2-2011

Statistical Assessment of Probability of Detection for Automated Eddy Current Nondestructive Evaluation Inspection

Ming Li
Iowa State University

Norio Nakagawa
Iowa State University, nakagawa@iastate.edu

Brian F. Larson
Iowa State University, blarson@iastate.edu

William Q. Meeker
Iowa State University, wqmeeker@iastate.edu

Statistical Assessment of Probability of Detection for Automated Eddy Current Nondestructive Evaluation Inspection

Abstract
Eddy current inspection is widely used in nondestructive evaluation to detect cracks in metal structures. The impedance plane measurement response collected using our motion controlled eddy current inspection system, are used in the analysis. A scalar reduction from the impedance plane response is used to minimize human-factor variation. We apply a noise interference model to the data from a large-scale experiment taking measurements on fastener rivet holes in multi-layer structures with fatigue cracks, and estimate the probability of detection (POD). The estimates of POD as a function of crack size will be valuable for future model-assisted POD study.

Keywords
Center for Nondestructive Evaluation, automated eddy current inspection, maximum likelihood, noise interference model, POD, rivet hole fatigue cracks

Disciplines
Statistics and Probability

Comments
This preprint was published as Mink Li, Norio Nakagawa, Brian F. Larson & William Q. Meeker, "Statistical Assessment of Probability of Detection for Automated Eddy Current Nondestructive Evaluation Inspection", *Research in Nondestruction Evaluation* (2013): 89-104, doi: 10.1080/09349847.2012.735352.

STATISTICAL ASSESSMENT OF PROBABILITY OF DETECTION FOR AUTOMATED EDDY CURRENT NONDESTRUCTIVE EVALUATION INSPECTION

Ming Li[1], Norio Nakagawa[2], Brian F. Larson[2], William Q. Meeker[1]

[1]Center for Nondestructive Evaluation and Department of Statistics, Iowa State University, Ames, Iowa, USA

[2]Center for Nondestructive Evaluation, Iowa State University, Ames, Iowa, USA

Abstract: Eddy current inspection is widely used in nondestructive evaluation to detect cracks in metal structures. The impedance plane measurement response collected using our motion controlled eddy current inspection system, are used in the analysis. A scalar reduction from the impedance plane response is used to minimize human-factor variation. We apply a noise interference model to the data from a large-scale experiment taking measurements on fastener rivet holes in multi-layer structures with fatigue cracks, and estimate the probability of detection (POD). The estimates of POD as a function of crack size will be valuable for future model-assisted POD study.

Keywords: maximum likelihood, noise interference model, POD, fatigue, rivet holes

1. INTRODUCTION

1.1 Background and Motivation

The eddy current inspection method is widely used in nondestructive evaluation (NDE) applications to detect surface and subsurface fatigue cracks in metal structures [1,2]. The operator commonly makes crack existence (Hit) or not (Miss) binary decisions for each eddy current measurement based on the overall impedance plane response. For a given inspection problem, efforts are made to devise a rigid set of procedures and decision criteria to follow, based on operational knowledge and experience. However, despite the efforts for inspection process control, such Hit/Miss results do not escape the issue of human-factors variation. Automated eddy current inspection methods can reduce the effect of human factors on inspection capability, while there has been only limited quantitative research concerning probability of detection (POD) for such measurements. Recently Larson, Madison, and Nakagawa [3] described the use of an eddy current method with a low frequency sliding probe to inspect for inner layer cracks at fastener rivet holes in simulated-lap splice airframe structures. These specimens were specially fabricated with fatigue cracks of known size and orientation. Laboratory-based measurements were conducted in approximate adherence to an OEM procedure, except that a slightly different probe from the specification was used, and that its motion was controlled by step motors. POD as a function of crack size was calculated based on Hit/Miss data obtained from the operator's interpretation of the impedance plane data.

As an alternative to the Hit-Miss analysis, the use of a continuous scalar response can be considered, based on characteristics of the pattern of the eddy current response in

the impedance plane (e.g., the maximum of the horizontal signal component in the signal trace). From such quantitative data, POD as a function of crack size can be estimated by using the â-versus-a method [4].

There are a wide variety of factors in the rivet-hole locations that can influence the measurement response. These factors include material thickness, roundness and angle of the rivet holes, coating conditions and conductivity. Thus the scalar response from the impedance plane generally involves a large amount of variation and noise. In applications such as this where there is a large amount of noise in the data, the traditional statistical methods to obtain â-versus-a POD can lead to non-conservative bias in POD estimates [5]. In this paper, we apply a noise-interference model (NIM) to the scalar reduction impedance plane data, providing statistically-consistent POD estimates as a function of crack size. These quantitative POD results will be valuable for the future model-assisted eddy current inspection POD studies.

