JOURNAL OF AIRCRAFT
Vol. 46, No. 6, November–December 2009

Optimization of Distribution Parameters for Estimating Probability of Crack Detection

Alexandra Coppe,* Raphael T. Haftka,† and Nam-Ho Kim‡
University of Florida, Gainesville, Florida 32611

DOI: 10.2514/1.43804

In the inspection of aircraft structures, the probability of detection has typically been determined based on the size of damage alone. However, the inspection process involves randomness due to variability in inspection conditions, including inspector competence, difficulties associated with the location, and type of damage. To account for these other factors, we develop a simple inspection model from the assumption that for each combination of crack location and inspector there is a threshold crack size, such that larger cracks will be detected and smaller ones will be missed. The proposed model fits the threshold crack sizes to 2603 detection events reported for 43 panels inspected by 62 inspectors from a U.S. Air Force study. The threshold values are obtained by maximizing the matching detection events between the model and the inspection data. First, when 62 inspector thresholds are used, the model matches 78% of detection events. It is further increased to 81% when both inspector and location thresholds are employed. For comparison, the matching percentage using crack size alone is only 55%. Cross-validation of the fitting process indicates that part of the improvement represents fitting noise in the data, and that only about 72% matching is repeatable. The proposed model is further extended by including randomness to represent inconsistent behavior of inspectors. Replicating the observed inconsistency for the same inspector reduces the matching to about the same 72%. Confidence intervals of the proposed model are presented. The proposed model may be used to assess the effect of better training for inspectors, and this is illustrated with a study that shows that, if the poorer half of inspectors is replaced by the top half, the crack size identified with 50% probability can be reduced by half.

Nomenclature

a	=	normalized crack size, a'/a'_m
a_{corr}	=	normalized corrected crack size
a'	=	crack size, inches
a'_m	=	crack size detected with 50% probability, mm
a'_{trs}	=	detection threshold, in.
a_{trs}	=	normalized threshold, mm
d	=	detection event
d_e	=	experimental detection event
d_s	=	simulated detection event
m_{ij}	=	detection margin
N_j^{det}	=	number of inspectors having detected the jth crack
P_d	=	probability of detection
P_e^j	=	experimental probability of detection for panel
x_{ij}	=	agreement margin
β	=	detection parameter in Palmberg equation
Δa^h	=	inspector competence increment (negative for high competence)
Δa^l	=	location difficulty increment (positive for high difficulty)
θ	=	penalty function

I. Introduction

MOST aircraft structural components are designed based on a damage-tolerance philosophy that uses inspection and maintenance to detect degradation before it can cause structural failure. In general, the inspection can be performed either manually or by using onboard equipment. In this paper, the former is referred to as manual inspection, and the latter as structural health monitoring (SHM). For manual inspections, different techniques have been used, such as radiographic inspection (Lawson and Parker [1]). Usually, SHM uses actuator-sensor technique (Giurgiutiu and Cuc [2]) to detect damages, such as ultrasonic and eddy current techniques (Pohl et al. [3]), comparative vacuum monitoring (Stehmeier and Speckmann [4]), or elastic wave propagation and electromechanical impedance (Giurgiutiu et al. [5]).

The effectiveness of various inspection techniques is typically characterized by probability of detection (POD) curves that relate the size of damage to POD (Zheng and Ellingwood [6]). The information on POD can be used for various purposes, including structural diagnosis and prognosis (Zheng and Ellingwood [6]). For example, Kale and Haftka [7] used POD curves to optimize the inspection schedule that can maintain a certain level of structural reliability. Although the POD curve is traditionally given in terms of damage size, in reality, POD depends not only on damage size but also on other variables. For example, damage in some locations is more difficult to detect than in other locations. The competence of inspector or inspection method can also be an important factor for determining POD curves.

Developing an accurate damage detection model that can take into account the effects of the location of damage and the competence of inspector is important, but it is not available in the literature. The objective of the paper is to demonstrate the development of a more complete characterization of POD curves based on the results of a large number of inspections. As a first step toward developing such a model, we propose a simple model based on a damage detection threshold size that is affected by both the damage location and the inspector competence. We further simplify the model by assuming that the damage detection process is deterministic, not probabilistic. The proposed model assigns a competence score to each inspector and location difficulty score to each panel. Then, the equivalent damage threshold size for a specific panel and inspector is obtained using the scores.

Although the proposed model can take into account location difficulty and human factors, it is still a deterministic model, which means that the detection event is completely determined with the

Presented as Paper 5943 at the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Victoria, British Columbia, Canada, 10–12 September 2008; received 13 February 2009; revision received 6 August 2009; accepted for publication 21 August 2009.

*Graduate Student; alex.coppe@ufl.edu. Student Member AIAA.
†Distinguished Professor; haftka@ufl.edu. Fellow AIAA.
‡Associate Professor; nkim@ufl.edu. Member AIAA

threshold crack size. However, there exists uncertainty in detecting a crack even if the same inspector inspects the same crack again. To model this randomness, we further improve the model using a traditional POD curve-based inspection process.

To demonstrate the performance of the proposed model, we use the U.S. Air Force study from the 1970s in which 43 panels with different crack sizes are inspected by 62 inspectors (2603 detection events) (Lewis et al. [8]). The eddy current inspection method is used for a C-130 center wing box section with surface fatigue cracks radiating from fastener sites undergone in service. The intact section is first stripped of paint for visual crack detection and measurement during test, a surface finish was then reapplied for the nondestructive inspection testing. We use two optimization techniques to find the location factors associated with 43 panels and the human factors associated with 62 inspectors. The first one uses the matching events as a discrete objective function, which can lead to numerous local optima, whereas the second one uses a regularized objective function, which is continuous and more stable. The robustness of the proposed models is demonstrated using cross-validation, in which the modeling process is repeated by using 90% of data chosen randomly.

