

# A Statistical Model for Estimating Probability of Crack Detection

Alexandra Coppe, Raphael T. Haftka, Nam-Ho Kim, and Christian Bes, *Member, IEEE*

**Abstract**— 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’s competence), as well as difficulties associated with the location and type of damages. To demonstrate the effects of these other factors, we present a simple model from the assumption 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 below that size will be missed. The proposed model adjusts the threshold crack size according to the difficulty associated with the crack location and the competence of inspectors. The model is then used to fit 2,603 detection events reported for 43 panels inspected by 62 technicians in an Air Force study. The threshold increments by location and inspector are obtained by maximizing the matching percentage in detection events between the model and the experiment. We first use 62 inspector thresholds only and find the best matching percentage of 78%. It is further increased to 81% when both inspector and location thresholds are considered. For comparison, the matching percentage using crack size alone is only 55%. We then add randomness to the process in order to include inconsistency on the part of inspectors. Replicating the observed inconsistency reduces the matching to about 72%. We conclude that most of the randomness in manual inspections is due to the circumstances of the inspections. We speculate that much of this randomness will be eliminated by automated structural health monitoring (SHM), which will be an important benefit of SHM.

**Index Terms**—Detection, Inspection, Health monitoring, Probability of detection, Optimization

## 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 an important task, but 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.

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 threshold crack size. However, there exists uncertainty in detecting a crack even if the same technician inspects the same crack. In order to model this randomness, we further improved the model using a traditional POD curve.

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

Manuscript received May 23, 2008. This work was supported by the U.S. Air Force under Grant FA9550-07-1- and by the NASA under Grant NNX08AC334.

Alexandra Coppe is Graduate Research Assistant with University of Florida, Gainesville, FL 32611 USA (e-mail: [alex.coppe@ufl.edu](mailto:alex.coppe@ufl.edu)).

Raphael T. Haftka, is Distinguished Professor with University of Florida, Gainesville, FL 32611 USA (e-mail: [haftka@ufl.edu](mailto:haftka@ufl.edu)).

Nam-Ho Kim is Associate Professor with University of Florida, Gainesville, FL 32611 USA (e-mail: [nkim@ufl.edu](mailto:nkim@ufl.edu)).

Christian Bes is Professor with Universite Paul Sabatier, Toulouse, France (email: [cbes@cict.fr](mailto:cbes@cict.fr)).

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.

## 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

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

where  $\beta$  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  $\beta$  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

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

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:

$$a_{trs} = \frac{a'_{trs}}{a'_m} \quad (3)$$

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 a 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  $\Delta a^l$  and that from the inspection variability  $\Delta a^h$ :

$$a_{trs} = 1 + \Delta a^l + \Delta a^h \quad (4)$$

A positive  $\Delta a^l$  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  $\Delta a^h$  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<sub>det</sub>' (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.

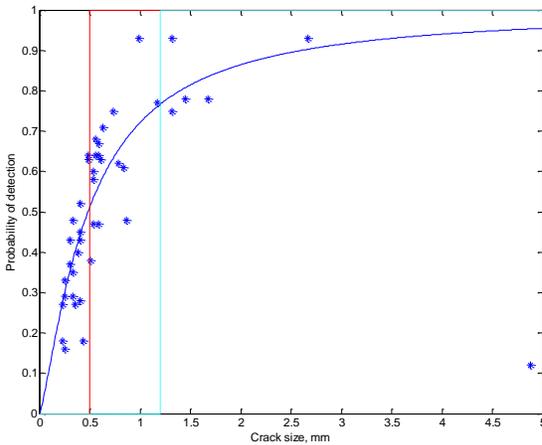
**Table 1 – Partial matrix of crack detection (1=detection, 0=non-detection), from [5]**

Flaw ID	Flaw size	Technician ID											
		0201	0202	0204	0207	0208	03E1	03E2	03E3	03E4	03E5	03E7	03E9
77a	0.09	0	0	1	0	0	0	0	1	0	0	0	0
122	0.09	1	0	0	0	0	0	1	1	0	0	0	0
132	0.10	1	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	1
75	0.10	1	0	1	0	0	0	0	0	0	1	0	0
76b	0.12	1	0	1	0	0	0	0	1	1	1	1	0
80	0.12	1	1	1	0	0	0	0	1	0	0	0	0
77b	0.13	1	1	1	0	0	0	0	1	1	1	0	1
79d	0.13	1	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	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

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  $j^{\text{th}}$  crack. Note that the  $j^{\text{th}}$  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).