1.2 Overview

The rest of this paper is organized as follows. Section 2 describes the multi-layer crack panel and the experimental setup of the computer-controlled eddy current inspection system. Section 3 presents the binary decision data and the Hit/Miss POD analysis for the rivet holes at different locations. Section 4 describes the noise interference model. Section 5 provides the detailed statistical analysis of the scalar reduction impedance plane data. Section 6 contains some concluding remarks and extensions for future research.

2. MULTI-LAYER CRACK PANEL AND EXPERIMENTAL SETUP DESCRIPTION

2.1 Multi-layer Crack Panel

The multi-layer crack panels used in our eddy current inspection were produced by the Airworthiness Assurance NDI Validation Center at Sandia National Laboratory to simulate the lap splices on the body of a Boeing 737 aircraft. The panels consisted of a top sheet, a bonded internal doubler, a lower skin, and mock tear straps. Three rows of 5D5 Al flush head rivets were used to fasten the sheets together. In tear strap areas rivets were alodined, and otherwise they were anodized. Prior to assembly of the panels, artificially induced fatigue cracks were grown in the inner skin around the rivet holes. The fatigue cracks have lengths ranging from 0.25 mm to 12.7 mm. Based on the rivet location, we separate the data population into three groups: at tear strap locations (TS), and outside tear strap locations subdivided into region one (NT-1) and region two (NT-2). Here, the data for the no tear strap locations split into two populations because the null location for the data set NT-1 was over a tear strap while the null location for the data set NT-2 was of the tear straps. Further details of sample panel and rivet configurations, along with corresponding signal behaviors, can be found in [3].

2.2 Eddy Current Inspection System

To recapitulate, a computer controlled eddy current inspection system with mechanical motion control and automated data collection was used in the experiment in order to minimize human-factors variation. Attached to the data acquisition instrument is the

sliding probe operated in the reflection mode at a frequency of 2.0 kHz with the instrument gain of 51.5 dB and the phase angle of 315 degrees. During the measurement, the surrounding areas of each rivet hole were scanned and the probe alignment and position were checked at each rivet-hole location. The probe was moved with a 0.1 mm increment and the voltage output from the instrument was recorded in the vector-voltage plane (impedance plane) for each increment. Each rivet was scanned three times with the sample dismounted and remounted between scan runs to provide an evaluation of measurement repeatability.

2.3 Typical Impedance Plane Response

The typical impedance plane responses acquired by the eddy current inspection system are shown in Figure 1. The top portion of each image at Figure 1 is the impedance trace when the probe was moved, and the bottom portion of each image is the conceptual illustration of the crack panel with the probe location and rivet locations indicated. When the probe was moved from left to right to scan the rivet-hole locations, the impedance trace was plotted automatically. When the probe was moved from the null point between two rivet holes (Figure 1 (a)), a concave signal trace was produced. The maximum horizontal deflection of the signal was achieved when the probe was at the top of the rivet (Figure 1 (b)). When the probe moved beyond the center of the rivet, the signal trace returned back towards the signal null point (Figure 1 (c)). Figure 1 (a), (b) and (c) are the typical impedance plane responses for the rivet holes not at tear strap locations. Similarly, Figure 1 (d), (e) and (f) are the typical impedance plane responses for the rivet holes at tear strap locations.

When there is a crack in the inner skin around the rivet hole, some particular characteristics will show up in the measured impedance plane trace and those types of characteristics vary as a function of crack size. Three different types of characteristics are considered in our eddy current analysis as illustrated in Figure 1 (c): (A) the vertical voltage at maximum horizontal positions (VV@MH), (B) the maximum vertical voltage (MVV), and (C) the relative opening between the two impedance traces (RO). The impedance traces behave differently for rivet location TS, NT-1 and NT-2. In the Hit/Miss analysis (Section 3), all three characteristics are used to make Hit or Miss decisions. For the scalar reduction NIM analysis (Section 5), only VV@MH is used to determine POD given the fact that VV@MH is the best single indicator among all three characteristics that might be used to decide whether a rivet hole in the inner skin has a crack or not.