The goal of the proposed model is not to replace the traditional POD curve. Rather, it improves the understanding of inspection process by modeling the effect of various parameters, such as inspector capability and location difficulty, on the quality of inspection. The method can be used as a tool to simulate the effect of changing inspection parameters, such as inspector training on the POD.

The remainder of the paper is organized as follows. Section II presents the proposed inspector-location-size (ILS) model based on deterministic thresholds. Two different optimization formulations are discussed in Sec. III. First, the optimization problem is solved for individual parameters in Sec. IV, followed by the automated optimization results in Sec. V. Section VI extends the proposed model to include the randomness of the process along with cross-validation. The confidence intervals and a parameter study of the proposed model are presented in Sec. VII, followed by conclusions in Sec. VIII.

II. Inspector-Location-Size Inspection Model

The detection process is conventionally modeled using a probability of detection curve, which describes the probability of detecting a crack with a specific size. A commonly used POD curve is the Palmberg equation (Palmberg et al. [9]). It specifies the probability of detecting a crack of size a' as

$$P_d(a') = \frac{(a'/a'_m)^\beta}{1 + (a'/a'_m)^\beta} \tag{1}$$

where a'_m is the crack size that corresponds to 50% probability of detection (hence, it measures the quality of the inspection process). As the exponent β increases, the detection process approaches a deterministic one; that is, all cracks larger than a'_m will be detected and smaller ones be missed. When $\beta = 4$, for example, the probability of detecting a crack size of $a' = 2a'_m$ is 94%. It is noted that the POD curve in Eq. (1) only accounts for crack size.

Although the Palmberg model has been widely used in manual inspections, it overlooks important aspects of the inspection process. For example, when damage exists in a location difficult to detect, the POD is relatively low even if the size of damage is large. Thus, the actual inspection results are often scattered around the POD curve and sometimes show inconsistent behavior in that a small crack may be detected while a larger one missed. The scatter in inspection results can be explained by differences in the competence of inspectors and differences in damage location. The former includes inspector's skill, inspection method, and inspection environment (such as fatigue and distractions).

In this paper, we seek a model that includes the aforementioned two effects, in addition to the traditional crack size effect. We assume that, when a panel is subjected to periodic inspections, the failure to detect a crack of size $a' = a'_m$ is due to the following two variables. The first variable, denoted by h, characterizes the circumstances of the inspection, such as the competence of the inspector and difficulties in the inspection process. The other variable, denoted by l, characterizes the difficulty associated with the location of the damage. These two variables are random by nature. For example, an inspector who missed a crack with size a' may detect the crack in the second trial. As a first step, we use a quasi-deterministic model that assumes that with sufficient knowledge there is no randomness in the detection process. We assume that, for given inspector and location, there is a threshold crack size so that every crack larger than this threshold will be detected and every crack below it will be missed. This model interprets the randomness as being entirely epistemic (lack of knowledge). That is, if we knew everything about the location of the damage and the inspection condition, then the randomness would disappear. Denoting the threshold value by a'_{trs}, the detection event d for a crack of size a' can be defined as

$$d = \begin{cases} 0 & \text{if } a' - a'_{\text{trs}} < 0 \\ 1 & \text{if } a' - a'_{\text{trs}} \geq 0 \end{cases} \tag{2}$$

In the preceding equation, $d = 0$ means that the crack has been missed, whereas $d = 1$ means that the crack has been detected. We simplify the following derivations by normalizing all crack sizes using the mean value a'_m of the threshold crack size over all locations and inspectors:

$$a_{\text{trs}} = \frac{a'_{\text{trs}}}{a'_m} \tag{3}$$

The same normalization is applied to a' such that $a = a'/a'_m$.

The objective is to develop a model of a_{trs} that can accurately represent the contributions from both location and inspection condition. In view of the deterministic model, if the damage is located in the neutral position and if the inspection conditions are the same, every crack larger than a'_m will be detected and those smaller than that will be missed. The proposed model adjusts the threshold based on the contribution from the location Δa^l and that from the inspection competency Δa^h as

$$a_{\text{trs}} = 1 + \Delta a^l + \Delta a^h \tag{4}$$

A positive Δa^l means that the crack is in a more difficult location than average such that its detection threshold is larger than a'_m. A positive Δa^h means the inspection circumstances are more difficult than average. Thus, a value of a_{trs} greater than one means that the crack is more difficult to detect than average because either its location is difficult to find or the inspector is not competent. The detection event can be determined using the normalized version of Eq. (2). We call this model the ILS_{det} (inspector-location-size) model.

The performance of this model will be tested by applying it to a matrix of tests where a series of 43 panels with cracks were inspected by 62 inspectors [8]. Part of the matrix (13 inspectors, 32 locations) is shown in Table 1.

To generate the traditional Palmberg equation from the experiments, we start by calculating the probability of detection for each panel: $P_e^j = N_{\text{det}}^j/62$, $j = 1, \ldots, 43$, where N_{det}^j is the number of inspectors that detect the jth crack. Note that the jth panel has a crack of size a'_j. We then fit the two Palmberg parameters to the 43 probabilities by minimizing the discrepancy defined in Eq. (5):

$$\min_{a'_m, \beta} \sum_{j=1}^{43} |P_d(a'_j) - P_e^j| \tag{5}$$

After minimization, we obtain $a'_m = 0.48$ cm and $\beta = 1.308$. Figure 1 shows the POD curve as function of crack size a', along with the 43 probability data P_e^j, used to fit it. It can be observed that the curve fits most of the points, but on the bottom right we can see that the largest crack has a very low probability of detection, which indicates that this crack might be located in a place where it is very difficult to detect. Figure 1 also shows two typical ILS_{det} curves when $a'_{\text{trs}} = 0.5$ and $a'_{\text{trs}} = 1.2$ (two vertical lines).

