993, Schmerr and Thompson (1993) expanded on the concept of POD models integrated into a unified life cycle engineering approach that was earlier reported by Gray, et al. Figure 4-11 shows the elements of a design environment that includes CAD, stress and materials analysis, NDE inspectability estimates (POD), and reliability calculations that are coupled together to form a new concurrent engineering design technology. These unified life cycle engineering approaches that incorporate POD modeling capability were also reported by Wormley, et al. (1993).

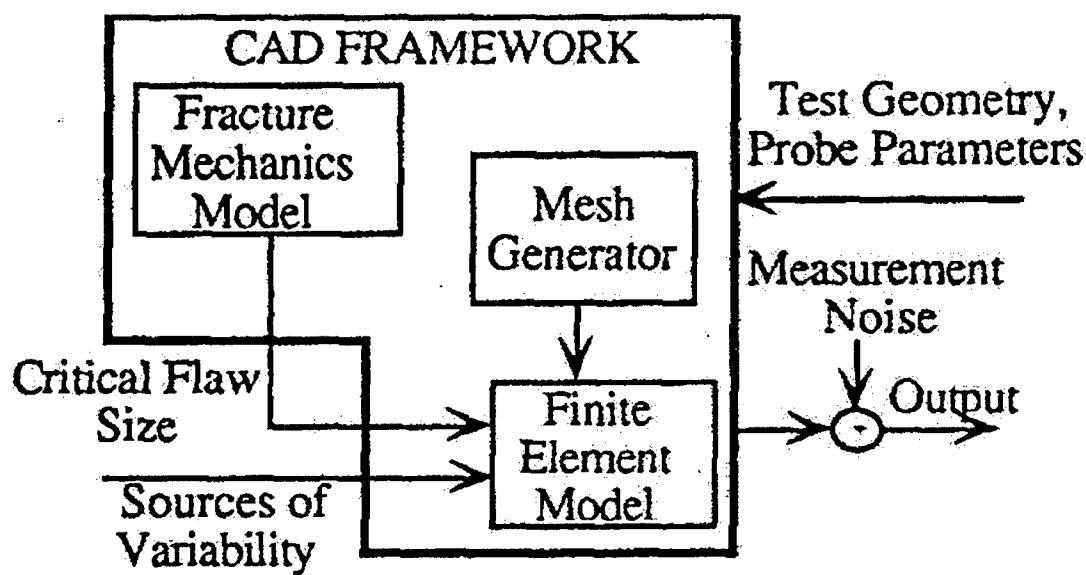


Figure 4-9 Schematic diagram a finite element based POD modeling approach (Rajesh, et al., 1991 and 1992)

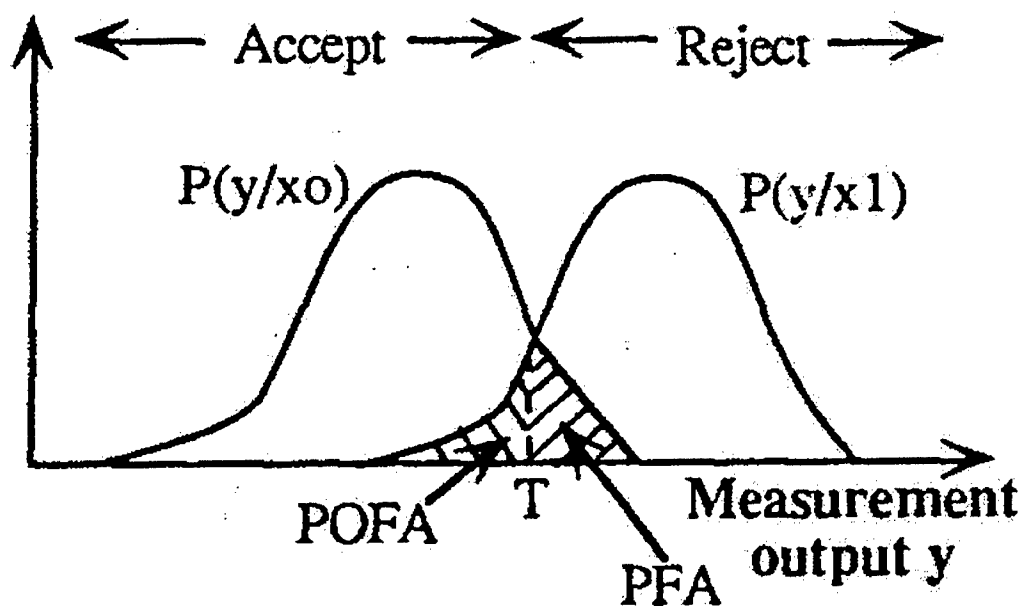


Figure 4-10 Probability density functions of eddy current signals with and without a flaw (Rajesh, et al., 1991 and 1992)

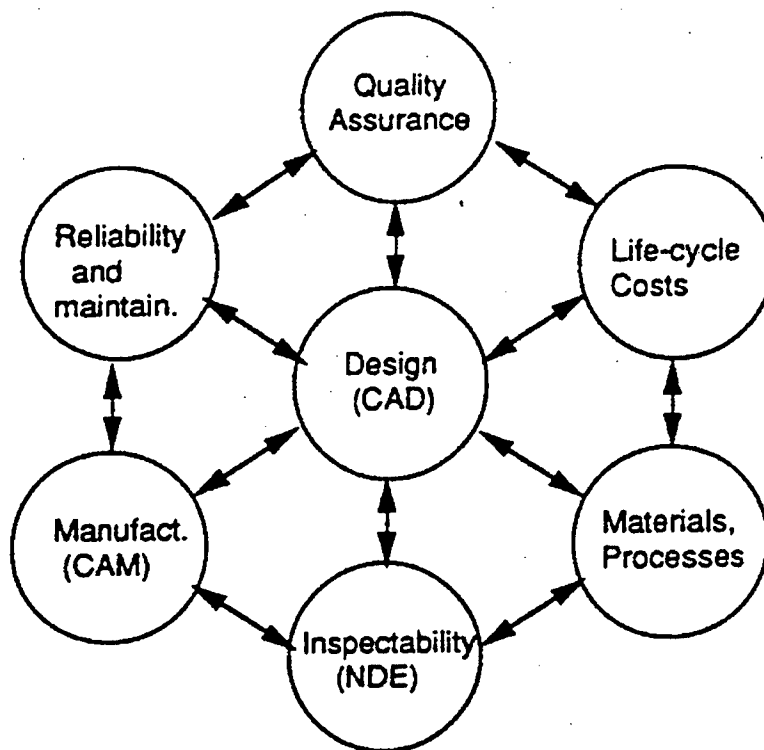


Figure 4-11 A general Unified Life Cycle Engineering Environment (Schmerr and Thompson 1993)

Thompson and Schmerr continued their efforts to describe the development and uses of model-based probability of detection curves in a 1993 review paper (Thompson and Schmerr, 1993). They pointed out that the modeling of POD for NDE measurements was based on the continuing advances being made in the capability to physically model a number of NDE techniques. They described a number of uses of model based POD curves including: optimizing NDE system design and procedures, defining NDE system performance capabilities, developing NDE calibrations and standards, allowing NDE system validation, separating NDE measurement capability from operator effects, bringing NDE into the early design process, and simulating for the training of NDE operators.

In parallel with the efforts in the 1990s on modeling of POD centered in the United States at Iowa State University, researchers at Harwell Laboratory in Oxford, United Kingdom were also pursuing development of modeling of POD. Ogilvy (1993) reported on his work that developed a mathematical model to predict the theoretical POD of planar buried defects from conventional ultrasonic pulse-echo inspection. He developed a physically based model for the scattering of ultrasound, with the amplitude of the signal based on an approximate Kirchhoff theory, by well-oriented planar defects. He combined this model with a noise theory model into a package designed for quick numerical evaluation to determine probability of detection. His overall model also addresses the problem of false indication of a defect.