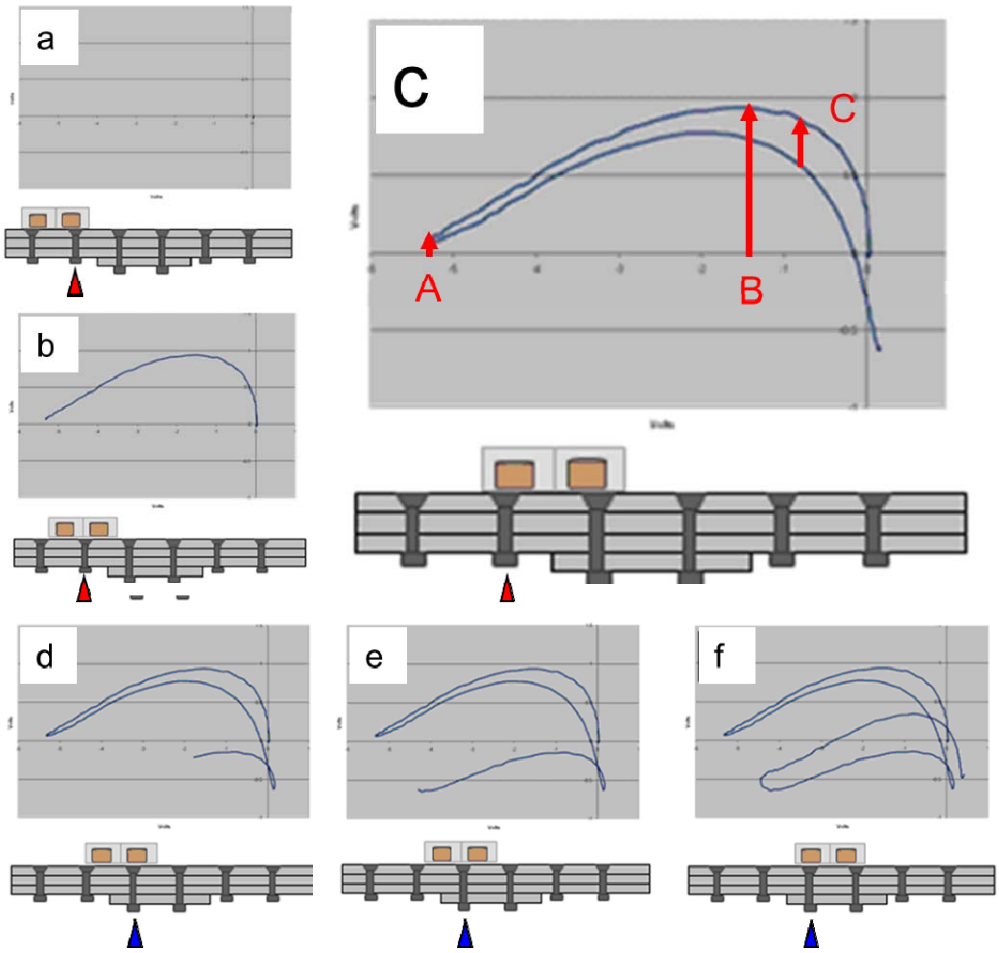

FIGURE 1. Typical impedance plane responses for the eddy current inspection system: the rivet positions and the relative detection probe position are illustrated at the bottom of each image; the impedance traces for rivet holes not at a tear strap location are displayed in (a), (b), and (c); the impedance traces for rivet holes at a tear strap location are displayed through (d), (e) and (f); three different types of character to determine crack existence are shown with arrows at (c): A (VV@MH), B (MVV) and C (RO).

3. HIT/MISS POD ANALYSIS

3.1 Standard Model

The name “Hit/Miss POD” is attached to the NDE reliability model for inspections that return binary decision results about whether a crack is thought to exist or not. The inspection operator uses knowledge and experience, or more commonly follows an experience-based decision procedure, and makes decisions based on a combination of the shape of the impedance trace and the signal deflection in the instrument display. The call is made in terms of being present (“hit”), or otherwise not present (“miss”), and the signal magnitude used in the decision is not used in the POD analysis. The probability of a hit (i.e. POD) as a function of crack size is calculated by fitting the decision data to the binary regression model. Such a binary regression analysis can be carried out as a special case of the generalized linear model, available in most modern statistical software packages. A typical binary regression POD with a logit link function can be written as

$$\mathrm{POD}(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}$$

where x is the crack size or transformation of crack size; and β_0 and β_1 are the binary-model regression parameters [4,5].

3.2 Probability of Detection

For our eddy current inspection system, the primary factors to determine the existence of a crack are the VV@MH, MVV, and RO characteristics extracted from the impedance plane signal. The relationships between VV@MH and MVV for the TS rivet holes and NT rivet holes are shown in Figure 2 (a) and (c), respectively, with filled symbols

representing rivet holes with a crack and open symbols representing rivet holes without a crack. There are no clear boundaries to separate the rivet holes with or without a crack based on the scatter plots of Figure 2. The operator's Hit/Miss decisions based on all three types of characteristics from the impedance plane data are plotted as symbols in Figures 2 (b) and (d) for the TS, NT-1 and NT-2 rivet holes with one for a "hit" and zero for a "miss". The logit link generalized linear regression model (also known as the logistic binary regression model) POD results for the TS, NT-1 and NT-2 rivet holes are also shown in Figure 2 (b) and (d), respectively. Figure 2 (c) indicates the difference of measurement responses for the NT-1 and NT-2 rivet holes and confirms the necessity of the separation of the NT-1 and NT-2 rivet holes for statistical analysis.

