QUALIFICATION AND VALIDATION OF THE PERFORMANCE CAPABILITY (POD) FOR NONDESTRUCTIVE INSPECTION PROCEDURES

Ward D. Rummel, Littleton, CO, United States

Abstract: This paper focuses on practical experience with the development and validation of the performance of nondestructive inspection procedures and automated nondestructive inspection systems as applied in the aerospace industry. Discussion includes automation of fluorescent penetrant, eddy current, ultrasonic and X-radiographic inspection systems; the logic, design of experiment and data required for performance validation; and the results of validation and implementation in industry.

Introduction: The introduction of probability of detection (POD) as a method of assessing the capability of nondestructive testing (NDT) procedures challenged some long held beliefs, myths and exaggerations concerning the output and confidence in NDT applications [1]. My first presentation of results using the POD method (in the early 1970's) was met with considerable disbelief [2,3] . There is still considerable reluctance to quantify NDT capabilities and both successes and shortfalls are frequently attributed to "Human Factors". The capability of an NDT procedure may be summarized as dependent on the following:

- **Flaw (Artifact) Variables**
- **Test Object Variables**
- **NDT Method Variables**
- **NDT Materials Variables**
- **NDT Equipment Variables**
- **NDT Procedure Variables**
- **NDT Process Variables**
- **Calibration Variables**
- **Acceptance Criteria / Decision Variables**
- **Human Factors**

Note that "Human factors is listed as the last item in influencing the preceeding variables are not controlled, the human operator at the end of the line has little chance.

Human factors remain and continue to be cited as the cause of a failure to detect. The widely held solution to removing human factors and thus improving detection capability and reliability, is automation. However, automation without attention to all application variables will not improve detection or reliability and can be a mode of applying a varying (uncontrolled) inspection at an increased speed level.

In applying POD assessments to various NDT procedures, we have learned much. The desired output from the applicable of an NDT procedure is in reliable detection. Detection RELIABILITY, in turn, is dependent on the:

- Capability - POD
- Reproducibility -"Calibration"
- Repeatability - Process Control

Process control is required in all NDT procedures applications and is a dominant part of non-instrumental NDT procedures such as X-radiography, liquid penetrant, magnetic particle and visual inspections. For instrumental NDT procedures, the dominant variable is in "calibration", but attention to process control remains as an important consideration.. Automation of an NDT procedure requires consideration attention to all variables, with special attention to "calibration".

For NDT methods that do not provide a scalar output (non- instrumental), characterization of process parameters and the interaction of process parameters may require several assessments and is typical of that required for any multiparameter process. POD provides a tool for NDT process parameters characterization and is applied by assessment of the effects of single parameter variances while holding all other parameters constant. Reproducibility and confidence levels in final process control depends on the number of representative independent tests performed to characterize process parameters and variances.

In like manner, process characterization and control required for instrumental (quantitative) NDT procedures are required. The primary difference is that output is quantified directly and may be directly related to a baseline “calibration”. Reproducibility is primarily dependent on the fidelity of the “calibration” to assure that NDT system response is replicated. Attention must be addressed to process application parameters to assure that the “calibration” fits each secific application. For example, response when the transducer (energy field) is large with respect to the flaw may have different slope than that where the transducer (energy field) is small with respect to the flaw.

Results: The recognized method of POD data analysis is the Beren’s model [4] that relates inspection signal output to flaw size. For most applications, a log linear response relationship is produced as shown in Figure 1. A decision threshold is applied to the signal produced to discriminate between signal and noise (noise is response produced from non flawed areas). In order to reproduce the inspection capability in subsequent applications, it is necessary to reproduce the log linear response relationship that was used in validation of the procedure and to establish the POD capability. The POD curve derived from the Beren’s model requires generation of a flaw size / flaw response relationship (by both “Hit / Miss” analysis and a versus a-hat methods) and is depends on reproduction of that relationship for each application of the validated NDT procedure (Figure 2.) .

A common practice and “calibration” method that is widely prescribed in specifications, standards and codes is the use of a single notch, or artifact, to set-up an instrument. A single notch does not provide either reproduction of the log linear response relationship used for procedure validation (POD) or confidence in reproducible function of the NDT measurement system as shown in Figure 3. A revolution in NDT “calibration” practice is required to improve the reliability of both manual and automated NDT applications.