Ogilvy used the model to quantify the uncertainty that results in POD predictions due to uncertainty in flaw orientation, roughness, aspect ratio, and flaw depth. The author pointed out that his model was based on similar principles that were used by others to model ultrasonic NDE probability of detection. However, he added that his model goes beyond the prediction of POD to address the issue of false calls, i.e. the apparent detection of a flaw when none is present. In practice, inspection is a trade-off between increasing the POD and retaining the POFA or probability of false indication (PFI) at an acceptably low level.

Although the model, as reported here, is limited to incident monochromatic waves detected through scattering, the author stated that incident pulses and other wave-defect models could be incorporated. The effects of altering the reporting threshold, of performing raster scanning and of altering the number of probe positions were also investigated. The author also concluded that POD models for calculating the *capability* for detection are a good starting point for incorporating potential human error effects. This could be accomplished by utilizing a parametric form for the effect of human error on POD.

In 1994, Wall and Wedgwood (1994) from the National NDE Centre, Harwell, United Kingdom, presented a review of the economic benefits to be derived from being able to improve and quantify speed, coverage and reliability of NDE measurements. As part of the review, the authors presented a substantial amount of information on probability of detection and probability of false indication. The authors stressed that although PFI has great economic importance, very little effort on PFI has been reported in the literature. Their review of POD and PFI included a survey of experimental and modeling efforts, and the modeling effort survey focused mainly on the work at Harwell.

Wall and Wedgwood presented a conceptual model for NDE reliability shown in Figure 4-12 and stated that it is unlikely that a theoretical mode for human factors will be developed in the near future. This is the case since human factors have been shown to be dependent on a large number of variables such as fatigue, environment and job complexity. The authors concluded

that a future trend would be the increasing emphasis in obtaining and using NDE performance parameters (sensitivity, speed, and reliability) when planning NDE inspections. This will lead to the continued develop of models and databases for this purpose.

Efforts continued at Iowa State University on developing and describing their efforts on POD modeling and the potential benefits from POD modeling capability. In 1994, Schmerr and Thompson (1994) presented an updated review of efforts to develop a unified life-cycle engineering approach that was based, in part on the growing capability to model NDE performance. The authors concluded that this is leading to a new technology that will allow NDE to be fully integrated in design and life cycle performance.

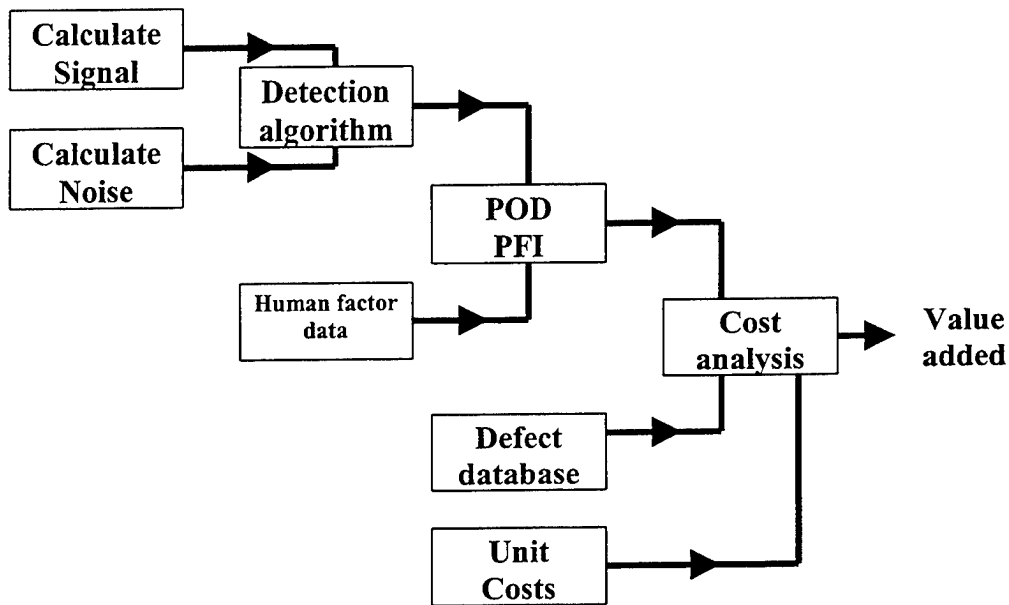


Figure 4-12 Conceptual model for NDT reliability. The information required to estimate reliability (POD, PFI) and the added value from an inspection is shown (Wall and Wedgwood 1994)

Also in 1994, Elshafiey, et al., (1994) reported on using parallel computers to speed up the calculations for modeling the performance of eddy current measurements. They used a finite element, three-dimensional approach that involves the solution of massively large matrix equations that normally took several hours of computational time. The authors concluded that they solved this problem with the use of parallel computers with hypercube architecture.

Klemmt and Tober (1995) applied a Monte-Carlo Simulation statistical method to calculate NDE inspection reliability for multiple site damage. Their effort at Deutsche Aerospace Airbus, Bremen, Germany was used to extend the use of POD curves from single defect applications to multiple site damage. The authors used the case of multiple fatigue cracks in an aircraft lap joint inner skin to illustrate their work. Although this effort apparently used experimental POD data, the methodology can apparently be applied to POD curves derived from modeling.

Chiou, et al. (1995) reported on their research that was in support of the efforts at Iowa State to develop modeling capability for determining the POD of ultrasonic inspection of titanium aircraft engine billet material and parts. The authors characterized flat-bottom holes in titanium alloy reference standards by both experimental measurements and physical modeling. For the modeling effort, the authors used a hybrid model centered around the reciprocity model developed by Thompson and Gray (1983). They used the method of optimal truncation (MOOT) as a plane wave scattering amplitude solution. This was combined with the high-frequency Kirchhoff approximation with numerical integration and a simplified reciprocity relationship for special cases. They solved this problem by modeling the response as a Gauss-Hermite series expansion, and they called this approach (FKIR). The authors stated that the FKIR approach is applicable to the normal incident case, and the MOOT approach is valid for all tilt angles. Figure 4-13 illustrates the accuracy of their model predictions. The figure displays three experimental rf waveforms and the corresponding model predictions for a flat-bottom hole (FBH). The authors concluded that they successfully demonstrated the use of ultrasonic models for predicting absolute flat-bottom hole response in full-range C-scan calibration experiments.

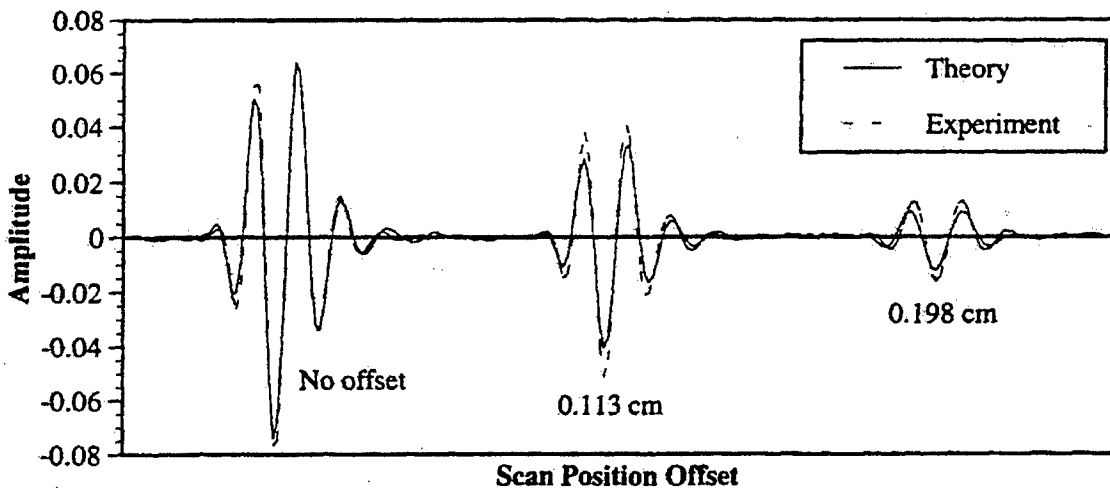


Figure 4-13 Absolute amplitude and phase comparisons between model and experiment for a typical #4 flat bottom hole at normal-incidence using a transducer focused on a hole (Thompson and Gray 1983)

The following year, Chiou, et al. (1996) extended their modeling work reported the previous year on the study of flat-bottom holes in titanium alloy engine billet material. They derived a volumetric formulation for Auld's electromechanical reciprocity relationship under the Born approximation. They also derived a surface formulation using ad-hoc boundary conditions at the flaw's (inclusion's) surface. They then used a numerical integration of the wave field over the flaw surface/volume to obtain the ultrasonic response. The authors once again evaluated their model using the response from flat-bottom holes in a titanium alloy sample. The authors concluded that they had obtained overall good agreement, and looked to future work to reduce the limitations on the model.

