Abstract

The goal of this paper is to make available a rather simple method for comparing the effectiveness of various NDE techniques based on the *Probability of Detection* (POD) and *False Alarm Rate* (FAR) parameters associated with the techniques. Concepts relating NDE to component reliability are introduced to facilitate evaluation of various ultrasonic NDE methods. NDE reliability is defined in terms of NDE decisions "Flaw" or "No Flaw." The concepts of *Success Probability Ratio* (SPR) and *Quality Factor* (QF) are developed in terms of the *Probability of Flaw* (P(f)) and the critical NDE factors, POD and FAR. Bayes Theorem is described and utilized to generate the relations necessary to describe important NDE considerations such as inspection based confidence in the integrity of a component and confidence in NDE decisions. An ultrasonic NDE weld inspection technique known as "FAST™ " is introduced and the physical basis for its effectiveness as a weld inspection tool is elaborated. Results obtained from a specialized and well-characterized collection of weld specimens using both conventional and FAST methods are presented. FAST results and conventional results are cast in terms of the reliability model used within this paper. The data resulting from a comparative study of methods is presented in a graphical fashion that describes the perceived reliability of an inspected component and the confidence afforded the NDE technique used. Appendices are included that review probability concepts, the FAST technique, and the underlying physical basis for FAST performance as an ultrasonic NDE technique.

Introduction

No nondestructive evaluation (NDE) technique is ideal for each and every application to which it may be applied. Flaws are missed and non-flaws are called flaws. Thus the parameters Probability of Detection (POD) and False Alarm Rate (FAR) have real meaning in the framework of NDE. The NDE manager, then, save going to redundant methodology [1], must exercise judgement regarding which of the options available to him is most fitting to the inspection task he has at hand. The goal of this paper is to make available a rather simple method for comparing the effectiveness of various NDE techniques based on the Probability of Detection and False Alarm Rate parameters associated with the techniques. The method is based on Bayes Rule, a statistical relationship that enables calculation of the probability of an event occurring given that some other event has already occurred.

Having a simple method available for comparisons of NDE techniques facilitates decisions regarding which NDE technique to use in light of the economic and safety issues involved. Safety issues are generally associated with the reliability of a component for continued service without failure. Economic issues are associated with the cost to implement the particular NDE technique, the time to perform the NDE, and the actions that follow after the NDE has been used (e.g., unnecessarily replacing a good component following a false call). Generally, both safety and economic issues are involved. For

™ FAST is a trademark of Sikorski & Pfannenstiel Innovative NDE example, in a nuclear power plant environment, a false call may cause engineering personnel to be unnecessarily exposed to radiation at the site of the supposed flaw.

The method presented here will allow the user to identify the strengths and weaknesses of various NDE techniques. Through the user's assessment of how these play into his inspection program, he can select the more appropriate NDE.

To illustrate the use of the method, performance data from a relatively new ultrasonic NDE technique known as FAST™ (See Appendix A) is compared with conventional NDE techniques. All techniques were evaluated on a specialized and well-characterized collection of welded pipe specimens containing cracks and other types of conditions. Data sets consisted of a POD and a FAR for each NDE technique per the specific type of sample examined.

Relevant Conditional Probabilities

Component reliability can be defined in a number of ways. Cast probabilistically, as a number between 0 and 1, it expresses our belief or confidence in the fitness of component for continued service use. To obtain some measure of the validity of a reliability estimate, NDE methods are used. Based on historical evidence, scientific analysis, and other considerations, the likelihood of a flaw occurring in a component can be estimated; NDE either supports this estimate or indicates that a modified estimate is more appropriate.

The link between component reliability and NDE can be established by introducing terms and mathematical relations that describe our characterization of the NDE process. The more familiar terms are Probability of a Flaw, P(f), Probability of Detection, POD, and False Alarm Rate, FAR. Less familiar terms are the terms Success Probability Ratio, SPR and Quality Factor, QF.

Implicitly, these terms are defined as conditional probabilities. A conditional probability is typically annotated a P(B | A) and interpreted as "The probability of condition B given that condition A is present. For example, POD can written as P(F | f) where "f" means a flaw is present, and "F" means NDE calls a flaw. POD is then interpreted as "the probability that NDE will call a flaw when, in fact, a flaw is present."

Developing a Relationship Between NDE and Reliability

Mathematically, and in probabilistic terms, reliability [2] can be written as

R = P(no flaw exists | NDE says no flaw)

To develop a relationship between reliability and NDE capability, Bayes Rule [3] can be used. Bayes Rule (see Appendix B) is formulated in the following manner:

$$P(A|B)\,P(B) = P(B|A)\,P(A)$$

The form *P(x)* is simply symbolism for the probability that condition *x* occurs.

To simplify notation, the following convention will be used.

Lower case letters - the actual condition or state of nature. For example, “nf” would indicate the condition that no actual flaw is present

Upper case letters - the result of a NDE evaluation, or “NDE says …..” For example, “NF” would indicate that NDE evaluation suggested that no flaw was present.

(1) POD - Probability of Detection. Given a flaw actually exists, the probability that NDE will indicate a flaw exists. It is defined mathematically as $P(F \mid f)$.

(2) FAR - False Alarm Rate. Given that a flaw does not exist, the probability that NDE will indicate that a flaw does exist. It is defined mathematically as $P(F \mid nf)$

(3) P(f) - Probability that a flaw actually exists in a component.

Relationships key to the development of the reliability concept for NDE are:

(4) $$P(f) + P(nf) = 1$$

(5) $$P(NF) = P(NF \mid f) P(f) + P(NF \mid nf) P(nf)$$

(6) $$P(F \mid f) + P(NF \mid f) = 1 ; P(NF \mid f) = 1 - POD$$

(7) $$P(F \mid nf) + P(NF \mid nf) = 1; P(NF \mid nf) = 1 - FAR$$

Development begins with the definition for reliability and application of Bayes theorem;

$$R = P(nf \mid NF)$$

$$P(nf \mid NF) P(NF) = P(NF \mid nf) P(nf)$$

or

$$R = \frac{P(NF \mid nf) P(nf)}{P(NF)}$$

Using (5),

$$R = \frac{P(NF \mid nf)\, P(nf)}{P(NF \mid f)\, P(f) + P(NF \mid nf)\, P(nf)}$$

Dividing through by P(NF | nf) P(nf) and rearranging terms,

$$R = \frac{1}{1 + \dfrac{P(NF \mid f)\, P(f)}{P(NF \mid nf)\, P(nf)}}$$

Noting that P(nf) = 1 - P(f) from (4), the term $\dfrac{P(f)}{P(nf)}$

can be written as $\dfrac{P(f)}{1 - P(f)}$

Defining a “Quality Factor” can summarize this ratio.

$$QF = \frac{1 - P(f)}{P(f)}$$

Note that for P(f) =1., QF = 0; for P(f) =0., QF is infinite. Figure 1 shows the variation of QF with the probability of a flaw, P(f).