Table 1 Partial matrix of crack detection events (1 = detection, 0 = nondetection) from [8]

		Inspector ID												
Flaw ID	Flaw length, in	0201	0202	0204	0207	0208	03E1	03E2	03E3	03E4	03E5	03E7	03E9	03E10
77a	0.09	0	0	1	0	0	0	0	0	1	0	0	0	0
122	0.09	1	0	0	0	0	0	0	1	1	0	0	0	0
132	0.10	1	0	0	0	0	0	0	0	0	0	0	1	0
121	0.10	1	0	1	0	0	1	0	1	1	0	0	0	1
75	0.10	1	0	1	0	0	0	0	0	0	0	1	0	0
76b	0.12	1	0	1	0	0	0	0	0	1	1	1	1	0
80	0.12	1	1	1	0	0	0	0	0	1	0	0	0	0
77b	0.13	1	1	1	0	0	0	0	0	1	1	1	0	1
79d	0.13	1	0	0	0	0	0	0	0	0	0	0	1	0
125	0.13	1	0	0	1	0	0	0	1	0	0	0	0	1
133	0.14	0	0	0	0	0	0	——	0	0	0	0	0	0
12a	0.15	1	1	1	1	0	0	0	1	1	0	0	0	0
78	0.16	1	1	1	0	0	0	0	0	1	0	1	0	0
9b	0.16	1	1	1	1	0	0	0	1	0	1	0	1	1
131	0.16	1	1	1	0	0	0	0	1	1	0	1	0	0
7	0.16	1	0	1	1	0	0	——	1	1	0	0	0	0
8a	0.17	0	0	1	0	0	0	——	1	0	0	0	0	0
130	0.19	1	1	1	1	0	0	1	1	1	1	0	0	1
10d	0.19	1	1	1	1	0	1	1	0	0	1	1	0	1
101	0.20	0	0	1	1	1	0	——	0	0	0	0	0	1
76a	0.21	1	0	1	0	1	0	1	0	1	1	0	0	1
81b	0.21	1	1	1	1	1	0	1	1	1	0	0	1	1
123	0.21	1	1	1	1	1	0	1	0	1	0	1	0	0
8c	0.21	1	0	1	0	0	0	0	1	0	1	1	0	1
9a	0.22	1	1	1	1	1	0	1	1	1	1	0	1	1
11b	0.22	1	1	1	1	0	0	1	1	1	1	0	1	1
79a	0.23	1	1	1	1	1	0	1	0	1	1	1	0	1
12b	0.23	1	1	1	1	1	0	1	1	1	1	0	1	1
8b	0.23	1	0	1	0	0	0	——	1	0	1	0	0	1
10a	0.24	1	1	1	1	0	0	1	1	0	1	0	1	1
81a	0.25	1	1	1	1	1	0	1	1	1	1	1	1	1
11a	0.29	1	1	1	1	0	1	1	1	1	1	0	1	1

III. Optimization Formulations

To test if the proposed ILS_{det} model describes well the results of the inspections, we seek 43 Δa_i^l corresponding to the 43 panels and 62 Δa_j^h corresponding to the 62 inspectors that fit best the observed inspection events. The threshold increments Δa_j^h associated with the inspectors are easy to estimate because for each inspector we have inspection results for 43 different crack sizes. On the other hand, for each location, we have only a single crack size, and therefore the estimate of the location difficulty must rely on the scores of the inspectors. If a crack in a panel is not found even by the most competent inspectors (lowest threshold values), then we can deduce that it is in a difficult location.

The inspection results from the U.S. Air Force study [8] are used, in which 2603 detections events out of $62 \times 43 = 2666$ possible events are reported. The objective is to find 105 values for the inspectors Δa_i^h and the panels Δa_j^l that will predict the large majority of the inspection results. An optimization problem is formulated such that the differences between the inspection results from the tests and that from the model in Eqs. (2) and (4) are minimized.

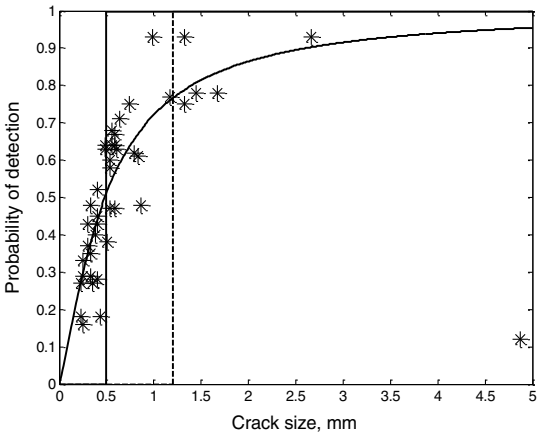


Fig. 1 Probability of detection ($a'_m = 0.48$ and $\beta = 1.308$) curves including the traditional Palmberg equation and two typical ILS curves corresponding to the data in [8].

We consider two ways of quantifying the differences between the prediction and the inspection results: discrete and continuous formulations. The former represents detection and nondetection as binary events, that is, 0 or 1. We denote by d_e the detected events from actual inspections and by d_s the detected events from the model. We define the detection margin m_{ij} and the detection event $d_{s_{ij}}$ resulting from the model for the crack in the ith panel and jth inspector as follows:

$$m_{ij} = a_i - (1 + \Delta a_j^h + \Delta a_i^l)$$
$$d_{s_{ij}} = \begin{cases} 0 & \text{if } m_{ij} < 0 \text{ (nondetection)} \\ 1 & \text{if } m_{ij} \geq 0 \text{ (detection)} \end{cases} \tag{6}$$

The objective function of the discrete formulation is then defined as

$$\min_{\Delta a_l, \Delta a_h} \sum_{i=1}^{43} \sum_{j=1}^{62} |d_{s_{ij}} - d_{e_{ij}}| \tag{7}$$

The objective function in Eq. (7) is obviously discontinuous as infinitesimal changes in the threshold values can switch a detected event to a nondetected event and vice versa. It is noted that the preceding optimization formulation is different from the conventional discrete optimization in which the variables are discontinuous but the objective function is usually continuous. In the preceding optimization formulation, the variables Δa_i^h and Δa_j^l are continuous, while the objective function is discrete.

