Recent Advances in Model-Assisted Probability of Detection

R. Bruce THOMPSON, L. J. BRASCHE,
Center for Nondestructive Evaluation, Ames, United States,
David FORSYTH, Texas Research Institute, Austin, United States,
E. LINDGREN, Air Force Research Laboratories, Dayton, United States,
P. SWINDELL, FAA William J. Hughes Research Center, Atlantic City, United States,
W. WINFREE, NASA Research Laboratories, Langley, United States

Abstract. The increased role played by probability of detection (POD) in structural integrity programs, combined with the significant time and cost associated with the purely empirical determination of POD, provides motivation for alternate means to estimate this important metric of nondestructive evaluation (NDE) techniques. One approach to make the process of POD estimation more efficient is to complement limited empirical experiments with information from physics-based models of the inspection process or controlled laboratory experiments. The Model-Assisted Probability of Detection (MAPOD) Working Group was formed by the Air Force Research Laboratory (AFRL), the Federal Aviation Administration (FAA) Technical Center, and National Aeronautics and Space Administration (NASA) to explore these possibilities. Since the 2004 inception of the MAPOD Working Group, 11 meetings have been held in conjunction with major NDE conferences. This paper will review the accomplishments of this group, which includes over 90 members from around the world. Included will be a discussion of strategies developed to combine physics-based and empirical understanding, draft protocols that have been developed to guide application of the strategies, and demonstrations that have been or are being carried out in a number of countries. The talk will conclude with a discussion of future directions, which will include documentation of benefits via case studies, development of formal protocols for engineering practice, as well as a number of specific technical issues.

1. Introduction

Probability of Detection (POD) [1-3] is a metric that is increasingly used to quantify the efficacy of an inspection in components designed and used in accordance with damage tolerant concepts [4]. In the previously developed safe life design practices as applied to metals, fatigue is treated as a nucleation process and there is no explicit consideration of the possibility for crack growth (failure is assumed when cracks are first formed). Damage tolerant design was developed to overcome some of the problems encountered with safe life design, such as premature failure of parts with unanticipated initial damage. It was assumed that structures contain cracks below some initially defined size and credit is given for the period of time that these cracks would take to grow to a critical size under the expected service conditions, i.e., the period of time in which cracks grow in a controlled fashion. The initially assumed crack size is related to inspection limits, i.e., the largest size flaw that the manufacturing inspection employed would not likely miss. POD is used to quantify this value, with the inspection limit often taken to be the flaw size for which the POD has a value of 90% with 95% confidence.
POD as a metric is also finding an increasing role in the quantification of in-service inspections. In one scenario, known as Retirement for Cause [4], a component is inspected after a fraction of the expected life (the time at which flaws of size equal to the inspection limit would be expected to grow to the critical size). If no defects are found, then the part can continue to be used for the additional time that it would take flaws of size equal to the inspection limit of the in-service inspection to grow to some conservative fraction of the critical size. Again, POD studies are used to quantify that inspection limit, not necessarily equal to that of the manufacturing inspection. In yet another mode, unanticipated durability problems associated with field use may require the design of new inspections. Again, POD studies are required to quantify their efficacy and the safe interval between such inspections.

In yet another example, one may wish to determine the POD in components which are simply too expensive to produce in the quantities required for POD studies. Again, an alternative is needed to the empirical approach.

As damage tolerant design techniques become more widely used, and as the structures that were designed in this way age, there becomes an increasing need for POD studies to quantify the efficacy of the resulting inspections. However, such studies can be quite expensive. POD is typically determined empirically, requiring the fabrication of costly specimens that replicate the physical situation of interest. Then inspections must be designed to possibly involve multiple operators, equipment and sites, with details depending on the circumstance. The overarching goal is to capture the effects of as many variables as possible that influence the inspection results, and hence the probability of detection for a given procedure. The associated cost in time and dollars can be large, often measuring in the hundreds of thousands of dollars, establishing the need for a more effective procedure.

In the same time frame that POD became an engineering metric critical to the management of structural integrity, the capability of computational models to simulate the results of inspections has increased dramatically, with industrial usage growing significantly [5]. These models can be used to quantify the effects of a number of variables on an inspection. The ability to theoretically predict the consequence of variables upon inspection results potentially reduces the need to capture those effects empirically, with an opportunity for reduced time and cost of POD studies [3, 6-7].