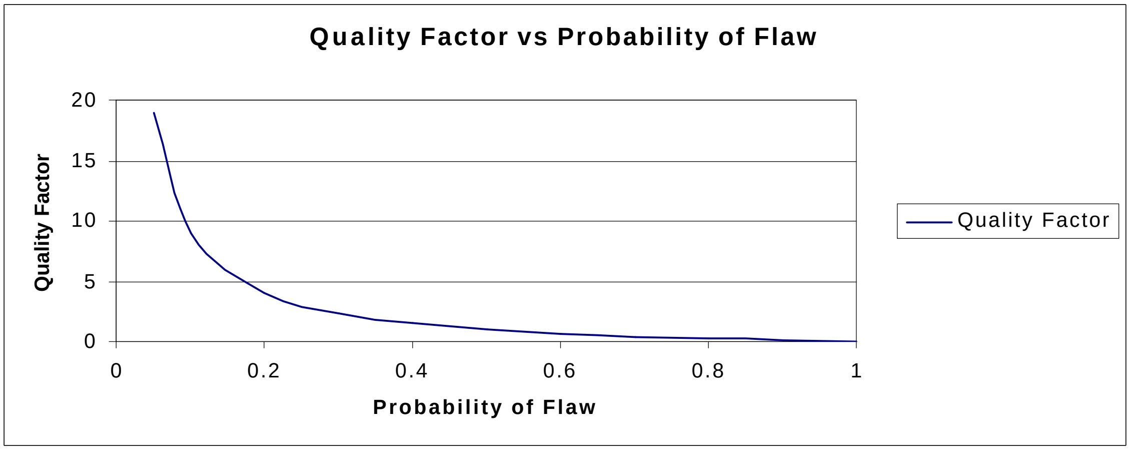

Figure 1 Quality Factor as a Function of Probability of a Flaw

Using QF,

$$R = \frac{1}{1 + \dfrac{P(NF \mid f)}{P(NF \mid nf)\, QF}}$$

Using relationships (6) & (7),

$$P(NF \mid f) = 1 - POD$$

$$P(NF \mid nf) = 1 - FAR$$

$$R = \frac{1}{1 + \frac{(1 - POD)}{(1 - FAR)\, QF}}$$

The ratio, $\frac{1 - POD}{1 - FAR}$ can be inverted as defined as a "Success Probability Ratio."

$$SPR = \frac{1 - FAR}{1 - POD}$$

Note that for POD = 1., SPR is infinite; for POD =0., SPR = 1 - FAR. Figure 2 shows how SPR varies with POD and FAR. As POD tends toward 1 and FAR tends toward 0, SPR tends toward ∞; SPR increases with POD and decreases with higher false alarm rates.

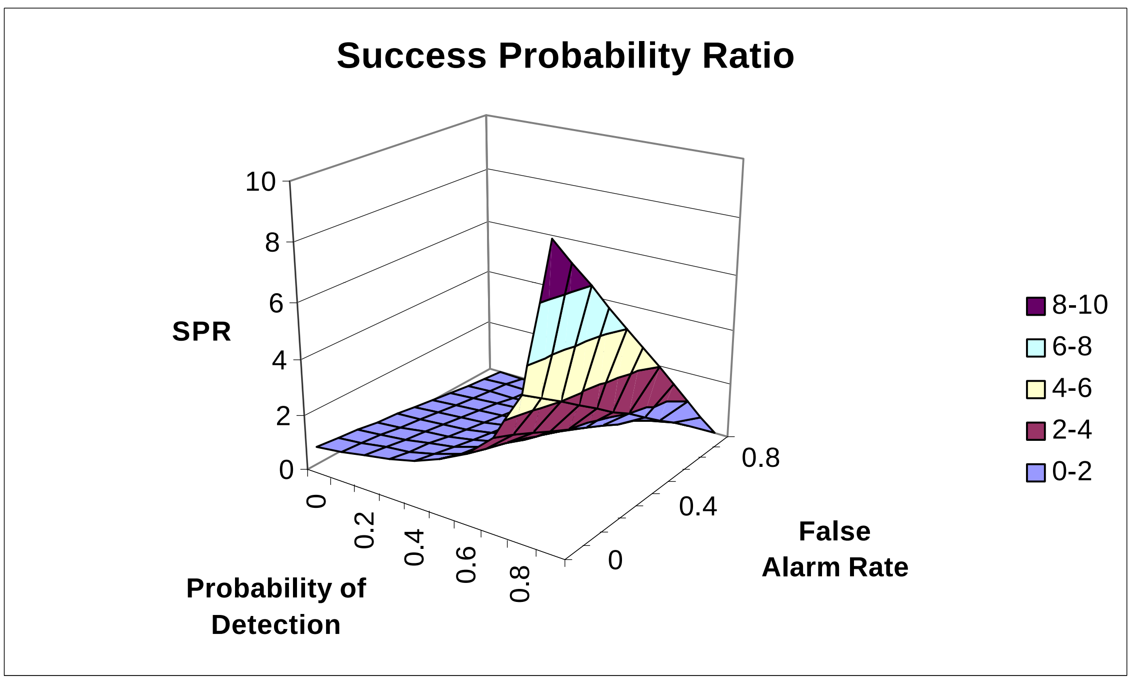

Figure 2. Variation of SPR with POD and FAR

The final form of R is then,

$$R = \frac{1}{1 + \frac{1}{SPR\ QF}}$$

A plot of reliability for two different NDE capabilities is shown in Figure 3. The x-axis is 1 - P(f). The right most side of the axis represents the condition P(f) = 0; the left most side, P(f) = 1. Since QF is generally not known to a high degree of confidence, all possibilities are considered via this approach.

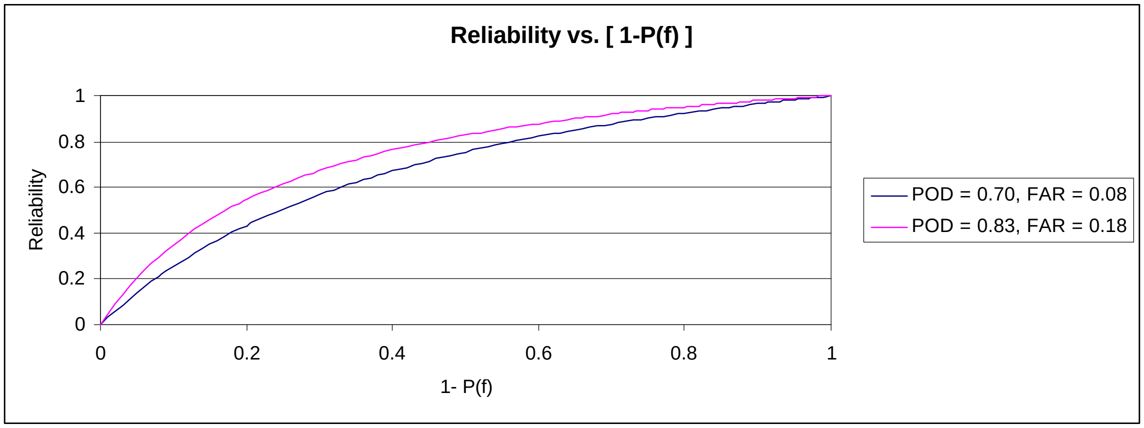

Figure 3 – Example of Reliability Curves as A Method for Comparing NDE Techniques

In general, reliability can be defined without consideration for component inspection. Classically, reliability is defined as 1-P(f) where P(f) is taken as the probability of flaw (that can cause component failure) as a function of time.

