SAND2015-4452C

2015 International Workshop on Structural Health Monitoring

Establishing the Reliability of SHM Systems Through the Extrapolation of NDI Probability of Detection Principles

Dennis Roach
Tom Rice
Stephen Neidigk
Sandia National Laboratories
FAA Airworthiness Assurance Center

Dave Piotrowski
Delta Air Lines

John Linn
Boeing

ABSTRACT

Extensive Structural Health Monitoring (SHM) studies have highlighted the ability of various sensors to detect common flaws found in composite and metal structures with sensitivities that meet or exceed current damage detection requirements. Reliable SHM systems can automatically process data, assess structural condition, and signal the need for human intervention. While ad-hoc efforts to introduce SHM into routine aircraft maintenance practices are valuable in leading the way for more widespread SHM use, there is a significant need for formal SHM certification efforts to exercise and define the process of producing routine use of SHM solutions. SHM certification must address the full spectrum of issues ranging from design to performance and deployment to continued airworthiness. Currently, there are no guidelines for SHM system designers or agreed-upon procedures for quantifying the performance of SHM systems. The FAA Airworthiness Assurance Center (AANC) at Sandia Labs, in conjunction with Boeing, Delta Air Lines, Structural Monitoring Systems and Anodyne Electronic Manufacturing, is conducting a study to develop and carry out a certification process for SHM. By conducting a focused assessment of a particular aircraft application, all aspects of SHM integration are being addressed. While it is important to recognize the unique validation and verification tasks that arise from distinct differences between SHM and nondestructive inspection (NDI) deployment and flaw detection, it should be recognized that some portions of the methodology needed to determine NDI performance can be adapted to the validation of SHM systems. In this study, statistical methods were applied to laboratory and flight test data to derive Probability of Detection (POD) values for SHM sensors in a fashion that agrees with current NDI requirements.

Dennis Roach, Sandia National Labs, PO BOX 5800, Albuquerque, NM 87185
Tom Rice, Sandia National Labs, PO BOX 5800, Albuquerque, NM 87185
Stephen Neidigk, Sandia National Labs, PO BOX 5800, Albuquerque, NM 87185
David Piotrowski, Delta TechOps, 1775 M H Jackson Service Road, Atlanta, GA 30354
John Linn, Boeing, 2600 Westminster Boulevard, Seal Beach CA, 90740

INTRODUCTION

Multi-site fatigue damage and hidden cracks in hard-to-reach locations are among the major flaws encountered in today's extensive array of aging structures and mechanical assemblies. The costs associated with the increasing maintenance and surveillance needs of aging structures are rising. The application of Structural Health Monitoring (SHM) systems using distributed sensor networks can reduce these costs by facilitating rapid and global assessments of structural integrity. These systems also allow for condition-based maintenance practices to be substituted for the current time- or cycle-based maintenance approach thus optimizing maintenance labor. Other advantages of on-board distributed sensor systems are that they can eliminate costly, and potentially damaging, disassembly, improve sensitivity by producing optimum placement of sensors with minimized human factors concerns in deployment and decrease maintenance costs by eliminating more time-consuming manual inspections. Through the use of in-situ sensors, it is possible to quickly, routinely, and remotely monitor the integrity of a structure in service [1]. This requires the use of reliable structural health monitoring systems that can automatically process data, assess structural condition, and signal the need for specific maintenance actions.

Current aircraft maintenance operations require personnel entry into normally-inaccessible or hazardous areas to perform mandated, nondestructive inspections. To gain access for these inspections, structure must be removed, sealant must be removed and restored, fuel cells must be vented to a safe condition, or other disassembly processes must be completed. These processes are not only time consuming but they provide the opportunity to induce damage to the structure. The use of in-situ sensors for monitoring the condition of aircraft structure, coupled with remote interrogation, can be employed to overcome a myriad of inspection impediments stemming from accessibility limitations, complex geometries, and the location and depth of hidden damage. Furthermore, prevention of unexpected flaw growth and structural failure could be improved if on-board health monitoring systems are used to more regularly assess structural integrity [2, 3]. The ease of monitoring an entire network of distributed sensors means that structural health assessments can occur more often, allowing operators to be even more vigilant with respect to flaw onset.