3.3 Probability of False Alarm

The distributions of VV@MH for rivet holes without cracks are shown in the insets of Figure 2 (b) and (d) among the TS and NT rivet holes, respectively. The VV@MH for the TS rivet holes can be described adequately with a normal distribution, while the distribution of VV@MH for the NT rivet holes has a double peak, indicating the need to perform separate analyses of the data for the NT-1 and NT-2 rivet holes. The Hit/Miss decisions for rivet holes without a crack are used to determine the Hit/Miss probability of false alarm (PFA) (i.e., the probability to have a "hit" decision for rivet holes without a crack). The observed Hit/Miss PFA for the TS, NT-1 and NT-2 rivet holes were 2.7%, 0.0%, and 1.2%, respectively.

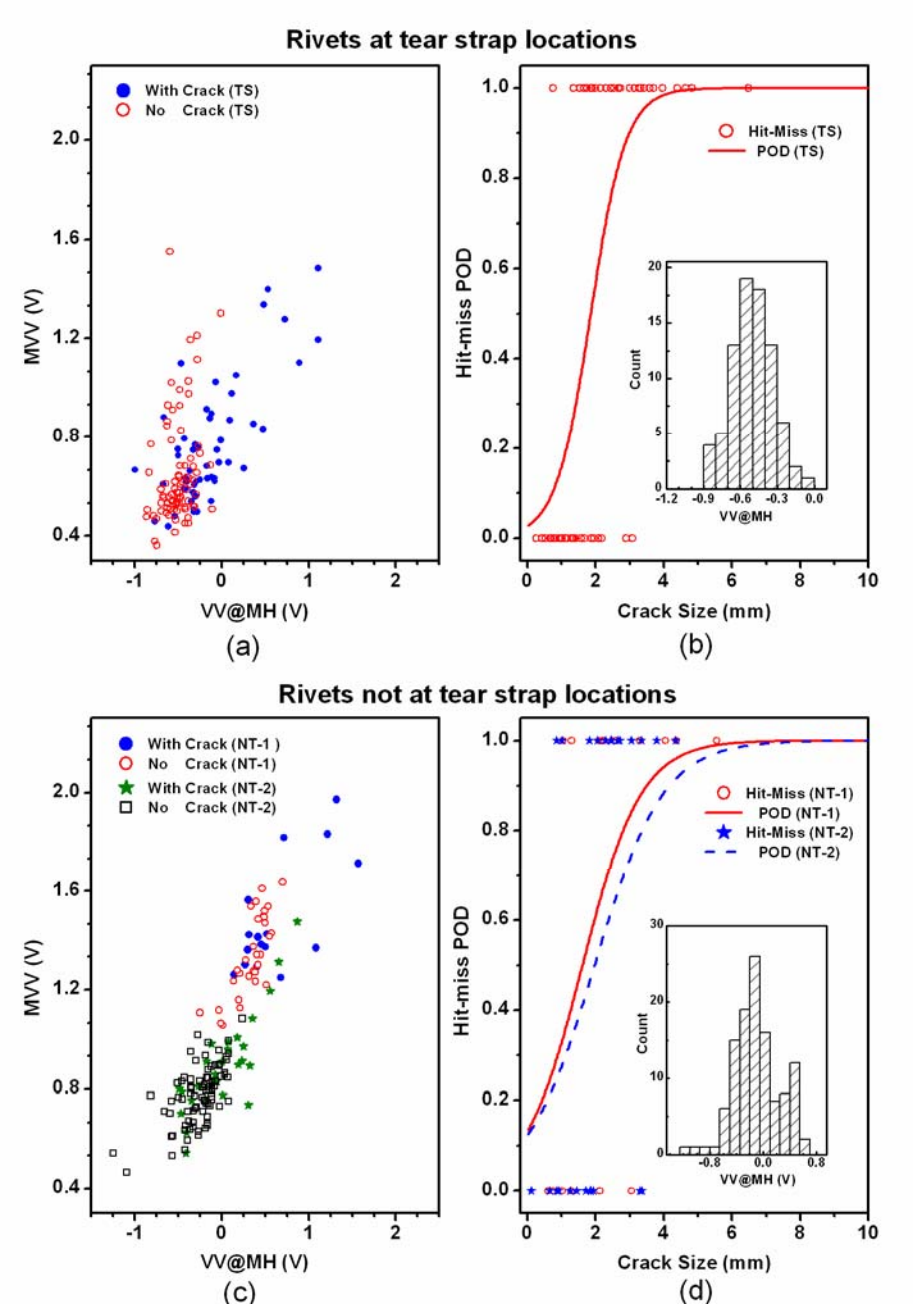

FIGURE 2. The relationships betweens MVV and VV@MH are shown for TS (a) and NT (c) with solid symbols representing rivet holes with a crack and open symbols representing rivet holes without a crack. The Hit/Miss PODs are shown in (b) for TS and in (d) for NT, where the Hit/Miss decisions are indicated as zero (miss) and one (hit). The noise distribution of VV@MH for both locations are plotted in the inset of (b) for TS and of (d) for NT.