To show the reliability measure defined here in consistent with the more traditional definition, the SPR can be set to 1. That is, setting POD = FAR =0 to emulate no inspection capability, SPR=1[1]. Doing this reduces reliability to

$$R = \frac{1}{1 + \frac{1}{QF}}$$

Which reduces to

$$R = 1 - P(f)$$

after writing QF in terms of P(f).

Obviously, inspecting a component does not change the quality of it. What is changed is our belief in the quality of the item inspected. In this sense, the axis in Figure 3 labeled as 1 - P(f) is reliability without inspection.

Note, at the extreme cases of P(f) = 1 , certainty of a flaw, and P(f) = 0, certainty of a flawless component, that inspection does change our reliability estimate. It is only when we are uncertain (assign $0<P(f)<1$) about the reliability of a component that inspection helps us.

[1] The condition POD = FAR also emulates no inspection capability.

As an example, let P(f) = 0.100 be an estimate of the probability of a flaw before inspection is performed; R_{before} = 0.900 . Then, using values from the data used to plot Figure 3, $R_{POD=0.67}$ = 0.950 and $R_{POD=0.53}$ = 0.943 after inspection. Converting to flaw probabilities [P(f) = 1-R], $P(f)_{POD=0.67}$ = 0.050 and $P(f)_{POD=0.53}$=0.057 after inspection.

Developing a Relationship Between NDE and Others Factors

As it may turn out, some NDE techniques can exhibit very similar reliability curves or, in some cases virtually no difference between curves for each of the techniques. This latter case would place the techniques on equal footing from a reliability point of view. Additional criteria would be necessary for an NDE manager to make a well-founded decision regarding the technique to use.

An example of this scenario, based on the confidence that can be attributed to the calls of a particular method, is presented below.

P(f | F) is the probability that a flaw is present given that the NDE method says there is a flaw present. It is actually a measure of confidence in the NDE method being used. Economically, false calls are very costly. The time and effort typically allotted to the repair or removal and replacement of a flawed component is also allotted to an incorrectly labeled flaw condition.

Using Bayes Theorem and rearranging,

$$P(f \mid F) = \frac{P(F \mid f)\,P(f)}{P(F)} = \frac{POD\ P(f)}{POD\ P(f) + FAR\ P(nf)}$$

Dividing through by POD P(f), we have

$$P(f \mid F) = \frac{1}{1 + \frac{FAR\ P(nf)}{POD\ P(f)}}$$

The ratio, P(nf) / P(f) = (1 - P(f)) / P(f) is easily identified as the previously defined "Quality Factor" of the component. Using QF to represent the quality factor,

$$P(f \mid F) = \frac{1}{1 + \frac{FAR\ QF}{POD}}$$

This expression agrees with intuition. Note as either FAR and/or QF decrease, the probability of a correct call (call flaw, is a flaw) increases. As POD increases, the probability of a correct flaw call also increases. To perform an evaluation, QF must be allowed to vary over its full range. The reason being the lack of knowledge about the true number of flaws and the true number of non-flaws. Figure 4 shows an example of variation of P(f | F) with 1-P(f).

Comparison of Ultrasonic NDE Techniques

To show how the comparison method presented here can be used to assess NDE techniques, the FAST ultrasonic technique is compared with conventional techniques on the basis of reliability and on the basis of confidence in correct calls. The Table below shows POD and FAR data obtained from collection pipes containing defects (and geometry) that are either flaws or non-flaws (flaw meaning a defect that may increase the probability of failure for a component).

Figures 5 through 11 show the comparisons of reliability results. Figures 12 through 18 show results of confidence in flaw calls. In those Figures, the dark curves represent the FAST technique and the light curves conventional techniques. Table 1 shows the NDE setting applicable to the comparison tests. IGSCC is an acronym for Inter-Granular Stress Corrosion Cracking.

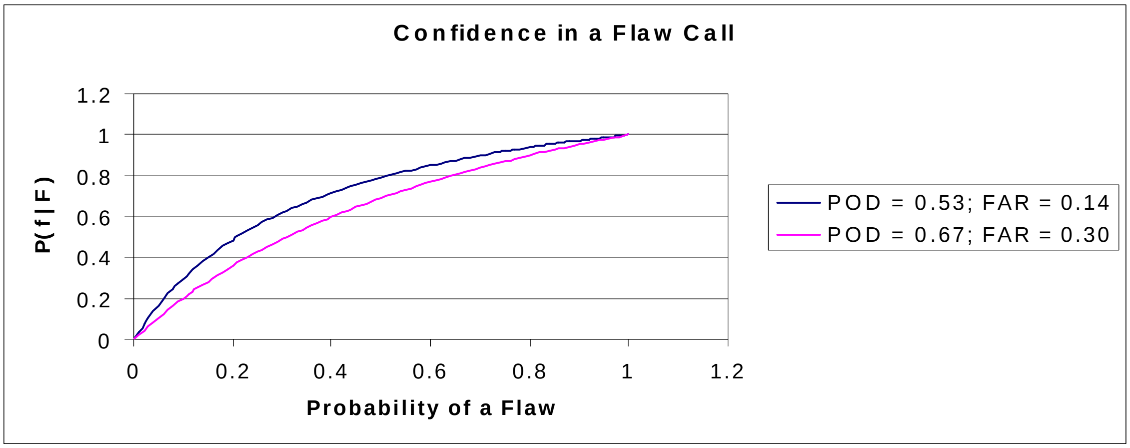

Figure 4 Confidence in Flaw Calls versa the Likelihood of a Flaw Being Present

Table 1 - POD and FAR Data used to Compare Techniques

Conditions	Conventional UT POD	Conventional UT FAR	FAST™ POD	FAST™ FAR
Austenitic Steel Pipe	0.83	0.18	0.81	0.08
Ferritic Pipe	0.97	0.14	0.95	0.08
Wall Thickness < 0.5”	0.89	0.17	0.89	0.05
0.5” < Wall Thickness < 1.5”	0.81	0.19	0.77	0.10
Wall Thickness > 1.5”	0.93	0.13	0.89	0.05
0.5” < Wall Thickness < 1.5” with no IGSCC	0.87	0.15	0.89	0.08
IGSCC Only	0.67	0.30	0.53	0.14

