INTRODUCTION

This paper presents a progress report on the development of model-based methodology to estimate Nondestructive Evaluation (NDE) capability. Previous work on this project was reported in Meeker et al. [3, 4]. The methodology uses combinations of physical modeling of an inspection process, along with laboratory and production data, to estimate Nondestructive Evaluation (NDE) capability. The methodology is based on a physical/statistical prediction model and will be used to predict NDE capability in terms of Probability of Detection (POD), Probability of False Alarm (PFA), Probability of Indication (POI) and Receiver Operating Characteristic (ROC) curves. The physical model explains and allow predictions for the effects of making changes to the inspection setup (e.g. transducer properties and scan increment). The statistical/empirical model will quantify unexplained variability, adjust for potential model bias, and provide a means for obtaining corresponding uncertainty intervals. The work has been motivated by the need for methods to predict ultrasonic (UT) inspection POD for detecting hard-alpha and other subsurface flaws in titanium using gated peak detection. This is a challenging problem because the inspection must detect very complex subsurface flaws in the presence of significant "material" noise. The underlying framework of the methodology should, however, be general enough to apply to other NDE methods. This paper describes recent work based on application of the new methodology to the detection of synthetic hard alpha flaws in titanium alloys. In particular we describe and illustrate methods to assess the effect that changes in scan plans and gate width will have on POD.

PHYSICAL/STATISTICAL MODEL FOR A UT SIGNAL

The methodology described here uses a physical model based on the theory of ultrasonic wave scattering to provide predictions for typical measurements from the flaw-signal distribution. A statistical model is used to quantify variability in the distributions of noise signals and the distributions of signals from flaws in noise. This is accomplished by developing a statistical model for the deviations between the physical model predictions and actual NDE measurements and using this distribution to describe the various sources of unexplained variability in NDE signals. This model for the deviations provides a framework for statistical estimation of PFA and POD and the related ROC and POI functions.
Distributions of Noise and Flaw&Noise Signals

NDE techniques are used to detect flaws by identifying a flaw’s signal. Let $Y$ denote an observed UT signal (maximum peak-to-peak voltage in a specified gate). Ideally there would be very little variability in the amplitude of a UT signal for a flaw of a given size. Operationally, however, the strength of a signal will have variability due to factors like flaw position relative to the beam, size and position of microstructural grain boundaries relative to the flaw, characteristics of the inhomogeneous medium between the transducer and the flaw as well as the depth, orientation, composition, shape and other characteristics of the flaw itself.

Conditional on a set of specified fixed factors that affect the signal strength, the cumulative probability distribution for the flaw&noise signal can be expressed as

$$\Pr(Y \leq y) = F(y; \mathbf{z}, \boldsymbol{\theta}).$$

Here $\boldsymbol{\theta}$ is a vector of parameters that is, for the most part, independent of $\mathbf{z}$ and $\mathbf{z} = (\mathbf{z}_{\text{FLAW}}, \mathbf{z}_{\text{NOE}}, \mathbf{z}_{\text{PART}})$ is a vector of factors that affect the ultrasonic signal response. In particular,

- $\mathbf{z}_{\text{NOE}}$ contains NDE system factors like transducer and electronic system characteristics.
- $\mathbf{z}_{\text{PART}}$ contains PART factors like part geometry, type of material being inspected, surface roughness, etc.
- $\mathbf{z}_{\text{FLAW}}$ contains flaw factors like size, density, shape, composition, and degree of voiding/cracking, and orientation/position relative to the ultrasonic beam.

Our model for noise-only signals (UT signal when there is no flaw illuminated by the beam) is similar except that the distribution would not depend on $\mathbf{z}_{\text{FLAW}}$.

Generalized Deviations

The physical model for ultrasonic NDE signals (UNDE model) will predict the flaw signal as a function of $\mathbf{z}_{\text{FLAW}}, \mathbf{z}_{\text{NOE}},$ and $\mathbf{z}_{\text{PART}}$. The model used here is described in Chiou et al. [1, 2]. For observations taken in controlled laboratory experiments, quantities like flaw position relative to the transducer and flaw characteristics $\boldsymbol{\theta}$ are fixed. As described in Meeker et al. [4], for production/field inspection observations, factors like position of a flaw relative to scan lines and flaw morphology are not under control. Thus in models for production/field inspection, they are considered to be random.