Meeker, et al. (1996) presented an updated overview of the efforts to develop modeling capability to determine the POD and PFA for hard inclusions in titanium engine billet material. The authors discussed the physical models based on theory of ultrasonic wave scattering that are used to predict measurements from the flaw signal distributions. They then presented a

statistical model to be used to quantify deviations between the physical model predictions and actual NDE measurements. Figure 4-14 indicates the path that their future modeling efforts will take; the model will be developed and refined first for synthetic flaws and then extended to real flaws.

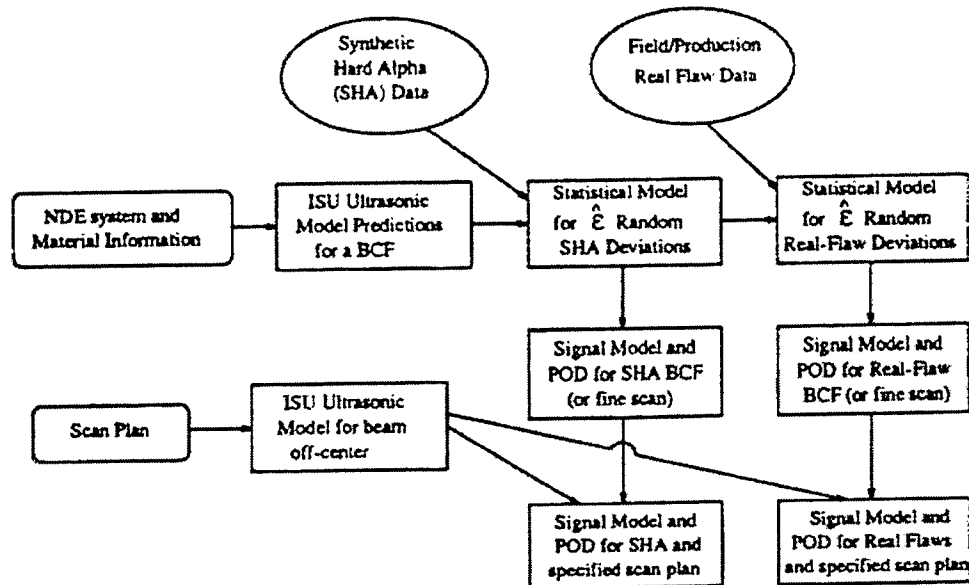


Figure 4-14 Proposed ETC POD estimation methodology (Meeker, et al. 1996)

In the following year, Wall (1997) at the National NDT Centre, Harwell presented a review on modeling of NDE reliability and the application of corrections for human factors. The focus was also on accomplishments on modeling at the National NDT Centre. The author presented the approaches that are presently available to make predictions of POD as follows:

- Physical models for POD and PFI
- Signal/noise models
- Image classification models (visual POD)
- Inspection simulators
- Statistical Models (curve fitting)
- Human reliability models
- Expert judgement

Wall presented Table 4-1 as a summary of the PC-based models currently available at the National NDT Centre, AEA Technology for NDE inspection reliability. He also presented the corrections for human factors that are being applied to the POD/PFI models. These corrections are based on experimental data since a theoretical model of human factors would be very complex and is not available.

Wall concluded that models should be seriously considered as an integral part of future POD trails. This would: 1) reduce the number of experimental samples required, 2) help gain acceptance and familiarity for the modeling approach, 3) provide validation and lead to improvements in the model predictions and correction methods used for human and environmental effects.

Table 4- 1 PC-Based models for inspection reliability available in AEA Technology, showing physical model used to calculate signal in calculating POD and PFI (Wall 1997)

Model	Technique	Details	Signal calculation
PODUT	Ultrasonic	Pulse-echo, buried crack	Kirchoff approximation (Ogilvy)
PODSURF	Ultrasonic	Pulse-echo, surface crack	Semi-empirical model (Silk)
PODTOFD	Ultrasonic	TOFD, cracks	Geometric diffraction (Temple)
XPOSE	Radiography (X-ray)	Film radiography	Geometrical model (Windsor and Wall, Helmshaw)
NNXPOSE	Radiography (X-ray)	Film radiography	Geometrical model (Windsor and Wall, Helmshaw). Neural networks for defect detection.
PODET	Eddy Current	Surface or near-surface cracks	OPERA, POD calculation (Holt) – developmental
HUMPOD	Human Reliability	Correction method	Empirical correction method for human and environmental effects
PODDATA	All	POD Database	Database of past experimental trial data. Input for integrity assessments and model validation.

Meeker, et al. (1997) published a review and update of their modeling efforts on predicting POD and POF for hard alpha inclusions in titanium aircraft engine billet material. In the new part of the effort reported here, the authors extended the model to include: 1) details related to the microstructure, 2) flaw morphology, and flaw position relative to the ultrasonic probe.

Also in 1997, Schmerr and Thompson (1997) presented a white paper on the role of Modeling in NDE standards. The authors recommended that future work related to modeling of NDE measurements should address:

- 1) the use of models to design, validate, and extend the measurement process
- 2) the use of models to calibrate and quantify the capability of NDE hardware
- 3) the use of model to train and educate NDE personnel
- 4) the validation of models themselves.

An overall review of the modeling methodology developed at Iowa State University for POD determination was presented by Thompson and Meeker (1997) at the European-American Workshop for Determination of Reliability and Validation Methods of NDE in Berlin.

4.2 Recent Modeling Efforts

Near the end of the 1990s, efforts continued on modeling of NDE reliability. In 1998, Sarkar, et al. (1998) reported on the development of a model to predict the probability of ultrasonic detection of cracks in heat exchanger tubes. Since it is difficult to produce synthetic subsurface flaws, the traditional experimental approach to POD /POFA determinations are difficult to apply. The authors treated the depth of the crack and some inspection modality parameters as known fix-effect factors and treated all other parameters as random-effect factors. They developed a modeling methodology to estimate POD as a function of the known-effect factors. They used a deterministic POD model to explain the effect that inspection modality factors have on the expected ultrasonic signal strength. They coupled this with a statistical model to quantify the variability in the ultrasonic signal strength. The combination of these two models allows one to predict POD at inspection conditions that are different from the original conditions for which experimental data is available.

The authors obtained experimental data for various heat exchanger tubes for signal amplitude and crack depth, from autopsy, at a number of different axial positions. For each given ultrasonic signal amplitude data point, they calculated and estimated crack size with just their physical/theoretical model. Figure 4-15 shows the experimental and calculated crack sizes. One can see that except for the three data points at zero crack size, the theoretical model predictions are always above the corresponding data points. This over prediction of crack size is due to the theoretical model prediction curve being computed for an ideal reflector while actual reflectors have less reflectivity. Most of the variability is most likely due to flaw morphology. The three points corresponding to zero flaw size are thought to be the results of false indications in the vicinity of the crack.

Sarkar, et al. also calculated the POD curves for different detection threshold levels and obtained the POD curves as shown in Figure 4-16. For higher detection thresholds, the asymptotes of the estimated POD curves are attained at higher crack size. Also as expected, for a given crack size the estimated POD is lower if the detection threshold is higher.