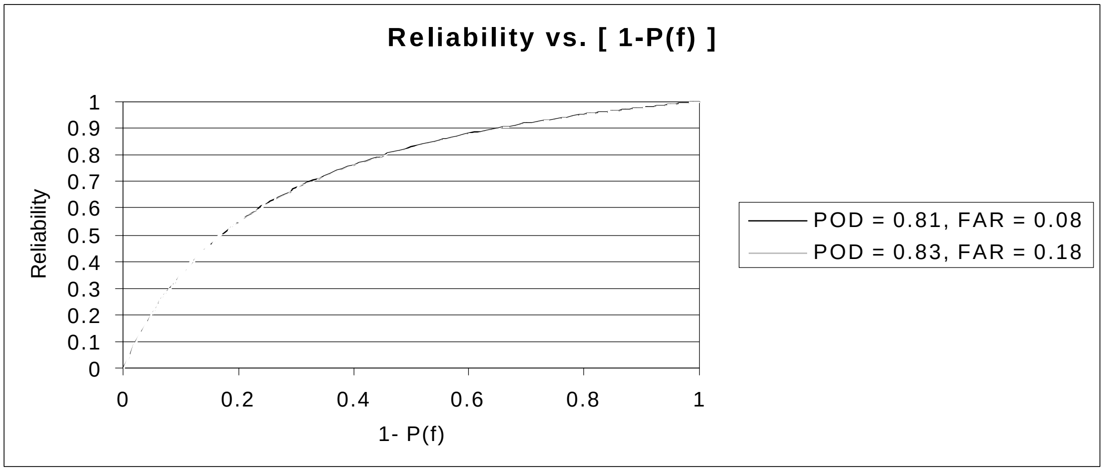

Figure 5. Reliability Comparison for Conventional Techniques and FAST for Category *All Austenitic*

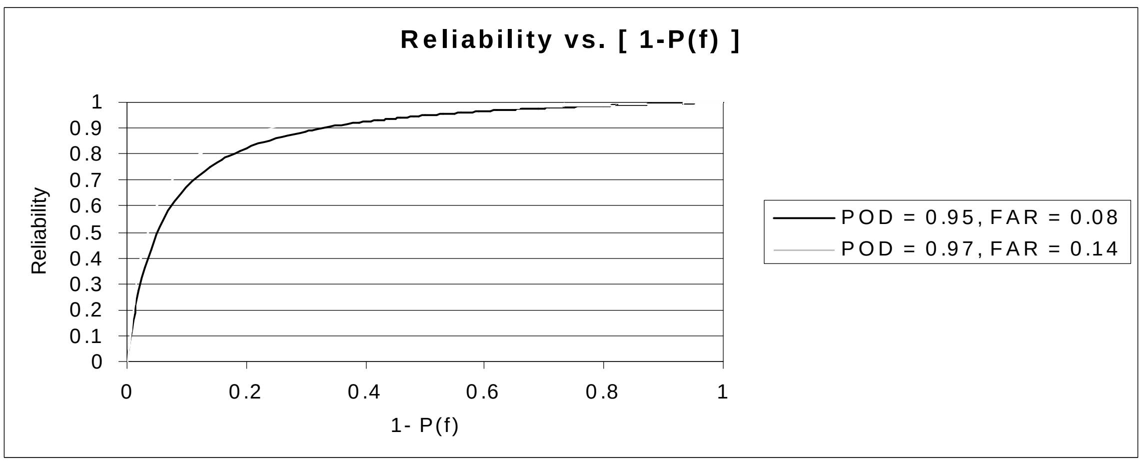

Figure 6. Reliability Comparison for Conventional Techniques and FAST for Category *All Ferritic*

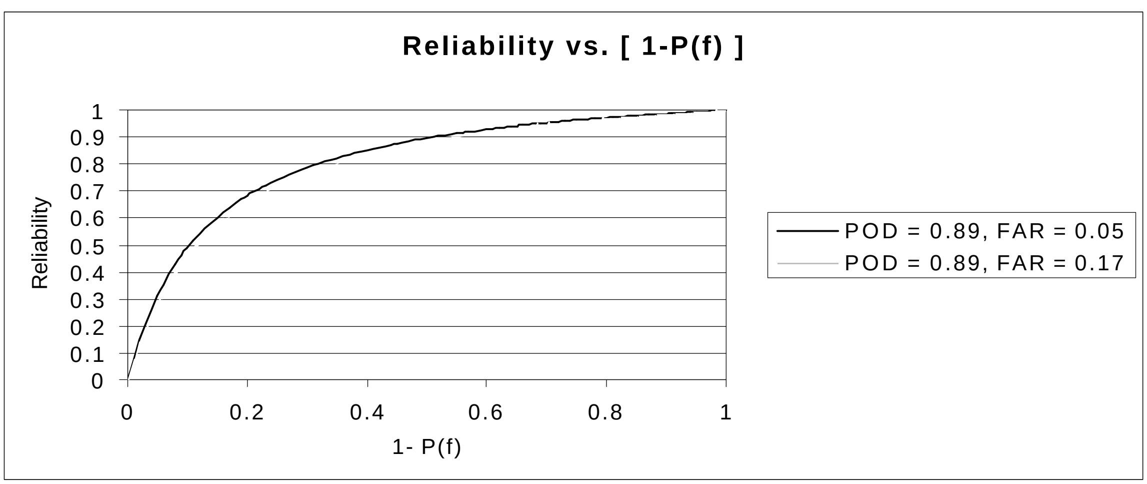

Figure 7. Reliability Comparison for Conventional Techniques and FAST for Category *All Materials & Flaw Types; Wall <0.5"*

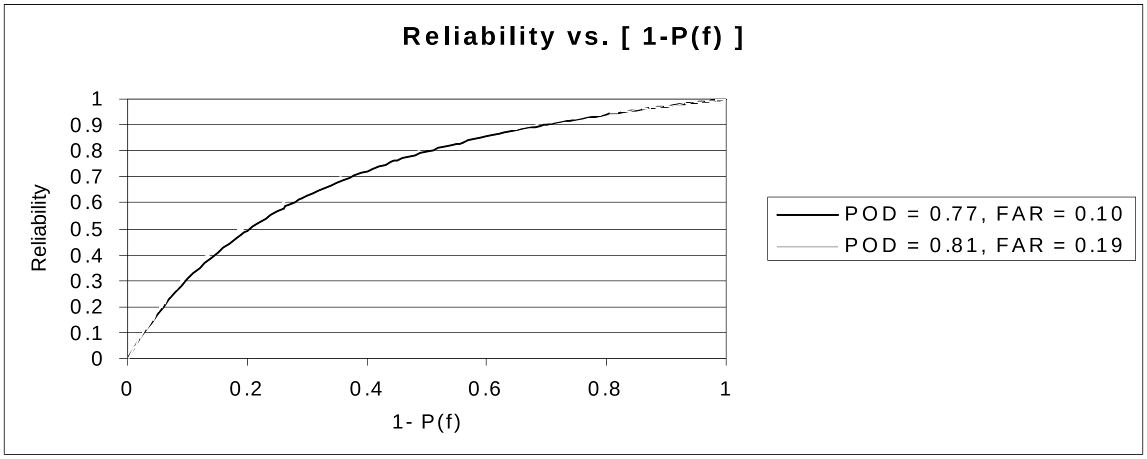

Figure 8. Reliability Comparison for Conventional Techniques and FAST for Category *All Materials & Flaw Types; 0.5" < Wall < 1.5"*