The continuous formulation takes into account the size of the margin m_{ij} in each detection event. When the detection events from

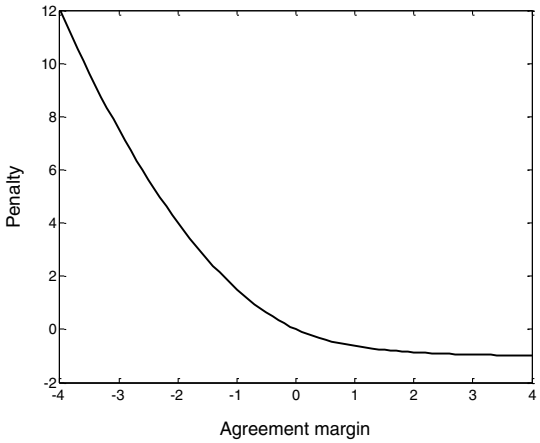


Fig. 2 Penalty function for matching margins (positive margins denote agreement between simulation and data).

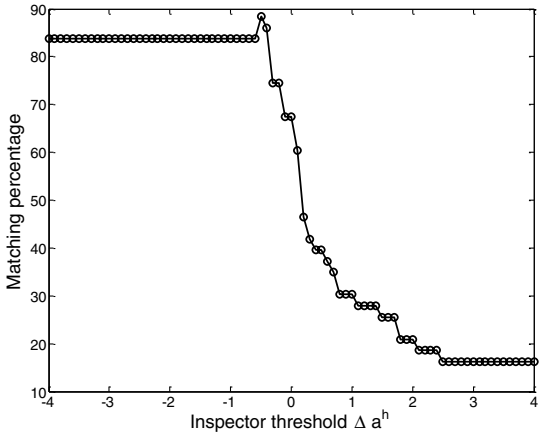


Fig. 3 Percentages of matching events for inspector 14 as a function of threshold.

inspection and model are not consistent, we penalize the event based on the size of the margin. On the other hand, when the detection events from inspection and model are consistent, we provide a small reward to the objective function that increases with the margin.

Using the margins in an objective function allows us to define a continuous objective function that may be easier to optimize. In addition, it will provide a more robust fit that is likely to be less sensitive to small changes in the threshold. However, it is desirable to define an objective function that will lead to an optimum that will not result in substantial deterioration in the matching objective of Eq. (7). The continuous optimization problem is defined as

$$\begin{cases} \min\limits_{\Delta a_l, \Delta a_h} \sum_{i=1}^{43} \sum_{j=1}^{62} \theta_{ij} \\ \theta_{ij} = \begin{cases} \frac{x_{ij}^2}{2} - x_{ij} & \text{if } x_{ij} < 0 \\ \exp(-x_{ij}) - 1 & \text{otherwise} \end{cases} \\ x_{ij} = (2d_{e_{ij}} - 1)m_{ij} \end{cases} \tag{8}$$

where x_{ij} is the agreement margin, which is negative when the results of the model do not match the experimental events and positive when they do. The penalty θ_{ij} is chosen such that it increases rapidly when the model and experiment are not matched, whereas it decreases slowly when the results are matched. In addition, the two expressions of θ_{ij} have the same value and slope at $x_{ij} = 0$. The penalty function for a single detection event is shown in Fig. 2. Although the individual penalty function is monotonic, the objective function is not unimodal in general because it is sum of all penalty functions.

IV. Initial Estimates of Thresholds

Because the optimization problem in Eq. (7) is not continuous and that in Eq. (8) may not be unimodal, it is important to start the optimization process with a good initial estimate. To obtain the initial estimate, the optimization problem is simplified by separating the inspector contributions from the location contributions. We first estimate the threshold Δa_j^h of each inspector by solving the problem column by column without associating any difficulty with the cracks (i.e., all Δa_i^l are zero). The next step is to estimate Δa_i^l using these Δa_j^h. The first step is illustrated graphically in Fig. 3. The figure shows the percentage of matches of 43 panels for inspector 14. For this inspector, the best threshold increment is $\Delta a_{14}^h = -0.5$, associated with which 38 of the 43 detection events are matched, or 88.37% match. That is, for average location difficulty ($\Delta a^l = 0$), this inspector will detect every crack longer than 0.5 times a'_m.

Figure 4 shows the continuous objective function θ_{i14} for the same inspector. Comparison of Figs. 3 and 4 shows a good agreement between the two optima.

By solving the optimization problem individually for each inspector, we find 62 Δa_j^h, which represent the competence of inspectors. Overall, these 62 Δa_j^h match 78.30% of the detection

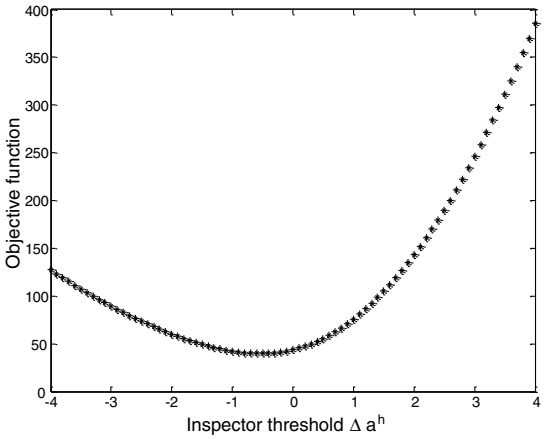


Fig. 4 Continuous objective function for one inspection and different matches.

events. Using the optimal Δa_j^h as initial estimates, and varying Δa_i^l, we obtain a matching percentage of 80.61%, a small improvement. The value of continuous objective function at these estimates is −458.97.