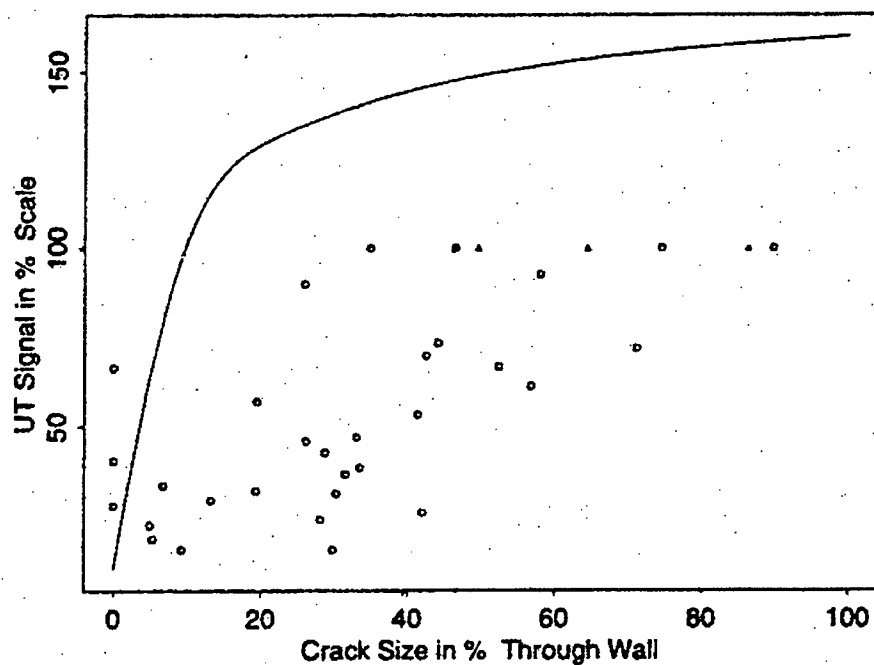


Figure 4-15 UT measurements and theoretical model prediction. Experimental ultrasonic measurements and theoretical model predictions for cracks in heat exchanger tubing (Sarkar, et al. 1998)

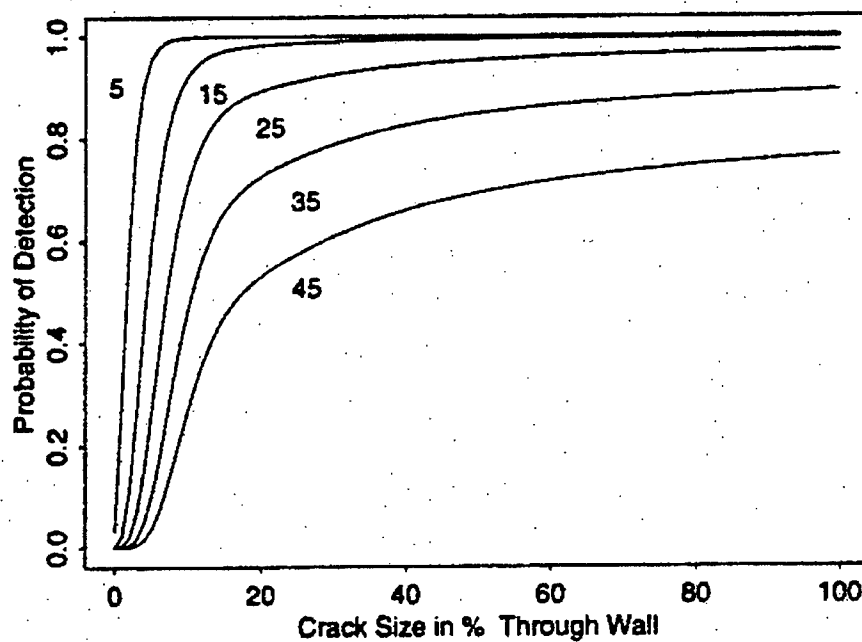


Figure 4-16 Calculated POD curves for different ultrasonic detection threshold levels (Sarkar, et al. 1998)

Meeker, et al. (1998) presented another update on their continuing efforts to improve modeling methodology for determining inspection reliability for detecting hard alpha inclusions in titanium engine billet material. This paper describes the authors' new efforts to develop methodology to detect synthetic hard alpha flaws in titanium alloys. They described and illustrated the methods they used to assess the effect that changes in ultrasonic scan speed and gate width will have on POD.

The authors obtained, in earlier work, experimental ultrasonic measurement data on the synthetic hard alpha inclusion samples. The data provide information about the effect on ultrasonic signal strength of flaw size, the distance between the beam and the center of the flaw, and the distance between the focal depth and the top of the flaw. The data also provide information on flaw-to-flaw variability for nominally identical flaws.

Meeker et al. calculated the POD for various flaw sizes as a function of threshold values, and looked at the effect of scan increments and gate width. They concluded by stating that they planned future research on real hard alpha inclusions.

In 1999, Thompson (1999) presented an updated review of the POD methodology that had been developed for NDE of titanium engine components. Figure 4-17 illustrates the three major sources of variability in the automated NDE inspection of titanium aircraft engine billet material. These are associated with micro-structural effects, instrumentation and scanning procedures, and flaw morphology. Figure 4-18 is a schematic of the general steps taken to develop the POD/PFA modeling methodology. Thompson presented a number of POD curves obtained via various models.

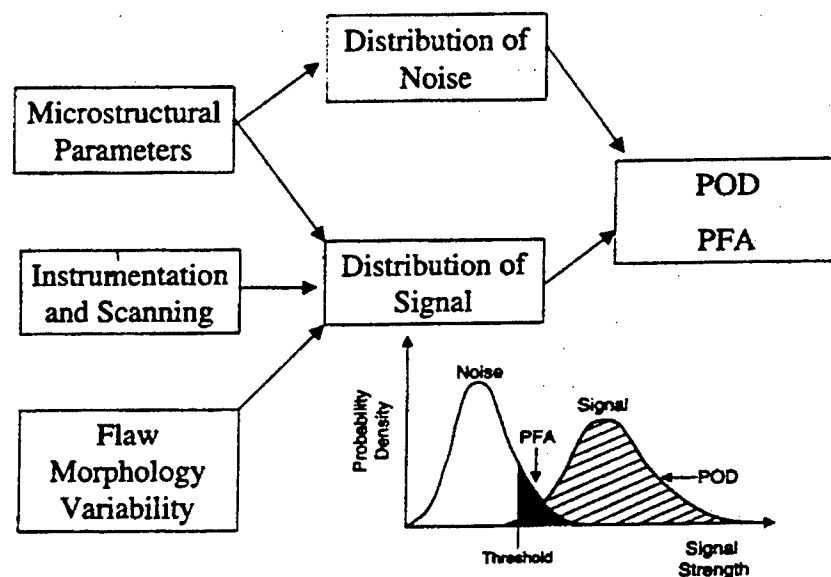


Figure 4-17 General strategy of POD methodology. Shaded areas correspond to POD and PFA, as deduced from noise and signal distributions, for a particular threshold (Thompson 1999)

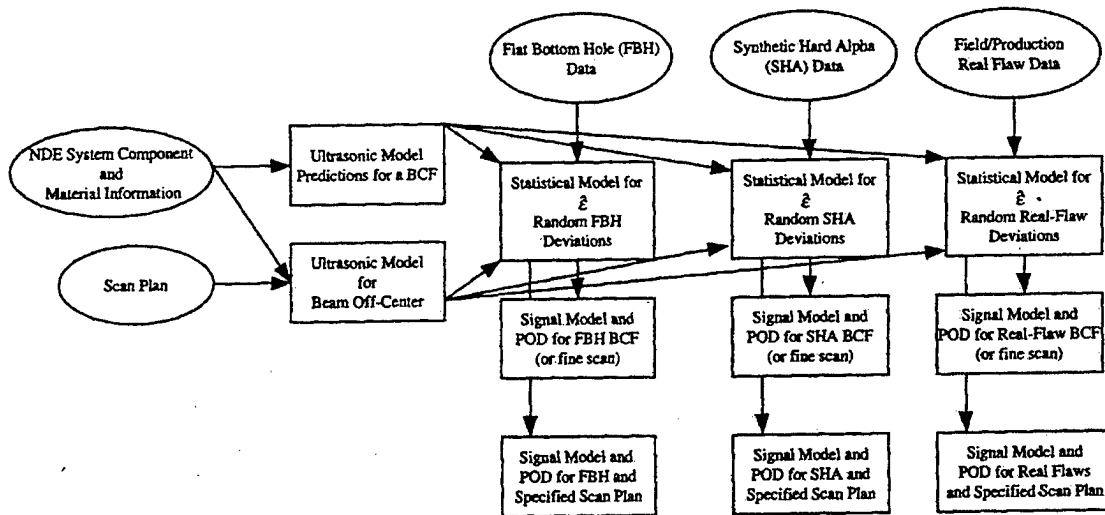


Figure 4-18 General steps in developing modeling methodology for POD predictions (Thompson 1999)