The Model-Assisted Probability of Detection (MAPOD) Working Group was established in 2004 by the Air Force Research Laboratory (AFRL) in cooperation with the Federal Aviation Administration (FAA) and the National Aeronautics and Space Agency (NASA), to explore those opportunities. The MAPOD Working Group has as its goal the promotion of the increased understanding, development and implementation of MAPOD methodologies. This is a voluntary activity in which working group members meet periodically in conjunction with an international meeting that many would be attending independent of this activity to

- Discuss strategies for model-assisted POD determination
- Discuss requirements for models to be used in POD studies
- Identify gaps that need to be addressed between state of the art models and real world problems
- Provide input regarding examples of specific problems that would demonstrate the utility of model-assisted POD activities
- Communicate the results of model-assisted POD demonstrations

Current membership includes over 90 individuals from around the world representing government, industry, national laboratories and academia.
To date, eleven meetings of the MA POD Working Group have taken place. Minutes, including copies of the PowerPoint slides presented, may be found on the internet [8].

A number of MA POD demonstrations have been undertaken under the support of both industrial organizations and government sponsors in a number of countries.

This paper provides a broad overview of the activities of the MAPOD Working Group and associated demonstrations.

2. Strategies and Procedures

It must first be emphasized that the MA POD strategies do not purport to make estimates of POD solely on the basis of physics-based simulations of the inspection. POD is a consequence of the variability of inspection results that may be controlled by many factors, including flaw morphology, operator, equipment and procedure variability, etc. Some of these are controlled by well understood physical phenomenon which can be described by physics-based models of the inspection process. Others, e.g., the effects of human variability, cannot be captured by such models and must be quantified by well designed, empirical experiments. The opportunity lies in developing the strategies that will allow the insights that can be gained from these two paths to be combined in a way that is both accurate and less costly in time and dollars than fully empirical approaches.

In the initial discussion of the MAPOD Working Group, two distinct approaches were proposed. The Transfer Function approach is illustrated in Figure 1 [8]. Suppose that one has determined a baseline POD curve, as shown on the left, in a fully empirical study. Then it might be possible to use physics-based models of the inspection process to transfer those results to another situation, e.g., a similar inspection on a part of a different material, a part with a different curvature, or the same part with a different flaw type (e.g., naturally occurring versus synthetic defects). As shown, the regression line and standard deviation about that line have changed. It would, of course, be possible that the slope of the line changes as well, not shown here. Such a transfer function could be determined using either physics-based models of the inspection or carefully controlled laboratory experiments.

The second approach that was identified was called the Full Model-Assisted approach, as illustrated in Figure 2 [8]. Here the factors that control the variability of an inspection are systematically identified. Physics-based models are used to predict the signal and noise distributions that are influenced by those factors that represent well understood physical phenomenon. This information is combined with empirical knowledge of the variability associated with other factors, e.g., those associated with
systems and operators. The results are combined to capture the total variability, which is used to predict the POD.

![Diagram of full model-assisted approach to POD](image)

**FIGURE 2.** Full Model-Assisted Approach to POD.

As thinking evolved, it became clear that these procedures were two special cases of a unified approach that is illustrated in Figure 3 [8-10]. The factors that control the variability of an experiment are partitioned into two sets, one whose effects must be evaluated empirically and one which is controlled by understandable physical factors. The effects of the latter can be assessed based on either controlled laboratory experiments or physics-based models of the inspection process. Draft protocols have been developed to implement this approach [8-9], which is also recognized in a draft military handbook [10].

![Diagram of unified approach to MAPOD](image)

**FIGURE 3.** Unified approach to MAPOD.

### 3. Demonstrations

MAPOD procedures have or are being used in a number of demonstrations [8-9]. Included are the following, with citations to where technical details may be found.

- Ultrasonic Detection of Flat Bottom Holes in Engine Disks of Different Alloys [12,13]
- Capability of Advanced Eddy Current Techniques to Detect Fatigue Cracks in Wing Lap Joints [14]
- Generic Bolt Hole Eddy Current Testing [15]
- Ultrasonics Inspection of Lower Wing Skin Fastener Holes [16, 17]
- Ultrasonic Detection of Cracks in Cold Worked Holes [18]
- Eddy Current Detection of Cracks in Airframe Fastener Holes [19]
- Eddy Current Detection of Cracks in Engine Bolt Holes [19]

As one example, Figure 4 shows a comparison of the prediction of the POD for Flat-Bottom Holes in an engine disk for the empirical and model-assisted approaches [12, 13]. Excellent agreement is seen.