4. THE NOISE INTERFERENCE MODEL

The traditional statistical method for estimating POD from an NDE study with a quantitative response is the *â*-versus-*a* method described in [4]. The traditional *â*-versus-*a* method has, for small cracks, an asymptotic limit for POD that approaches zero. This characteristic is in contradiction to the fact that for zero crack length (i.e., rivet holes without crack) the POD should be approximately equal to the PFA. When NDE measurements are taken in locations where there are no target cracks, the reading can still be of some value to quantify measurement and background noise. Such noise data are usually used to estimate PFA and set the detection threshold. In locations where there are very small cracks, the observed response could be the result of a noise-causing artifact rather than the small crack. Based only on the experimental measurements, we cannot be sure whether the measurement came from a crack or a noise-causing artifact.

To account for possible mixture of flaw and noise responses, we extend the *â*-versus-*a* POD model by using the NIM. A detailed derivation of NIM can be found in reference [5]. We define the observed measurement response or some transformation as Y, the signal response as Y_{signal}, and the noise response as Y_{noise}. The NIM components are as follows:

- The signal response is modeled as $Y_{signal} = \beta_0 + \beta_1 x + \varepsilon_s$ where β_0 and β_1 are regression parameters, x is the crack length or some transformation, and ε_s is a random variation term, assumed to be normally distributed with mean zero and variance σ_s^2, i.e. $\varepsilon_s \sim N\left(0, \sigma_s^2\right)$.

- The noise response Y_{noise} is assumed to be normally distributed with mean μ_n and variance σ_n^2; that is, $\varepsilon_n \sim N\left(\mu_n, \sigma_n^2\right)$.
- The actual observed response (i.e. the experimental measurement) is the maximum of the signal and noise: $Y = \max\left(Y_{noise}, Y_{signal}\right)$.
- With the available measurement data, estimates of the model parameter vector $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\mu}_n, \hat{\sigma}_n^2, \hat{\sigma}_s^2\right)$ and the estimated variance covariance matrix of these estimates can be obtained through standard maximum likelihood methods [6].
- The detection threshold y_{th} is set to control the PFA, based on the measurement responses from specimens without crack.
- Finally the POD estimate, as a function of the explanatory variable x, can be calculated from the formula

$$\mathrm{POD}(x) = 1 - \Phi\left(\frac{y_{\mathrm{th}} - \hat{\beta}_0 - \hat{\beta}_1 x}{\hat{\sigma}_{\mathrm{s}}}\right)\Phi\left(\frac{y_{\mathrm{th}} - \hat{\mu}_{\mathrm{n}}}{\hat{\sigma}_{\mathrm{n}}}\right)$$

where y_{th} is the detection threshold and $\Phi\left(x\right)$ is the standard normal cumulative distribution function.

It is easy to show analytically that the NIM POD approaches PFA as the crack size approaches 0. The standard error of the estimated NIM POD is smaller than the traditional model, which indicates that the NIM model fits the data better and provides better statistical inferences [5].

5. NIM APPLIED TO EDDY CURRENT DATA

From our eddy current measurements, we observe that the noise responses from the rivet holes without a crack are similar in magnitude to the signal responses from the rivet holes with a small crack. Thus the responses are a mixture of signals from rivet holes with a crack (filled symbols) and rivet holes without a crack (open symbols), as shown in Figures 2 (a) and 2 (c). To better describe the signal response from the noisy data, the NIM can be used for more efficient and reliable statistical analysis.

All three types of characteristics (VV@MH, MVV, and RO) contain useful information to make crack existence decisions. For scalar variable analysis, such as the traditional â–versus-a model and the NIM method, one response variable has to be chosen. The RO characteristic has an undesirably large variation with respect to the different rivet coatings, generally leading to a higher false alarm rate for a given amount of sensitivity. The VV@MH and MVV characteristics are similar and mutually correlated, as shown in Figure 2 (a) and (c). We choose VV@MH as the response variable in the NIM POD analysis based on the fact that it gives better predictions when compared with MVV.

5.1 Rivet Holes at Tear Strap Locations

We first apply the NIM to the TS rivet holes with VV@MH as the response variable and crack size as the explanatory variable. The detection threshold is set such that there is a 10% PFA. We chose a relatively high PFA for POD analysis because the noise response is relatively high compared to the signal response and there are many measurements from the rivet holes with cracks below the 10% PFA detection threshold.