Tow and Reuter (1998) tackled the question of how to include inspection results into a probabilistic model of the reliability of a structure. They used a very simple probabilistic fracture mechanics (PFM) model for pressure vessel reliability considering only the variability in applied stress; all materials properties were considered deterministic. The distinctive feature of the model is the way in which inspection results and the POD curve are used to calculate a probability of density function (PDF) for the number of flaws and the distribution of those flaws among the various size ranges. In combination with the PFM model, this density function is used to estimate the probability of failure (POF) of a structure in which flaws have been detected by NDE. In applying their probabilistic failure analysis methodology, Tow and Reuter used three POD curves corresponding to three levels of NDE reliability previously identified by Kahleel and Simonen (1994). These three curves are shown plotted on Figure 4-19. Tow and Reuter applied the methodology to exponential crack depth distributions for 476, 234, and 118 cracks, respectively. The overall POFs were calculated with the results summarized in Table 4-2. The benefits of more reliable NDE are clear from this table. As NDE becomes more reliable the calculated POF of the structure decreases and approaches the POF that would be calculated if the actual crack population were known. As NDE becomes less reliable the probability of larger undetected cracks becomes greater and the calculated POF increases. Results of the study show that NDE can play a significant role in reducing the probability of failure of structures if the inspection procedures satisfy strict minimum performance requirements.

Table 4- 2 POF calculations for three different NDE reliability levels and three crack populations (Tow and Reuter 1998)

Inspection Performance	476 actual cracks	234 actual cracks	118 actual cracks
Marginal	8.0×10^{-2}	4.0×10^{-2}	2.0×10^{-2}
Very good	1.3×10^{-2}	3.5×10^{-3}	3.9×10^{-4}
Advanced	4.9×10^{-4}	1.2×10^{-6}	8.2×10^{-8}

Khaleel and Simonen (2000) developed a numerical approach to predict the probability that a fabrication flaw in a reactor pressure vessel will extend by fatigue crack growth mechanisms and become a through-wall flaw. A probabilistic fracture mechanics (PFM) model was formulated for the growth of surface and buried cracks and was implemented into the probabilistic computer code VESSEL. Characteristic POD curves for the detection of reactor pressure vessel fatigue cracks based on expert judgement were input into the PFM model. The approach was to establish four POD curves that represented widely differing levels of NDE performance in ultrasonic examination of welds. The curves were intended to bound the performance expected from inspection teams operating in the field. Results show that POD capability appears to be the most limiting factor with regard to the overall capability of in-service inspections to reduce leak probabilities. The effects of flaw sizing errors are relatively small when calculations are based on inputs for flaw sizing capabilities and acceptance standards that are representative of current NDE capabilities and code requirements. However, the calculations show that gross errors in flaw sizing or significant departures from current flaw acceptance standards could negate the expected benefits of inspection methods that exhibit outstanding performance in the area of flaw detection.

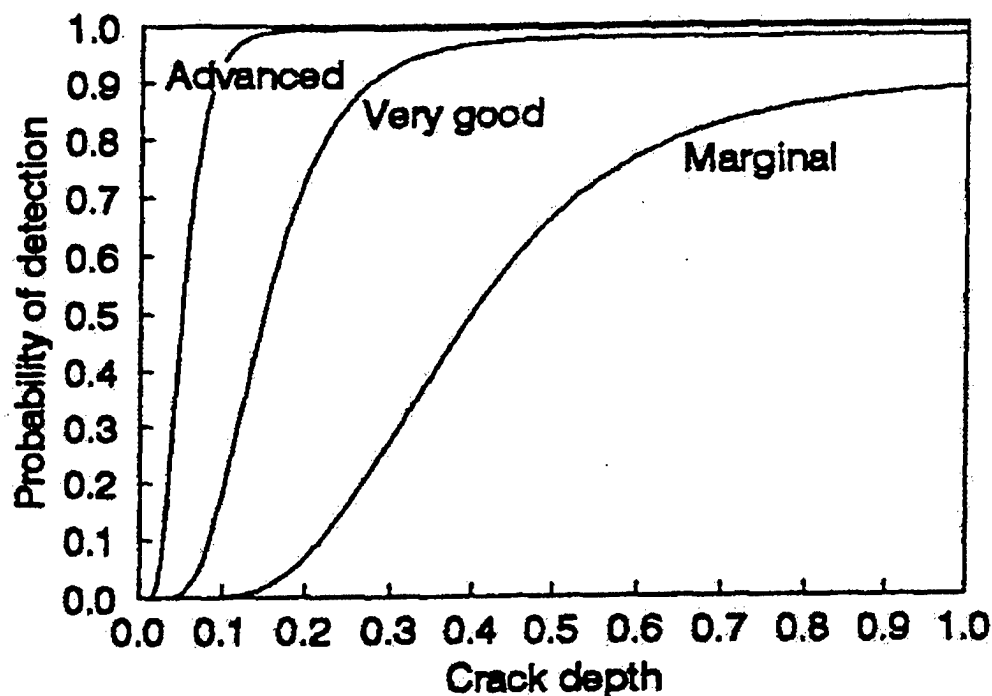


Figure 4-19 POD curves for three levels of NDE reliability
(Kahleel and Simonen 1994)

The use of statistical models for the evaluation of the reliability of nondestructive inspections was examined by Simola and Pulkkinen (1998). Models for flaw sizing on the basis of statistical logarithmic or logit transformations were constructed. The POD was then modeled as a function of flaw depth and length again based on logarithmic and logit transformations of the relative flaw sizes. Models for Bayesian updating of the flaw size distributions were also developed. Using these models it is possible to take into account the prior information of the flaw size and combine it with the measurement results.

Meeker, et al. (2001) have recently described a new modeling methodology that combines the use of physical modeling of factors that can be modeled adequately and statistical modeling of limited NDE inspection data to account for other important factors that are not included in the physical model (e.g., human factors, deliberate or unavoidable changes in inspection system properties, or complex changes in flaw morphology). The resulting physical/statistical model provides predictions of POD or other inspection-evaluation metrics as a function of specified inspection plans and flaw characteristics that extend the range of predictions that could be made empirically based on the available data. This methodology is especially important where there is not any satisfactory NDE referee technique by which to determine the "true" flaw characteristics, such as ultrasonic POD for detection of hard alpha defects in titanium billets and forgings.

It has been recently pointed out by Thompson (2001) that there has been a steady increase in the ability of physical models to accurately predict the results of tests of real parts and also in the ability to make rapid computer based simulations based on these models. This development of accurate simulation tools provides the opportunity to combine the ability of simulations to predict physical phenomena that are well understood with empirical measurements where such understanding does not exist to determine POD. Thompson concludes that by explicitly including the effects of many controlling variables, the use of simulation tools in POD determination can greatly reduce the need for empirical data, thereby enabling the cost effective determination of POD in a much greater range of geometries, materials, and discontinuity types than would otherwise be possible. Included is the ability to predict mean response and broadening produced by many of the sources of variability that influence POD, such as scan plans, variations in probe parameters and microstructural effects.

5.0 CONCLUSIONS AND PROGNOSIS

As discussed in previous sections of this Technology Assessment, the issues of POD and NDE reliability have received considerable attention from the research community over the last few years. Certainly part of the impetus for this movement is the realization that successful life management of the increasing inventory of aging systems depends on the ability to quantify NDE procedure capability and reliability. The accepted method of NDE capability quantification, the POD curve, shows the probability of a flaw's detection as a function of flaw size for a specific inspection technique. While a complete and successful POD determination is quite expensive, it has been pointed out by Grills (2001) that the return on investment can be very high in terms of improved safety, lower life cycle cost and extended system life. Recent research efforts involving POD and NDE reliability have been directed toward broadening application by using actual in-service inspection data, using smaller data sets, transferring information from simple shapes to complex structures, and using neural networks, among others.