Comparative Vacuum Monitoring (CVM) is a simple pneumatic sensor technology developed to detect the onset of cracks. CVM sensors are permanently installed to monitor critical regions of a structure. The CVM sensor is based on the principle that a steady state vacuum, maintained within a small volume, is sensitive to any leakage [4]. A crack in the material beneath the sensor will allow leakage resulting in detection via a rise in the monitored pressure. Figure 1 shows top-view and side-view schematics of the self-adhesive, elastomeric sensors with fine channels etched on the adhesive face along with a sensor being tested in a lap joint panel. When the sensors are adhered to the structure under test, the fine channels and the structure itself form a manifold of galleries alternately at low vacuum and atmospheric pressure. Vacuum monitoring is applied to small galleries that are placed adjacent to the set of galleries maintained at atmospheric pressure. If a flaw is not present, the low vacuum remains stable at the base value. If a flaw develops, air will flow from the atmospheric galleries through the flaw to the vacuum galleries. When a crack develops, it forms a leakage path between the atmospheric and vacuum galleries, producing a measurable change in the vacuum level. This change is detected by the CVM monitoring system shown in Figure 2. It is important to note that the sensor detects surface breaking cracks once they interact with the vacuum galleries.

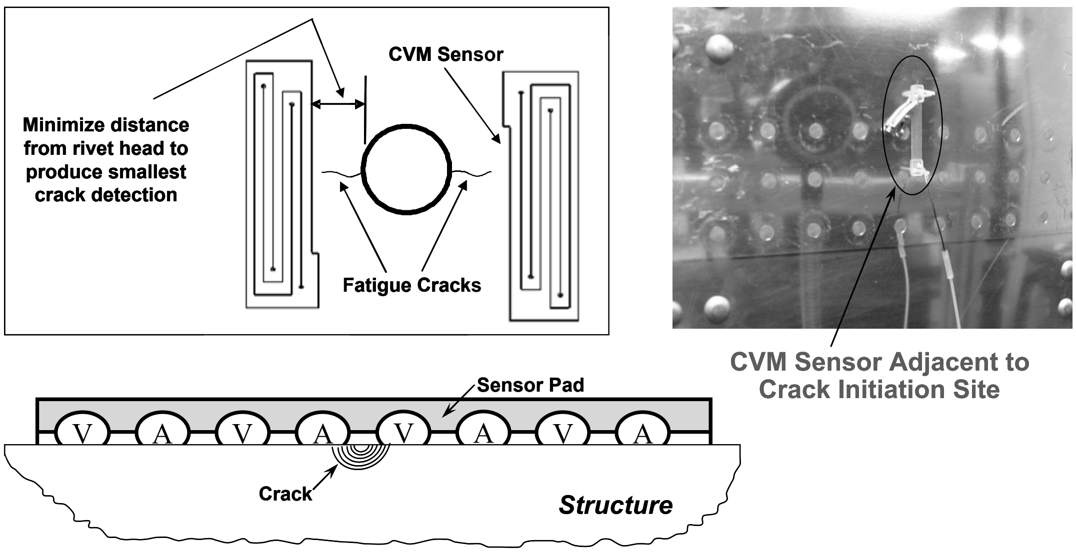


Figure 1: Schematics Depicting Operation of CVM Sensor and Polymer Sensor Mounted on the Outer Surface of a Riveted Lap Joint