Let $Y$ denote the experimental voltage signal and let $\hat{Y}$ denote the UNDE model prediction for $Y$. The prediction $\hat{Y}$ is a function of $\mathbf{z}_{\text{FLAW}}, \mathbf{z}_{\text{NOE}},$ and $\mathbf{z}_{\text{PART}}$ and provides a prediction of the center or location of the flaw&noise signal distribution. Generalized deviations between the UNDE predictions and the actual data are defined, using the Box-Cox family of transformations, as

$$\text{Deviation} = g(Y; \lambda, \mathbf{z}) = \begin{cases} 
\frac{(Y)^\lambda - 1}{\lambda} - \frac{(\hat{Y})^\lambda - 1}{\lambda}, & \lambda \neq 0 \\
\log(Y) - \log(\hat{Y}), & \lambda = 0
\end{cases}$$

(1)

These generalized deviations provide a statistical model for the shape and spread of the flaw&noise signal distribution. The value of $\lambda$ is chosen empirically to equalize variance (with respect to flaw size) and otherwise make distributions, as much as possible, independent of the factors $\mathbf{z} = (\mathbf{z}_{\text{FLAW}}, \mathbf{z}_{\text{NOE}}, \mathbf{z}_{\text{PART}})$ that serve as inputs to the UNDE model.
Basic POD (Probability of a Detection)

There is a detection when \( Y > y_{\text{thresh}} \), where \( Y \) is the maximum reading in the gate of an A-scan and \( y_{\text{thresh}} \) can be set according to specified user criteria (e.g., to make the probability of a false alarm essentially 0 or to minimize expected risk).

For some applications it may be of interest to compute POD values for one or more sets of fixed values for all of the components in \( \mathbf{z} = (z_{\text{Flaw}}, z_{\text{Ne},}, z_{\text{Part}}) \). Under the general model the probability of a detection on any given reading is

\[
\text{POD}(\mathbf{z}) = \Pr(Y > y_{\text{thresh}} \mid \mathbf{z}) = 1 - \Phi\left(\frac{g(y_{\text{thresh}}) - \hat{\mu}_g}{\hat{\sigma}_g}\right)
\]

where \( \Phi \) is the standard normal (Gaussian) cumulative distribution function, \( \hat{\mu}_g \) and \( \hat{\sigma}_g \) are estimates from the available deviation data. As in Meeker et al. [3, 4], we call (2) the "Basic POD."

POD for Production Inspection

For predicting POD for production inspection, it will be necessary to account for random factors in the inspection process such as flaw position relative to the beam. Such evaluation will require a joint distribution of the random factors. To illustrate this we will show how to evaluate the effect on POD of using different scan increments and gate widths. To keep the example simple we will assume that the cylindrical synthetic hard-alpha flaw is, as in the experiment, vertically oriented and one inch below the surface, and that the beam is focused, with normal incidence, at that depth. Then, to get POD as a function of size \( a \) and scan increment we repartition as \( \mathbf{z} = (a, z_{\text{Flaw}}, z_{\text{Ran}}) \), where \( z_{\text{Ran}} \) is the three-dimensional position of the flaw in the block and \( z_{\text{Flaw}} \) is a vector of all of the other factors in \( \mathbf{z} \), assumed to be fixed (e.g., transducer characteristics). Then to compute POD for fixed values of size \( a \) and \( z_{\text{Flaw}} \), we integrate (2) with respect to \( z_{\text{Ran}} \) over the entire range of \( z_{\text{Ran}} \).

\[
\text{POD}(a, z_{\text{Flaw}}) = \int f_{z_{\text{Flaw}}}(z_{\text{Ran}}) \text{POD}(a, z_{\text{Flaw}}, z_{\text{Ran}}) \, dz_{\text{Ran}}
\]

where \( f_{z_{\text{Flaw}}}(z_{\text{Ran}}) \) is the joint density function of \( z_{\text{Ran}} \). For example, to account for random position of a flaw in three dimensions, POD can be computed by integrating between scan lines (accounting for change in signal strength as the beam moves away from the flaw) and across the gate width (accounting for change in signal strength as the the focal point moves away from top of the flaw surface). Then, as a special case of (3) we have

\[
\text{POD}(a, z_{\text{Flaw}}) = \int \int f(x, y, z) \text{POD}(a, z_{\text{Flaw}} | x, y, z) \, dx \, dy \, dz
\]

where \( f(x, y, z) \) is the joint density function of the position of the flaw.