Until recently, essentially all applications of the POD concept have been empirical, i.e., a statistically significant number of samples are prepared with artificial flaws and then experiments are performed by a number of operators to test the NDE technique. With the advent of measurement models, however, the POD can be calculated for a specific set of conditions and verified with a few experimental samples at a considerable cost saving and with a capability for predictable extension to other inspection conditions. Emerging technology in recent years has provided a firm foundation for the development of modeling technology to determine the reliability of NDE measurements. Current NDE models for determining the reliability of NDE measurements are, in general, based on the combination of two or more computational models. The first model is based on a physical model of the NDE measurement, and this is combined with the second model, a statistical model that corrects for signal noise. The modeling of human factors based on a physical model is currently deemed to be too difficult to accomplish. However, a parametric correction for human factors has recently been applied to some of the models.

The rapid increase in both hardware and software computational capability has provided part of the methodology needed for modeling NDE measurements. The other major advance, the physical models for various NDE methods including ultrasonics, eddy currents, and x-ray radiography in a format for computer simulation of measurements, has provided the additional methodology needed for modeling NDE measurements. Physical models for other types of electromagnetic measurements are also available, as are physical models for thermography and thermal wave imaging. These physical models most likely need to be put into a computational format so that computer simulation of the measurements is available.

The impetus to advance the capability to model NDE reliability is increasing. This is due to the demands for a wide variety of POD/POFA data needed for a wide variety of NDE measurement methods applied to many different materials systems. Materials systems include metals,

ceramics, polymers, and composites with polymer or metal or ceramic matrices. The cost of obtaining this data by purely experimental means is becoming prohibitively expensive and time consuming. POD/POFA modeling capability can be used to:

- Replace costly and time consuming experimental programs for prediction of NDE reliability
- Improve component design and definition of the life cycle
- Provide an NDE simulator for training and education
- Optimize and validate NDE hardware and procedures
- Develop and quantify improved physical calibration standards.

The seminal papers on modeling of NDE measurement reliability began to appear in the late 1970s-early 1980s, and during the 1990s advances continued. During the 1990s, there have been major research efforts on modeling of NDE reliability centered at Iowa State University in the United States, and at the National NDT Centre, Harwell, United Kingdom. Other researchers at various organizations have also contributed to the knowledge base. Efforts to date on modeling of NDE reliability have been concentrated on ultrasonic, eddy current, and x-ray radiography measurements.

Looking to the future, one would expect the development and application of NDE models to continue as the need for NDE reliability data continues to grow. Perhaps the best way to enhance this modeling capability and its' widespread acceptance and use, would be for an organization to serve as a focal point for encouraging the development and use of POD models. This organization would be a focal point for:

- Encouraging the development of additional POD/POFA models
- Encouraging the validation of POD/POFA models
- Providing a repository of existing POD/POFA models
- Encouraging the development of models for POD/POFA for NDE data fusion method
- Encouraging the training and application of POD/POFA models
- Providing a focal point for planning to meet the needs for additional efforts related to POD/POFA modeling.

NTIAC has recently completed a Computational NDE and POD Modeling planning effort under sponsorship of the US Air Force, NDE Program Office, Air Force Research Laboratory at Wright Patterson Air Force Base to enhance modeling for NDE capability and stimulate its widespread acceptance and use. The proposed plan involves forming a consortium administered and coordinated by NTIAC with technical leadership provided by a board of directors composed of consortium members. Technical staff from a variety of research organizations would participate in distributed modeling activities and NTIAC would serve as a repository to distribute and maintain the software and to disseminate results of the consortium's efforts. NTIAC would also be the consortium's focal point for training on the use of POD modeling tools with consortium participants and others serving as instructors. The initial goal of the consortium would be to develop, validate, demonstrate and transfer to users a selected set of modeling capabilities for NDE reliability. The modeling capabilities would be broad enough to demonstrate the great value of modeling of NDE reliability.

References

- Aldrin, J. C. "Overview of Mathematical Modeling in Nondestructive Evaluation (NDE)." Nondestructive Testing Information Analysis Center (NTIAC), NT-SP-01-03. 2001.
- Aldrin, J. C., Cheng, A., Achenbach, J. D., Andrew, G. A., Mullis, R. T., editors; Thompson, D. O., Chimenti, D. E., "Detection of Cracks in Weep Holes Using Neural Networks." *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 19B, 2000, pp. 1979 – 1986.
- Beissner, R. E. and Graves, III, J. S., "Computer Modeling of Eddy Current Probability of Crack Detection". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 9, Plenum Press, New York, 1990, pp. 885-891.
- Beissner, R. E., Bartels, K. A. and Fisher, J. L., "Prediction of the Probability of Eddy Current Flaw Detection" *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 7b, Plenum Press, New York, 1988, pp. 1753-1760.
- Child, F. R., D. H. Phillips, L. W. Liese and W. D. Rummel. editors; Thompson, D. O., Chimenti, D. E., "Quantitative Assessment of the Detectability of Ceramic Inclusions in Structural Titanium Castings by X-Radiography," *Review of Progress in Quantitative Nondestructive Evaluation*, Volume 18B, 1999, pp. 2311 – 2317.
- Chiou, Chien-Ping, Margetan, Frank J. and Thompson, R. Bruce, "Ultrasonic Signal Characterization of Flat-Bottom Holes in Titanium Alloys: Experiment and Theory". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 14b, Plenum Press, New York, 1995, pp. 2121-2128.
- Chiou, Chien-Ping, Margetan, Frank J. and Thompson, R. Bruce, "Modeling of Ultrasonic Signals from Weak Inclusions". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 15a, Plenum Press, New York, 1996, pp. 49-55.
- Coffey, J. M., Chapman, R. K. and Hanstock, D. J., "The Ultrasonic Detectability of a Postulated Worst Case Flaw in a PWR Pressure Vessel", CEGB Report NW/SSD/82/0045/R April, 1982, CEGB, London, England, 1982.
- Elshafiey, I., Udpa, L. and Udpa, S. S., "Implementation of Probability of Detection Models for Eddy Current NDE Techniques on Parallel Computers". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 13, Plenum Press, New York, 1994, pp. 2167-2174.
- Fertig, K. W. and Richardson, J. M., *Computer Simulation of Probability of Detection*, Rockwell International Science Center, Thousand Oaks, CA, 1983, pp. 147 – 169.
- Forsyth, D. S., Fahr, A., editors; Thompson, D. O. Chimenti, D. E., "On the Independence of Multiple Inspections and the Resulting Probability of Detection," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 19B, 2000, pp. 2159 – 2166.

Forsyth, D. S., Fahr, A., Leemans, D. V., McRae, K. I., "Development of POD From In-Service NDI Data," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D. E., Vol. 19B, 2000, pp. 2167 – 2174.

Gray, J. N., Gray, T. A., Nakagawa, N. and Thompson, R. B., "Models for Predicting NDE Reliability", *Nondestructive Evaluation and Quality Control, ASM Handbook*, vol. 17, ASM International, Metals Park, OH, 1989, pp. 702 – 715.

Grills, R., "Probability of Detection – An NDT Solution", *Materials and Evaluation*, Vol. 59, No. 7, July 2001, pp. 847-854.

Jeffrey, B. and M. L. Peterson. "Assessment of Rail Flaw Inspection Data." Mountain-Plains Consortium Final Report 98-93. North Dakota State University, Fargo, ND. 1999.

Khaleel, M. A. and F. A. Simonen. "The Effects of Initial Flaw Sizes and Inservice Inspection on Piping Reliability." 1994 ASME Pressure Vessel and Piping Conference, American Society of Mechanical Engineers, PVP-Vol. 288. 1994. pp. 95-107.

Khaleel, M. A., Simonen, F. A., "A Model for Predicting Vessel Failure Probabilities Including the Effects of Service Inspection and Flaw Sizing Errors," *Nuclear Engineering and Design*, Vol. 200. 2000. pp.353-369.