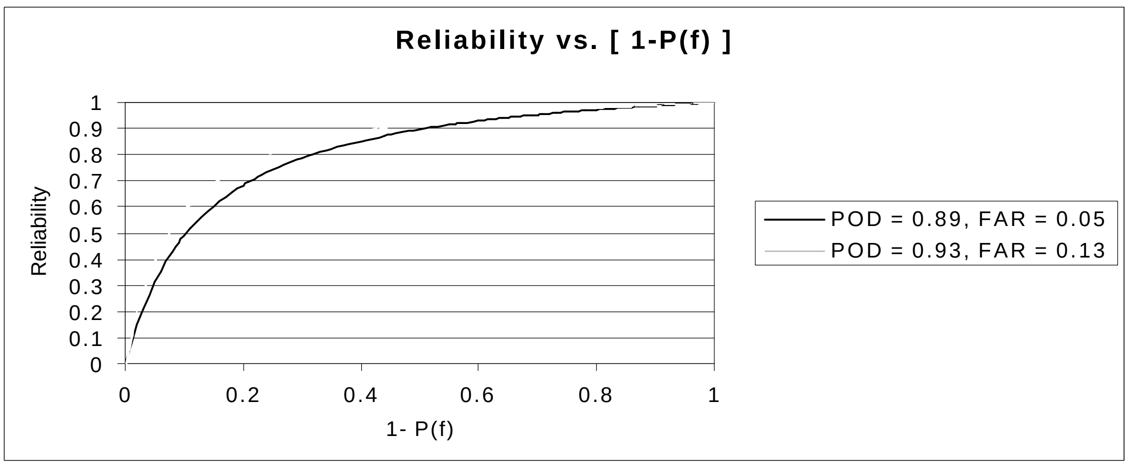

Figure 9. Reliability Comparison for Conventional Techniques and FAST for Category *All Materials & Flaw Types; Wall> 1.5"*

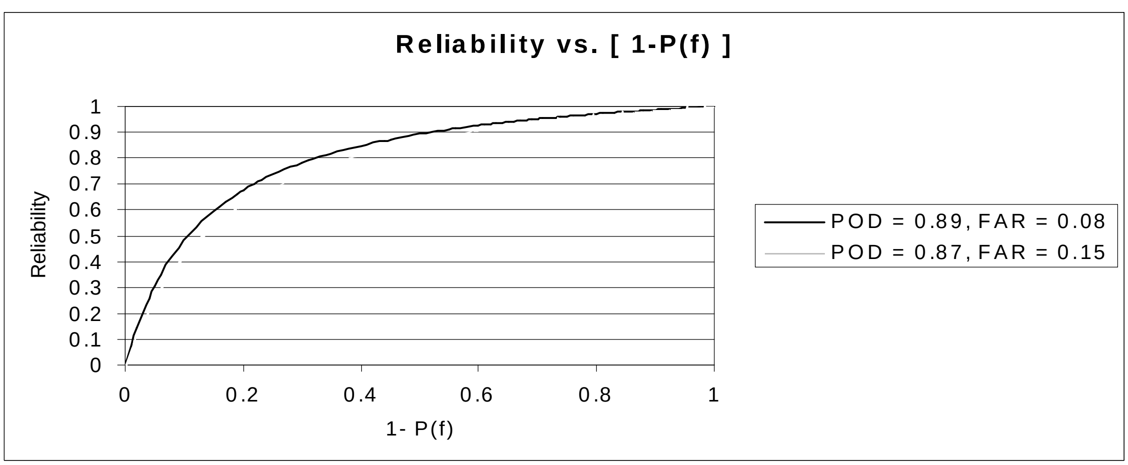

Figure 10. Reliability Comparison for Conventional Techniques and FAST for Category *All Materials with No IGSCC; 0.5" < Wall < 1.5"*

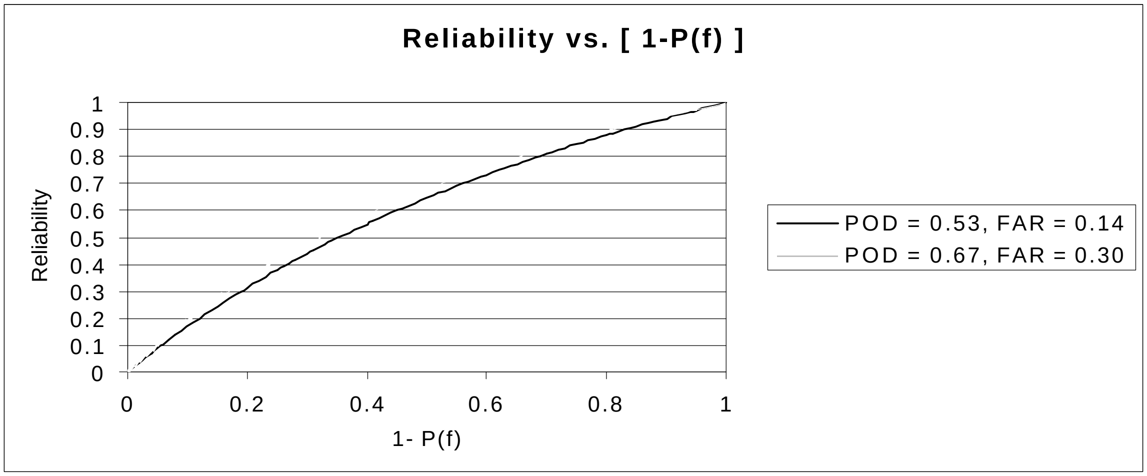

Figure 11. Reliability Comparison for Conventional Techniques and FAST for Category *IGSCC Only*