![Figure 4. Comparison of fully empirical and Model-Assisted POD curves.](image)

### 4. Guide to Extrapolation

Empirical determination of POD always is based on flaws in a finite size range. However, extrapolations are often made to smaller and larger sizes, since a POD curve is generally desired for all sizes. Extrapolation of the regression fit to the available data may not be valid. In a recent study of the POD of the ultrasonic detection of hard-alpha inclusions in aircraft engine rotating components, such a case was encountered. Figure 5 shows an $\hat{a}$ versus $a$ analysis based on field finds for such an inspection [20]. Because of the complex morphology of these naturally occurring defects, the slope of the regression fit (the right segment of the solid line) is very small. The implication of extrapolating the regression line to small flaw sizes is that much smaller flaws would also produce significant signals, inconsistent with the theory of Rayleigh scattering. That theory has been used to replace the regression line at smaller flaw sizes, leading to an abrupt change in the slope of the $\hat{a}$ versus $a$ analysis. The fact no naturally occurring flaw were present in the field find data for flaw sizes below this change in slope is consistent with this analysis.
5. Quantifying Accuracy

When a POD is determined, the user needs to have a sense of the accuracy of that prediction. In empirical techniques, this is captured by the statistical confidence intervals (sometimes called the uncertainty intervals), which are related to the number of experiments that are performed. The essence of POD determination is the quantification of the distributions that describe the variability. Different data sets (sampling the same distribution) will lead to different estimates of POD and the confidence interval quantifies the uncertainties in those estimates. As is well known, the more times one samples a distribution, the more accurately one can estimate its properties, e.g., the mean and standard deviations that control the ultimate POD curve. One can thus shrink the statistical confidence intervals to zero by increasing sample size (ignoring the practical issues of time and cost associated with such an approach).

On the other hand, when using a model, one can generate a large number of data points [21]. Therefore, statistical uncertainty, as traditionally measured by confidence bounds, can be driven to zero. However, uncertainty in model predictions will affect the accuracy of predictions of POD. As an example, in a program aimed at the use of models to determine the POD of ultrasonic detection of defects in aircraft engine billets [7], the ultrasonic simulation models were taken to be accurate to ±3 dB, believed to be on the order of the reproducibility of typical ultrasonic experiments.

In an application of the unified approach to MAPOD, one needs to consider both limits to accuracy, statistical uncertainty and uncertainty in modeling predictions to make an overall assessment of the accuracy of the POD predictions. Hence the accuracy required of the model will be a consequence of the particulars of the problem and the accuracy needed in the POD. Depending on the application, one might want a highly precise prediction for life management studies or an order of magnitude estimate for initial design studies.

6. Current Status and Future Directions

When the MAPOD Working Group was formed in 2004, it established metrics for its success that the MAPOD Working Group activities would lead to

- Draft protocols for model-assisted POD
- Draft requirements for model qualification for use in POD determination
- Model-assisted POD demonstrations
All three metrics have been met as discussed above. Hence, the MAPOD Working Group has concluded that its initial metrics have been satisfied, the viability of the MAPOD approach has been established, and its first phase has been completed.

A second phase of activities has been initiated, aimed at implementation issues. Included are the documentation of benefits via case studies and the development of formal protocols for engineering practice. In addition, there are a number of more detailed technical issues that require further maturation of understanding.

The majority of initial applications have been aimed at aerospace problems, but interest in growing elsewhere, such as in the nuclear power industry [22].

7. References

[12] Smith, K., Thompson, R. B., and Brasche, L., “Model-Based POD: Successes and Opportunities,” presented to the MAPOD Working Group at the September 24-25, 2004 meeting in Albuquerque, New Mexico. Slides may be found at the web site identified in Ref. 8.


[21] Thompson, R. B. and Meeker, W. Q., “Confidence Bounds on Model-Based Data,” presented to the MAPOD Working Group at the November 16, 2007 meeting in Las Vegas, Nevada. Slides may be found at the web site identified in Ref. 8.