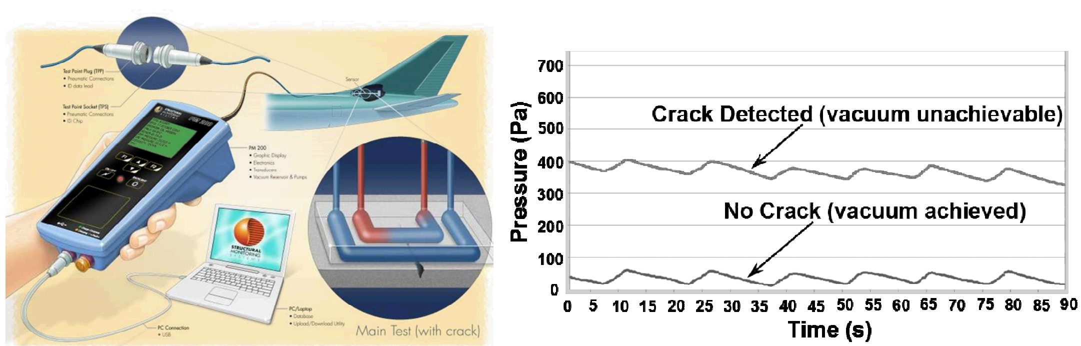


Figure 2: Crack Detection Monitoring with CVM System and Pressure Response Used to Indicate the Presence of a Crack

PERFORMANCE TESTING OF CVM SENSORS

The goal of this project is to produce sufficient data and to conduct the proper interface with regulatory agencies to certify CVM sensor technology for specific aircraft applications. Towards that end, probability of flaw detection assessments were coupled with on-aircraft flight tests to study the performance, deployment, and long-term operation of CVM sensors on aircraft. Statistical methods using one-sided tolerance intervals were employed to derive Probability of Detection (POD) levels for SHM sensors. The result is a series of flaw detection curves that can be used to propose CVM sensors for aircraft crack detection. The test specimens were wing box fittings from the Boeing 737 which was the chosen CVM application from Delta's fleet. Figure 3 shows the details of the wing box fitting application and installation of CVM sensors for the flight test program. Fatigue tests were completed on the wing box fittings using flight load spectrums (see Fig. 4) while the vacuum pressures within the various sensor galleries were simultaneously recorded. A fatigue crack was propagated until it engaged one of the vacuum galleries such that crack detection was achieved and the sensor indicated the presence of a crack by its inability to maintain a vacuum.

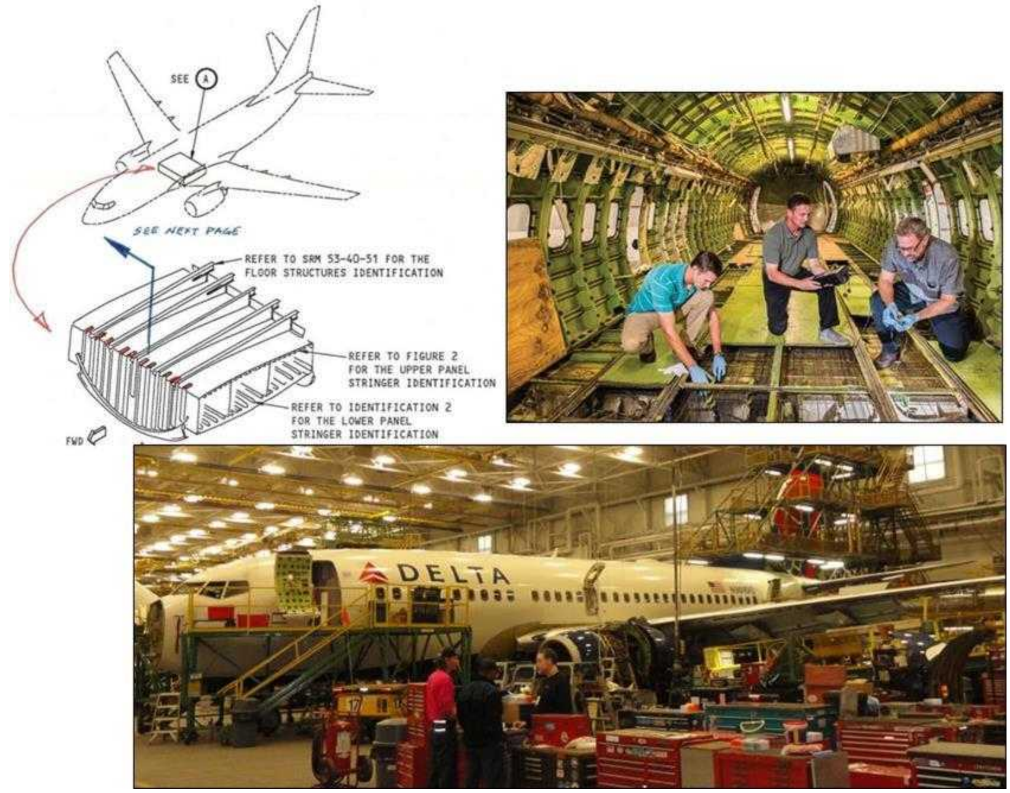


Figure 3: Wing Box Fitting Application and Installation of CVM Sensors on Delta Air Lines Aircraft for Flight Tests

