Optimization of Distribution Parameters for Estimating Probability of Crack Detection

Alexandra Coppe[1], Raphael T. Haftka[2] and Nam-Ho Kim[3]
University of Florida, Gainesville, FL, 32611

and

Palani Ramu[4]
Belcan Engineering Group, Inc.,
Caterpillar Champaign Simulation Center, Champaign, IL. 61820

Probability of detection curves used for manual inspections are typically based entirely on damage size. The randomness in the process is likely to be due to variability in the circumstances of the inspection (including the competence of inspector), as well as objective difficulties associated with the location and type of damage. In order to shed light on the relative contributions of these two sources of randomness, we analyze a large US Air Force study from the 1970s that distributed 43 panels with cracks to 62 inspectors. We develop a simple model that assumes that for each combination of crack location and inspector there is a threshold crack size such that all cracks above this size will be detected and all cracks below that size will be missed. The model is fitted to 2602 detection events by finding 43 location threshold increments and 62 inspector increments. With the 62 inspector increments we match 78% of the inspections, and location increments increase this value to 81%. For comparison, the traditional approach of using only the crack size matches on average only 55% of the inspections. We conclude that most of the randomness in manual inspections is due to the circumstances of the inspections. We speculate that most of such randomness will be eliminated by automated structural health monitoring (SHM), which will be an important benefit of SHM.

Nomenclature

a	=	crack size, inches
a_m	=	crack size detected with 50% probability
a_{trs}	=	detection threshold, inches
a'	=	normalized crack size
a'_{trs}	=	normalized threshold
d	=	detection event
d_e	=	experimental detection event
d_s	=	simulated detection event
m	=	detection margin
P_d	=	probability of detection
x	=	agreement margin
β	=	detection parameter in Palmberg equation
Δa_h	=	inspector competence
Δa_l	=	location difficulty
θ	=	penalty function

[1] Graduate research assistant, Mechanical and Aerospace Engineering department, P.O. Box 116250, Gainesville, FL, 32611, AIAA Student Member.
[2] Distinguished Professor, Mechanical and Aerospace Engineering department, P.O. Box 116250, Gainesville, FL, 32611, AIAA Fellow.
[3] Associate Professor, Mechanical and Aerospace Engineering department, P.O. Box 116250, Gainesville, FL, 32611, AIAA member.
[4] Engineer, Optimization Technology, Caterpillar Champaign Simulation Center, Champaign, IL, 61820

I. Introduction

MOST aircraft structural components are designed based on a fail-safe philosophy that uses inspection and maintenance in order to detect damage before it can cause structural failure. In general the inspection can be done either manually or by using onboard equipment. In this paper, the former is referred to manual inspection, while the latter to structural health monitoring (SHM). For manual inspections different techniques have been used, such as radiographic inspection (Lawson *et al.* [4]). Usually, SHM uses actuator-sensor technique (Giurgiutiu *et al.* [1]) to detect damages such as ultrasonic and eddy current techniques (Pohl *et al.* [7]), comparative vacuum monitoring (Stehmeier *et al.* [8]), elastic wave propagation and electromechanical impedance (Giurgiutiu *et al.* [2]).

The effectiveness of various inspection techniques is typically characterized by probability of detection (POD) curves that relate the size of damage to POD (Zheng *et al.* [9]). The information on POD can be used for various purposes, including structural diagnosis and prognosis (Zheng *et al.* [9]). For example, Kale *et al*. [3] 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. 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 process by assuming that 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. As a second step, we model the randomness in human performance by combining the traditional POD curve with our threshold model.

In order to demonstrate the performance of the proposed model, we used the US Air Force study from the 1970s in which 43 panels with different crack sizes are inspected by 62 technicians (2,603 detection events) (Lewis *et al*. [5]). We use optimization techniques to find the location factors associated with 43 panels and the human factors associated with 62 technicians.

More complete characterization of the probability of detection becomes important when SHM is used instead of manual inspections. The location may be much more important than in the case of manual inspection. The objective of the paper is to demonstrate the development of a more complete characterization of POD curves based on the results of large number of inspections. We propose to fit a model to the inspection results that assigns competence scores to inspectors and location difficulty scores to panels by maximizing the agreement with the results of inspections. We demonstrate the approach by using inspection results available in the literature [5].