The scatter plot in Figure 3 (a) shows VV@MH as a function of crack size with crosses for the rivet holes without cracks and dots for the rivet holes with cracks. The NIM is used to estimate the regression relationship between VV@MH and the crack size, based on the data from both rivet holes with and without a crack. The NIM regression line and its 95% confidence bounds are shown in Figure 3 (a). Figure 3 (a) also shows the NIM estimates of the noise mean and corresponding 99% percentile using the horizontal dashed and dotted lines, respectively. Following the steps described in Section 4, the NIM POD estimate based on the TS rivet holes data and the corresponding 95% lower confidence bound are shown in Figure 3 (b). A common metric used in NDE applications is the crack length associated with 90% POD with 95% lower confidence bound for POD. This is known as the a90/95 value. The a90/95 for the TS rivet holes is 3.23 mm and it is indicated by a vertical dotted line in Figure 3 (b). The NIM POD for zero crack size is approximately equal to 0.10, the PFA that was used to define the detection threshold.

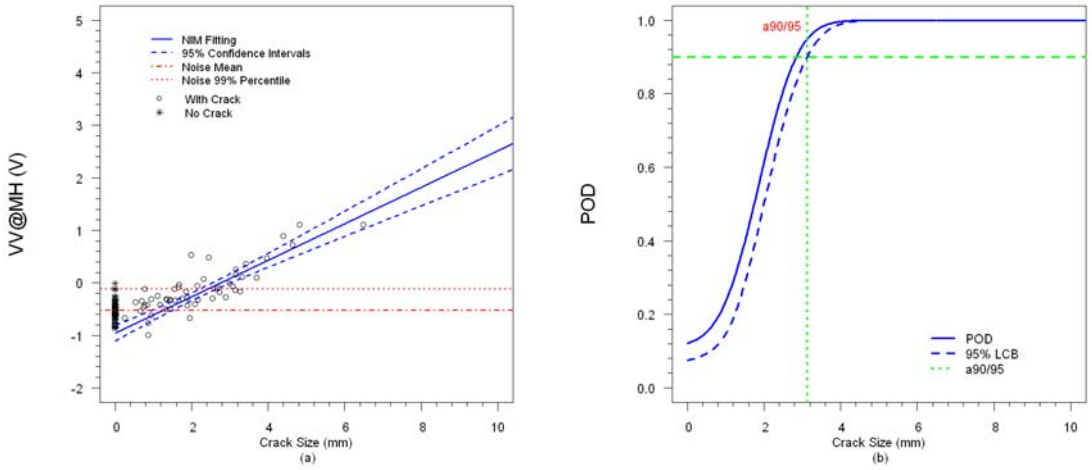

FIGURE 3. (a) NIM regression line and 95% pointwise confidence intervals with VV@MH response for TS rivet holes with a crack (dots) and without a crack (crosses). The estimated noise mean and 99% noise upper bound are indicated by horizontal dashed and dotted lines, respectively. (b) NIM

POD as a function of crack size with 95% POD lower confidence bound for TS rivet holes. The a90/95 value (3.23mm) is indicated by the vertical dotted line.

5.2 Rivet holes not at Tear Strap Locations

We applied the same procedure to the two populations with the rivet holes outside the tear strap (NT-1 and NT-2). The VV@MH responses as a function of crack size for the NT-1 rivet holes and NT-2 rivet holes are shown in Figure 4 (a) and (c) with crosses for the rivet holes without a crack and the dots for the rivet holes with a crack. The NIM regression line estimates and corresponding 95% pointwise confidence intervals are shown in Figure 4 (a) and (c) with the estimated noise mean (horizontal dashed lines) and estimated 99% noise percentile (horizontal dotted lines). The NIM PODs and their 95% lower bounds as function of crack size are shown in Figure 4 (b) and (d) for NT-1 and NT-2, respectively. The a90/95 value is 6.55 mm for the NT-1 rivet holes and 4.42 mm for the NT-2 rivet holes.

The rivet holes at NT-1 are the most difficult to detect because of the null location and edge effect. The VV@MH response variation for NT-1 rivet holes is larger than that for NT-2 rivet holes, and the noise response for the NT-1 rivet holes is high. The POD for the NT-1 rivet holes is smaller than the POD for the NT-2 rivet holes, and the 95% POD lower confidence bound for the NT-1 rivet holes is the widest among all the rivet hole groups. The NT-1 rivet holes have the largest a90/95 value, more than twice of the TS rivet holes.