In order to properly consider the effects of crack closure in an unloaded condition (i.e. during sensor monitoring), a crack was deemed to be detected when a permanent alarm was produced and the CVM sensor did not maintain a vacuum even if the fatigue stress was reduced to zero. Figure 5 shows the fatigue test set-up used to grow cracks and a close-up photo of a fatigue crack as it engages the first vacuum gallery of a CVM sensor. Crack detection lengths ranged from 0.145" to 0.245" in length for the wing box fitting application. The crack detection lengths correspond to permanent alarm levels for cracks engaging CVM sensors and the structure in an unloaded condition.

In addition to the lab-based certification tests, a series of 68 sensors were mounted on wing box fittings in seven different B-737 aircraft in the Delta Air Lines fleet. All sensors have been monitored every 90 days for the past 15 months, producing over 400 sensor response data points. These flight tests demonstrated the successful, long-term operation of the CVM sensors in actual operating environments. This environmental durability study compliments the laboratory flaw detection testing described below as part of an overall CVM certification effort.

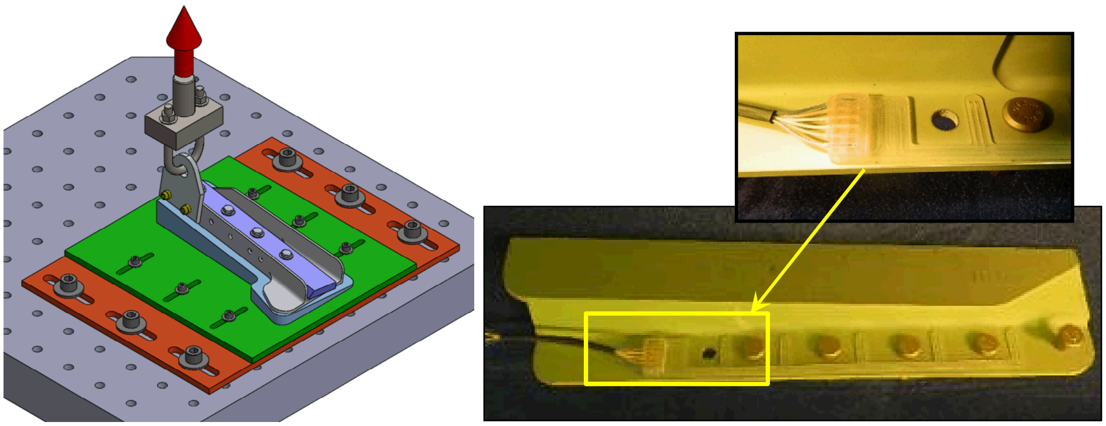


Figure 4: B-737 Wing Box Fitting with CVM Sensor Installed and
Test Set-Up to Produce Fatigue Crack Growth Along Rivet Row