II. Proposed Statistical Model (Inspector-Location-Size model)

The detection process is often conventionally modeled using a POD 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.* [6]). 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 β is the exponent, and 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; i.e., 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 $a' = 2a_m'$ will be 94.12%. 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, there are many cases that the model is unable to describe the inspection situations. For example, when damage exists in a difficult location 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. The scatter in inspection results can be explained by differences in the competence of inspectors and differences in damage location. The former includes technician's skill, inspection method, and inspection environment (such as fatigue and distractions).

We seek a model that includes the above 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. It is noted that by introducing these two variables, we move the uncertainty in detection process from POD to these two random variables.

As a first step we use a quasi-deterministic model that removes the randomness associated with the two variables. 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 aleatory (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

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

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:

$$a_{trs} = \frac{a_{trs}'}{a_m'} \tag{3}$$

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 variability Δa^h:

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

A positive Δa^l means that the crack is positioned in a more difficult location than average such that it will be detected when it becomes larger than a_m'. A positive Δa^h means the inspection condition is more difficult to detect the damage 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 technicians ([5]). Part of the matrix (13 inspectors, 32 locations) is shown in Table 1.

In order to generate the traditional Palmberg equation from the experiments, we start by calculating the probability of detection for each panel. $P_e^j = N_{det}^j/62,\ j = 1, \dots, 43$, where N_{det}^j is the number of inspectors that detect the j[th] crack. Note that the j[th] 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).

$$\underset{a_m, \beta}{\text{minimize}} \sum_{i=1}^{43} \left| P_d\left(a_j'\right) - P_e^j \right| \tag{5}$$

After minimization, we obtain $a_m' = 0.48\text{cm}$ and $\beta = 1.30$. Figure 1 shows the POD curve as function of crack size, a', along with the 43 probability data 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.

Table 1. Partial matrix of crack detection events(1=detection, 0=non-detection), from [5].

Flaw ID	Flaw size	Technician ID												
		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

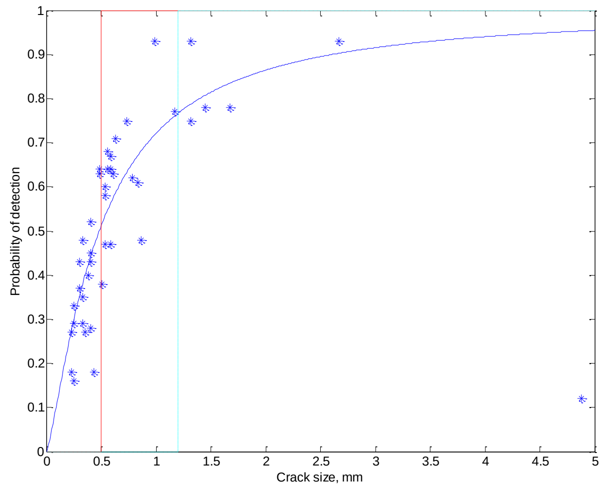

Figure 1. Probability of detection$a_m' = 0.48$cm and $\beta = 1.308$) curves including the traditional Palmberg equation and two typical ILS curves corresponding to the data in [5].

III. Optimization formulations

In order to test if the statistical model describes well the results of the inspections, we seek the 43 $\Delta a_{\mathrm{i}}^{l}$ corresponding to the 43 panels and the 62 $\Delta a_{\mathrm{j}}^{h}$ corresponding to the 62 technicians that fit best the observed inspections success. The threshold increments $\Delta a_{\mathrm{j}}^{h}$ associated with the technicians are easy to estimate because for each technician we have 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 technicians. If a crack in a panel is not found even by the most competent technicians (lowest threshold values), then we can deduce that it is in a difficult location.