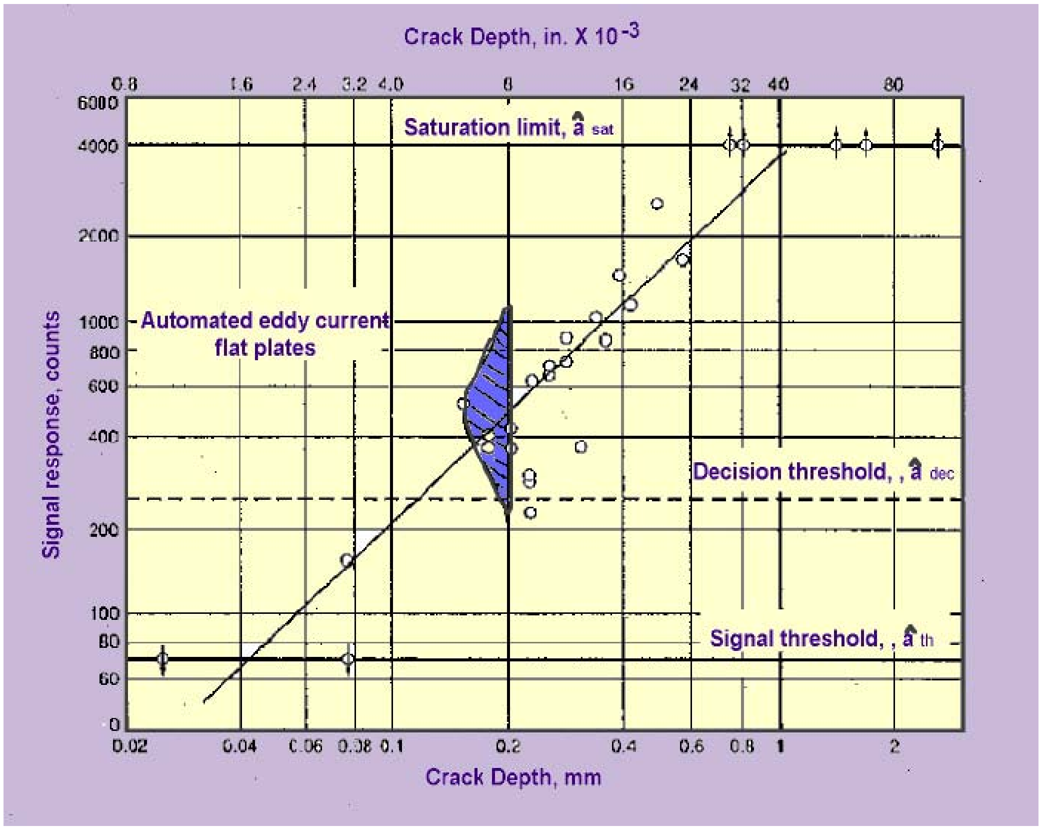

Figure 1. The Beren’s Model [4]