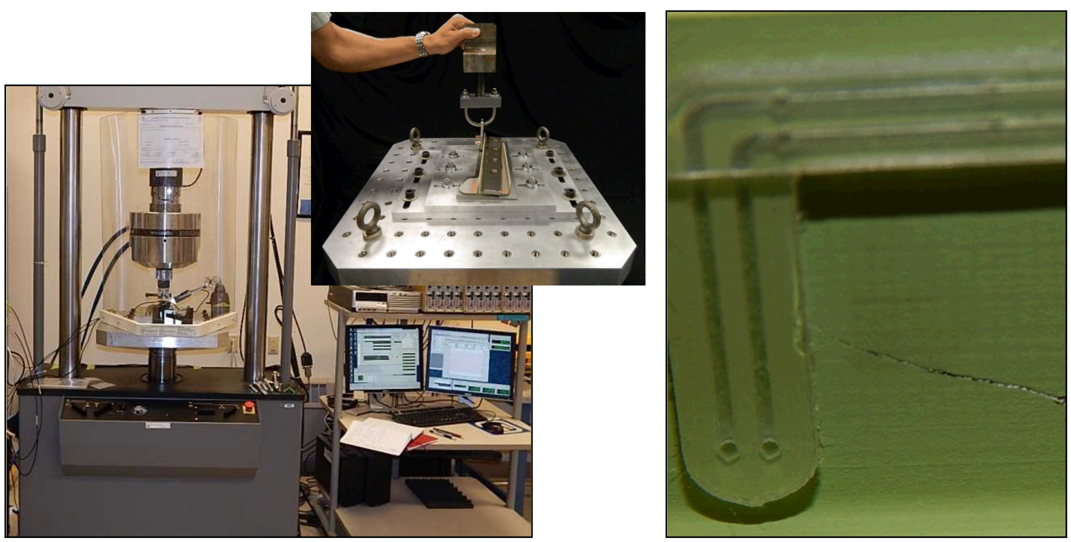


Figure 5: Overall Set-Up for Monitoring Crack Growth with CVM Sensor System and
Close-Up Showing Fatigue Crack Crossing into CVM Sensor

DATA ANALYSIS USING ONE-SIDED TOLERANCE INTERVALS

Some portions of the methodology needed to quantify NDI performance can be adapted to the validation of SHM systems. However, it is important to recognize the unique validation and verification tasks that arise from distinct differences between SHM and NDI deployment and flaw detection. SHM reliability calculations will depend greatly on the complexity of the structure and geometry of the flaw profile. For example, corrosion damage has a widely-varying flaw shape, both in the surface dimensions and in the changing depth. Contrast this with a fatigue crack that grows in a known propagation path such that the damage scenario can be described in a single parameter: crack length. In this latter case, the simplicity of such a one-dimensional entity allows for a more direct calculation of the reliability of the SHM system detecting such damage. The Probability of Detection for a fixed sensor detecting a crack which is propagating in a known direction in the vicinity of the sensor can be determined using the One-Sided Tolerance Interval (OSTI) approach. The OSTI estimates the upper bound which should contain a certain percentage of all measurements in the population with a specified confidence. Since it is based on a sample of the entire population (n data points), the confidence is less than 100%. Thus, the OSTI is greatly affected by two proportions: 1) the percent coverage which is the percent of the population that falls within the specified range (normally chosen as 90%), and 2) the degree of confidence desired (normally chosen as 95%).

Because of physical, time or cost constraints, it is often impractical to inspect an entire population. Instead, a small sample of the total population is tested and the data is used to gauge how well the entire population conforms to specifications. In traditional statistical process control, a significant number of data points are required in order to get a reasonably accurate estimate of process capability. This is because capability is usually calculated to cover a fixed multiple standard deviations. But this percentage only holds true for larger sample sizes; that is, greater than 50. As the sample size decreases, there is greater uncertainty in knowing the true location of the mean and the true magnitude of the population variance. Therefore, the estimate of the range of values encompassing a given percentage of the population must necessarily increase to compensate. In order to maintain a reasonably accurate estimate of the capability of a process for smaller sample sizes, it is necessary to adjust the number of multiple sample standard deviations used to define the region covering the desired proportion of the population distribution with a given confidence. An OSTI can be used for this purpose.