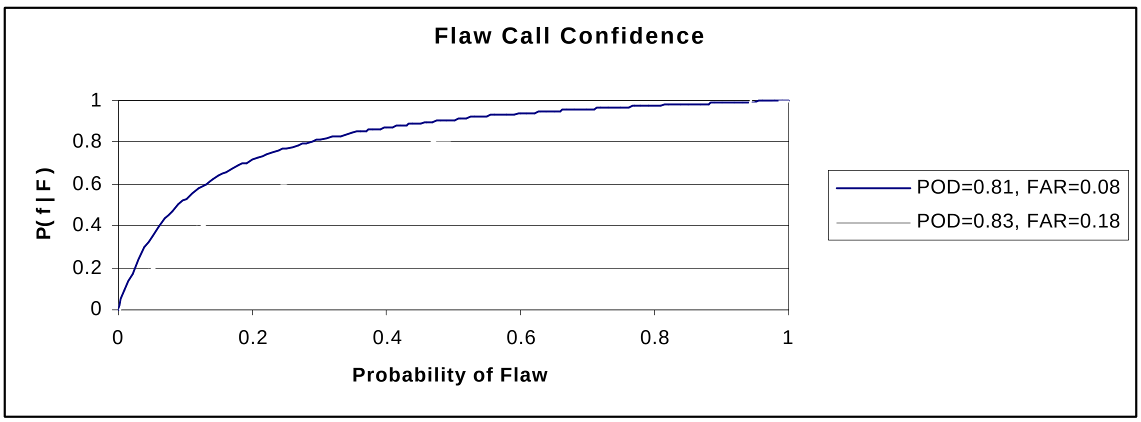

Figure 12 Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *All Austenitic*

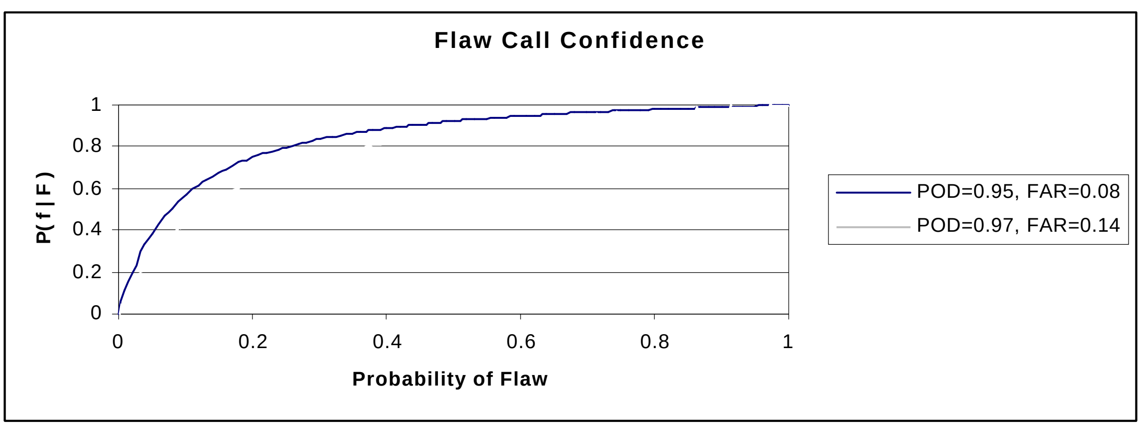

Figure 13 Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *All Ferritic*

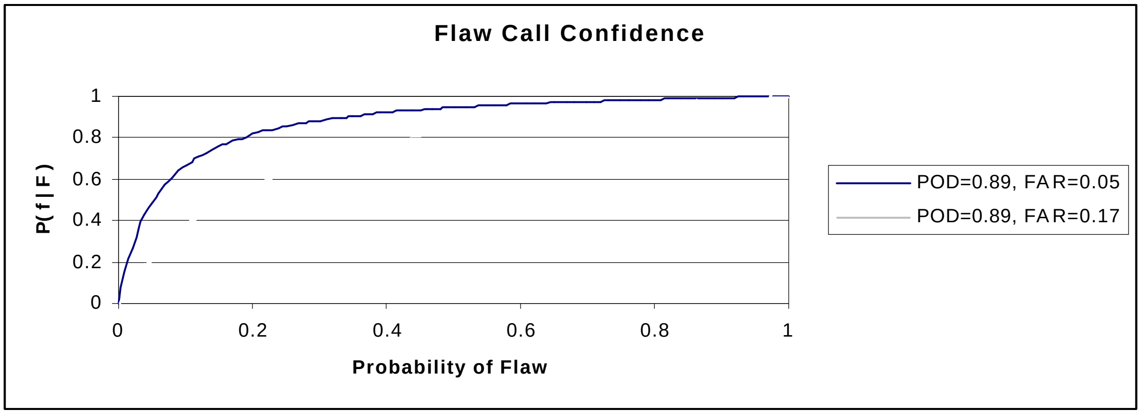

Figure 14. Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *All Materials & Flaw Types; Wall <0.5"*

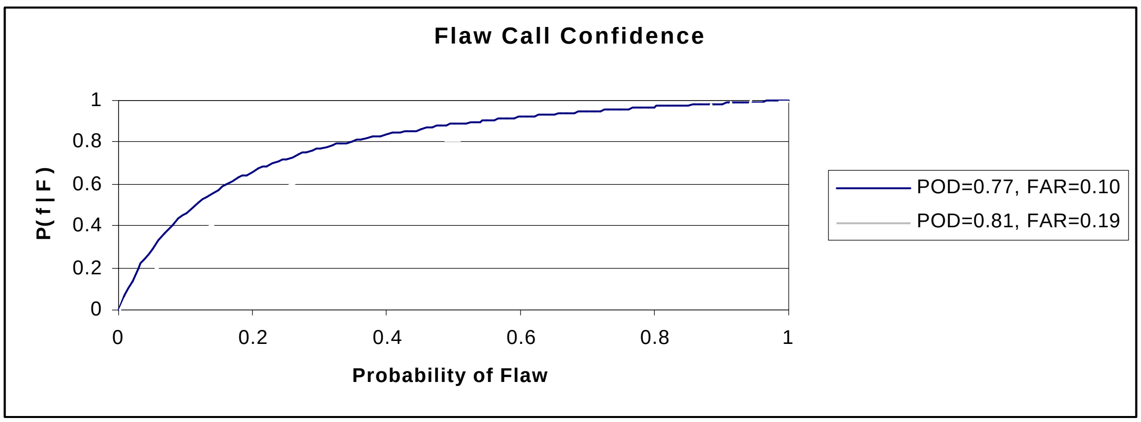

Figure 15. Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *All Materials & Flaw Types; 0.5"<Wall <1.5"*

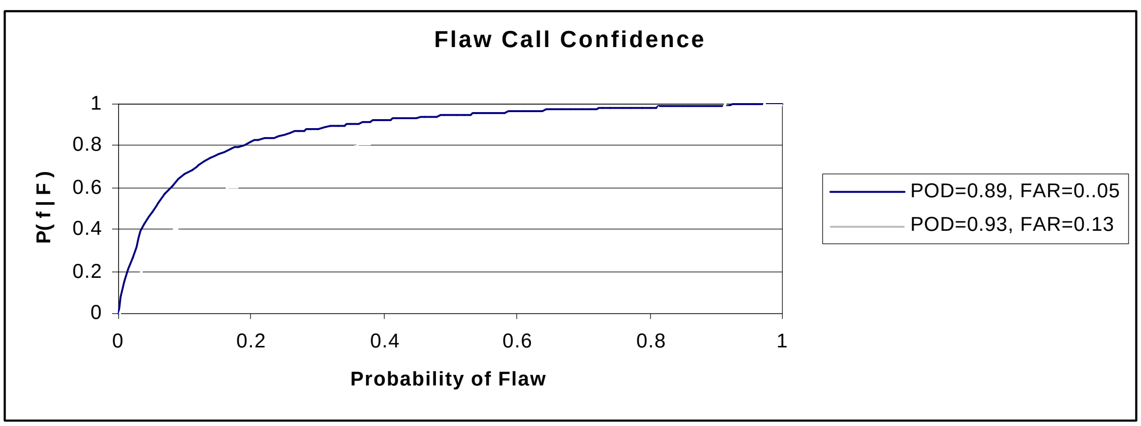

Figure 16. Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *All Materials & Flaw Types; Wall >1.5"*