The Air Force study [5] reports on 2,603 detections events out of $62 \times 43 = 2{,}666$ possible events. The objective is to find 105 values for the technicians, $\Delta a_{\mathrm{j}}^{h}$, and the panels, $\Delta a_{\mathrm{j}}^{h}$, 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. (1) and (4) is minimized. We consider two different ways of quantifying the differences between the prediction and the inspection results: binary and continuous formulations. The former represents detection and non-detection as binary events. We denote by d_e detected events from actual inspections and by d_s detected events from the model. We define the detection margin m_{ij} and the detection event d_{sij} resulting from the model for the crack in the i^{th} panel and j^{th} inspector as follows:

$$m_{ij} = a_i - (1 + \Delta a_{l_i} + \Delta a_{h_j})$$

$$d_{s_{ij}} = \begin{cases} 0 & \text{if } m_{ij} < 0 \text{ (non detection)} \\ 1 & \text{if } m_{ij} \geq 0 \text{ (detection)} \end{cases} \tag{6}$$

The objective function of the binary formulation is

$$\text{minimize} \sum_{i=1}^{43} \sum_{j=1}^{62} \left| d_{s_{ij}} - d_{e_{ij}} \right| \tag{7}$$

The objective function in Eq. (7) is obviously discontinuous as infinitesimal changes in the threshold values can switch a detected event to a non-detected event and vice versa.

The continuous formulation takes into account the size of the margin m_{ij} in each detection event. When the detection events from inspection and model are not consistent, we penalize the event based the size of the margin.

On the other hand, when the detection events from inspection and model are consistent, we provide a small bonus 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. 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). For this reason, we chose to associate higher penalty, θ, with non-matching margins than with matching margins. This is done by the following optimization problem:

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

where x_{ij} is the agreement margin. It is negative when the results of the statistical model do not match the experimental event and positive when there is a match. The objective function for a single detection event is shown in Figure 2. Although the individual penalty function is monotonic, the objective function may not be unimodal because it is sum of all penalty functions.

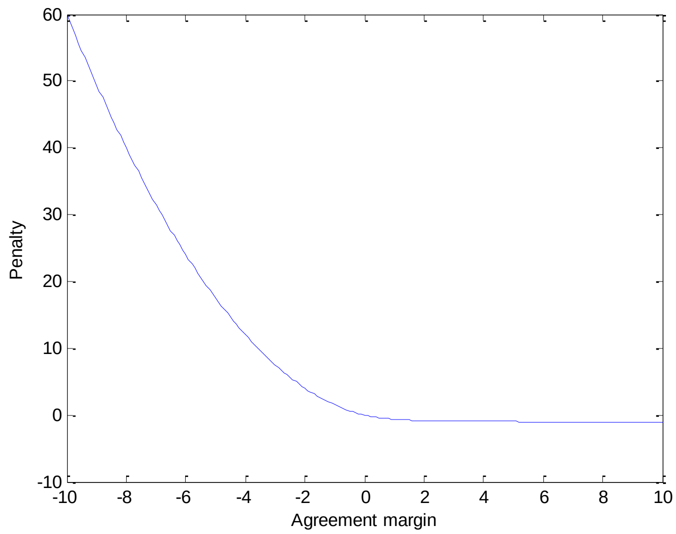

Figure 2. Penalty function

IV. Initial Estimates

Because the first objective function is not continuous and the second 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 $\Delta a_{\mathrm{j}}^{h}$ of each inspector by solving the problem column by column without associating any difficulty with the cracks (i.e., all $\Delta a_{\mathrm{i}}^{l}$ are zero). The next step is to estimate $\Delta a_{\mathrm{i}}^{l}$ using these $\Delta a_{\mathrm{j}}^{h}$. The first step is illustrated graphically in Figure 3. The figure shows the percentage of matches of 43 panels for inspector 14. For this inspector, a qualification of -0.5 is associated with matching 38 of the 43 detection events or 88.37% match. That is, for average location difficulty, this inspector will detect every crack longer than 1.5 times a_h.

Figure 4 shows the continuous objective function for the same inspector. Comparison of Figure 4 and Figure 3 shows a good agreement between the optima.

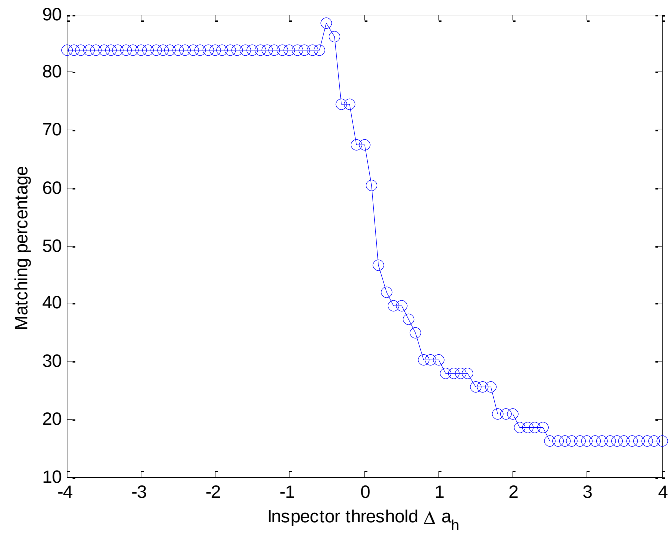

Figure 3. Percentages of matching events for technician 19 as a function of threshold