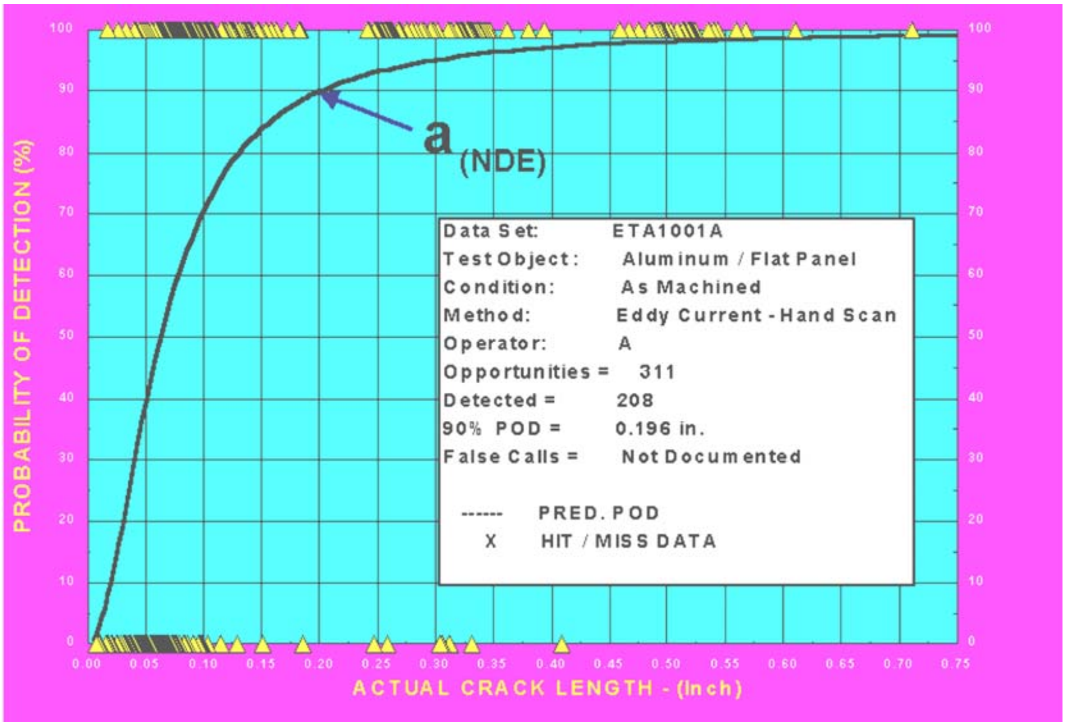

Figure 2. Typical Probability of Detection (POD) Output

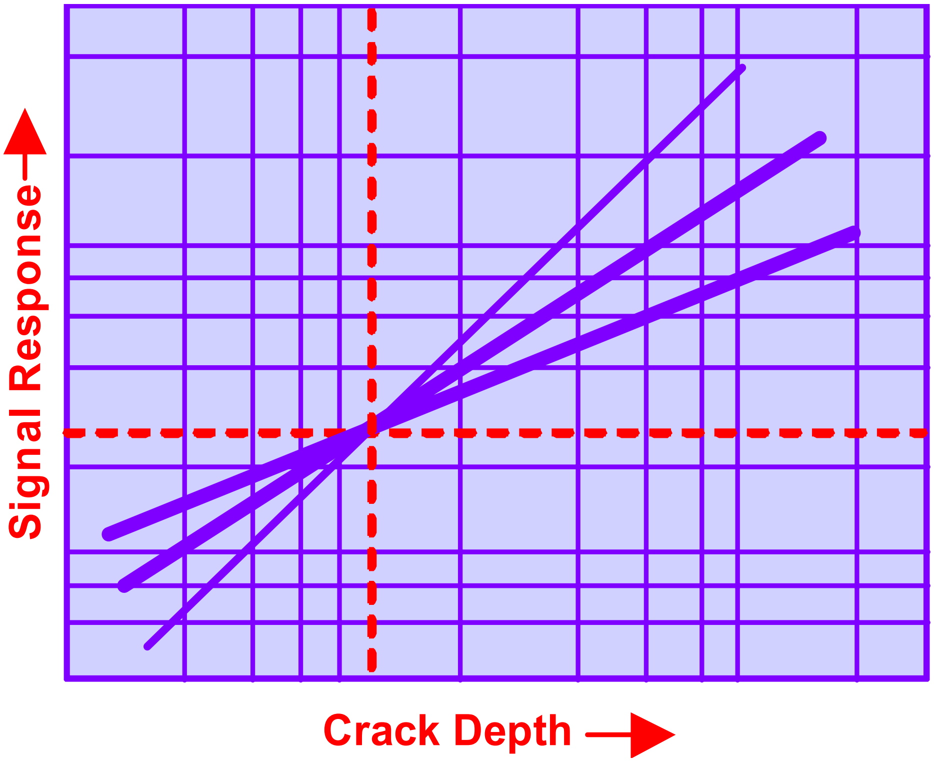

Figure 3 Same “Calibration , Different PODs

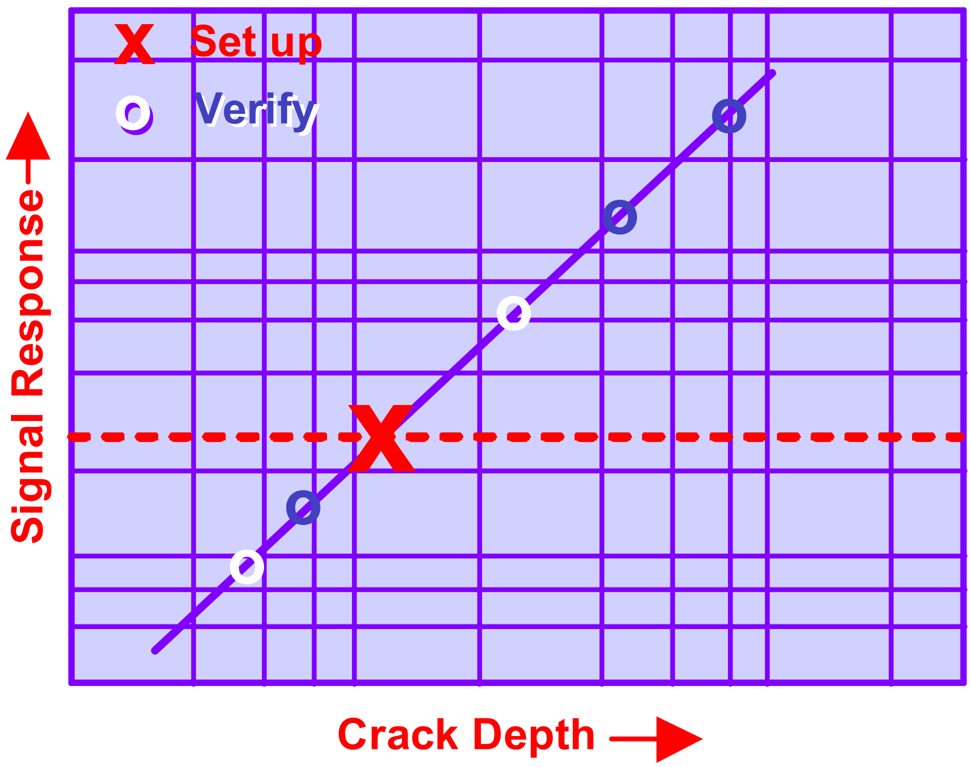

Figure 4. Multiple – Point Calibration

Discussion: The single point “calibration**”** is nominally valid for flaws that have an identical response to that from the single notch, but loses validity as the response slope approaches the horizontal condition and may not reproduce the validation (POD) response at any other flaw size. Changes in slope may be due to probe variables, cable variable, location / scanning variable and / or part surface condition and material properties variables. It is important to note that we are addressing total NDT system response reproducibility.

Reproduction of system response may be approached using a multiple point “calibration”. A multiple point “calibration” is shown schematically in Figure 4. Such a relationship is established after the NDT system and procedure are stable and ready for validation. A single notch is used to set-up the instrument in the established manner (Shown as an “X” symbol). System response measurements are then made using multiple notches of a size that is smaller than and larger then the set-up notch and in the size range of flaws that the NDT procedure is expected to detect. (Note that notches that produce responses above the saturation level and below the noise level cannot be used). Responses from the verification notches are shown as “O” symbols in Figure 4. This is the system condition that will be used to validate the detection capability (POD) of the NDT procedure.

Since variance is a part of the measurement process, multiple measurements at the “X” and “O” points must be made during procedure validation to establish an allowable variance. The boundaries established are then used as a part of the multipoint “calibration” procedure (See Figure 5). In like manner, the signal / noise relationship must be verified to assure that the response from the “calibration” artifacts is similar to that in the test article.