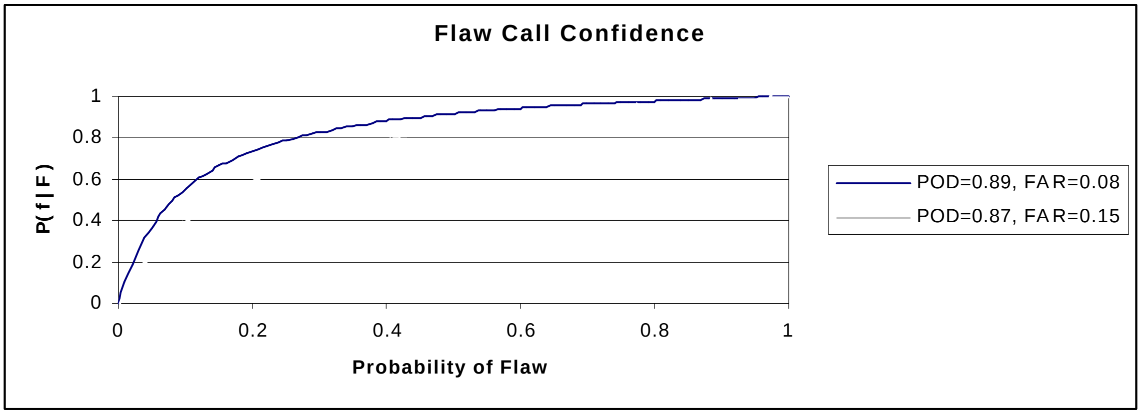

Figure 17. Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *All Materials with No IGSCC; 0.5"<Wall<1.5"*

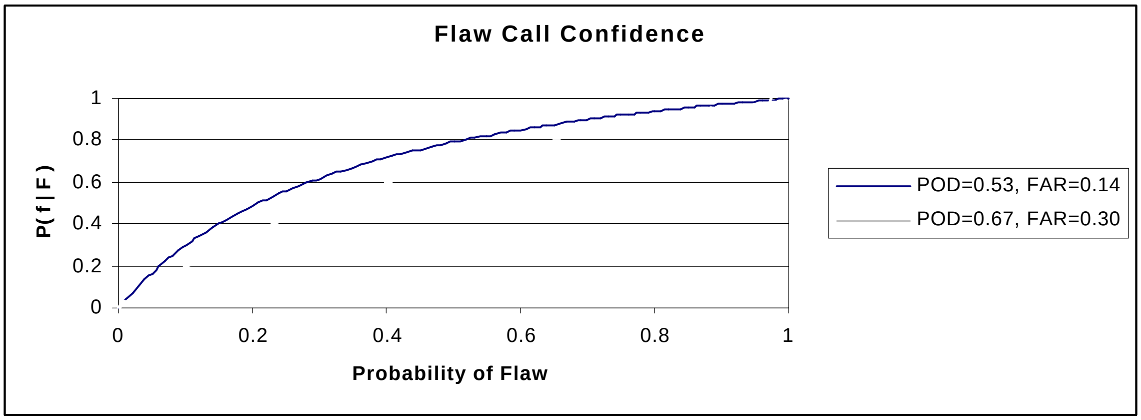

Figure 18 Flaw Call Confidence Comparison for Conventional Techniques and FAST for Category *IGSCC Only*

Interpretation

To aid the interpretation of the above results in a NDE inspection context, it is useful to consider the effect on reliability or on flaw call confidence that a change of technique has. Additionally, an overall look at flaw probability over time is helpful.

Figure 19 shows the absolute change and the percentage change obtained by going from one technique to another. Note that flaw call confidence is greatly increased by a technique change when the probability of a flaw is relatively low.

Figure 20 shows the general concept of probability of flaw as a function of time. The middle region is basically a region of low flaw probability. Throughout this region, a method with a low FAR would be desirable.

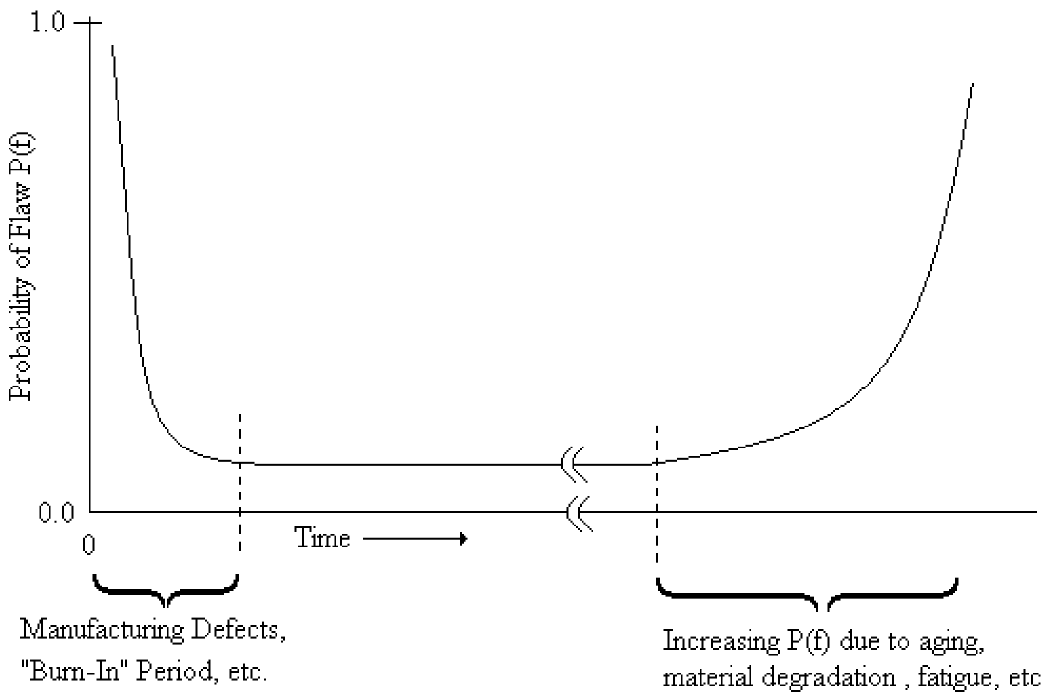

Figure 20 The "Bathtub" Curve Concept of Variation of Probability of Flaw Throughout a Component's Life

Conclusions

A method has been presented that allows NDE managers to evaluate various NDE techniques based on the characterizations of POD and FAR. Because the true probability of a flaw being present is rarely known exactly, the method addresses all possible probabilities. Flaw call confidence, P(f | F), and the ratios, Success Probability Ratio and Quality Factor can be used to index NDE performance and confidence in the reliability of a component for continued service.

The case study used in this paper was for the sole purpose of illustrating the method. Conclusions that are drawn from the use of the method should be based on the user's perception of the relative importance of the safety and economic factors involved. Although the case study showed the "FAST" technique superior regarding FAR, and comparable (slightly less than) with conventional methods regarding POD and reliability, flaw call confidence, P(f | F) was superior in every case.

References

[1] Rose, Joseph L., Ngoc, Tran D.K., and Avioli, Michael J. "Multi-Mode Ultrasonic Piping Inspection Potential for Improved Reliability," Proceedings of the 10th International Conference on NDE in the Nuclear and Pressure Vessel Industry, ASM International. Pages 409-416, Glasgow, Scotland, June 11-14, 1990.