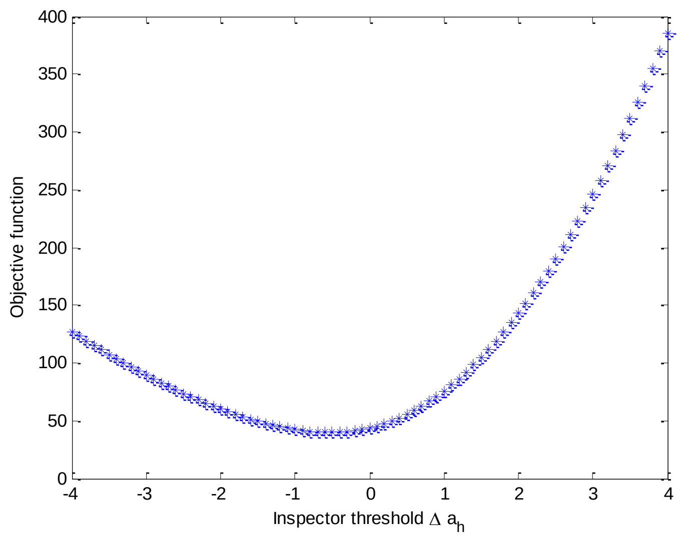

Figure 4. Continuous objective function for one inspection and different matches

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 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.

In order to evaluate the quality of the optimization results, we can 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 i^{th} panel has a crack with size $a_i^{'}$. Using the two-parameter Palmberg equation, the POD of the crack can be calculated by (Palmberg *et al.* [6]) as shown in Eq. (1).

By performing a Monte-Carlo simulation using the Palmberg equation to calculate the POD of the crack sizes used in [5] 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 78% in the proposed model. Thus our very 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 initial point for both optimization problems discussed previously (continuous and discrete). The main concern for both optimization problems is the number of variables (105) but both are constraint free.

The continuous optimization, Eq. (8), is solved using the Matlab function, *fminunc*, which uses the BFGS [11-14] Quasi-Newton method with a mixed quadratic and cubic line search procedure. The optimization problem converges to a solution that yields 77.79% matches. It can be observed that this result is slightly lower that obtained by performing a sequential optimization for both Δa_i^l and Δa_j^h. This can be explained by the fact that the continuous objective function is different from the discrete one.

The discrete optimization problem in Eq. (7) cannot solve using gradient-based optimization algorithms because the objective function is discontinuous. 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 algorithm (Schutte *et al.* [10]). 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 82.75% matches. However, there are many design points that yield the same value of optimum objective function. Figure 5 shows how the objective function varies along a line in 105^{th} dimensional space drawn between two optima, We see that the discrete objective function (non-matching percentage) changes rapidly near the optima and that the two designs yield the same matching percentage.