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



**Figure 1 - Probability of detection ( $a'_m = 0.48\text{cm}$ ,  $\beta = 1.30$ ) curves using the Palmberg equation and two typical ILS curves corresponding to the data in [1]**

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.

### III. LOSS FUNCTION FOR THE DETERMINISTIC ILS MODEL

For the Palmberg equation, we have only two parameters to fit 43 data. For the deterministic ILS model, we can fit 43 location  $\Delta a_i^l$  ( $i = 1, \dots, 43$ ) corresponding to 43 panels and 62  $\Delta a_j^h$  ( $j = 1, \dots, 62$ ) corresponding to 62 technicians that fit best the observed inspections events (detected or not). The Air Force study [5] reports on 2,603 detections events out of  $62 \times 43 = 2,666$  possible events. If there is no uncertainty in the detection events and all cracks are located in the position with same difficulty level, it is possible to estimate the threshold increments  $\Delta a_j^h$  associated with the technicians by arranging 43 panels in the order of crack size and finding the threshold value between the largest missed crack and the smallest detected crack. However, as can be found in Table 1, technician 0201 found the crack 122 (size = 0.9 inch), but missed crack 133 (size = 1.4 inch). This can be caused by either these two cracks being located in positions with different levels of difficulty or the inspection condition being changed. On the other hand, since each panel only has a single crack size, the estimation of location difficulty must rely on the scores of the technicians. If the crack in a panel is not found even by the most competent technicians (smallest threshold value), then we can deduce that it is in a difficult location. However, this may fail to provide deterministic value of location difficulty.

In the proposed study, we use optimization to find the threshold increments  $\Delta a_i^l$  and  $\Delta a_j^h$  by minimizing the difference between the detection events from experiment and that from the model. We present two different objective (or loss) functions as follows.

The loss function,  $f_{binary}$ , is defined as a sum of the differences between the two events. That is for each detection event, we can have agreement (zero loss) or disagreement (loss of one event). That is,

$$f_{binary} = \sum_{i=1}^{43} \sum_{j=1}^{62} |d_{ij}^s(\Delta a_i^l, \Delta a_j^h) - d_{ij}^e| \quad (6)$$

where  $d_{ij}^e$  is the detection event (0 or 1 in Table 1) for the  $i^{\text{th}}$  panel and  $j^{\text{th}}$  inspector, and  $d_{ij}^s$  is the predicted inspection result from the proposed model, defined by

$$d_{ij}^s = \begin{cases} 0 & \text{if } m_{ij} < 0 \text{ (not detected)} \\ 1 & \text{if } m_{ij} \geq 0 \text{ (detected)} \end{cases} \quad (7)$$

In Eq. (7), the detection margin  $m_{ij}$  is defined for the crack in the  $i^{\text{th}}$  panel and  $j^{\text{th}}$  inspector as

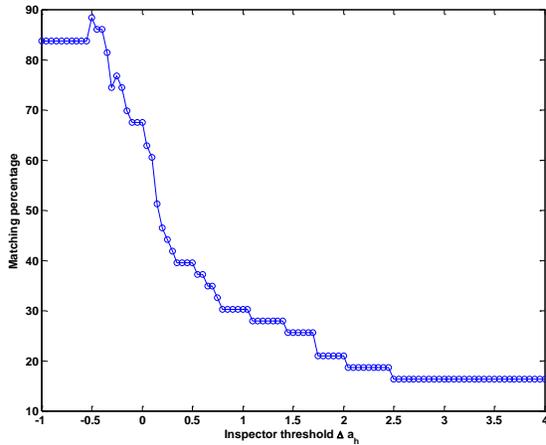
$$m_{ij} = a_i - (1 + \Delta a_i^l + \Delta a_j^h) \quad (8)$$

The loss function in Eq. (6) is obviously discontinuous because it can only have integer values. However, the design

variables are continuous. It is different from the conventional type of discrete optimization problems in which the objective function is continuous and design variables are integer or discrete.

#### IV. FITTING THE DETERMINISTIC ILS MODEL

Figure 2 shows the percentages of matching events for technician 19 as a function of  $\Delta a_{19}^h$ . The inspector detected 36 cracks and missed 7 cracks. The minimum of Eq. (6) is obtained with  $\Delta a_{19}^h = -0.5$  in which 26 cracks are detected. A total of 38 detection events (88.37%) matched out of possible 43. That is, for average location difficulty, this inspector will detect every crack longer than  $1 + \Delta a_{19}^h = 0.5$  times  $a_m^h$ . It is noted that the matching percentage decreases quickly as the inspector threshold increases, which means that the particular inspector performs better than average.