PFA (Probability of a False Alarm) as a Function of Model Parameters

The probability of a false alarm is the probability of an above-threshold reading when there is no flaw. Under our model, the probability of such a false alarm on any given reading is

\[
\text{PFA} = \Pr(Y > y_{\text{thresh}} \mid \text{no flaw, } \mathbf{z}) = 1 - \Phi\left(\frac{g(y_{\text{thresh}}) - \hat{\mu}_n}{\hat{\sigma}_n}\right)
\]

where the statistical model and parameters \( \hat{\mu}_n \) and \( \hat{\sigma}_n \) were identified, as with the flaw-noise signal distribution, but using an average of the voltage signals over several regions containing no flaws.
POI (Probability of Indication)

In some situations it is more meaningful to compute the probability of indication (POI) instead of POD. POI is the probability that the UT signal exceeds the threshold, irrespective of whether the signal is a reflection from a flaw or a result of noise. Computation of POI requires information on the distribution of noise signals. POI is computed as

$$POI(a) = PFA + (1 - PFA)POD(a).$$

(6)

For $a$ approaching 0, $POI(a) \approx PFA$. For large flaw size $a$, $POI(a) \approx POD(a)$. Plots of the POI function can be particularly useful for comparing different sets of inspection parameters. For example, it shows the effect on PFA of choosing different thresholds.

EXAMPLE IMPLEMENTATION FOR SYNTHETIC HARD-ALPHA FLAW DETECTION

Experimental Results

Chiu et al. [1, 2] describe a factorial experiment that was conducted to obtain information on the distribution of flaw&noise signals for synthetic hard-alpha flaws (SHAs) in titanium. Voltage readings were taken on each of 8 nominally similar #2, #3, #4, and #5 cylindrical synthetic hard-alpha inclusions, 1 inch deep in a titanium block. Here the flaw size measure was adapted from the flat-bottom-hole standard where #2 is 2/64 in., #3 is 3/64 in., etc. The part of the experiment used for the computations in this paper was conducted at focal depths of .5, 1, and 1.25 inches, incident angles of −2.5, 0, and 5°, with scan increments of 5 mils in both the $x$ and $y$ directions. C-scans were made with both a 5 MHz and a 10 MHz focused transducer. The data provide information about the effect on UT signal strength of flaw size, the distance between the beam and the center of the SHA flaw, and the distance between the focal depth and the top of the flaw. The data also provide information about flaw-to-flaw variability for nominally identical flaws and deviations from model predictions.

Statistical Model for Deviations in the Synthetic Hard-Alpha Flaw Experiment

Different values of the transformation parameter $\lambda$ were used to compute the generalized deviations in (1) and the shape and spread of the resulting distributions were investigated. Meeker et al. [4] give examples of graphical displays of the distributions of the observed generalized deviations and discuss sensitivity analyses done to assess the effect of using different values of $\lambda = .3$. For the SHA experimental data with both the 5 MHz and 10 MHz transducers, using $\lambda = .3$ suggests that the generalized deviations follow, approximately, a normal distribution with constant mean $\mu_g$ and standard deviation $\sigma_g$. This same value of $\lambda$ was reported in Meeker et al. [3] as appropriate for stabilizing the distribution of UT signals from flat bottom holes.

Basic POD for SHA Flaws

Figure 1 shows the basic POD for different thresholds, computed from (2) for a focused, 10 MHz transducers as function of specified size of the SHA flaw. This POD is based on the estimated distribution of the generalized deviations (1) and the predictions of UT signal from the UNDE model.

POD for SHA Flaws as a Function of Inspection Parameters

For the SHA experiment, with random $x$ and $y$ flaw position in the plane and fixed focal depth, we assume that flaw position is uniformly distributed between scan lines and
within the specified gate (if a uniform distribution does not provide an adequate description of position, it is easy to substitute another distributional form for the joint distribution $f(x, y, z)$ in (4)).