Klemmt, Wulf-Bernhelm and Tober, Gustav, "Statistic Approach Based on 'Monte-Carlo Simulation' to Calculate Inspection Reliability for Multiple Site Damage". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 14, Plenum Press, New York, 1995, pp. 2383-2390.

Leemans, D. V. "Probability of Detection Based on Field Inspection Data." Final Project Report for Department of National Defence (Canada) Air Vehicles Section. 1998.

Meeker, W. Q. ,Victor Chan, R. Bruce Thompson, and Chien-Ping Chiou. "A Methodology for Predicting Probability of Detection for Ultrasonic Testing." *Review of Progress in Quantitative Nondestructive Evaluation*, ed. by D. O. Thompson and D. E. Chimenti, Vol. 20, 2001. pp.1972-1978.

Meeker, William Q., Jeng, Shuen-Lin, Chiou, Chien-Ping, and Thompson, R. Bruce, "Improved Methodology for Inspection Reliability Assessment for Detecting Synthetic Hard Alpha Inclusions in Titanium". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 17b, Plenum Press, New York, January, 1998, pp. 2061-2068.

Meeker, William Q., Thompson, R. Bruce, Chiou, Chien-Ping, abd Jeng, Shuen-Lin, "Methodology for Estimating Nondestructive Evaluation Capability". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 15b, Plenum Press, New York, 1996, pp. 1983-1990.

Mueller, C., T. Fritz, G. Tillack, C. Bellon, M. Scharmach, "Theory and Application of the Modular Approach to NDT Reliability", *Materials and Evaluation*, July 2001.

Munns, I. J., Schneider, C. R. A., "The Reliability of Radiography of Thick Section Welds," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 19B, 2000, pp. 2151 – 2158.

Nakagawa, N. and Beissner, R. E., "Probability of Tight Crack Detection Via Eddy-Current Inspection", *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O. Chimenti, D. E., Volume 9, Plenum Press, New York, 1990, pp. 893 – 899.

Nakagawa, N., Kubovich, M. W. and Moulder, J. C., "Modeling Inspectability for an Automated Eddy Current Measurement System". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 9, Plenum Press, New York, 1990, pp. 1065-1072.

Ogilvy, J. A., "Model for Predicting Ultrasonic Pulse-Echo Probability of Detection". *NDT&E International*, vol. 26, no. 1, Butterworth-Heinemann LTD., Oxford, UK, 1993, pp. 19-29.

Okure, Mackay A. and Peshkin, Michael A., "Quantitative Evaluation of Neural Networks for NDE Applications Using the ROC Curve". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 14b, Plenum Press, New York, 1995, pp. 2405-2412.

Peterson, M. L., Jeffrey, B. D., Gutkowski, R. M., "Limitations in Size and Type of Detectable Defects in Rail Flaw Inspection," *INSIGHT*, Volume 42, No. 5, May, 2000, pp. 306 – 311.

Rajesh, S. N., Udpa, L., Udpa, S. S. and Nakagawa, N., "Probability of Detection (POD) Models for Eddy Current NDE Methods", *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 11a, 1991, pp. 265-272.

Rajesh, S. N., et al, in IEEE Conference on Electromagnetic Field Computations, June 1992.

Rajesh, S. N., Udpa, L. and Udpa, S. S., "Estimation of Eddy Current Probability of Detection (POD) Using Finite Element Method". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 12b, Plenum Press, New York, 1993a, pp. 2365-2372.

Rajesh, S. N., Udpa, L. and Udpa, S. S., "Numerical Model Based Approach for Estimating Probability of Detection in NDE Applications". *IEEE Transactions on Magnetics*, Institute of Electrical and Electronic Engineers, Inc., New York, Vol. 29, No. 2, 1993b, pp. 1857-1858.

Rummel, P. E., W. D., "Transfer of POD Performance Capabilities From Simple Shapes to Complex Shapes," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D. E., Volume 18, 1999, pp. 2305 – 2310.

Rummel, W. D., "Requirements for Demonstrating Nondestructive Evaluation Capabilities," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D. E., Vol. 19B, 2000, pp. 2175 – 2181.

Rummel, W. D. and G. A. Matzkanin. "Nondestructive Evaluation (NDE) Capabilities Data Book, Third Edition." NTIAC-DB-97-02, Nondestructive Testing Information Analysis Center, 1997.

Sarkar, Pradipta, Meeker, William Q., Thompson, R. Bruce, Gray, Timothy A. and Junker, Warren, "Probability of Detection Modeling for Ultrasonic Testing". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 17b, Plenum Press, New York, July, 1998, pp. 2045-2052.

Schmerr, Jr., L. W., Song, S-J., "The Use of Probabilistic Neural Networks for Discontinuity Classification Problems," *Materials Evaluation*, Vol. 58, No. 1, January, 2000, pp. 80 – 84.

Schmerr, Jr., Lester W. and Thompson, Donald O., "NDE and Design – A Unified Life-Cycle Engineering Approach". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 12b, Plenum Press, New York, 1993, pp. 2325-2331.

Schmerr, Jr., Lester W. and Thompson, Donald O., "NDE Models and Design – A Unified Life-Cycle Engineering". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 13, Plenum Press, New York, 1994, pp. 2183-2190.

Schmerr Jr., Lester W. and Thompson, Donald O., *The Role Of Modeling in NDE Standards*, The Center for Nondestructive Evaluation, Iowa State University, Ames, Iowa, 50011, September, 1997 25 pp.

Segal, Y., Notea, A. and Segal, A., *Research Techniques in NDT*, Vol. 3, Chapter 9, Ed. R. S. Sharpe, Academic Press, London England, 1977

Singh, R., "Three Decades of NDT Reliability Assessment and Demonstration," *Materials and Evaluation*, Vol. 59, No. 7, July 2001, pp. 856-860.

Simola, K., Pulkkinen, U., "Models for Non-destructive Inspection Data," *Reliability Engineering and System Safety*, Volume 60, No., 5, 1998, pp. 1 – 52.

Song, S. J., Park, H. J., Shin, Y. K., Lee, H. B., editors; Thompson, D. O., Chimenti, D. E., "Flaw Characterization in Tubes by Inversion of Eddy Current Signals Using Neural Networks Trained by Finite Element Model-Based Synthetic Data," *Review of Progress in Quantitative Nondestructive Evaluation*, Volume 18, 1999, pp. 881 – 888.

Spencer, F. W., "Identifying Sources of Variation for Reliability Analysis of Field Inspections," *RTO Meetings Proceedings 10: Airframe Inspection Reliability Under Field/Depot Conditions*, 1998, pp. 11-1 – 11-8.

Spencer, F., "Estimating Probability of Detection Curves from Regression Data," *Materials Evaluation*, Vol. 59, No. 7, July 2001, pp. 866-870.

Thompson, R. Bruce, "Use of Elastodynamic Theories in the Stochastic Description of the Effects of Microstructure on Ultrasonic Flaw and Noise Signals," Submitted to a special issue of *Wave Motion* in honor of contributions by Achenbach.

Thompson, R. Bruce, "Overview of the ETC POD Methodology". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 18b, Plenum Press, New York, July 19-24, 1998, pp. 2295-2304.

Thompson, R. Bruce. "Using Physical Models of the Testing Process in the Determination of Probability of Detection." *Materials Evaluation*, Vol. 59, No. 7, 2001. pp.861-865.

Thompson, R. B. and Achenbach, J. D., EPRI Report NP-3822. 1985.

Thompson, R. B. and Gray, T. A., "A Model Relating Ultrasonic Scattering Measurements Through Liquid-Solid Interfaces to Unbound Medium Scattering Amplitudes", *J. Acoustical Soc. Of America*, 74 (4), 1983, 1279-1290

Thompson, R. B. and Meeker Jr., W. Q., "A Methodology for POD Determination Incorporating Insight From Physical Models", *European-American Workshop, Determination of Reliability and Validation Methods of NDE*, Berlin, June, 18 – 20, 1997, pp. 107 – 114.