**Figure 2 - Percentages of matching events for technician 19 as a function of threshold**

By solving the optimization problem individually for each inspector, we find 62  $\Delta a_j^h$ , which represent the competence of inspectors. Overall these 62  $\Delta a_j^h$  match 78.3% of the detection events. Starting from the optimal  $\Delta a_j^h$  as initial estimates, the optimization problem in Eq. (6) is solved by varying  $\Delta a_i^l$ , to obtain a matching percentage of 80.6%, 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  $i^{\text{th}}$  panel has a crack with size  $a_i^l$ . Using the two-parameter Palmberg equation, the POD of the crack can be calculated by equation (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 have a matching percentage between the simulated data and the experimental ones of 55.5% (with a standard deviation of 0.96%) which is much lower than 80.61% in the proposed model. Thus our very simple model accounts much better for the actual inspection results than a model that takes into account only the crack size.

#### V. RANDOM ILS MODEL

The results of the inspections reported in [1] have some randomness in them, since they are dictated by the circumstances of the inspections (e.g., human fatigue, distraction, temperature, etc) as well as by objective difficulty (location) and inspector competence. This means that if the process were repeated, we would expect somewhat different set of inspection effects. An indication of the magnitude of the randomness can be obtained by inconsistency of an inspector. That is, if an inspector finds a difficult crack and misses an easy one, we have evidence of randomness. In order to test for this randomness, we first define a corrected crack size that accounts for the location difficulty

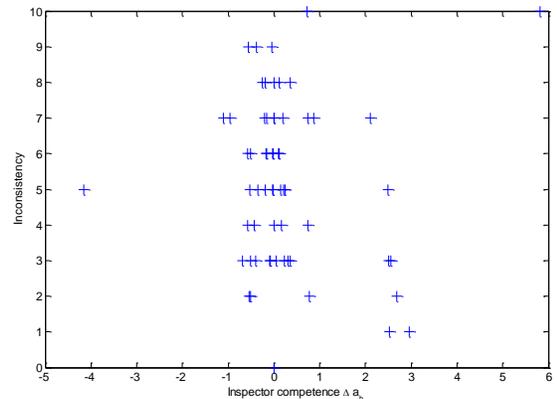
$$a_{corr} = a - \Delta a^l \quad (9)$$

Next we arrange the plates in increasing order of corrected crack size and calculate the number of times the detection and non-detection events alternate

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

With  $\langle d_{i-1,j} - d_{i,j} \rangle = \max(0, d_{i-1,j} - d_{i,j}, 0)$

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 Figure 3. Note that  $\Delta a_h$  is interpreted, for an average location difficulty ( $\Delta a_l = 0$ ), as follows: a value of 1 means that the inspector's threshold is twice the average, while a -1 means that the inspector can identify any size of crack of average difficulty.



**Figure 3 - Inspectors inconsistency with respect to their competence**

To account for inconsistency, the threshold in the deterministic model is used in the traditional POD curve as the size  $a_m$  corresponding to 50% POD, obtaining the new  $ILS_{random, a_m}$ . That is, the Palmberg equation is rewritten as

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

A very large value of  $\beta$  in Eq. (11) would correspond to the deterministic ILS model. In ILS<sub>random</sub> model, the exponent  $\beta$  is selected such that the inconsistency from the model matches with that of the experimental data. In order 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 in order to reach the same level of inconsistency. A summary of the results is given in Table 2.

**Table 2 - Matching percentage and average inconsistency for the 3 models presented in this paper. Values in parenthesis are standard deviations**

Method	Palmberg	ILS <sub>random</sub>	ILS <sub>det</sub>
Matching percentage	55.5 (0.96)	72.07(0.68)	80.61
Average inconsistency	10.1 (0.2)	5.2 (0.1)	0

Unlike the ILS<sub>det</sub> model, the ILS<sub>random</sub> model will have different results if the same process is repeated. The values in the parenthesis in the third column of Table 2 show the standard deviation of the matching percentage. In order 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%).

This shows that the match with the experimental result is a reasonable manifestation of the inconsistency of the inspectors.

## VI. CONCLUSIONS AND FUTURE WORK

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 80.61% of the detection events compared to 55% achieved by the commonly used model based on crack size alone.

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.

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.

## 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.