[2] S.H. Bush, NUREG/CR-3110, PNL-4584, Volume 3, "Reliability of Nondestructive Examination," Chapter 13A., August 1983, Division of Engineering Technology, Office of Nuclear Regulatory Research, U.S. Nuclear Regulatory Commission, Washington, D.C., 20555, NRC FIN B2289

[3] Chou, Ya-lun. "Statistical Analysis," Holt, Rinehart, and Winston, Inc., New York, 1969. Pages 138-142

[4] Probability of Detection Tutorial Workshop Sessions
1993 ASNT Spring Conference
March 29-April 2 pages 10, 143-51
Speaker Handouts
The American Society for Nondestructive Testing, Inc.
P.O. Box 8518, Columbus, OH, USA, 43228

Appendix B

Probability of Detection, False Alarm Rate, and Bayes Theorem

Two criteria for assessing the value of a particular NDE method are *probability of detection* (POD) and *false alarm rate* (FAR). The purpose of this discussion is to clarify the meaning of these terms and to show the basis for Bayes Theorem.

First, it should be noted that POD and FAR are independent of each other. To see this, consider a component that has a population of M flaws and N non-flaws. A non-flaw could be geometry common to the component (e.g., a weld). If the NDE method under evaluation correctly identifies X of the flaws and incorrectly calls Y of the non-flaws as flaws, then

$$POD = \frac{X}{M} \qquad FAR = \frac{Y}{N}$$

For example, let M = 100, X =90, and N = 300 with Y = 60.

POD = 0.90 and FAR = 0.20

or, for the same M and X, let N = 400, and Y = 50,

POD =0.90 and FAR =0.125

In terms of the mathematics of probability, POD and FAR are defined as

$$POD = P(\,F \mid f\,), \qquad FAR = P(\,F \mid nf\,)$$

The symbolism is interpreted as:

Probability of Detection = the probability that the method *calls a flaw* (F) when in fact, a *flaw exists* (f). [Upper Case - method indicates; Lower Case - actual condition]

False Alarm Rate = the probability that the method *calls a flaw* (F) when in fact no flaw exists (nf).

For use further on in this discussion, the probability of a flaw can be defined as P(f) and the probability of no flaw as P(nf); P(nf) = 1 - P(f). In relation to the discussion above,

$$P(f) = \frac{\text{No. of Flaws}}{\text{Total population}} = \frac{M}{M + N}$$

To, for instance, find out how often a method is apt to call a flaw P(F) [P(F) is the probability of indicating a flaw condition], the following relation is used,

$$P(F) = P(F \mid f) P(f) + P(F \mid nf) P(nf)$$

This relationship says that the probability of calling a flaw is the combination of the only two possible situations; there is a flaw and a flaw is indicated, or there is a flaw indicated and there is no flaw present. In terms of probability mathematics, the "+" sign denotes "or."

Returning to the example,

$$P(F) = POD\ P(f) + FAR\ P(nf)$$

$$P(F) = \frac{X}{M} \frac{M}{M+N} + \frac{Y}{N} \frac{N}{M+N}$$

$$P(F) = \frac{X + Y}{M + N}$$

This (intuitive) result says that the probability of calling a flaw, F, is simply the ratio of everything called a flaw to the total population.

Bayes Theorem

. An expression of the form

$$P(B \mid A)$$

is called a conditional probability. It is interpreted as:

> "The probability of condition B given that A is already in evidence is P(B | A) [some number between 0 and 1]."

Bayes theorem relates the conditional probability P(A | B) to conditional probability P(B | A). Formally,

$$P(A \mid B) P(B) = P(B \mid A) P(A).$$

The Venn diagram shown in Figure B-1 provides an intuitive feeling .for the theorem. The shaded ellipses, labeled A and B in Figure C-1 can be considered as a collection of states of nature (A), e.g., cracks, and collection of NDE calls such as "Cracked" (B). In the context of NDE we are interested in those times when the NDE call agrees with the state of nature, $A \cap B$.

In Figure C-1, note that the intersection region is actually a percentage of the larger region B (or the region A). This percentage can be written as $P(A \cap B)/P(B)$ or $(P(A \cap B)/P(A))$. This percentage of either A or B also corresponds to the condition where both A and B, in a sense "agree;" e.g., a crack is called a crack. If A is considered as being present, then the possibility of B also being present is written as $P(B \mid A)$. $P(B \mid A)$ is actually the percentage of A that also coincides with a part of B, or

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

If B were considered as being present, the possibility for A being present also is given by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Because the intersection region, $A \cap B$, is common to either expression, the relationship,

$$P(A \mid B)\,P(B) = P(B \mid A)\,P(A)$$

is easily obtained. This relationship is known as Bayes Rule (Theorem).

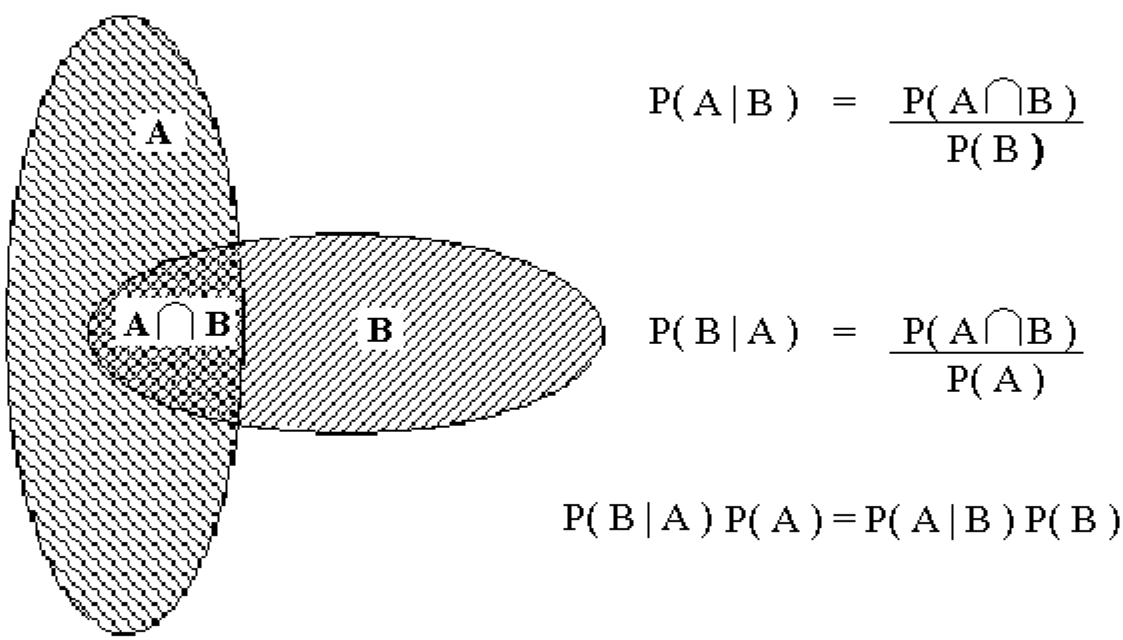

Figure C-1 Venn Diagram Representation of Bayes Theorem