To evaluate the quality of the optimization results, we compare the matching percentage result with the traditional model in which POD is determined based entirely on the crack size. Let us consider that the ith panel has a crack with size a'_i. Using the two-parameter Palmberg equation in Eq. (1), the POD of the crack can be calculated (Palmberg et al. [9]). By performing a Monte Carlo simulation using the Palmberg equation to calculate the POD of the crack sizes used in [8] for 62 inspectors, we obtain a matching percentage between those simulated data and the experimental ones of 55.5% (with a standard deviation of 0.96%) which is much lower than 81% in the proposed model. Thus, the proposed simple model accounts much better for the actual inspection results than a model that takes only the crack length.

V. Optimization Results

The estimation results are used as the initial point for both optimization problems discussed previously (continuous and discrete). Both optimization formulations are unconstrained optimization with 105 variables.

First, the continuous optimization, Eq. (8), is solved using the Matlab function *fminunc*, which uses the Broyden–Fletcher–Goldfarb–Shanno [10–13] quasi-Newton method with a mixed quadratic and cubic line search procedure. The optimization problem using the evaluated Δas discussed in Sec. IV converges to a solution

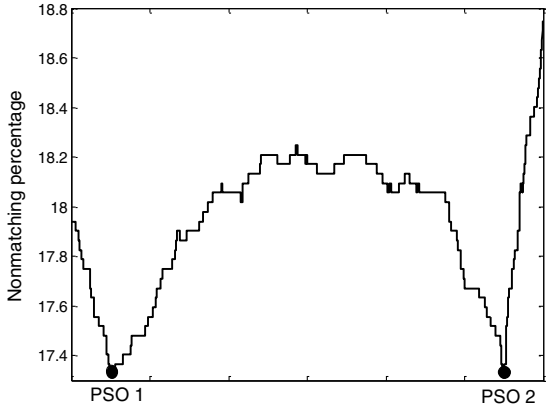


Fig. 5 Nonmatching percentage variation between two PSO optima.

that yields 77.79% matches and the penalty function value of −733.19. Note that because *fminunc* is a local optimizer, we first estimate the Δas to have a result which is presumably close to the global optimum. It can be observed that the matching percentage is slightly lower than that from the sequential optimization for both Δa_i^l and Δa_j^h, but having a much lower penalty function means that the continuous optimum is more stable than the sequential one. The slight increase of the objective function can be explained by the fact that the continuous objective function (penalty function) is different from the discrete one (matching percentage).

The discrete optimization problem in Eq. (7) cannot be solved using gradient-based optimization algorithms because the objective function is discontinuous. Indeed, the discrete objective function is a piecewise step function in 105 dimensional space. It is different from conventional discrete optimization problems, in which the objective function is continuous while the design variables are discrete. The discrete optimization problem is solved using Particle Swarm Optimization (PSO) algorithm (Schutte et al. [14]). This algorithm suits well our purpose because it does not require the gradient of objective function, and the design variables are continuous. The discrete optimization problem converges to a percentage of 82.79% matches, which is slightly better than the initial estimate. It turns out that there are many design points that yield the same value of optimum objective function, even if different initial stating points are used. We did find that PSO converges to the global optimum better than the gradient-based method. To show how the objective function varies between two optimum points, two optimum points are chosen. By connecting these two points in 105-dimensional space and by evaluating objective functions along the line, Fig. 5 plots variation of the objective function. PSO1 and PSO2 represent two optima, and the values of objective function at these two points are minimal. The discrete objective function (nonmatching percentage) changes rapidly near the optima, which indicates that the two optimum points are unstable. Having many optimum points means that different combinations of Δa_j^h and Δa_i^l can yield the same level of matching percentage with the experiments.

A summary of the different optimization results, matching percentage, and continuous objective function can be found in Table 2.

To study the stability further for continuous and discrete optimizations, two optimum points from the discrete formulation and one optimum point from the continuous formulation are chosen. By locating them in the three corners in a plane, we can project 105-dimensional space into 2-dimensional space. Then, the values of objective function are plotted on the plane as a contour. Figure 6a plots the contour of the continuous objective function, whereas Fig. 6b plots that of the discrete objective function. It can be observed that the continuous objective function is smooth and converges to a single optimum, whereas the discrete one has numerous local optima. This illustrates clearly why the continuous objective function yields a single and more stable optimal model. Note that Fig. 5 is the value of the discrete objective function along the hypotenuse.

VI. Probabilistic $\mathrm{ILS}_{\mathrm{rand}}$ Inspection Model

As we discussed in the earlier sections, the proposed $\mathrm{ILS}_{\mathrm{det}}$ model neglects the variability in inspection environment and human factors. That is, the same inspector inspecting the same panel may detect the damage at the first trial but miss it the second time. Thus, we cannot expect a 100% match even if the margin m_{ij} is positive. In fact, it is quite possible that the optimization process overfitted the variability in the data. If the 2603 inspection events were repeated, the match between the two repetitions would be less than 82.79%.