Thompson, Donald O. and Schmerr, Jr., Lester W., "Uses for Model-Based Probability of Detection Curves". *Nondestructive Inspection of Aging Aircraft*, SPIE, vol. 2001, 1993, pp. 121-132.

Tow, David M. and Reuter, Walter G., "Probabilistic Model for Pressure Vessel Reliability Incorporating Fracture Mechanics and Nondestructive Examination". *Nondestructive Evaluation of Utilities and Pipelines II*, SPIE, vol. 3398, April 1, 1998, pp. 168-176.

Wall, M., "Modelling of NDT Reliability and Applying Corrections for Human Factors", *European-American Workshop, Determination of Reliability and Validation Methods of NDE*, Berlin, June 18-20, 1997, pp. 87-98.

Wall, M. and Wedgwood, F. A., "NDT – Value for Money". *Insight*, vol. 36, no. 10, British Institute of Nondestructive Testing, Northhampton, UK, 1994, pp. 782-790.

Wormley, S. J., Shah, H., Sheth, D., Forouraghi, B. and Schmerr, Jr. L. W., "An Application of Object-Oriented Technology to Unified Life Cycle Engineering". *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 12a, Plenum Press, New York, 1993, pp. 811-817.

Zhang, Z., Zhang, Y., Udpa, L. Udpa, S. S., "Probability of Detection Model for Pipeline Inspection," "Review of Progress in Quantitative Nondestructive Evaluation, edited by; Thompson, D. O., Chimenti, D. E., Volume 16B, 1997, pp. 1307 – 1314.

Bibliography

Achenbach, J. D., "Quantitative Nondestructive Evaluation," *International Journal of Solids and Structures*, Volume 37, No. 5, 2000, pp. 13 – 27.

Berens, Alan P., "NDE Reliability Data Analysis". *Nondestructive Evaluation and Quality Control, ASM Handbook*, vol. 17, ASM International, Metals Park, OH, 1992, pp. 689-701.

Burkel, R. H., Sturges, D. J., Tucker, W. T., Gilmore, R. S., "Probability of Detection for Applied Ultrasonic Inspections," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D. E., Volume 15B, 1996, pp. 1991 – 1998.

Chiou, Chien-Ping, R. Bruce Thompson, Victor Chan, and William Q. Meeker, "Ultrasonic and Statistical Analyses of Hard-Alpha Defects in Titanium Alloys," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 20, Plenum Press, New York, 2001, pp. 1979-1986.

Ewing, Anthony, "SQUID NDE and POD Using a BEM Measurement Model," Vanderbilt University, Nashville, Tennessee, April, 1998.

Fahr, A., Forsyth, D., "POD Measurement Using Actual Components," *SPIE*, Volume 3397, 1998. pp. 194 – 202.

Fahr, A., Forsyth, D. S., "POD Assessment Using Real Aircraft Engine Components." *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D.E., Volume 17B, 1998, pp. 2005 – 2012.

Flood, J. T., Pfau, R. E., Grauvogl, E., Regler, F., "Field Inspection Results and Damage Analysis of F-4F Horizontal Stabilizer Internal Structure," *RTO Meetings Proceedings 10: Airframe Inspection Reliability Under Field/Depot Conditions*, 1998, pp. 16-1 – 16-8.

Forsyth, D. S., Fahr, A., Martineau, N., "Inspection Reliability Assessment," *Canadian Aeronautics and Space Journal*, Volume 43, No. 5, 1997, pp. 50 – 55.

Heida, J. H., Grooteman, F. P., "Airframe Inspection Reliability Using Field Inspection Data," *RTO Meetings Proceedings 10: Airframe Inspection Reliability Under Field/Depot Conditions*, 1998, pp. 5-1 – 5-5.

International Standard, Second Edition, ISO 9712:1999(E), *Non-destructive Testing – Qualification and Certification of Personnel*, International Organization for Standardization, Case postale 56, CH-1211 Geneve 20, Switzerland, 1999, 22 pp.

Joyner, M. L., Banks, H. T., Wincheski, B., Winfree, W. P., editor; Mal, A. K., "A Reduced Order Computational Methodology for Damage Detection in Structures," *Nondestructive Evaluation of Aging Aircraft, Airports, and Aerospace Hardware IV, SPIE*, Volume 3994, 2000, pp. 10 – 17.

Meeker, William Q., Jeng, Shuen-Lin, Chiou, Chien-Ping, and Thompson, R. Bruce, "Improved Methodology for Predicting POD of Detecting Synthetic Hard Alpha Inclusions in Titanium". *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 16, Plenum Press, New York, 1997, pp. 2021-2028.

Mendoza, Y., Osegueda, R., "Theoretical Explanation for the Empirical Probability of Detection (POD) Curve: A Neural Network-Motivated Approach," *42nd Midwest Symposium on Circuits and Systems*, Volume 5, 2000, pp. 435 – 438.

Moulder, John, C., Degeratu, Alexandru, and Nakagawa, Norio, "Predicting the Performance of Eddy Current Probes". *1992 JANNAF Nondestructive Evaluation Subcommittee Meeting*, CPIA-584, Johns Hopkins University, Chemical Propulsion Information Agency, Columbia, MD, July, 1992, pp. 171-179.

Nagem, R. J., Seng, J. M., Williams, Jr., J. H., "Residual Life Predictions of Composite Aircraft Structures Via Nondestructive Testing, Part 1: Prediction Methodology and Nondestructive Testing," *Materials Evaluation*, September, 2000, pp. 1065 – 1074.

Nagem, R. J., Seng, J. M., Williams, Jr., J. H., "Residual Life Predictions of Composite Aircraft Structures Via Nondestructive Testing, Part 2: Degradation Modeling and Residual Life Prediction," *Materials Evaluation*, November, 2000, pp. 1310 – 1319.

Rudlin, J. R. and Wolstenholme, L. C., "Development of Statistical Probability of Detection Models Using Actual Trial Inspection Data". *British Journal of Non-Destructive Testing*, vol. 34, no. 12, British Institute of Non-Destructive Testing, Northampton, UK, December, 1992, pp. 583-589.

Rummel, W. D., "Overview of Nondestructive Evaluation (NDE) Capabilities Assessments and Probability of Detection (POD) Concepts," *14th World Conference on Nondestructive Testing (14th WCNDT)*, New Delhi, India, December 8 – 13, 1996, pp. 1-5.

Rummel, W. D., "Probability of Detection As a Quantitative Measure of Nondestructive Testing End-To-End Process Capabilities," *Materials Evaluation*, Vol. 59, No. 1, January, 1998, pp. 29 – 35.

Rummel, W. D., Matzkanin, G. A., edited by; Thompson, D. O., Chimenti, D. E., "How Good is Your Inspection? How Do You Measure Up?," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D. E., Volume 18, 1999, pp. 2353 – 2359.

Rummel, W. D., Matzkanin, G. A., "NDE Capabilities Data, What Might I Expect?, What Is The Prior Art?," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by; Thompson, D. O., Chimenti, D. E., Vol. 19B, 2000, pp. 2145 – 2150.

Sweeting, Trevor, "Statistical Models for Nondestructive Evaluation". *International Statistical Review*, vol. 63, no. 2, International Statistical Institute, Voorburg, Netherlands, 1995, pp. 199-214.

Thompson, D. O. and Schmerr, Jr., L. W., "The Role of Modeling in the Determination of Probability of Detection". *Advances in Signal Processing for Nondestructive Evaluation of Materials; NATO ASI Series, Series E, Applied Sciences* vol. 262, Kluwer Academic Publishers, Dordrecht, Netherlands, 1994, pp. 285-301.

Woolridge, A. B. Chapman, R. K., Woodcock, G. S., Munns, I., Georgiou, G., "Reliability of Radiography for Detection of Planar Manufacturing Defects in Thick-Section Welds," *INSIGHT*, Volume 39, No. 3, March, 1997, pp. 139 – 147.

Worrall, G. M., "Practical Assessment of Inspection Reliability," *INSIGHT*, Volume 39, No. 3, March, 1997, pp. 153 – 155.