Figure 2 shows POD for a 10 MHz transducer, normal incidence, assuming 30 mil scan increments, and focused at the same depth as the synthetic hard alpha flaw and with a very narrow gate width. As expected, the POD curves are a bit lower when compared to the "Basic POD" curves in Figure 1. Figure 3, in contrast, illustrates the difference of integrating in the depth dimension using a gate width of .5 inches. As expected, the POD curves are slightly lower in comparison with those in Figure 2.

![Figure 1. Basic POD for 10 MHz transducer, normal incidence, focused directly on top of and at the same depth as synthetic hard alpha flaw for threshold values of 50, 75, 100, 125, and 150 mV.](image)

**Comparison of SHA POI for Two Different Transducers**

It is generally possible to improve POD by lowering the detection threshold. This can, however, have an adverse effect on PFA. When comparing two different inspections, it is therefore important to compare inspections having the same POD. Graphically, POI, defined in (6), combines POD with PFA. Figure 4 presents a comparison of POI for Transducer 2 (the 5 MHz transducer) and Transducer 4 (the 10 MHz transducer) used in our SHA experiments. The thresholds were chosen such that PFA=.02 for both of the transducers. The comparison shows that the 10 MHz transducer provides considerable improvement in POD with the same PFA.

**Uncertainty Bounds**

Given the large amount of data used to estimate the distribution of generalized deviations, the dominant source of uncertainty in the predictions arises from error in the UNDE model. With improvements that have been made to the UNDE model, the expected error in the prediction is expected to be no more than ±3dB, a target selected because it is comparable to the reproducibility of typical industrial experiments. Thus, to account for uncertainty the POD evaluation algorithm is evaluated for all possible signal predictions within a ±3 dB uncertainty band. Figure 5 shows POD for a 10 MHz transducer showing uncertainty bands for ±3dB uncertainty in the model predictions.
Figure 2. POD for 10 MHz transducer, normal incidence, assuming 30 mil scan increment, focused at the same depth as the synthetic hard alpha flaw for threshold values of 50, 75, 100, 125, and 150 mV.

Figure 3. POD for a 10 MHz transducer, normal incidence, assuming 30 mil scan increments and a gate width of .5 inches for threshold values of 50, 75, 100, 125, and 150 mV.
Figure 4. Comparison of POI for Transducer 2 (5 MHz) and Transducer 4 (10 MHz), using normal incidence, assuming 20 mil scan increments and a gate width of .25 inches with the threshold adjusted such that PFA = .02 (56.4 mV for Transducer 4 and 144 mV for Transducer 2).

Figure 5. POD for a 10 MHz transducer, normal incidence, assuming 30 mil scan increments and a gate width of .5 inches, showing uncertainty bands for ±3dB uncertainty in the model predictions for threshold values of 50 and 100 mV.
ON-GOING AND FUTURE WORK

In future research we expect to extend the work presented here to allow prediction of POD for real hard-alpha flaws. This will be done by applying the methodology by estimating the generalized deviations from limited real-flaw and other field-find data. One result of this extension is that, because such data will be limited (relative to the amount of synthetic-flaw data), there will unignorable uncertainty in model parameters. This will require adding such uncertainty to the model uncertainty to construct uncertainty bounds. The model used in Yalda et al. [5] might be adapted to replace the empirical Box-Cox transformation, providing a firmer basis for the extrapolation needed to predict POD and other capability functions for situations where experimental data are not available.

ACKNOWLEDGEMENTS

This material is based upon work performed at Iowa State University, General Electric, Pratt & Whitney, and Allied Signal, as part of the Engine Titanium Consortium (ETC) operated by Iowa State University. The ETC is supported by the Federal Aviation Administration Technical Center, Atlantic City, New Jersey, under Grant number 94G048. Other members of the ETC POD Working Group who participated in meetings leading to the development of this methodology include: Chuck Annis, Richard Burkel, Bob Gilmore, Frank Margetan, Jim Rose, Derek Sturges, and Bill Tucker.

REFERENCES