To check the overfitting case, the technique of cross-validation [15] is used. The cross-validation can be done by leaving out 10% of the data, fitting the model to the remaining 90%, and checking the fit with the left out 10% events. This process is repeated 100 times. We find that the matching percentage reduces to 72.6% (with standard deviation of 2.7%) for the sequential approach and 75.6% (with standard deviation of 2.4%) for the continuous approach. The 72.6% match in sequential approach indicates that about 10% of the match in the discrete optimization captures the instantaneous performance of the inspectors rather than their average competence, and this confirms the fact that the continuous optimum is more stable.

The results of the optimization provide additional insight into the magnitude of this randomness in the form of inconsistency in the performance of the inspectors. That is, even with the optimum assignments of difficulty to the 43 panels, the experimental results still show inspectors identifying a difficult crack while missing an easier one. Thus, the present model can be improved by modeling the randomness responsible for this inconsistency.

To quantify the inconsistency in the inspector performance, we first define a corrected crack size that accounts for the location difficulty

$$a_{\mathrm{corr}} = a - \Delta a^l \tag{9}$$

The crack that is difficult to find is considered to be a smaller-sized crack in the nominal location.

Next, we arrange the plates in increasing order of corrected crack size and calculate the number of times the detection and nondetection events alternate. For that purpose, we define inconsistency indicator as

$$I_j = \sum_{i=2}^{43} \langle d_{i-1,j} - d_{i,j} \rangle \tag{10}$$

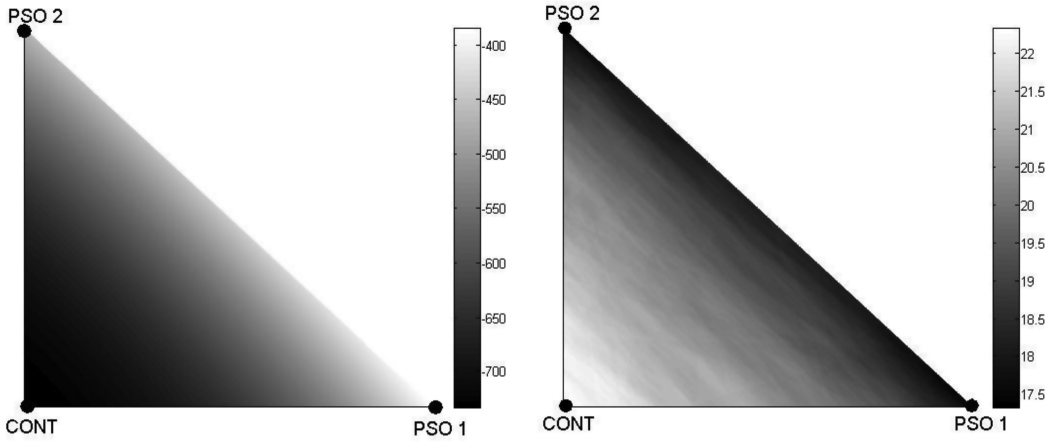
with $\langle d_{i-1,j} - d_{i,j} \rangle = \max(d_{i-1,j} - d_{i,j}, 0)$. Because all 43 cracks are ordered in increasing corrected size, the detection events should be $d_{i-1,j} \leq d_{i,j}$ if the inspector is consistent. Whenever the inspector detects a smaller crack and misses a larger crack, the inconsistency indicator in Eq. (10) increases by one.

A perfectly consistent inspector will have $I_j = 0$, with zeros for small cracks and ones for large ones. The average value of I_j for the 62 inspectors is 5.24 with the values ranging from 1 to 10. Surprisingly, the inconsistency levels did not appear to be related to competence as seen in Fig. 7. Competent inspectors ($\Delta a^h < 0$) have as large inconsistency indicator as incompetent inspectors $\Delta a^h > 0$.

To account for inconsistency, a new $\mathrm{ILS}_{\mathrm{rand}}$ model is proposed where the threshold in the deterministic model is used as the crack

Table 2 Optimization results (note that robust agreement leads to negative penalties as seen in Fig. 2)

	Sequential optimum	Discrete optimum	Continuous optimum
Matching percentage	80.61	82.79	77.79
Penalty function	−458.97	−164.41	−733.19



a) Continuous objective function **b) Discrete objective function**

Fig. 6 Contour plots of continuous and discrete objective functions on the projected plane of three optima (two discrete and one continuous).

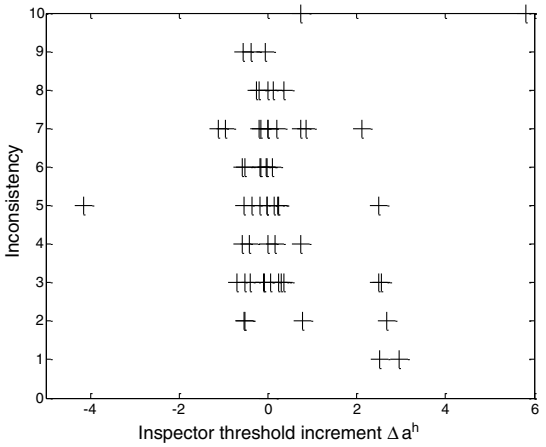


Fig. 7 Inspectors inconsistency versus their competence. Large threshold increments correspond to incompetent inspectors.

size that has 50% probability of detection; that is, a_m in the traditional POD curve. Thus, in the ILS_{rand} model, the Palmberg equation is rewritten as

$$P_d^{i,j} = \frac{(a^i/a_{\text{trs}}^{i,j})^\beta}{1 + (a^i/a_{\text{trs}}^{i,j})^\beta} \tag{11}$$

A very large value of β in Eq. (11) would correspond to the deterministic ILS model. In the ILS_{rand} model, the exponent β is selected such that the inconsistency from the model matches with that of the experimental data. To match the average inconsistency of 5.24, we need $\beta = 2.7$. In addition, we obtain a matching percentage about 72% using Eq. (11). This means that we lose about 10% matching percentage to reach the same level of inconsistency. This confirms the results obtained in the cross-validation study. A summary of the results is given in Table 3.