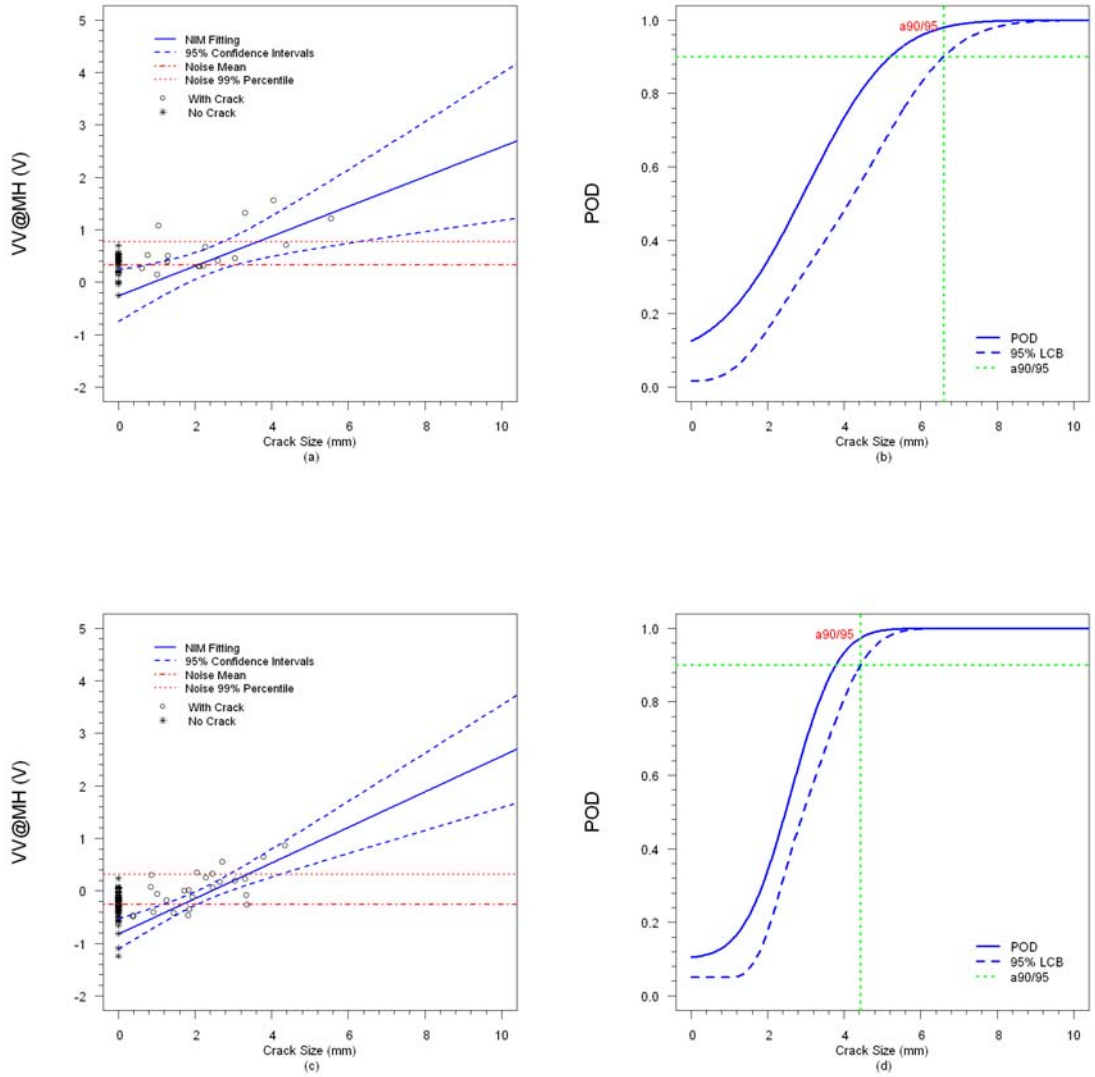

FIGURE 4. (a) NIM regression line estimate and 95% pointwise confidence intervals with VV@MH response for NT-1 rivet holes with crack (dots) and without crack (crosses), and estimated noise mean and 99% noise percentiles are indicated by the horizontal dashed and dotted lines, respectively. (b) NIM POD estimate as function of crack size with 95% POD lower confidence bound for NT-1 rivet holes. The a90/95 value (6.55 mm) is indicated by the vertical dotted line. Note that (c) is the same as (a) except for NT-2 rivet holes and (d) is the same as (b) except for NT-2 rivet holes with a 90/95 value of 4.42 mm.

5.3 Comparisons of NIM and Hit/Miss POD

After obtaining the PODs and their lower confidence bounds for both the Hit/Miss and NIM analyses, we then compared the performance of the two methods for the rivet holes at the TS, NT-1 and NT-2 locations respectively in Figure 5. For the rivet holes at NT-1, where the signals are affected by the edge effect due to the proximity of the rivets to the nearby tear strap, the Hit/Miss POD is uniformly higher than the NIM POD. Given the large variance of VV@MH and high noise response for the NT-1 rivet holes, the single variable VV@MH alone is insufficient to make adequate crack existence decision, and the operator's experience and knowledge of the overall features of the impedance plane data exhibit better performance. For the rivet holes at TS and NT-2, the Hit-Miss and NIM PODs are similar, indicating that automatic crack decision criterion based on VV@MH can be adopted in place of operator's detailed examination of each impedance plane data.