The data captured is that of the flaw length at the time for which the CVM provided sustainable detection. With these assumptions there exists a distribution on the flaw lengths at which detection is first made. In this context, the probability of detection for a given flaw length is just the proportion of the flaws that have a detectable length less than that given length. That is, the reliability analysis becomes one of characterizing the distribution of flaw lengths and the cumulative distribution function is analogous to a Probability of Detection (POD) curve. Assuming that the distribution of flaws is such that the logarithm of the lengths has a Gaussian distribution, it is possible to calculate a one sided tolerance bound for various percentile flaw sizes. To calculate a one sided tolerance bound, it is necessary to find factors $K_{n,\gamma,\alpha}$ to determine the confidence γ such that at least a proportion (α) of the distribution will be less than $X + (K_{n,\gamma,\alpha})S$ where X and S are estimators of the mean and the standard deviation computed from a random sample of size *n*. There may also be situations where the process capability is measured relative to a single-sided limit. These situations arise when a product characteristic need only meet a minimum specification limit or remain below a maximum specification limit. In this case, the desired POD value is the maximum crack length associated with the 90% POD level so the one-sided tolerance interval is used. The K factor for an OSTI can be obtained from standard statistical tables.

From this reliability analysis a cumulative distribution function is produced to provide the maximum likelihood estimation (POD). This stems from the one-sided tolerance bound for the flaw of interest using the equation:

$$T_{\text{POD}(90,\,95)} = X + (K_{n,\gamma,\alpha})(S) \tag{1}$$

Where,

- T = Tolerance interval for crack length corresponding to 90% POD with a 95% confidence
- X = Mean of detection lengths
- K = Probability factor (~ sample size and confidence level desired)
- S = Standard deviation of detection lengths
- n = Sample size
- α = Detection level
- ɣ = Confidence level

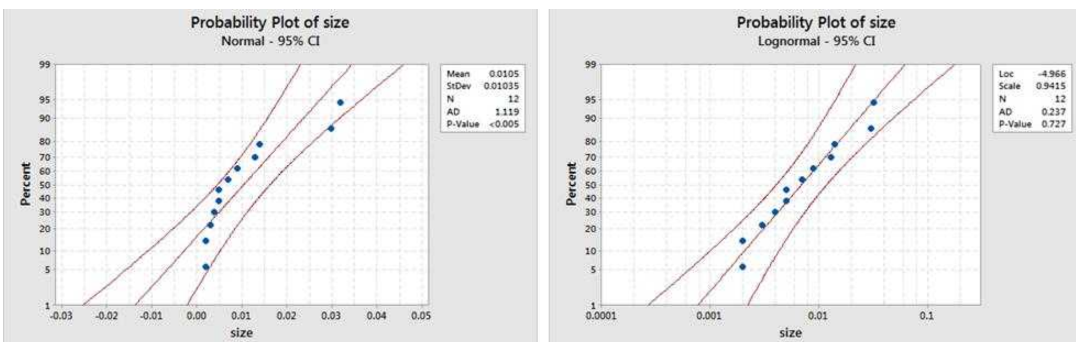
The formula in equation (1) is set-up to produce the upper bound for the tolerance interval which represents the actual POD value.

In order to ensure the validity of a log-normal, or Gaussian, distribution on the flaw lengths, the data should plot linearly on a semi-log scale and the data should be clustered near the 50th percentile. The assumption of normality can also be tested by applying the Anderson-Darling test [5]. The Anderson-Darling test yields a P-value that can be compared to the chosen significance level to determine whether or not the assumption of normality should be rejected. The significance level, *a*, is chosen to be 0.05. Any value of P less than *a* = 0.5 indicates that there is sufficient evidence to reject the assumption of normality. A normal probability plot was created using Minitab® statistical software. Figure 6 shows two plots of sample CVM crack detection data which indicates that a log-normal distribution is a correct assumption. In addition, the Anderson-Darling test returns the required value of P > 0.05.

With the same parameters described above, the maximum likelihood estimate describing the upper bound or optimal performance on the Probability of Detection for the OSTI approach can be calculated as:

$$\textbf{POD}(\text{Max Likelihood Est}) = \frac{1}{xS\sqrt{2\pi}} EXP\left(\frac{-(\ln(x) - X)^2}{2S^2}\right) \tag{2}$$