Unlike the ILS_{det} model, the ILS_{rand} model will have different results if the same process is repeated. The values in the parenthesis in Table 3 show the standard deviation of the matching percentage and average inconsistency. To estimate the effect of this randomness in the model, we find that the percentage of matches between two different trials is 67.4% (standard deviation of 0.9%).

VII. Using Model for Assessing Impact of Parameter Changes

An important aspect of the proposed model is that it allows one to quantify the effect of the parameters on the conventional POD curves. For example, it is possible to measure the improvement of the POD curve when the inspectors have better training. To perform a parameter study, it is necessary to build a POD curve first. Both methods presented earlier can be used to simulate inspection results for a given set of inspectors and location thresholds. The idea is similar to fitting the conventional POD, that is, Palmberg equation, to inspection data as shown in Fig. 1. Figure 8 shows the Palmberg equation fitted to the data simulated using both models. Note that, due to the random nature of ILS_{rand}, 1000 samples of simulated inspection data are generated, and the Palmberg equation is fitted to the mean of the probability of detection for each location. In the previous section, ILS_{rand} does not provide matching results as good as ILS_{det} but, when it comes to the fitting conventional POD curves, it agrees better with the experimental results. This is due to ILS_{det} ignoring the inconsistency in inspector performance, thus obtaining more deterministic (steeper) dependence of POD on crack size.

Because the ILS_{rand} model includes randomness in identifying the POD curve, it is necessary to estimate the confidence bounds for the POD curve associated to the model (MIL-HDBK-1823 [16]). Figure 9 shows 5 and 95% confidence levels, along with the mean POD curve. They are calculated the same way as the mean presented in Fig. 8. The 95% confidence level is close to the Palmberg equation fitted to the actual inspection data but on the conservative side.

Once the POD curve is calculated, it is possible to find the effect of model parameters on the POD. An important possible issue is the effect of inspector competence on the POD. In practice, inspector competence can be improved by training. This is simulated by replacing the inspectors with scores in the bottom half by duplicating the inspectors in the top half. Using the improved inspector competencies, a new Palmberg equation is fitted to the mean probability of detection values. Figure 10 shows significant improvement of POD resulting from improvement of inspector competences. The crack size that has 50% probability of detection is reduced by half.

Table 3 Matching percentage and average inconsistency for the three models presented in this paper (values in parentheses are standard deviations)

Method	Traditional Palmberg model	ILS_{rand}	ILS_{det}
Matching percentage	55.5 (0.96)	72.07 (0.68)	82.79
Average inconsistency	10.1 (0.2)	5.2 (0.1)	0

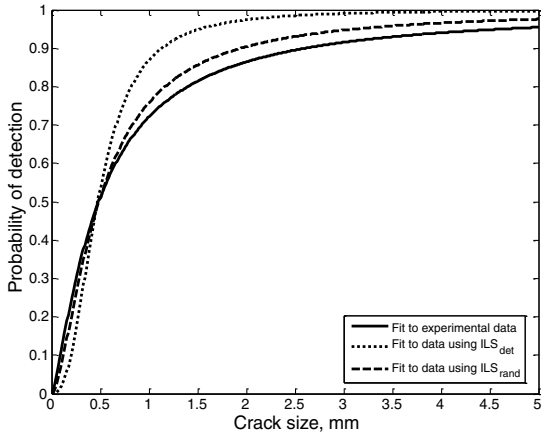


Fig. 8 Fitted Palmberg equation for the experimental data and both models.

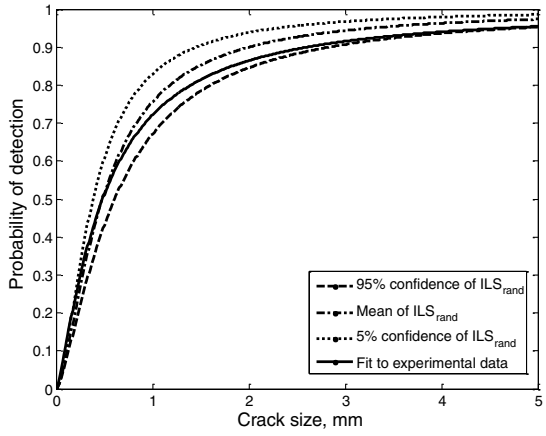


Fig. 9 Confidence level of POD curves fitted to data simulated using the $\mathbf{ILS_{rand}}$ model.

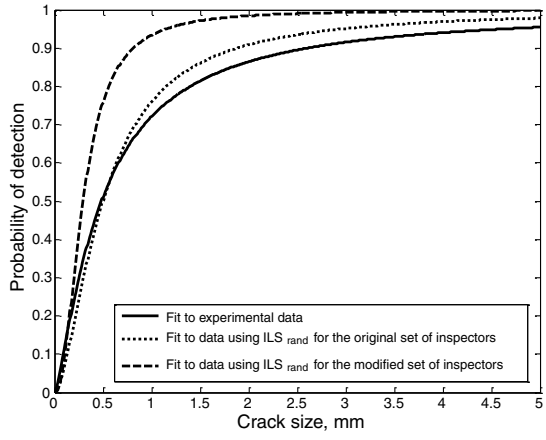


Fig. 10 Effect of inspector's training on the conventional POD curve.

In a more complex model, other parameters could be added by following the same idea as the one presented here.

VIII. Conclusions

We developed a simple model that accounts for inspector competence and location difficulty to explain the randomness in detecting cracks by manual inspection. By fitting 105 parameters to 2602 experiments, we were able to match about 77.8% of the detection events using continuous optimization and 82.8% using discrete optimization. This is a significant improvement from 55% matching by the commonly used model based on crack size alone. Through the cross-validation, we showed that the higher performance of the discrete optimization is related to the fact that the optimization fits randomness of the inspection process.