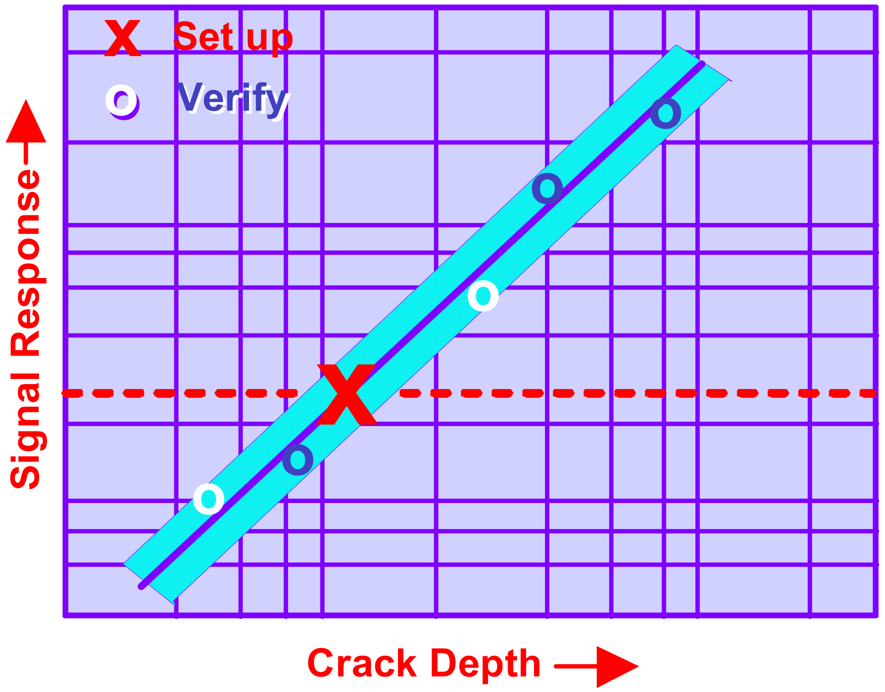

Figure 5. allowance for Measurement Variance

The multipoint "calibration" method has been shown to provide significant improvement in reproducibility in aircraft maintenance field applications and in reproduction of baseline (POD) capabilities following system upgrades in automated systems [5]. The method provides traceability to the POD capability demonstrated during procedure validation. The multipoint "calibration" improvement is based on experience in quantitative NDT applications using the POD method and principles of assessment and on long established metrology principles and practices. The result is significant improvements in reproducibility (and hence reliability) and increased confidence in NDT applications and NDT technology. To the NDT community, it may be viewed as a significant step in moving NDE from an art to a disciplined science.

Good metrology practices require traceability to the artifacts that were used in procedure validation. The artifacts used for procedure validation may be used for all subsequent application of the procedure or may be used as a "Master Gage" for relating to responses from secondary artifacts (with offsets) to enable reproduction of original system response values at various inspection sites and using multiple inspection systems. The secondary "standards" are then used at various inspection stations, at field locations, at subcontractors sites, etc. where quantification of capabilities and treaceability to POD validation are required. The "Master Gage" and secondary standards discipline becomes increasingly important for NDE procedures that are operated near their detection threshold and for procedures involving detection of small flaws. Unless "Master Gage" traceability is used, a separate POD validation is required for each machine or site application.

Conclusions: Development and application of probability of detection (POD) capabilities assessment methods have increased NDT engineering understanding of NDT applications and expected results. POD assessment provides a useful method for quantifying detection capabilities and for NDE procedure characterization, optimisation and validation. For non-instrumental NDT procedures, it provides a quantitative metric for characterization of process parameter variances in terms of desired process output. Quantification of NDT capabilities has been used for assessment of multiple contributing parameters including the effects of equipment, transducers, process control and human factors variations.

In like manner, a multiple point "calibration" not only improves reproducibility for manually scanning operations, but aids in reducing human factors variance. In multiple studies of the effect of "Human Factors" on NDT detection capabilities, the single most important variable identified has been "recency of experience" with the inspection procedure [6]. This is similar to that found and periodically reinforced in airline pilot skills and capabilities. NDT trainers would aid in improving the "Human Factors" part of inspection capability and reliability.

The added benefit of multiple point "calibration" is in positive feedback and reinforcement of operator skills in reproducing the system response that was established during procedure validation.

Thirty years of progress in NDT technology and in POD application have provided increasing NDT engineering knowledge and knowledge of contributing application parameters including the effects of equipment, transducers, process control and human factors variations. A revolution in NDT system "calibration" is proposed based on that experience and on demonstrated improvements using multiple point "calibration" procedures. This proposal challenges many existing specifications, standards, codes, etc. and it is anticipated that considerable effort and time must be allowed to effect the proposed and demonstrated improvements. Presentation at this "World Forum" is considered to be an important step in effecting improvements. The importance of multipoint "calibration" and system response reproducibility is greatly increased in automation of NDT systems. It is not difficult to build such capability into a system, but is difficult to inject an improvement into a system that has been implemented in a production mode.

References:

1. NDE Capabilities Data Book, 3rd Edition, DB-2, 1997, available through NTIAC, (512) 263-2106.
2. Rummel, Ward D., Paul H. Todd Jr., Sandor A, Frecska and Richard A. Rathke, "The Detection of Fatigue Cracks by Nondestructive Testing Methods", Paper presented to the Spring Conference, American Society for Nondestructive Testing, Los Angeles, CA, March 1972**.**
3. Rummel, Ward D. and Richard A. Rathke, "Detection and Measurement of Fatigue Cracks in Aluminum Alloy Sheet by Nondestructive Evalation Techniquies", Prevention of Structural Failure, edited by Thomas D. Cooper, Paul F. Packman and B.G. W. Yee, American Society for Metals, 1975.
4. Berens, A.P. and Hovey, P.W. (1984), "Flaw Detection Reliability Criteria, Volume I – Methods and Results," AFWAL-TR-84-4022, Air Force Wright Aeronautical Laboratories, Wright-Patterson Air Force Base, Ohio, April, 1984.
5. Private communication.
6. Rummel, Ward D., Steven J. Mullen, Brent K. Christner, Frank B. Ross and Robert E. Muthart, Reliability of Nondestructive Inspection (NDI) of Aircraft Engine Components, SA-ALC/MM-8151, January 1984.