Data acquired from CVM fatigue tests were used to calculate the 90% POD level for CVM crack detection on 0.1" thick 2024-T3 aluminum structure subjected to tension-tension fatigue loading. Table I summarizes the crack detection data and shows the calculated quantities for equation (1) in the log transform. Twelve data points (bare surface) and ten data points (primer surface) were used in lieu of the 51 or greater that are required in conventional POD calculations. Due to the limited number of data points, the reliability calculations induce a penalty by increasing the magnitude of the K (probability) factor. As a result, while most of the crack detection levels were less than 0.015", the overall POD value (95% confidence level) for CVM crack detection was calculated from equation (1) as 0.023". The K values correspond to the desired ɣ (confidence level) of 95%. This POD curve, representing the 95% confidence level, is plotted in Figure 7. The maximum likelihood estimated POD function, representing the optimum performance for CVM crack detection, was calculated from equation (2) and is plotted alongside the 95% confidence bound. As the number of data points increases, the K value will decrease and the POD numbers could also decrease. In this particular instance, it was desired to achieve crack detection before the crack reached 0.1" in length so this goal was achieved. In over 150 fatigue tests conducted using CVM sensors there were no false calls produced by the sensors in any of the tests.



Normal Distribution **Lognormal Distribution**

Figure 6: Plots of CVM Crack Detection Data where Linear Plots Show that the Data Does Not Follow a Normal Distribution (linear scale) but Does Adhere to a Log-Normal Distribution (semi-log scale)

TABLE I: CVM CRACK DETECTION VALUES FROM 0.1" THICK ALUMINUM PLATE

CVM Crack Detection Data (0.040" th)

Bare Metal		Over Primer	
Flaw size (inch)	Log (flaw size)	Flaw size (inch)	Log (flaw size)
0.003	-2.52	0.002	-2.70
0.007	-2.15	0.007	-2.15
0.002	-2.70	0.010	-2.00
0.030	-1.52	0.009	-2.05
0.009	-2.05	0.004	-2.40
0.005	-2.30	0.006	-2.22
0.004	-2.40	0.010	-2.00
0.002	-2.70	0.009	-2.05
0.014	-1.85	0.011	-1.96
0.005	-2.30	0.007	-2.15
0.013	-1.89		
0.032	-1.49		

Statistic Estimates on Log Scale

Statistic	Over Bare metal	Over Primer
Mean	-2.1566	-2.1679
Stnd deviation	0.40889	0.22809

POD Detection Levels
(ɣ = 95%, n = 12 for bare, n=10 for primer)

Detection level $(1-\alpha)$	$K_{n,0.95,\alpha}$		$\bar{X}+K_{n,0.95,\alpha}\cdot S$ (log scale)		Flaw size in inches	
	bare	primer	bare	primer	bare	primer
0.75	1.366	1.465	-1.598	-1.834	0.025	0.015
0.90	2.210	2.355	-1.253	-1.631	0.056	0.023
0.95	2.736	2.911	-1.038	-1.504	0.092	0.031
0.99	3.747	3.981	-0.624	-1.260	0.237	0.055
0.999	4.900	5.203	-0.153	-0.981	0.703	0.104