The deterministic model is simple but it is not robust because it fits noise. The random model is closer to the traditional POD and is actually a mix between the deterministic model and the traditional POD. The random model does not match as well the experimental results but it is more robust.

The experiments revealed inconsistency in the performance of inspectors, and we defined a measure of inconsistency and matched it by adding randomness to our model. With this new model we found that the matches reduced to about 72%, which was similar to the matches between two realizations of the simulated inspections. This indicates that our model captured well both the deterministic and random components of the inspection process.

The procedure revealed that most of the randomness in the detection process is due to inspector competence rather than due to the crack location. This implies that automated structural health monitoring, which will eliminate most of the variability due to the circumstance of the inspection, is likely to provide substantial improvement in the probability of detection. Note that there will still be some variability in inspection that needs to be modeled differently, such as error in the model, initial location of the sensor, and various reading errors related to the sensors.

To demonstrate the utility of the model, we showed that it can be used to examine the benefits of training better inspectors. We found that, if only the top 50% of inspectors were used, the size of the crack that can be detected with 50% probability would be decreased by half.

The proposed model is a first step toward a more advanced inspection model that can include many other variables of inspection process. When more information about damage is available, such as crack morphology, it can be incorporated in a more advanced model. However, this is out of the scope of this paper because the U.S. Air Force report (Lewis et al. [8]) does not provide the required information.

Acknowledgments

This work was supported by the U.S. Air Force under Grant FA9550-07-1 and by NASA under Grant NNX08AC334.

References

[1] Lawson, S. W., and Parker, G. A., "Intelligent Segmentation of Industrial Radiographs Using Neural Networks," *Proceedings of SPIE: The International Society for Optical Engineering*, Vol. 2347, Nov. 1994, pp. 245–255. doi:10.1117/12.188736

[2] Giurgiutiu, V., and Cuc, A., "Embedded Non-Destructive Evaluation for Structural Health Monitoring, Damage Detection, and Failure Prevention," *Shock and Vibration Digest*, Vol. 37, No. 2, 2005, pp. 83–105. doi:10.1177/0583102405052561

[3] Pohl, R., Erhard, A., Montag, H. J., Thomas, H. M., and Wüstenberg, H., "NDT Techniques for Railroad Wheel and Gauge Corner Inspection," *NDT & E International: Independent Nondestructive Testing and Evaluation*, Vol. 37, No. 2, 2004, pp. 89–94.

[4] Stehmeier, H., and Speckmann, H., "Comparative Vacuum Monitoring (CVM) Monitoring of fatigue cracking in aircraft structures," *2nd European Workshop on Structural Health Monitoring*, 2004, http://atlas-conferences.com/c/a/n/b/49.htm.

[5] Giurgiutiu, V., Redmond, J., Roach, D., and Rackow, K., "Active Sensors for Health Monitoring of Aging Aerospace Structures," *Proceedings of the SPIE Conference on Smart Structures and Integrated Systems*, Vol. 3985, Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, 2000, pp. 294–305.

[6] Zheng, R., and Ellingwood, B., "Role of Non-Destructive Evaluation in Time-Dependent Reliability Analysis," *Structural Safety*, Vol. 20, No. 4, 1998, pp. 325–339. doi:10.1016/S0167-4730(98)00021-6

[7] Kale, A., and Haftka, R., "Tradeoff of Structural Weight and Inspection Cost in Reliability Based Optimization Using Multiple Inspection Types," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA Paper 2004-4404, Aug. 2004.

[8] Lewis, W. G., Dodd, B. D., Sproat, W. H., and Hamilton, J. M., "Reliability of Nondestructive Inspections: Final Report," U.S. Air Force, San Antonio Air Logistics Center, Kelly AFB, TX, SAALC/MME 76-6-38-1, Dec. 1978.

[9] Palmberg, B., Blom, A. F., and Eggwertz, S., "Probabilistic Damage Tolerance Analysis of Aircraft Structures," *Probabilistic Fracture Mechanics and Reliability*, M. Nijhoff, Dordrecht, The Netherlands, 1987.

[10] Broyden, C. G., "The Convergence of a Class of Double-Rank Minimization Algorithms," *Journal of the Institute of Mathematics and its Applications*, Vol. 6, No. 1, 1970, pp. 76–90. doi:10.1093/imamat/6.1.76

[11] Fletcher, R., "A New Approach to Variable Metric Algorithms," *Computer Journal*, Vol. 13, No. 3, 1970, pp. 317–322. doi:10.1093/comjnl/13.3.317

[12] Goldfarb, D., "A Family of Variable Metric Updates Derived by Variational Means," *Mathematics of Computation*, Vol. 24, 1970, pp. 23–26. doi:10.2307/2004873

[13] Shanno, D. F., "Conditioning of Quasi-Newton Methods for Function Minimization," *Mathematics of Computation*, Vol. 24, 1970, pp. 647–656. doi:10.2307/2004840

[14] Schutte, J. F., Reinbolt, J. A., Fregly, B. J., Haftka, R. T., and George, A. D., "Parallel Global Optimization with the Particle Swarm Algorithm," *International Journal of Numerical Methods in Engineering*, Vol. 61, No. 13, 2004, pp. 2296–2315. doi:10.1002/nme.1149

[15] Kohavi, R., "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, Morgan Kaufmann, San Mateo, Vol. 2, No. 12, 1995, pp. 1137–1143.

[16] "Nondestructive Evaluation System Reliability Assessment," MIL-HDBK-1823, U.S. Dept. of Defense, 30 April 1999.