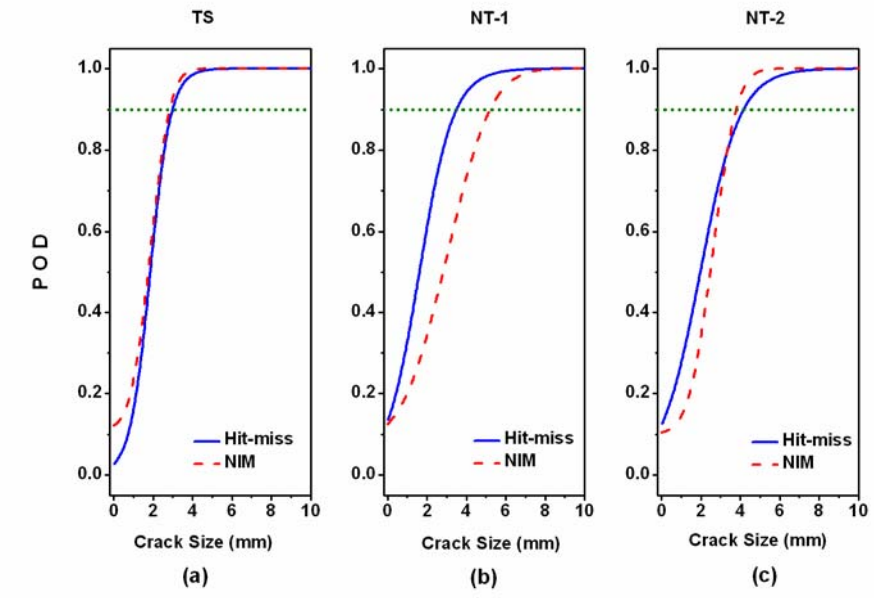

FIGURE 5. POD comparisons between hit/miss and NIM methods for rivet holes at TS (a), NT-1 (b) and NT-2 (c).

6. CONCLUSIONS

In this paper, we applied the noise interference model to estimate POD for a large data set taken by a computer controlled eddy current inspection system. The scalar reduction of the impedance plane data (denoted by VV@MH) was used in the noise interference model (NIM) for crack existence decisions. The POD results were compared between the Hit/Miss method and the noise interference model. It was found that the automatic crack decision criterion based on VV@MH can largely replace operator's Hit/Miss calls with similar POD results, except for those inspections of the rivets outside but near tear straps where the signals were affected by the edge effect. The proposed crack decision criterion based on NIM intends to minimize human-factor variation, and the quantitative results are valuable for future model-assisted POD studies.

There are a number of extensions for the methodology presented in this paper that suggest future research directions. These include the following:

1. We used a scalar reduction of the impedance plane data for the noise interference model. We can develop a bivariate noise interference model to use two types of characteristics that might be expected to provide a better POD for a given PFA.
2. Beside the three types of characteristics used in this paper, we could develop a detection algorithm based on the overall shape of the impedance trace, using image-analysis techniques.
3. We could develop physical model to describe the impedance traces difference for rivet holes at TS, NT-1 and NT-2 then use these physical-model results to formulate detection criteria for each type of rivet hole.

Acknowledgment

This material is based upon work supported by the Federal Aviation Administration under Contract #DTFA03-98-D-00008, Delivery Order #0058 and performed at Iowa State University's Center for NDE as part of the Center for Aviation Systems Reliability program.

References

1. B. Auld and J. Moulder, “Review of Advances in Quantitative Eddy Current Nondestructive Evaluation,” *Journal of Nondestructive Evaluation*, Vol. **18**, 3-36 (1999).
2. J. Knopp, J. Aldrin, and P. Misra, “Considerations in the Validation and Application of Models for Eddy Current Inspection of Cracks Around Fastener Holes,” *Journal of Nondestructive Evaluation*, Vol. **25**, 123-137 (2006).
3. B. Larson, E. Madison, and N. Nakagawa, “Laboratory POD Data Acquisition from Inner Layer Cracks in Simulated Airframe Structures,” *Review of Quantitative Nondestructive Evaluation*, American Institute of Physics, Vol. **28**, 1840-1847 (2009).
4. MIL-HDBK-1823A, Nondestructive Evaluation System Reliability Assessment, Standardization Order Desk, Philadelphia, PA (2009).
5. M. Li and W. Meeker, “A Noise Interference Model for Estimating Probability of Detection for Nondestructive Evaluations,” *Review of Quantitative Nondestructive Evaluation*, American Institute of Physics, Vol. **28**, 1769-1776 (2009).

6. Y. Pawitan, *In All Likelihood: Statistical Modelling and Inference Using Likelihood*, Oxford University Press, New York (2001).