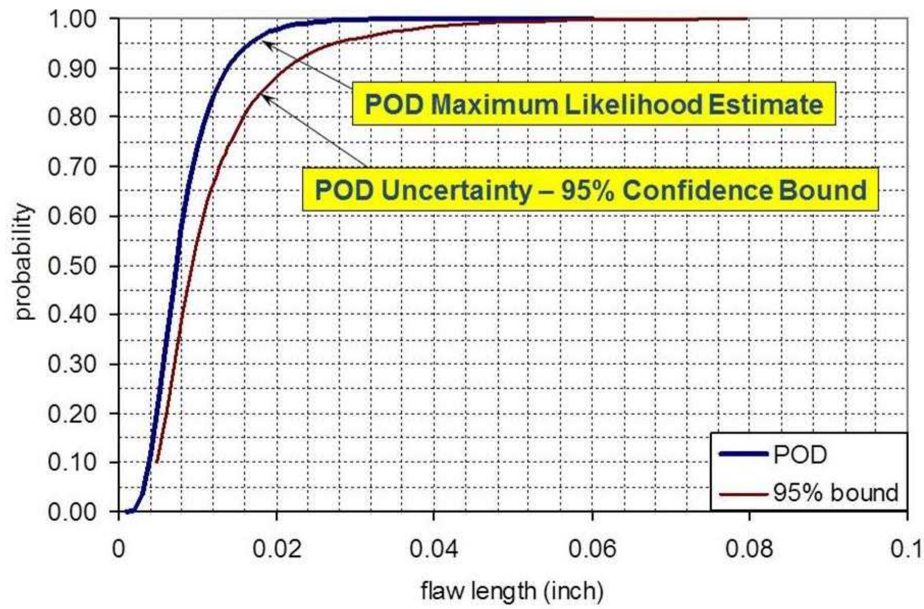


Figure 7: Probability of Crack Detection Curves Showing Detectable Flaw Lengths for CVM Sensor - Data Analysis Using One-Sided Tolerance Intervals

CONCLUSIONS

The effect of structural aging and the dangerous combination of fatigue and corrosion has produced a greater emphasis on the application of sophisticated health monitoring systems. In addition, the costs associated with the increasing maintenance and surveillance needs of aging structures are rising. Corrective repairs initiated by early detection of structural damage are more cost effective since they reduce the need for subsequent major repairs and may avert a structural failure. Global SHM, achieved through the use of sensor networks, can be used to assess overall performance (or deviations from optimum performance) of large structures such as aircraft, bridges, pipelines, large vehicles, and buildings. The ease of monitoring an entire network of distributed sensors means that structural health assessments can occur more often, allowing operators to be even more vigilant with respect to flaw onset.

Through the use of in-situ CVM sensors, it is possible to quickly, routinely, and remotely monitor the integrity of a structure in service and detect incipient damage before catastrophic failures occur. These sensors can be attached to a structure in areas where crack growth is known to occur. On a pre-established engineering interval, a reading will be taken from an easily accessible point on the structure. Each time a reading is taken, the system performs a self-test. This inherent fail-safe property ensures the sensor is attached to the structure and working properly prior to any data acquisition.

This study showed the viability of using the One-Sided Tolerance Interval (OSTI) approach to determine the Probability of Detection for a fixed sensor detecting a crack which is propagating in a known direction in the vicinity of the sensor. The OSTI approach yields a reasonable estimate for the CVM crack detection capability even with small data sets. In several structural categories studied, the CVM sensors provided crack detection well before the crack propagated to the critical length determined by damage tolerance analyses. In addition, there were no false calls experienced in the fatigue crack detection tests. The sensitivity, reliability, and cost effectiveness of the CVM sensor system was demonstrated in both laboratory and field test environments.

This program is also establishing an optimum OEM-airline-regulator process and determining how to safely adopt SHM solutions. Close consultation with regulatory agencies is being used to produce a process that is acceptable to both the aviation industry and the FAA. The activities conducted in this program facilitate the evolution of an SHM certification process including the development of regulatory guidelines and advisory materials for the implementation of SHM systems via reliable certification programs. Formal SHM validation will allow the aviation industry to confidently make informed decisions about the proper utilization of SHM.

ACKNOWLEDGEMENTS

This work was sponsored by the FAA William J Hughes Technical Center under the program lead of Paul Swindell. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

REFERENCES

1. Roach, D., “Health Monitoring of Aircraft Structures Using Distributed Sensor Systems,” DoD/NASA/FAA Aging Aircraft Conference, March 2006.
2. Beral, B., Speckman, H., “Structural Health Monitoring for Aircraft Structures: A Challenge for System Developers and Aircraft Manufacturers,” 4th International Workshop on Structural Health Monitoring, September 2003.
3. Roach, D., “Use of Distributed Sensor Systems to Monitor Structural Integrity in Real-Time,” Quality, Reliability, and Maintenance in Engineering, Professional Engineering Publishing Ltd., Oxford, UK, 2004.
4. Roach, D, Kollgaard, J, and Emery, S., “Application and Certification of Comparative Vacuum Monitoring Sensors for In-Situ Crack Detection,” Air Transport Assoc. Nondestructive Testing Forum, October 2006.
5. Anderson, T.W. and Darling, D.A., "A Test of Goodness-of-Fit,". Journal of the American Statistical Association, Vol 49, Sept. 1954.