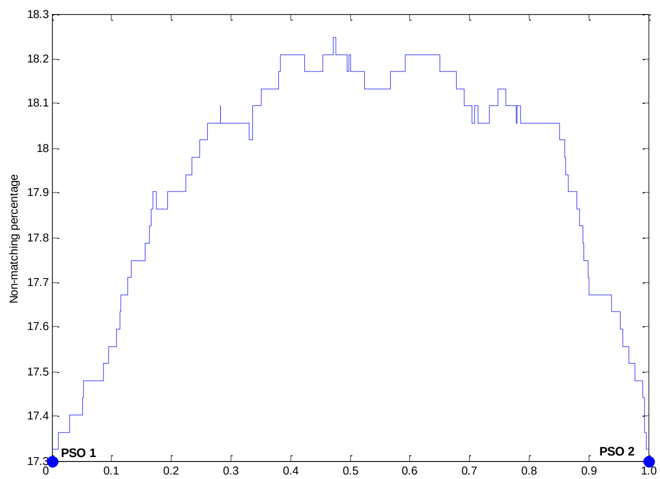

Figure 5. Non-matching percentage variation between two PSO optima

Our ILS_{det} model neglects the variability in human performance. That is, the same technician inspecting the same panel may detect the damage at the first trial but miss it in the second time. Thus we cannot expect 100% match. In fact, it is quite possible that the optimization process over fitted the data, in that if the 2603 inspection events were repeated, the match between the two repetitions would be less than 82.75%. The results of the optimization provide some insight into the magnitude of this randomness in the form of inconsistency on the part of the technicians. That is, even with the optimum assignments of difficulty to the 43 panels, the experimental results still show technicians identifying a difficult crack while missing an easier one. Thus the present model can be improved by modeling the randomness responsible for this inconsistency.

VI. Conclusions

We developed a simple model that accounts for inspector competence and location difficulty in order to explain the randomness in detecting cracks by manual inspection. By fitting 105 parameters to 2,602 experiments we were able to match more than 82% of the detection events compared to 55% achieved by the commonly used model which is based only on crack size. The procedure revealed that most of the randomness in the detection is due to inspector competence rather than due to the location of the crack. 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.

Additional work is being done to model the randomness in human performance. This will account for the fact that even with the optimum results some detection events are inconsistent in that a technician identifies a difficult crack while missing an easier one.

Acknowledgments

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

References

[1] Giurgiutiu V. and Cuc A., "Embedded non-destructive evaluation for structural health monitoring, damage detection, and failure prevention" Schock Vibr. Dig. **37**, 83105 (2005).

[2] Giurgiutiu V, Redmond J, Roach D, Rackow K. "Active sensors for health monitoring of aging aerospace structures". Proceedings of the SPIE Conference on Smart Structures and Integrated Systems, New Port Beach, 2000, 3985:294–305.

[3] Kale A. and Haftka R. , "Tradeoff of Structural Weight and Inspection Cost in Reliability Based Optimization Using Multiple Inspection Types", AIAA-2004-4404, 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Albany, New York, Aug. 30-1, 2004.

[4] Lawson, S.W., Parker G.A., 1994. "Intelligent segmentation of industrial radiographs using neural networks", Proceedings of SPIE, Vol. 2347, pp. 245–255.

[5] Lewis, W. G., Dodd, B. D., Sproat, W. H., and Hamilton, J. M., "Reliability of Nondestructive Inspections - Final Report," SAALC/ MME 76-6-38-1, Dec. 1978.

[6] Palmberg, B., Blom, A. F., Eggwertz, S., "Probabilistic Damage Tolerance Analysis of Aircraft Structures". Probabilistic Fracture Mechanics and Reliability, 1987.

[7] Pohl R., Erhard A., Montag H.J., Thomas H.M. and Wüstenberg H., NDT techniques for railroad wheel and gauge corner inspection, *NDT & E Int* **37** (2004) (2), pp. 89–94.

[8] Stehmeier H., Speckmann H., "Comparative Vacuum Monitoring (CVM) Monitoring of fatigue cracking in aircraft structures", 2nd European Workshop on Structural Health Monitoring, 2004 - atlas-conferences.com.

[9] Zheng R. and Ellingwood B., "Role of non-destructive evaluation in time-dependent reliability analysis". Structural Safety **20** 4 (1998), pp. 325–339.

[10] Schutte, J.F., Reinbolt, J.A., Fregly, B.J., Haftka, R.T., and George, A.D. (2004) "Parallel global optimization with the particle swarm algorithm". *International Journal of Numerical Methods in Engineering* **61**, 2296-2315.

[11] Broyden, C.G., "The Convergence of a Class of Double-Rank Minimization Algorithms," *J. Inst. Math. Applic*., Vol. 6, pp. 76-90, 1970.

[12] Fletcher, R.,"A New Approach to Variable Metric Algorithms," *Computer Journal*, Vol. 13, pp. 317-322, 1970.

[13] Goldfarb, D., "A Family of Variable Metric Updates Derived by Variational Means," *Mathematics of Computing*, Vol. 24, pp. 23-26, 1970.

[14] [4] Shanno, D.F., "Conditioning of Quasi-Newton Methods for Function Minimization," *Mathematics of Computing*, Vol. 24, pp. 647-656, 1970.